# South African Journal *of* Science

volume 119
*number 9/10*


Measuring physical activity in South African children

Discussions on load shedding

Economic impact of SA's public universities

ASSAf
ACADEMY OF SCIENCE OF SOUTH AFRICA

Volume 119
Number 9/10
September/October 2023



# South African Journal of Science

eISSN: 1996-7489

**Published by**
the Academy of Science of South Africa (www.assaf.org.za) with financial assistance from the Department of Science & Innovation.

**Design and layout**
diacriTech Technologies



## Review Article



## Research Article

**Correspondence and enquiries**
sajs@assaf.org.za



**Submissions**
Submissions should be made at www.sajs.co.za



**Cover caption**
South African children playing. Physical activity has positive effects on cardiovascular health that can be observed already in childhood. In their Research Article, Arnaiz and colleagues compare two methods to assess physical activity levels in children in Gqeberha, South Africa. And in their Commentary, Breet and colleagues reveal that South African children are wearing school shoes that do not properly fit as they are not designed to fit the unique foot shape of habitually barefoot South African children.

# Responding to load shedding

There can be no scientist based in South Africa who has not felt impacted by load shedding in a variety of ways. For many of us, as is the case with others, the disruptions have affected our home lives, our travel to and from work, and our leisure activities, as Marchetti-Mercer notes in a Commentary in this issue. There are also distinctive challenges for scientists – load shedding affects all university activities, the functioning of laboratories and field trips, and the viability of storage of samples, to name a few.

In keeping with the vision and mission of our Journal, we have a responsibility to highlight issues of broad concern, not just across academic disciplines but also for the public at large, and for policymakers. The problem of load shedding in many ways encapsulates both the potential and the challenges implicit in the Journal's commitment to multidisciplinarity and to being a forum through which scholars from a range of disciplines can interact, debate, learn, and develop strategies to address problems of common concern. The series of commentaries featured in this issue is by no means comprehensive, and is very far from the last word on an issue for which, as in some definitions of 'wicked problems'[1], may not ultimately be solvable to the satisfaction of all. Nevertheless, we are pleased that our call for commentaries yielded responses from a range of disciplines and perspectives. Load shedding is a technical problem to which there may be a range of technical responses and solutions, an opportunity for those with resources to work around, but with probable negative knock-on effects, a problem with complex and contested socio-economic antecedents and costs, a challenge to important climate-related research and action, and an opportunity for broader methodological reflection on issues of health research and monitoring. This list is definitely not exhaustive, and we hope to receive more submissions on load shedding in the future. It is an issue which brings together the range of disciplines represented in this Journal, from politics to engineering, and from sociology to climatology. There is no one understanding and experience of the problem, and there is no single way forward. Load shedding also requires scientists to engage with a range of groups and publics outside the academic community, and, in this way, it becomes an issue of science communication and engagement as well.

In the weeks leading to the writing of this Leader, it has become increasingly clear (not that it was not before), that load shedding is by no means the only problem demanding complex thinking and solutions. There is the omnipresent existential threat of climate change and environmental breakdown, well covered by many contributions to the Journal, and in *Rock|Water|Life: Ecology and Humanities for a Decolonial South Africa* by Lesley Green, the 2023 ASSAf Humanities Book Award winning book, which was reviewed in a special issue of the Journal late in 2022. There have been gas explosions in the streets of Johannesburg, and the horrific deaths of over 70 people in a fire in the same city. Many of those who died were migrants living precarious lives before the tragedy itself, vulnerable to criminal exploitation, and living, like so many others, at the wrong end of a vastly unequal society, where terms like 'housing shortage' and 'crumbling infrastructure' take on an embodied meaning much more consequential than for those of us who are more privileged. One of the ironies of load shedding is that the inconvenience and loss to productivity it has caused has given some (albeit much more limited) experiential weight to the most privileged in our society of daily, ongoing, problems experienced by poor and marginalised communities. The question of access to clean, reliable energy, just like that of access to clean water and other resources, is nothing new for many in our country and on our continent. And the threats to our existence, as we have seen recently, have an immediacy for those people who come to the attention of the more privileged only sporadically, commonly when there are headlines showcasing acute events rather than the daily struggles of life on the margins.

In the most recent Leader in this Journal, we discussed issues of unequal access to developing scientists and contributors to researching and debating solutions to complex problems. Those of us who have the privilege to be scientists and researchers should not be constrained in our work by political interference or a narrow understanding of what kinds of thinking may help make things better in our country and our world. But this relative freedom, which goes along with being protected from much of the hardship experienced by many in our country and world, must be balanced by a commitment to using our privilege to help find solutions for all. It is a good thing, as our commentaries on load shedding exemplify, that there are many ways to address difficult problems, and that there are debates and contestations. As a Journal, we hope to continue to be a forum for those debates. We believe that contestation is key to opening new ways of thinking and addressing challenges.

## Reference

1. Lönngren J, Van Poeck K. Wicked problems: A mapping review of the literature. Int J Sustain Dev World Ecol. 2021;28(6):481–502. https://doi.org/10.1080/13504509.2020.1859415

**HOW TO CITE:**
Swartz L. Responding to load shedding. S Afr J Sci. 2023;119(9/10), Art. #16836. https://doi.org/10.17159/sajs.2023/16836

**AUTHOR:**
Bruce Rubidge[1]

**AFFILIATION:**
[1]Evolutionary Studies Institute, University of the Witwatersrand, Johannesburg, South Africa

**CORRESPONDENCE TO:**
Bruce Rubidge

**EMAIL:**
Bruce.Rubidge@wits.ac.za

**HOW TO CITE:**
Rubidge B. Charles Kimberlin (Bob) Brain (1931–2023): Naturalist, scientific leader, and family man. S Afr J Sci. 2023;119(9/10), Art. #16660. https://doi.org/10.17159/sajs.2023/16660

**ARTICLE INCLUDES:**
☐ Peer review
☐ Supplementary material

**PUBLISHED:**
28 September 2023



# Charles Kimberlin (Bob) Brain (1931–2023): Naturalist, scientific leader, and family man

Bob Brain's career as a naturalist was almost predestined, as he was born into a family of scientists with an entomologist father and botanist mother. He devoted his life to science, and his multidisciplinary scientific pursuits gained international acclaim. Bob was born on 7 May 1931 in what was then Salisbury, Rhodesia, and matriculated at Pretoria Boy's High School at the age of 16 in 1947. He then proceeded to the University of Cape Town where he obtained his BSc degree in Zoology and Geology, and later a PhD as well.

While still a scholar at Pretoria Boy's High School, Bob began what was to become a lifelong association with the Transvaal Museum when he went there to learn how to mount birds.[1] Apart from a short spell as a geologist with the National Building Research Institute at the CSIR, he had a long and distinguished career at the Transvaal Museum. He started in the Palaeontology Department where he was employed on contract as a research associate from 1954 to 1957, with a grant from the Wenner-Gren Foundation for Anthropological Research in New York.[2] Here, at the suggestion of John Robinson, then curator of palaeontology at the Museum, he undertook the first systematic investigation into the stratigraphy of the South African fossil hominid-bearing cave deposits. Here he established that each deposit was of a different age and reflected a different climatic regime. An important consequence of this research at both Sterkfontein and Makapansgat caves, was the discovery, for the first time, of stone artifacts associated with *Australopithecus.*[3] As his contract in the Palaeontology Department had come to an end, Bob was appointed curator of Lower Vertebrates in 1957, the only position then available at the Transvaal Museum, during which time he published on frogs, snakes and lizards. This position was previously occupied by Dr Vivian FitzSimons, the noted South African herpetologist who had taken up the directorship of the Museum. Bob worked closely with FitzSimons in completing the landmark book *The Snakes of South Africa.* A highlight of this period was demonstrating that species-specific behaviour patterns in reptiles could be used as taxonomic criteria in the same way as morphological features, and he applied this pioneering approach to chameleons and barking geckos.

In the early 1960s three new museums were being built in Zimbabwe (then Rhodesia) and Bob participated in this exciting development by simultaneously taking up the positions of Keeper of Zoology and Deputy Director of the Queen Victoria Museum in Harare from 1961 to 1964. Here he undertook pioneering comparative behavioural research on monkeys and was also responsible for the spectacular zoological display at this museum, which remains today. He returned to the Transvaal Museum in 1965 as Curator of the Palaeontology Department, and, 3 years later was appointed Director, remaining in this position for 23 years until his retirement in 1991.[4] Here he planned and co-ordinated three major display halls and, through his own productive research example, established the Transvaal Museum as an internationally renowned research institution.

With a remarkable multidisciplinary approach, Bob undertook groundbreaking research at the Swartkrans fossil hominid cave over a 28-year period. He began by first sorting through the dumps left behind by the earlier lime miners over a period of 7 years and later continued with excavations which resulted in the cataloguing of at least 170 fossil hominid specimens, and 4800 faunal fossils, stone and bone artefacts. His broadly focused taphonomic study resulted in numerous publications and culminated in the production of two books: *The Hunters or the Hunted? An introduction to African Cave Taphonomy* and later *Swartkrans: A Cave's Chronicle of Early Man.* A dramatic finding from Swartkrans was the oldest evidence for the controlled use of fire by hominids dating to about one million years ago which was presented in a paper published in *Nature.*[5]

To determine the mode of accumulation of the Swartkrans bone assemblage, Bob collected and studied bony food remains scattered around Nama settlements in Namibia to gain an understanding of the effects of human behaviour on bone accumulations. Coupled with this he researched the feeding behaviour of different carnivores, especially cheetahs which were kept on a farm in Namibia. This work demonstrated that bone assemblages from Makapansgat that had originally been attributed by Raymond Dart to early hominid activity[6], possibly had other explanations and were probably the result of carnivore activity. These innovative ideas led to the new discipline of African cave taphonomy, which enables the reconstruction of early hominid behaviour and palaeoenvironments. In this field Bob was an international leader.

From his palaeontological and geological research at Swartkrans, Bob was further able to identify cycles of deposition within the Quaternary period and linked habitat changes to global variability in temperature which he correlated with hominid evolutionary events in Africa.[7] For the results of this multifaceted research, he was awarded a DSc by the University of the Witwatersrand in 1981.

Bob Brain was one of the founders of the Namib Desert Research Station at Gobabeb in 1959, together with Bernard Carp, Vivian FitzSimons and Charles Kock. Later, as Director of the Transvaal Museum, he was responsible for guiding Gobabeb to establish itself as a research station of international repute for work on desert ecology.[8]

Apart from his busy research and administrative schedule, Bob was also a dedicated museum educationalist, greatly expanding the education department of the Museum under the guidance of O.P.M. Prozesky. He also planned and co-ordinated the completion of three display halls which are still the backbone of the museum exhibits: Austin Roberts Bird Hall, Life's Genesis I and Life's Genesis II. In the latter two exhibits, he successfully experimented with a narrative concept leading visitors through a succession of displays depicting the development of life through the ages so that the visitor experienced a unified storyline rather than a set of disjointed displays.[9,10]

An inseparable partnership – Bob and Laura at their home in Irene, Pretoria

After retiring as Director in 1991, Bob shifted his research focus and was appointed Curator of Lower Invertebrates at the Museum. Initially he worked on the taxonomy of rotifers which led him to consider the origins of multicellularity and also predation as a mediator in the evolution of sense organs and intelligence in both prey and predators.[11] In the process, he undertook pioneering fieldwork which led to the discovery of a new fauna of early metazoan fossils from the Late Precambrian Otavi Group of Namibia.[12]

As a gifted public speaker, Bob was invited to present at many prestigious lectures: Alex du Toit Memorial Lecture (Geological Society of South Africa); Sidney Haughton Memorial Lecture (Royal Society of South Africa); Schonland Memorial Lecture (Royal Society of South Africa); Meester Memorial Lecture (Kaffrarian Museum); Sidney Rubidge Memorial Lecture (Graff-Reinet Training College); 70th James Arthur Lecture (American Museum of Natural History); Robert Broom Memorial Lecture (Ditsong Museum).

Bob belonged to numerous scientific organisations which he served with distinction: Royal Society of South Africa (Fellow); Museums Association of Southern Africa (Fellow; Life Member; President); Academy of Science of South Africa (Founder Member); Archaeological Society of South Africa (Founder Member; Councillor; President) Transvaal Branch (Patron); Zoological Society of Southern Africa (Founder Member; Councillor; President); South African Biological Society (Founder Member; Councillor; President); South African Association for Quaternary Research (Founder Member; Councillor; President); Palaeontological Society of Southern Africa (Founder Member; Councillor; first President, Honorary Life Member); South African Association for the Advancement of Science (Founder Member; Councillor; President); Royal Society of South Africa (Councillor; Vice-president); Speleological Association of South Africa; (Honorary Life Member); Zoological Society of London; Southern African Society of Aquatic Scientists (Member); Anatomical Society of Southern Africa (Member); Microscopy Society of Southern Africa (Member); Queckelt Microscopical Club, London (Member); International Society of Protozoologists; International Society of Cryptozoology (Member); Spider Club of South Africa (Honorary Member).

Bob received many awards, including honorary DSc degrees from the Universities of Cape Town, Natal, Pretoria, and Witwatersrand; Gold medal of the Zoological Society; Honorary membership of the Palaeontological Society of southern Africa and the Geological Society of South Africa; Senior Captain Scott Memorial Medal of the South African Biological Society; Achievement Award of the Claude Harris Leon Foundation; John FW Hershel Medal of the Royal Society of South Africa; one of the four outstanding young South Africans; and the South African Medal of the South African Association for the Advancement of Science for exceptional contributions to science (1997); Certificate of Merit from the Southern African Association for the Advancement of Science (2002); A-rated Researcher by the FRD (NRF) from inception in 1984 to 1997.

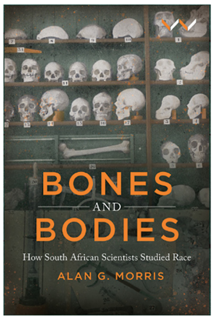
As is evidenced by more than 160 publications and several books, Bob Brain had an extraordinarily diverse and productive research career in which he was loyally supported by his wife Laura, and their four children. All research projects undertaken by him were imaginative and innovative, and produced significant results with universal applicability which have withstood the test of time. He inspired the younger generation to undertake research and supervised the theses of at least 18 local and international master's and doctoral students.

Bob's scientific pursuits have brought great honour to South Africa. While Director of the Transvaal Museum, he established a happy environment with staff who were passionate about their work. The institution gained an international reputation for its diverse natural history research endeavours, achieved through allowing staff considerable personal freedom, provided productivity was maintained. Bob stressed the importance of fun in research. He maintained that if this was in place, creativity and productivity would follow naturally. He constantly warned against the danger of obsession with performance to the detriment of creativity. Bob was naturally warm-hearted, and the welcoming Brain household was an extremely happy one with much fun and banter which Bob once described as "children inter-stratified with rocks and fossils"[1]. His wife Laura and their children Rosemary (Mel), Virginia (Ginni), Tim and Conrad (Nad) were all accommodated in the "Brain family research endeavour" from which they all derived great enjoyment and personal fulfilment.

## References

1. Helme N, editor. Irene personalities. Pretoria: Gutenberg Book Printers; 1982.
2. Rubidge BS. Charles Kimberlin (Bob) Brain – A tribute. Palaeontol Afr. 2000;36:1–9.
3. Brain CK, Van Riet Lowe C, Dart RA. Kafuan stone implements in the post australopithecine breccia at Makapansgat. Nature. 1955;175:16–20. https://doi.org/10.1038/175016a0
4. Brain CK. The Transvaal Museum 1968–1991: A place of quiet scholarship. Suppl Bull Tvl Mus. 1991;22:1–8.
5. Brain CK, Sillen A. Evidence from the Swartkrans Cave for the earliest use of fire. Nature. 1988;336:464–466. https://doi.org/10.1038/336464a0
6. Ardrey R. The territorial imperative. London: Collins Clear-Type Press; 1967.
7. Brain CK. The evolution of man in Africa: Was it a consequence of Cainozoic cooling? Alex du Toit Memorial Lecture 17. Proc Geol Soc S Afr. 1981;LXXXIV:1–19.
8. Seely MK, editor. Namib ecology: 25 years of Namib research. Tvl Mus Monogr. 1990;7:5–12.
9. Brain CK. Life's genesis: A display concept in museum display. South African Museums Association Bulletin. 1976;12:137–145.
10. Brain CK. The narrative concept in museum display at the Transvaal Museum. South African Museums Association Bulletin. 1976;13:90–95.
11. Brain CK. The importance of predation in the course of human and other animal evolution. S Afr Archaeol Soc Bull. 1995;50:93–97. https://doi.org/10.2307/3889057
12. Brain CK, Hoffman K-K, Prave AR, Fallick AE, Coetzee J, Botha AJ. Interpretive problems in the search for micro-invertebrate fossils from a Neoproterozoic limestone in Namibia. Palaeontol Afr. 2001;37:1–12.

**BOOK TITLE:**
Bones and bodies: How South African scientists studied race




**AUTHOR:**
Alan G. Morris

**ISBN:**
9781776147236 (paperback, 304pp)
9781776147243 (hardback, 304pp)

**PUBLISHER:**
Wits University Press, Johannesburg; ZAR375

**PUBLISHED:**
2022

**REVIEWER:**
Himla Soodyall[1]

**AFFILIATION:**
[1]Academy of Science of South Africa, Pretoria, South Africa

**EMAIL:**
himla@assaf.org.za

**HOW TO CITE:**
Soodyall H. Physical anthropology through the eyes of the scientists. S Afr J Sci. 2023;119(9/10), Art. #16804. https://doi.org/10.17159/sajs.2023/16804

**ARTICLE INCLUDES:**
☐ Peer review
☐ Supplementary material

**PUBLISHED:**
28 September 2023



# Physical anthropology through the eyes of the scientists

In *Bones and Bodies*, Alan Morris brings to life the historic narrative of how "the skeletons I study were once the frames of living people" (p. 2) and how "biological information can have no value without understanding the social and cultural context of the people" (p. 2). Morris ploughed through collections of correspondence between early researchers in the field and volumes of literature, and conducted interviews with seminal figures in the fields of physical anthropology, archaeology and palaeontology, among others, to eloquently narrate the contribution of scholars who shaped the physical anthropology landscape in South Africa.

Morris takes us on a deep dive covering the field of physical anthropology with insights into the people who collected, examined, and described skeletal remains, sharing the stories not usually found in scientific publications. One gets to experience science from the early 1900s with narratives on how science at that time set the foundation for examining the relationships of early inhabitants at the Cape, albeit through the eyes of mainly foreign scientists. Morris does this by including eight anthropological vignettes: Dr Louis Péringuey's well-travelled skeletons (Chapter 1), Boskop: The first South African fossil human celebrity (Chapter 2), Matthew Drennan and the Scottish influence in Cape Town (Chapter 3), The age of racial typology in South Africa (Chapter 4), Raymond Dart's complicated legacy (Chapter 5), Ronald Singer, Phillip Tobias and the 'new physical anthropology' (Chapter 6), Physical anthropology and the administration of apartheid (Chapter 7), and The politics of racial classification in modern South Africa (Chapter 8). Throughout these narratives we see how typology influenced classification of indigenous people in southern Africa, and how terminology used to refer to South African people changed over time.

The South African Association for the Advancement of Science ($S_2A_3$) meeting hosted in Cape Town in 1905 brought many foreign scientists to South Africa, which afforded scholars with an interest in the living people of southern Africa opportunities to collaborate with South African scientists. At the time there was already a growing body of anthropometrical data, comparative physiology, and psychology of the local people. There was also a keen interest in the link between behaviour and race. According to Morris, the 1905 meeting "had a special significance for developing the field of physical anthropology" (p. 18) in South Africa.

Starting with Dr Louis Péringuey, who served as the Director of the South African Museum from 1906 until his death in 1924, Morris acknowledges Péringuey for "breaking new ground in Africa" (p. 12) in the field of archaeology and for acknowledging the living San for their contribution to the archaeology in the region. This was perhaps the first time that a local population outside of Europe, from Africa, was seen as important in unravelling questions related to human origins. This sparked Péringuey's interest in collecting human skeletons and he encouraged specialists and non-specialists to send him specimens.

Since there were no specialist anatomists in the Cape prior to 1911 when the University of Cape Town's Medical School was launched, Péringuey sent specimens from his collection to Frank Shrubsall in England to analyse. Sending specimens from South African collections to scientists overseas was a common practice and many of the scientists who held positions at museums or taught at medical schools were foreigners, predominantly of European ancestry. Consequently, methodologies used when analysing local specimens and interpreting them were based on European practices, and often resulted in reference to the African specimens in a derogatory manner.

After examining specimens from 43 "Bushmen" and 30 "Hottentots" and the "cave-dweller" specimens sent to him by Péringuey, Shrubsall concluded that the "Strandlooper-Bush-Hottentots, formed a single group with great antiquity probably throughout Africa" (p. 24). Morris also noted that the article published by Shrubsall made mention of "post-cranial bones" used to assess variation of Khoesan peoples (p. 24). This sparked an interest in craniological features among scholars, resulting in several publications describing the specimens and their interpretation of the relationship among the different collections.

In subsequent chapters in the book, Morris goes on to narrate how other scholars and their discoveries have built on the early work in the Cape, and elsewhere in South Africa, and how these studies have contributed to our understanding of the history of the indigenous people of southern Africa. We are reminded about seminal discoveries like the Boskop specimen discovered in 1913, some controversies that resulted with interpretation of data, and how the use of single specimens without proper comparative analysis can lead to misinterpretations. Morris also shares interesting accounts of the personalities and work of seminal scholars like Drennan, Dreyer, Singer, Dart, Broom, Tobias, and many more who contributed to the advancement of physical anthropology in South Africa (including his own research).

Morris addresses the question of the origin of 'races' by introducing us to the discovery of human remains from Chancelade in France in 1888 and the skeletons from the Grimaldi site in Italy in 1901 which were said to represent 'Caucasoid', 'Mongoloid' and 'Negroid' races, "suggesting that the origin of all modern humans must have been in Europe" (p. 40). Throughout the book Morris highlights and juxtaposes how interpretations based on description of skeletal remains (typology) had 'race' and racial implications, including a harrowing account on Dart's interpretation (see Chapter 5), and later how Tobias tried to cover up Dart's account and add his own thoughts on the subject (Chapter 6).

The first six chapters of the book set the stage for the last two chapters that delve into the reality of racial ideologies, racism and the way in which race classification supported the advancement of the political landscape of South

Africa. Morris outlines how race classification came about and how legislation around this classification system impacted on the lives and livelihood of non-white people and makes a case that those "responsible for the dirty work of classification were not scientists" (p. 251). Sometimes scientists were invited to have input on cases before the Race Classification Board; some accepted the invitation while others did not, due to ethical considerations. Those who did, like Tobias and Jenkins, would often use the broad spectrum of variation to not place subjects into specific categories, claiming that typology could not be used to classify subjects into one or another race.

With time, some physical anthropology scholars started to use multivariate computational methodologies like biological distance measure to analyse data, and it soon became apparent that "none of these techniques identified race" (p. 276). As serogenetic markers like blood groups and blood proteins were identified as useful tools to study population variation, these studies added to the knowledge of the affinities of southern African populations.[1] More recently, DNA studies have been used to study genetic variation among sub-Saharan African populations[2], lending further support for the 'Out of Africa' theory concerning modern human origins advanced by Rebecca Cann, Mark Stoneking and Alan Wilson[3] using human mitochondrial DNA variation. Other studies, now including whole genome analysis, provide refinement on this theory.[4]

I had the privilege of engaging with Morris on this book during a webinar hosted by the Academy of Science of South Africa (ASSAf) as part of ASSAf's Heritage Month celebrations on 20 September 2022. Our paths first crossed at an Archaeological Society meeting in Cape Town in the late 1980s and since then we have engaged on several topics of mutual interest. It was quite amusing to hear his account of how the idea of the book came about (p. 3). He had submitted an abstract to the American Association of Physical Anthropologists to attend their physical anthropology meeting in 1991 and it was rejected because it "lacked originality" (p. 3). Apparently, the reviewer felt that the topic was adequately covered in a paper published by Tobias in 1985.

This book is timely for a few reasons. Firstly, it reminds us that the post-apartheid era should be celebrated, given the plight of most people who were subjected to legislation that violated basic human rights of individuals following the colonial era. Secondly, scientific information has shown that there are no 'pure' populations and that we are all connected to the same origin as a point back in time, and that human variation is a consequence of adaptation over time to changing environments. Thirdly, while contemporary science cannot be separated from the politics of the time, every endeavour should be made for science to be based on evidence rather than succumbing to political pressures.

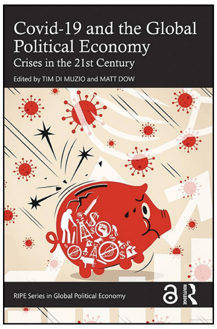
Morris is a bold and insightful scholar, whose attention for detail meticulously brings life to the narrative of physical anthropology from the early 1900s to the present in *Bones and Bodies*.

## References

1. Nurse GT, Weiner J, Jenkins T. The peoples of southern Africa and their affinities. Oxford: Oxford University Press; 1985.
2. Schlebusch CM, Malmström H, Günthe T, Sjödin P, Coutinho A, Edlund H, et al. Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago. Science. 2017;358:652–655. https://doi.org/10.1126/science.aao6266
3. Cann RL, Stoneking M, Wilson AC. Mitochondrial DNA and human evolution. Nature. 1987;325:31–36. https://doi.org/10.1038/325031a0
4. Schlebusch CM, Sjödin P, Breton G, Günther T, Naidoo T, Hollfelder N, et al. Khoe-San genomes reveal unique variation and confirm the deepest population divergence in *Homo sapiens*. Mol Biol Evol. 2020;37:2944–2954. https://doi.org/10.1093/molbev/msaa140

Check for updates

**BOOK TITLE:**
Covid-19 and the global political economy: Crises in the 21st century




**AUTHORS:**
Tim Di Muzio and Matt Dow

**ISBN:**
9781032168197 (paperback, 316 pp, GBP36)

9781032168210 (hardback, 316 pp, GBP130)

9781003250432 (ebook, 316 pp, GBP32)

**PUBLISHER:**
Routledge, Abingdon, UK

**PUBLISHED:**
2023

**REVIEWER:**
Jonathan Jansen[1]

**AFFILIATION:**
[1]Department of Education Policy Studies, Stellenbosch University, Stellenbosch, South Africa

**EMAIL:**
jonathanjansen@sun.ac.za

**HOW TO CITE:**
Jansen J. Understanding human health in relation to the global political economy. S Afr J Sci. 2023;119(9/10), Art. #15870. https://doi.org/10.17159/sajs.2023/15870

**ARTICLE INCLUDES:**
☐ Peer review
☐ Supplementary material

**PUBLISHED:**
28 September 2023



# Understanding human health in relation to the global political economy

Long gone are the days of centring medical maladies within the human body itself, or making simple connections beyond it in descriptions such as the social determinants of health. This book offers something much more ambitious: a systematic account of the global political economy of health. Its central question is how, in what ways, and to what extent, the recent COVID-19 pandemic has changed or challenged the terms of the international political economy that governs human health.

It is not a book for the fainthearted. In a strident collection of chapters, the editors and authors examine close-up the pandemic effects and consequences on human health. The opening gambit is familiar: that the pandemic has not only laid bare the deep inequalities within health systems but also exacerbated those divisions between rich and poor. The topics covered are wide-ranging – from energy and carbonisation to race and meat as well as vaccines and the precarity of work. Stringing together these seemingly diverse chapters is a relentless focus on the workings of neoliberalism, capitalism, and financial markets over the course of the pandemic. The outcomes are gloomy in that the pandemic, by the book's argument, was actually good for the rich in that it created new opportunities for unbridled wealth accumulation even as the death toll mounted among the more vulnerable sections of human population – the elderly in frail care, the desperate meatpackers, black people and the poor.

The authors of the 14 chapters represent a collection of senior professors and emerging scholars, such as postdoctoral fellows, and so there is, as can be expected, an unevenness in the quality of the arguments and the depth of the available data. Nevertheless, from Theodor Adorno to Michel Foucault, there is a rich referencing of critical theory in reference to knowledge production and how institutions work, all informed by arguments familiar to those who work with the tools of political economy. One walks away with a very sharp sense of the power of the international political economy that governs human health and the opportunism that the pandemic provided to, in fact, strengthen the system of capital accumulation rather than create a more equitable and more just healthcare system given the ravages of COVID-19.

The strengths of the book lie in the systematic application of political economy thinking across the different chapters, the rich mix of data, theory and argument that sustains the narrative line, and the synthesis at the end of the book that keeps the different contributions in conversation with each other. The critical lens applied sometimes veers in the direction of dogmatic argument, but that is compensated for by the open-mindedness of the editors to the idea of "re-evaluating certain features of the global political economy". For these reasons alone, the book is a vital resource for those who study human health *relationally* and not simply on its own terms as if pandemic effects speak for themselves. It is, in any event, delightful to have a book that takes advantage of this once-in-a-century pandemic to find out what we can discover anew about health, money, power and politics in the world we inhabit where the costly and uneven distribution of medicine is often taken as given.

The main limitation of the book is its surprisingly narrow focus on international political economy as if the Global North alone matters; except for Chapter 6, co-authored by South African professor of medicine, Solly Benatar, the references to the world outside of the dominant capitalist nations are spotty. It is not difficult to see that the global political economy only exists in relation to the Global South. Those interdependencies have a deep history that stretches across the colonial period, and so it is quite difficult to fathom why scholars describing themselves as "the critical left" would leave out a systematic analysis of how health inequities are sustained across national borders; the scandal over what was dubbed 'vaccine apartheid' during the pandemic is a case in point. In the eloquent language of the book itself, it should have given much more attention in its 300 pages to "how global pandemics are constituted in historical structures" (p.5) between North and South.

Still, this book should be recommended for all health science students, for without it, there is the real risk of their training still being dominated by a biomedical understanding of human health without the critical instincts and indeed global orientations that would make for a more just and equitable world.

Check for updates

# South African grasslands and ploughing: Outlook for agricultural expansion in Africa


**AUTHORS:**
Elaine Slooten[1,2]
Eulalia Jordaan[1,2]
Joseph D.M. White[3,4]
Sally Archibald[4]
Frances Siebert[1]

**AFFILIATIONS:**
[1]Unit for Environmental Sciences and Management, North-West University, Potchefstroom, South Africa
[2]Mpumalanga Tourism and Parks Agency, Lydenburg Research Station, Mashishing, South Africa
[3]Royal Botanic Gardens, Kew, Richmond, UK
[4]School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg, South Africa

**CORRESPONDENCE TO:**
Elaine Slooten

**EMAIL:**
eslooten98@gmail.com



**HOW TO CITE:**
Slooten E, Jordaan E, White JDM, Archibald S, Siebert F. South African grasslands and ploughing: Outlook for agricultural expansion in Africa. S Afr J Sci. 2023;119(9/10), Art. #15540. https://doi.org/10.17159/sajs.2023/15540

**ARTICLE INCLUDES:**
☐ Peer review
☐ Supplementary material

**KEYWORDS:**
biodiversity offsets, carbon credits, resilience, vegetation mapping, climate action

**FUNDING:**
Oppenheimer Generations Research and Conservation

**PUBLISHED:**
28 September 2023





**Significance:**

Grasslands were often viewed as successional vegetation, a precursor to a possible forest, and were thought to have low diversity. However, a growing number of research projects have shown that these ecosystems boast high biodiversity and massive carbon storing potential, and require at least a century to recover after agricultural ploughing which severely disrupts the environment and consequently lowers plant diversity and carbon stocks. Conserving natural grasslands should be prioritised, as they provide us with important ecosystem services beneficial to not only natural systems, but also the longevity of humans through food security and buffering of the effects of climate change.


## South African grasslands

Grasslands are ancient ecosystems that spread across the planet during the Miocene – approximately 8–20 million years ago.[1] They are characterised by a continuous herbaceous layer of grasses, a high species diversity of herbaceous forbs, an abundance of long-lived perennial plants, and well-developed belowground structures such as large storage organs and dense root structures.[2,3] These characteristics make grassland species resilient to endogenous disturbances, such as megafaunal herbivores and frequent fires.[4] Grasslands are globally distributed and provide a unique suite of ecosystem services, including, amongst others, water supply and its regulation, carbon storage, and forage production.[3] With extensive historical and rapid contemporary grassland transformation primarily due to agriculture and silviculture, grasslands urgently require prioritisation of their conservation and management.

A third of South Africa's land surface is covered by the Grassland Biome, with the broad vegetation units consisting of the Drakensberg grassland, and the dry- and mesic highveld grasslands.[5] A total of 34% of grasslands have been irreversibly transformed,with less than with only 2% being formally protected. The remaining 64% is mainly used for grazing by livestock and game[5] but is also subject to land use change through agricultural expansion and intensification, leading to increased fragmentation. Grasslands are permanently transformed by surface mining, intensive agriculture and silviculture, as well as infrastructure projects. These exogenous disturbances in grasslands severely disrupt soil structure, destroy belowground root and organ structures, and decrease species diversity.[3,6] Land use transformation in grasslands occurs mainly for agricultural expansion and economic growth, but with the gain of food security, other important ecosystem services are lost. Some of these ecosystem services include the loss of carbon stock and the carbon storage capabilities that are linked with soil structure, as well as faunal and floral biodiversity which plays a critical role in regulating the global biogeochemical cycles.[7,8] Moreover, as there is no penalty or consequence for ploughing intact grassland, and no current incentive in South Africa to protect these ecosystems, often these lands are abandoned after just a few years of cultivation.

Although exogenous disturbances disrupt the soil surface and ultimately lead to transformed grasslands, they are not the only drivers that degrade a grassland ecosystem. It is the combination of the removal or mismanagement of naturally occurring endogenous disturbances such as fire and/or herbivory, paired with intensive anthropogenic influence like ploughing that severely degrade grasslands.[9] Increasingly, farmers are finding woody encroachment to be a concern; species like *Seriphium plumosum, Lantana camara,* and *Campuloclinium macrocephalum* have especially increased in many South African grasslands in recent years.

Grasslands are fascinating in that their maintenance and conservation depends on complex feedbacks between soil, plants, and the herbivores and fire that consume them. After a disturbance like fire, drought or herbivory, herbaceous plants[10] can resprout from various underground storage systems such as woody or tuberous root systems, which store carbon and other resources.[11] However, when a grassland is ploughed – an exogenous disturbance – these hardy herbaceous plants with their high underground storage system diversity and large carbon storage capacity are permanently removed from the soil.[4] Ploughing severely damages the feedback loop where plant species are dependent on the soil carbon stock for growth. Removing this carbon stock consequently lowers the probability that plant communities will re-establish in that area, which negatively impacts ecosystem functionality and productivity.[12]

Grasslands are thought to be resilient in that the halting of invasive agricultural practices, such as ploughing, will allow the disturbed biodiversity to recover. Unfortunately, studies that investigated this assumption show that even when restorative measures are implemented, species recovery takes much longer than anticipated, with no guarantee of full recovery.[11] Current studies show that disturbed grasslands will take at least a century to recover to former species richness, which means that the carbon pool that is irreversibly linked to aboveground vegetation will take just as long to recover to its original storage capabilities.[4]

In South Africa, we have good knowledge of the carbon stored in our natural grasslands. We also know, from groundbreaking work in the 20th century, that there is no evidence that grass composition in the Highveld ever recovers from ploughing. Evidence gathered by Eddie Roux[11] at the Frankenwald research station was the first evidence that the ideas of succession, first put forward by Clements in 1916 and adopted by ecologists globally, were not sufficient to explain dynamics of disturbance-driven temperate and tropical ecosystems. This is supported by the identified 'purple veld' or 'climax grasslands' that never returned in the ploughed plots, even after decades of resting. However, these early researchers considered the recovery ofgrass species only, and did not document the loss or recovery of other herbaceous plants.[11] Recent demonstrations[11,12] showed that certain herbaceous functional groups were permanently lost from grasslands exposed to exogenous disturbance like silviculture, and that this impacted their belowground carbon reserves and functioning. This work raises questions about the degree to which our current legislation protects grasslands, and whether sufficient systems and efforts are in place to protect what remains, putting more effort into restoration, and planning our land use and biodiversity offsets accordingly.

## Steps for the future: Biodiversity offsets and carbon credits

The limitations involved in restoring transformed grasslands places huge importance on protecting existing old-growth grasslands. In parallel to protected area networks, two current market-based mechanisms in use for conserving old-growth ecosystems are carbon and, more recently, biodiversity credits. These credits aim to incentivise landowners to be more mindful in the way they manage their grasslands, with the goal of firstly protecting what is left, and secondly restoring previously disturbed fields.[13,14]

In response to the increasing loss of native vegetation and biodiversity, a growing number of countries have adopted 'offsetting' policies that seek to balance local habitat destruction by restoring, enhancing and/or protecting similar but separate habitats. Biodiversity offsetting from the National Biodiversity Offset Guidelines[15] refers to conservation efforts implemented to counteract the residual, harmful effects to biodiversity in an area due to activities (developmental or otherwise). The objective is to ensure that high-value ecosystem services, many of which are found in grasslands, that are being harmed in the process of development, are being compensated to maintain ecological integrity.

Not only are grasslands home to various threatened animal and plant species, but grasslands themselves are a critically threatened ecosystem.[13] Using biodiversity offsets as a conservation tactic is a feasible option as priority biodiversity areas are being taken into account as the offset option, instead of simply allocating any piece of land available as the trade-off.[13,15] Although biodiversity offsetting is recognised as a policy option published in the *NEMBA Act of 2004* to curb the decline of South Africa's biodiversity, it is yet to be officially included in South Africa's environmental law, unlike Mozambique and Madagascar that have both introduced such legislation. However, biodiversity offset legislation needs to be carefully thought out if it is to achieve its goal – there are examples that have resulted in perverse or undesirable outcomes due to the complexity of implementation – ultimately, this legislation requires that landowners and communities sacrifice some of their land with economic potential, and who sacrifices, and what is saved, may be complicated by power imbalances.[16] At its best, however, it can help farmers, developers, and communities to rethink their priorities, and come together to achieve healthier ecosystems. One example of such a collaboration is between the conservation organisation Overberg Renosterveld Conservation Trust (ORCT) and private landowners to protect what is left of the heavily fragmented Renosterveld vegetation type in the Western Cape. The ORCT approaches landowners that have Renosterveld on their property with a voluntary conservation easement contract that allows the organisation to aid in managing the natural remaining patches. These policies often have a stated aim of producing a 'net gain' or 'no net loss' in environmental benefits. It is challenging to determine the potential impacts of a policy and if, or when, it will achieve its objectives.[13,17]

Hand-in-hand with offsetting policies for biodiversity, the Paris Climate Agreement was founded in 2015 and was created as part of the movement for natural climate solutions with the goal to meet the climate goals outlined. In grasslands, carbon storage can best be achieved by avoidance of ploughing and land use change[18], but also through more sustainable grazing methods which influence both above- and below-ground soil and plant structures. Land managers can adopt more sustainable land management practices, but can also monetise from these climate change mitigation efforts. This links with the new concept of carbon credits and how the avoidance of old-growth, perennial grassland conversion mitigates climate change because of the regulating services grasslands provide in sequestering carbon in soil. Old-growth grasslands are estimated to store 20% of the global carbon, but in the case of land conversion up to 50% of that stored carbon could be lost.[15] Currently, financial mechanisms to avoid this carbon loss in grassland soils are not as sophisticated as programmes like REDD+ (Reducing Emissions from Deforestation and Forest Degradation) that focuses on incentives for reducing deforestation, i.e. there is no similar mechanism for reducing soil carbon loss in non-forest ecosystems. There are potential mechanisms through the voluntary carbon market, however, and if land managers incorporate both offsetting policies with regards to biodiversity and carbon credits for soil health, they can mitigate climate change and receive rewards for their efforts.

Although there are mechanisms by which funding for grassland conservation can be sought through the voluntary carbon market, a recent report on carbon mitigation options for Africa[18] highlights avoided land conversion of grasslands, a key carbon-mitigation activity with co-benefits for biodiversity and livelihoods. These authors call for REDD+ initiatives in grasslands.[18]

## Resilience within grasslands

Unsustainable management practices can negatively impact herbage accumulation, fodder quality, soil attributes, nutrient cycles, plant diversity and plant population density, as well as greenhouse gas emission.[7,8] Sustainable farming links to resilience, where agroecosystems are recently being exposed to external shocks like that from climate change. If ecological resilience of a system is increased, the capacity for a system to recover after disturbances also increases, which directly links to greater security for social and economic stability through food security.

Table 1 shows four of the biggest causes of land use change in grasslands and gives a brief overview of the resilience of grasslands. Resilience refers to both natural resilience and how the management of grasslands can increase this natural resilience.

## Future research for resilient grasslands

The first step to achieve Sustainable Development Goals is to quantify different aspects of grassland biodiversity and ecosystem service conservation in South Africa. For this, we identify three major research questions across African grasslands: (1) What are the rates of biodiversity and soil carbon recovery after grasslands are ploughed? (2) What is the extent of agricultural transformation across African grasslands? (3) What are the measures that could be put in place to minimise further transformation of our remaining grasslands?

**Table 1:** Managing for resilience with different land use types

| | Land use type | Resilience |
|---|---|---|
| **Ordered by severity of land use change[19]** | Overgrazed | Managing plant heterogeneity to maintain structure, composition, function, and diversity of plants. |
| | Woody-plant encroached | Adopting appropriate disturbance regimes (herbivory and fire) to maintain desired state. |
| | Intensive cropland | Mixed cropland- or plantation-livestock systems to maintain semi-natural pastures. Spatial arrangements to contribute to landscape integrity and connectivity.[19] |
| | Silviculture | Adopting appropriate disturbance regimes (herbivory and fire) in remnants to maintain desired state. |

A joint research project aims to quantify the extent of transformed grasslands in Africa, as well as the rates at which biodiversity and total carbon recover after grasslands are ploughed. The research aims to determine whether plants and soil faunacan re-establish in a ploughed grassland that has been left for a long time, coupled with how the carbon stock in previously ploughed land compares with an old-growth grassland which still contains the perennials with dense root systems that have carbon storage capabilities. By investigating the effects of exogenous disturbances (specifically ploughing) on both the floristic diversity and carbon stock, a holistic view begins to form as we look at, not only the visible effects above-ground, but the unseen below-ground impacts too.

There is currently conflicting information on both the historical and current extent of old-growth grasslands. Because grasslands often occur interspersed with forests (such as in the Drakensberg or the Miombo ecoregion), they can be hidden in global maps. Moreover, maps of potential vegetation used for restoration projects will often indicate that ancient old-growth grasslands are degraded forests because the environmental conditions can support forests.[20] Without a clear understanding of where grasslands occur, we cannot meaningfully conserve them. Fortunately, novel methods in vegetation mapping, that make use of high-performance computing as well as global biodiversity databases, provide new options for grassland mapping in Africa. We aim to (1) review and refine current grassland maps for Africa using comparative approaches to produce up-to-date, biologically meaningful vegetation maps. This will allow the identification of both core and transitional regions between grasslands and other vegetation types. Importantly, once core grassland regions are identified, we aim to (2) monitor the extent of land use conversion in African grasslands. With constantly improving access to remote-sensing imagery, we are now able to track land use change both historically and in near-real time. We will focus on developing open-access tools to monitor agricultural expansion into grasslands to conserve their ecosystem services and biodiversity.

The tools that can identify grasslands and near-real time land use change can be used in identifying priority restoration sites of important grassland fragments to improve landscape integrity and connectivity, and are essential for effectively implementing the legislation and financing mechanisms to promote conservation and restoration of grasslands.

## Outlook

Aldo Leopold stated in his novel, *A Sand County Almanac*:

> *We abuse land because we regard it as a commodity belonging to us. When we see land as a community to which we belong, we may begin to use it with love and respect.*

The global extent and value of grasslands for human livelihoods strongly support the sustainable management and restoration of biodiversity and ecosystem services provided by old-growth grasslands. Improved evidence of biodiversity and soil carbon recovery rates and the extent of current grassland transformation will inform the selection of areas to be targeted for ecological restoration, and increase public support for grassland conservation.

Conservation is not meant to only be a follow-up action to be taken when an ecosystem is disturbed, but rather a series of proactive interventions to ensure ecosystem resilience. Conservation means protecting what is there, following best management practice that has been through trial and error, fostering biodiversity between croplands and ensuring that there is sufficient ecological feedback between components in an ecosystem. Grassland conservation acknowledges food security through advocating effective management and the reuse of previously ploughed land and adding ancient, old-growth grasslands on the conservation agenda to prevent further losses to agricultural expansion and to combat ill-thought out tree planting initiatives. Despite evidence of large-scale transformation, it is not too late to join hands with landowners and the government. The quantification of carbon stocks and rates of biodiversity change under different land uses will be used to justify and inform biodiversity offset legislation and help fund restoration through the voluntary carbon market. Moreover, with lessons learned here in South Africa, coupled with novel mapping techniques produced by this programme, it can be used to contribute to land use planning activities in parts of Africa where agricultural expansion in grasslands is occurring rapidly, sometimes with limited recognition of the long-term consequences.

## Acknowledgements

We acknowledge the The Future Ecosystems For Africa programme in partnership with Oppenheimer Generations Research and Conservation.

## Competing interests

We have no competing interests to declare.

## References

1. Edwards E, Osborne CP, Strömberg CAE, Smith SA, C4 Grasses Consortium, Bond WJ, et al. The origins of C4 grasslands: Integrating evolutionary and ecosystem science. Science. 2010;328(5978):587–591. https://doi.org/10.1126/science.1177216
2. Nerlekar AN, Veldman JW. High plant diversity and slow assembly of old-growth grasslands. Proc Natl Acad Sci USA. 2020;117(31):18550–18556. https://doi.org/10.1073/pnas.1922266117
3. Sanford W, Wangari E. Tropical grasslands: Dynamics and utilization. Nat Resour J. 1985;21(3):12–27.
4. Buisson E, Archibald S, Fidelis A, Suding KN. Ancient grasslands guide ambitious goals in grassland restoration. Science. 2022;377(6606):594–598. https://doi.org/10.1126/science.abo4605
5. Carbutt C, Tau M, Stephens A, Escott B. The conservation status of temperate grasslands in southern Africa. Grassroots. 2011;11(1):17–23. Available from: https://grassland.org.za/publications/grassroots/issues/february-2011/5%20201102%20Carbutt.pdf
6. Yellajosula G, Cihacek L, Faller T, Schauer C. Soil carbon change due to land conversion to grassland in a semi-arid environment. Soil Syst. 2020;4(3):43. https://doi.org/10.3390/soilsystems4030043
7. Bardgett RD, Bullock JM, Lavorel S, Manning P, Schaffner U, Ostle N, et al. Combatting global grassland degradation. Nat Rev Earth Environ. 2021;2(10):720–735. https://doi.org/10.1038/s43017-021-00207-2
8. O'Mara FP. The role of grasslands in food security and climate change. Ann Bot. 2012;110(6):1263–1270. https://doi.org/10.1093/aob/mcs209
9. Muller M, Siebert SJ, Ntloko BR, Siebert F. A floristic assessment of grassland diversity loss in South Africa. Bothalia. 2021;51(1):Art. #a11. https://doi.org/10.38201/btha.abc.v51.i1.11
10. Siebert F, Dreber N. Forb ecology research in dry African savannas: Knowledge, gaps, and future perspectives. Ecol Evol. 2019;9(13):7875–7891. https://doi.org/10.1002/ece3.5307
11. Roux E, Roux WM. Grass: A story of Frankenwald. Cape Town: Oxford University Press; 1969. https://doi.org/10.2307/3896812
12. Zaloumis NP, Bond WJ. Grassland restoration after afforestation: No direction home? Austral Ecol. 2011;36(4):357–366. https://doi.org/10.1111/j.1442-9993.2010.02158.x
13. Prinsloo H, Reilly B, Myburgh W. Identifying potential protected areas in the Grassland Biome of South Africa. S Afr J Sci. 2021;117(3/4):Art. #7507. https://dx.doi.org/10.17159/sajs.2021/7507
14. Dean C, Kirkpatrick JB, Harper RJ, Eldridge DJ. Optimising carbon sequestration in arid and semiarid rangelands. Ecol Eng. 2015;74:148–163. https://doi.org/10.1016/j.ecoleng.2014.09.125
15. National Biodiversity Offset Guideline. Section 24J of the National Environmental Management Act, 1998 (Act No. 107 of 1998) (NEMA). South Africa; 2022.
16. Bidaud C, Schreckenberg K, Jones JP. The local costs of biodiversity offsets: Comparing standards, policy and practice. Land Use Policy. 2018;77:43–50.

17. Brownlie S, Von Hase A, Botha M, Manuel J, Balmforth Z, Jenner N. Biodiversity offsets in South Africa – challenges and potential solutions. Impact Assess Proj Apprais. 2017;35(3):248–256. https://doi.org/10.1080/14615517.2017.1322810

18. Armani M, Asefa M, Zakari S, Agyekum EO. Enhancing climate change adaptation and mitigation actions on land in Africa. In: Archibald SA, Pereira LM, Coetzer KL, editors. Future ecosystems for Africa (FEFA). Johannesburg: University of the Witwatersrand; 2022. p. 92.

19. O'connor TG, Kuyler P. Impact of land use on the biodiversity integrity of the moist sub-biome of the grassland biome, South Africa. J Environ Manage. 2009;90(1):384–395. https://doi.org/10.1016/j.jenvman.2007.10.012

20. Bastin JF, Finegold Y, Mollicone D, Rezende M, Routh D, Zohner CM, et al. The global tree restoration potential. Science. 2019;365(6448):76–79. https://doi.org/10.1126/science.aax0848

# GYPWORLD Africa: Setting an agenda for gypsum ecosystem research in southern Africa

**AUTHORS:**
Stefan J. Siebert[1]
Sara Palacio[1,2]
Arantzazu L. Luzuriaga[3]
Gillian Maggs-Kölling[1,4]
Eugene Marais[4]
Silvia Matesanz[3]
María Prieto[3]
Yolanda Pueyo[5]
Nishanta Rajakaruna[1,6]
Ana M. Sánchez[3]
Sarina Claassens[1]

**AFFILIATIONS:**
[1]Unit for Environmental Sciences and Management, North-West University, Potchefstroom, South Africa
[2]Pyrenean Institute of Ecology, Higher Council for Scientific Research, Huesca, Spain
[3]Biodiversity and Conservation, Department of Biology and Geology, Physics and Inorganic Chemistry, Rey Juan Carlos University, Madrid, Spain
[4]Gobabeb – Namib Research Institute, Walvis Bay, Namibia
[5]Pyrenean Institute of Ecology, Higher Council for Scientific Research, Zaragoza, Spain
[6]Biological Sciences Department, California Polytechnic State University, San Luis Obispo, California, USA

**CORRESPONDENCE TO:**
Stefan Siebert

**EMAIL:**
stefan.siebert@nwu.ac.za

**HOW TO CITE:**
Siebert SJ, Palacio S, Luzuriaga AL, Maggs-Kölling G, Marais E, Matesanz S, et al. GYPWORLD Africa: Setting an agenda for gypsum ecosystem research in southern Africa. S Afr J Sci. 2023;119(9/10), Art. #15308. https://doi.org/10.17159/sajs.2023/15308

**ARTICLE INCLUDES:**
☐ Peer review
☐ Supplementary material

**KEYWORDS:**
edaphic endemics, Namib Desert, Namaqualand, soil, plants, lichens

**FUNDING:**
Fulbright US Scholar Program, North-West University, Horizon 2020 (H2020-MSCA-RISE-777803)

**PUBLISHED:**
28 September 2023



**Significance:**

This paper introduces gypsum ecosystem research to southern Africa. It is the result of current joint efforts to compare African gypsum ecosystems with those in other parts of the world. We highlight the expansion of an international network through joint projects and training of young scientists. We propose a research agenda to sensitise the ecological community in Africa to the significance of life on gypsum and to demystify the existence of gypsum soil and associated ecosystems in southern Africa.

Gypsum ecosystem research is rapidly moving up the international research agenda to better understand the dynamics and resilience of the life systems associated with atypical soils that are frequent in semi-arid to arid ecosystems. The southern African soil classification system does not sufficiently recognise the presence of gypsum as a differentiating criterion, hence not much is known about the region's gypsum soil, and gypsum ecology has subsequently been largely neglected. This neglect is unfortunate, as the livelihoods of people are dependent on these gypsum ecosystems and these areas are worthy of protection due to the rare biotas that are adapted to survive in these harsh environments.

One of the key players in gypsum ecosystem research, GYPWORLD (a European H2020-MSCA-RISE GYPWORLD project), approached the GeoEco Lab at the North-West University (NWU) to arrange a first research expedition to major gypsum areas of southern Africa, with a specific focus on the central Namib Desert of Namibia and Namaqualand in the northern Cape, South Africa. The goals of GYPWORLD are to (1) assess plant and lichen diversity associated with gypsum soils; (2) determine the origin of gypsophilic biotas worldwide; (3) understand the processes that regulate plant and lichen ecosystem function; (4) promote the study of gypsum ecosystems; and (5) communicate the ecological and conservation value of these ecosystems to the public.

The expedition from 12 August to 4 September 2022 was attended by 26 experts from Brazil, Namibia, South Africa, Spain, Turkey, and the USA. Participating researchers exchanged skills and knowledge during the expedition and in two public seminars. The seminars provided opportunities for scientists to network and discuss research related to plant, lichen and community dynamics and conservation. The first was hosted by the Unit for Environmental Sciences and Management at NWU on 11 August 2022, and the second by the Gobabeb Namib Research Institute on 23 August 2022. Delegates provided background information on southern Africa's gypsum ecosystems and discussed leading international research dealing with gypsum ecology and associated lichen and plant communities. Two invited Fulbright US Scholars, who are specialists on life in harsh environments, contributed to the discussion about the state of gypsum research in southern Africa, including current knowledge gaps. Young scientists participated to stimulate their thinking on the links between diversity, function, evolutionary history, and disturbance in these ecosystems and contributed interesting perspectives to the discussions.

The seminar participants emphasised that significant knowledge gaps in gypsum characteristics preclude assumptions about the age, origin and ecological significance of the pedogenic gypsum soils along Africa's southwestern coast. A lack of in-depth local knowledge regarding gypsum indicator species, particularly distinguishing between gypsophiles (endemic – restricted to gypsum) and gypsovags (those on both gypsum and non-gypsum soils), when studying gypsum ecology and community assembly, were also highlighted.

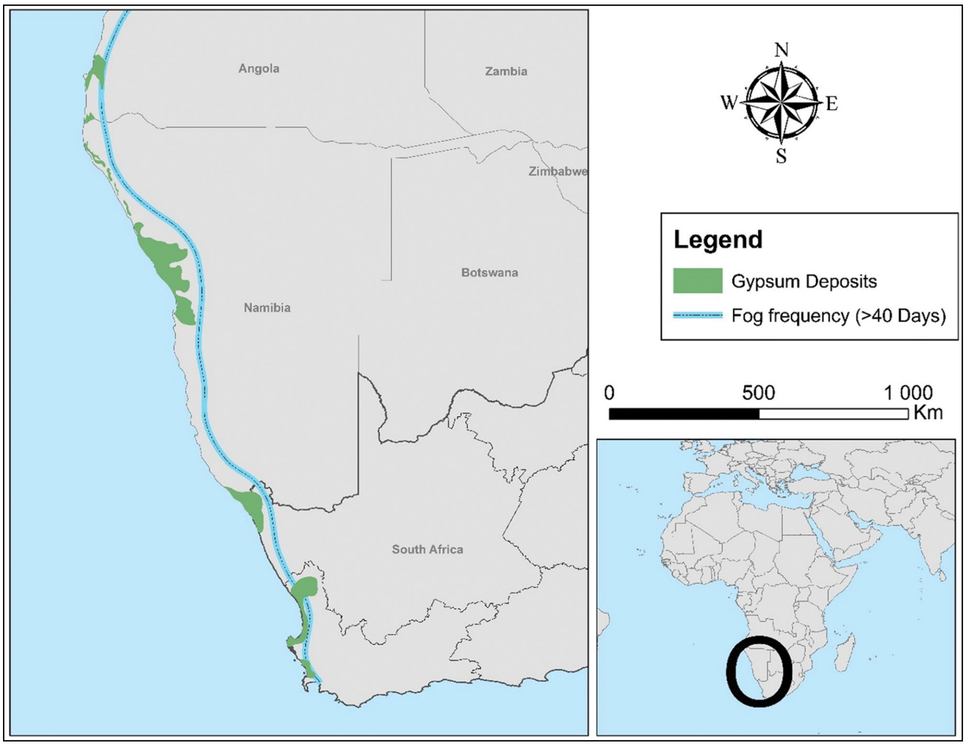
## Implications for the broader southern African ecological community

Prioritising gypsum ecosystem research in southern Africa will further increase the opportunities for international collaboration. Already, new opportunities for joint efforts to compare African and European gypsum ecosystems have been identified. There is also potential for greater access to international funding given the global relevance of the research. Such funding initiatives will afford young scientists with a unique and valuable opportunity to work with world leaders in the field and lay the foundations for future networking and collaboration. It is our aim that knowledge transfer will continue and not be limited to the 2022 engagement. We will achieve this by expanding our international network through joint projects and training of young scientists to sensitise the ecological community to the significance of research of life on gypsum and demystify the existence of gypsum soil and ecosystems in southern Africa.

## Priority research required to bolster gypsum ecology research in southern Africa

### *Gypsum characterisation*

Gypsum is a soft sulfate mineral ($CaSO_4 \cdot 2H_2O$) that forms deposits along the coast from southern Angola to South Africa and into the interior of South Africa (Figure 1). In some areas of the Namib Desert, pure gypsic deposits of up to 4 m deep and extending up to 100 km inland have been reported.[1] Detailed geological maps for gypsiferous deposits in southern Africa are not readily available[2], thus characterisation of gypsum distribution, depth of the soil profile, and gypsum content (%), is a crucial first step to study the region's gypsum ecosystems.




*Source: Data obtained from various sources*[2,3,4]

**Figure 1:** Pedogenic gypsisols are associated with the Benguela Upwelling System in southern Africa.[2,3] Mean annual rainfall for most of this coastal region with >40 fog days per year is less than 50 mm.[3,4]

Most South African deposits are derived from bedrock materials,[2] with those along the Namaqualand coast of similar pedogenic origin as deposits along the coast of Namibia and Angola. It is likely that pedogenic gypsum in the Namib Desert originates from marine $H_2S$ derived from phytoplankton decomposition in the highly productive Benguela Upwelling System (BUS) that is transported inland together with advective fog.[1] The mixing of marine sulfates with calcareous dust from the interior results in the formation of calcium sulfate minerals[2], primarily gypsum. The close association between the formation of extensive gypsum deposits and the evolution and northwards extension of the BUS and Namib Desert aridity is clear, but debate about the process, age and rate of deposition is ongoing.

## *Precipitation*

The Benguela Current draws surface water away from the west coast of southern Africa and carries it northwards, which results in the upwelling of cold, but nutrient-rich, deep ocean water. Cold, coastal waters inhibit the development of convective rain clouds and inflow of atmospheric moisture over the southwest coast of Africa. The arid conditions imposed by the proximity of the BUS, which are alleviated by distance from the coast, are a distinctive feature of the region's pedogenic gypsum soils. Water availability, provided by meagre rainfall and advective marine fog closer to the shore, is the most important factor driving environmental processes. Along the coast, fog (usually formed at night and dissipated during the day) and low clouds are advected over the adjacent coastal plains by onshore sea breezes from the South Atlantic Ocean[4], depositing varying quantities of fog water (Figure 1). The water has a low salt content and is a relatively predictable, though temporary, source of water for biota. Rainwater is rarely available on the surface, yet sporadic rain events are critical for many life forms and are the trigger for plant germination and growth. An improved understanding is required of how changes in the patterns of rainfall, likely resulting from global change, would alter seasonal and daily soil moisture levels and influence patterns of plant and lichen community composition.

## *Floristics*

Although lichen and plant species growing in gypsisols in southern Africa have been described in regional floras and ecological studies, a comprehensive list of those species strictly linked to gypsum soils is still lacking. A first step is to identify gypsophiles and gypsovags[5] by performing an exhaustive survey of gypsum and surrounding non-gypsic areas (Table 1). A special effort should be devoted to annual plants that are only intermittently present and may not have been observed during our recent survey. Similarly, microlichens and biological soil crusts should be described to better document cryptic diversity. Once the diversity is better documented, identification keys for both plants and lichens should be prepared. This floristic work will greatly benefit further research and global comparisons of diversity of gypsum ecosystems. Specifically, comparing gypsum diversity of southern Africa with those of other world regions would aid in completing the list of families, genera and species adapted to gypsum and to determine if there is a characteristic gypsum flora in southern Africa. This would be an important step to advance the study of evolutionary processes driving adaptation to gypsum soils and ecological processes determining gypsum community assembly.

## *Community ecology*

An exhaustive review of studies (on the Web of Science database) that deal with plant and lichen community composition on gypsum soils in southern Africa (south of latitude 17°S), rendered a bleak framework for gypsum community ecology in the region (Table 2). The only study that explicitly considered the species composition of natural vegetation growing on gypsum soils was performed almost 50 years ago in Botswana.[6] This study reported distinctive floristic composition, but no clear evidence of exclusive species. The lack of studies is very striking considering that plant communities have been profusely surveyed in Namibia[7] and in arid and semiarid regions of South Africa, such as the diverse Succulent Karoo biome.[8] Given the large land extensions with some degree of gypsum content in soils, we suggest the likelihood that many vegetation studies were performed on gypsum soils without an explicit mention. Similarly, no studies were found dealing with plant–plant interactions or seed germination of plants growing on gypsum soils (Table 2). Therefore, an imperative first research step is to characterise and compare plant community composition, diversity, and structure between gypsum and non-gypsum soils under similar environmental conditions. This would allow identification of gypsum indicator species and a better understanding of the relevance of gypsophily. Since gypsum outcrops confer additional harsh conditions compared to non-gypsum soils, especially in arid environments, we might expect a

**Table 1:** Families and genera of vascular plants and lichens recorded from gypsum soil during the 2022 expedition

| Vascular plant families | Number of species | Lichen families | Number of species |
|---|---|---|---|
| *Aizoaceae* | 18 | *Teloschistaceae* | 10 |
| *Asteraceae* | 9 | *Ramalinaceae* | 7 |
| *Amaranthaceae* | 5 | *Parmeliaceae* | 4 |
| *Crassulaceae* | 4 | *Acarosporaceae* | 3 |
| *Fabaceae* | 4 | *Caliciaceae* | 2 |
| *Solanaceae* | 4 | *Lecanoraceae* | 2 |
| *Zygophyllaceae* | 4 | *Lichinaceae* | 2 |
| *Euphorbiaceae* | 3 | *Physciaceae* | 2 |
| *Poaceae* | 3 | *Verrucariaceae* | 2 |
| 13 other families | 13 | 12 other families | 13 |
| Total | 67 | Total | 47 |

**Table 2:** Search terms included in the Web of Science for all databases from 1901 to 2022 (search data 21 October 2022)

| Search terms | Reference | Observation |
|---|---|---|
| Plant **community** AND Africa AND gyps* NOT bird NOT disease NOT Algeria NOT Somalia NOT Egypt NOT Tunisia* NOT Morocco* NOT Ethiopia* | Wild[6] | None studied plant community composition |
| Plant AND **germination** AND Africa AND gyps* NOT bird NOT disease NOT Algeria NOT Somalia NOT Egypt NOT Tunisia* NOT Morocco* NOT Ethiopia* | Atlas of Namibia Team[3] | None studied germination biology |
| Plant AND **interact*** AND Africa AND gyps* NOT bird NOT disease NOT Algeria NOT Somalia NOT Egypt NOT Tunisia* NOT Morocco* NOT Ethiopia* | Burke and Strohbach[7] | None studied plant–plant interactions |
| Plant AND **compet*** AND Africa AND gyps* NOT bird NOT disease NOT Algeria NOT Somalia NOT Egypt NOT Tunisia* NOT Morocco* NOT Ethiopia* | Eckardt and Spiro[1] | None studied plant competition |
| Plant AND **facilit*** AND Africa AND gyps* NOT bird NOT disease NOT Algeria NOT Somalia NOT Egypt NOT Tunisia* NOT Morocco* NOT Ethiopia* | Greyling and Van Rooy[2] | None studied facilitation |

strong bottleneck in plant establishment and regeneration that deserves investigation. Also, understanding the role of soil seed banks, plant–plant interactions (i.e. interference and facilitation), and cross-kingdom interactions on gypsum plant community dynamics would be relevant for providing specific recommendations to land managers.

### *Functional strategies*

Key to analysing plant life on gypsum is to understand the mechanisms of plants to cope with the characteristics of gypsum soils and potential convergence with gypsophiles in other parts of the world. Due to the subsurface accumulation of pedogenic gypsum in southern Africa, it would be crucial to evaluate the root penetration of potential gypsophiles to ascertain if root distribution within the soil profile concentrates on, or avoids, gypsum soils. Taxa that are specific to gypsum soils should be characterised for functional traits of known significance that allow them to cope with the environmental restrictions typical of gypsum soils. For example, gypsophiles from other regions of the world accumulate elements found in excess on gypsum soils (S, Ca and Mg) in their leaves[9], while gypsovags seem to block the uptake of these elements at the root level[10] showing lower foliar concentrations. Identifying functional traits related to water and nutrient use that have an impact on plant fitness and how they vary within and among populations is vital for understanding plant adaptation to gypsum soils and how they would respond to global change.[11]

### *Way forward*

During the recent expedition, 24 long-term monitoring plots were established in Namibia and South Africa. Continued sampling of these plots and long-term collaborative research will improve our understanding of gypsum ecosystem dynamics and help us identify potential threats to both species and functional diversity.

Further, funding for skills development and capacity building of African researchers is available through the GYPWORLD project. This enables young and established researchers to visit partner countries for research and collaboration on gypsum ecology.

The recent signing of a Memorandum of Understanding between NWU and Gobabeb signified the commitment of researchers in Namibia and South Africa to invest time and resources in the advancement of gypsum ecosystem research in southern Africa.

## Acknowledgements

We gratefully acknowledge funding to N.R. from the Fulbright US Scholar Program (South Africa). S.S., N.R. and S.C. received support from North-West University (Potchefstroom). All Spanish and South African authors were funded by European Union's Horizon 2020 (H2020-MSCA-RISE-777803).

## Competing interests

We have no competing interests to declare.

## References

1. Eckardt FD, Spiro B. The origin of sulphur in gypsum and dissolved sulphate in the Central Namib Desert, Namibia. Sediment Geol. 1999;123:255–273. https://doi.org/10.1016/S0037-0738(98)00137-7

2. Greyling M, Van Rooy JL. Hydrogeological properties of gypseous soils in South Africa. S Afr J Geol. 2019;122:389–396. https://doi.org/10.25131/sajg.122.0029

3. Atlas of Namibia Team. Atlas of Namibia: Its land, water and life. Windhoek: Namibia Nature Foundation; 2022. p. 151. https://atlasofnamibia.online/chapter-5/soil-types

4. Olivier J. Spatial distribution of fog in the Namib. J Arid Environ. 1995;29:129–138. https://doi.org/10.1016/S0140-1963(05)80084-9

5. Palacio S, Escudero A, Montserrat-Martí G, Maestro M, Milla R, Albert MJ. Plants living on gypsum: Beyond the specialist model. Ann Bot. 2007;99(2):333–343. https://doi.org/10.1093/aob/mcl263

6. Wild H. The natural vegetation of gypsum bearing soils in south central Africa. Kirkia. 1974;9:279–292. https://www.jstor.org/stable/23502022

7. Burke A, Strohbach BJ. Vegetation studies in Namibia. Dinteria. 2000;26:1–24. https://www.nbri.org.na/sites/default/files/dinteria26A.pdf

8. Schmiedel U, Kühne N, Twerski A, Oldeland J. Small-scale soil patterns drive sharp boundaries between succulent "dwarf" biomes (or habitats) in the arid Succulent Karoo, South Africa. S Afr J Bot. 2015;101:129–138. https://doi.org/10.1016/j.sajb.2015.05.001

9. Palacio S, Cera A, Escudero A, Luzuriaga AL, Sánchez AM, Mota JF, et al. Recent and ancient evolutionary events shaped plant elemental composition of edaphic endemics: A phylogeny-wide analysis of Iberian gypsum plants. New Phytol. 2022;235:2406–2423. https://doi.org/10.1111/nph.18309

10. Cera A, Montserrat-Martí G, Drenovsky RE, Ourry A, Brunel-Muguet S, Palacio S. Gypsum endemics accumulate excess nutrients in leaves as a potential constitutive strategy to grow in grazed extreme soils. Physiol Plant. 2022;174(4):e13738. https://doi.org/10.1111/ppl.13738

11. Matesanz S, Ramos-Muñoz M, Blanco-Sánchez M, Escudero A. High differentiation in functional traits but similar phenotypic plasticity in populations of a soil specialist along a climatic gradient. Ann Bot. 2020;125:969–980. https://doi.org/10.1093/aob/mcaa020

Check for updates

**AUTHORS:**
Pedro M.S. Monteiro[1]
Guy F. Midgley[1]

**AFFILIATION:**
[1]School for Climate Studies, Stellenbosch University, Stellenbosch, South Africa

**CORRESPONDENCE TO:**
Pedro Monteiro

**EMAIL:**
pmsmonteiro@sun.ac.za

**HOW TO CITE:**
Monteiro PMS, Midgley GF. Building a carbon dioxide removal science–policy partnership for southern Africa. S Afr J Sci. 2023;119(9/10), Art. #16320. https://doi.org/10.17159/sajs.2023/16320

**ARTICLE INCLUDES:**
☐ Peer review
☐ Supplementary material

**KEYWORDS:**
carbon sink, climate feedback, mitigation, sequestration, transient climate response

**PUBLISHED:**
28 September 2023



# Building a carbon dioxide removal science–policy partnership for southern Africa

**Significance:**

Carbon dioxide removal (CDR) negative emissions interventions are needed to steer the planet to a safe climate by the end of the century. In this Commentary, we frame the rationale and likely challenges for a regionally focused and coordinated CDR-centred science–policy platform with a global reach to support the opportunities and minimise the risks associated with CDR in southern Africa. We make a first attempt to frame a new CDR-centred strategic compact in the science–government–innovation–business nexus that is required to enable South Africa to provide regional and global climate leadership and impact over the 21st century.

All IPCC emission scenarios that aim to avoid dangerous climate change require carbon dioxide removal (CDR) negative emissions interventions to steer the planet to a safe climate by the end of the century.[1-5] Effective negative $CO_2$ emissions will be required on a scale that matches the order of magnitude of present positive emissions from oil, gas and coal.[4-6] For South Africa, this translates to CDR interventions in the order of 0.1–0.5 Gigatons of $CO_2$ per year ($GtCO_2$/year), which at the present carbon price range of USD100 per ton of $CO_2$ would equate to an annual USD10–50 B/year industry, excluding the required infrastructure and skills investments and avoided costs of climate damage.[7] CDR is thus a key element of South Africa's emerging economic development strategy, re-shaping science, mitigation, adaptation and financing policies and investments to strengthen the transition towards and beyond net-zero.[8,9] However, there are as yet very limited relevant governance mechanisms in place, technologies are in their infancy, and the scientific and tertiary training capabilities are largely unprepared to support these developments.[10] CDR represents a significant innovation, development and educational opportunity, but only if academic, science and policy (public and business) communities can build consensus on the efficacy and prioritisation of regionally suited selected approaches and coordinate efforts around this key regional–global challenge to avoid dangerous climate change.[9-11]

## What is CDR?

Carbon dioxide removal (CDR) comprises a variety of anthropogenic interventions that, directly or indirectly, **remove $CO_2$ from the atmosphere and store it durably** in geological, terrestrial, or ocean reservoirs, or in long-life products[3] (Figure 1). CDR is not carbon dioxide reduction, or emissions reductions, which is needed to achieve the policy goal of net-zero emissions.[3] CDR interventions can range from ecosystem process enhancements on land and in the ocean, through enhanced geochemical cycles to technology-intensive interventions such as direct air capture[2,3,5,6] (Figure 1). All these interventions aim to sequester and store the $CO_2$ removed from the atmosphere for the long term (>10 K years) in appropriate terrestrial and ocean reservoirs (Figure 1).[3,5,6] Thus CDR is a system of interventions and feedbacks, which determine its efficacy. Its integrated totality involves not just the interventions and their ecosystem trade-offs, but also the planetary–regional carbon-concentration and carbon-climate feedbacks in both land and ocean carbon sinks (Figure 1). It is the integrated nature of the entire CDR 'system', including the feedbacks, that determines the efficacy and 'final' magnitudes of CDR interventions.[3,6]

There are four main sources of risk and uncertainty that influence the efficacy and scalability of CDR. Firstly, there are economic, human livelihood, and biodiversity trade-offs directly or indirectly linked to the interventions; for example, interventions in land (includes agriculture) and ocean ecosystem processes, additional natural resource requirements and far-field unanticipated impacts.[3,5,6,11-14] These trade-offs are to a large extent within current – policies and scientific capabilities. Secondly, there are societal and science concerns about the ideas of further intervention in the climate system, which is where the as-yet underdeveloped CDR governance policy is critical to building trust and confidence through greater transparency.[10] Thirdly, there are significant knowledge gaps relating to the response of carbon-climate (heat) feedbacks from the ocean and land reservoirs to the interventions, as well as on and the nature and trajectories of the reversibility of climate change impacts on land and in the ocean (Figure 1).[5,6,12,14,15] These are projected to have the biggest influence on the uncertainty of CDR efficacy, which will impact society through carbon pricing, management of climate risk, and the costs of adaptation and mitigation. Finally, there are risks related to the as-yet uncertain feasibility and scalability of the technological interventions, such as direct air capture of $CO_2$, biomass energy with carbon capture (BECCS), and geochemical-, biogeochemical- and nature-based enhancements of the carbon cycle.[10,13]

CDR is likely to emerge and grow in two main phases. In the short term (±10–30 years), CDR is urgently needed to support the global policy aim of achieving net-zero emissions for greenhouse gases by 2050, especially in the context of an increasingly likely short-term temperature target overshoot emissions scenario.[4] During this initial period, the global response needs to accommodate not only slower emissions reduction trajectories in some countries, but also the weakly constrained but likely countervailing effects of improving air quality on radiative forcing, non-$CO_2$ emissions trajectories ($CH_4$ and $N_2O$), and recalcitrance in sectors with hard-to-abate emissions, such as air travel.[13] In South Africa, this may involve the integration of CDR into emissions reduction policies such as the Long Term Mitigation Strategy, its nationally determined contributions and its contribution to the global stocktake, also with the aim of meeting the policy objective of net-zero emissions by 2050.[8] However, this also raises the jeopardy of CDR being used to offset rather than complement emissions reductions, which

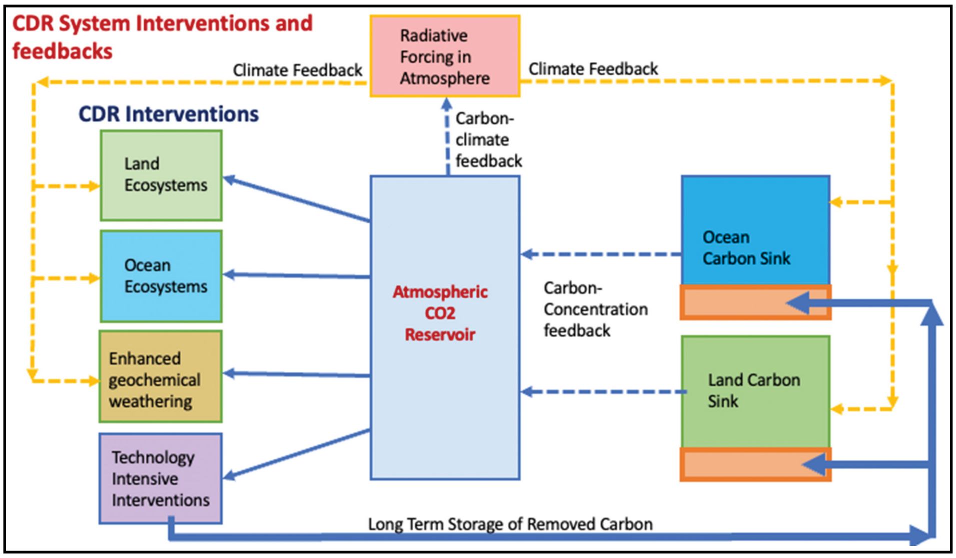



**Figure 1:** Schematic of the integrated key intervention–feedback nexus elements of a carbon dioxide removal (CDR) system. The main CDR intervention categories draw down atmospheric $CO_2$ (solid blue arrows) and store it in long-term land and ocean reservoirs (solid bold blue arrows). The active land and ocean reservoirs modulate the carbon-concentration feedback (dashed blue arrows), which sets the net decrease in atmospheric $CO_2$ and radiative forcing. Hence the effectiveness of the CDR interventions on the carbon-climate (heat and water) feedbacks in both the land and ocean ecosystems on both the intervention and response sides (dashed orange arrows). See Figure 3 for specific categories.

again highlights the urgency of well-founded science–governance policy capabilities.[9-11] In the medium to long term (20–100 years), CDR's growing and most likely largest impact will be to address the warming commitment from historical emissions, which remains a challenge to reducing global radiative forcing down to 1.9 W/m$^2$ or 2.6 W/m$^2$ by 2100 to meet 1.5 °C and 2 °C targets, respectively.[2,3,16] The existential challenges to achieving these objectives at the global scale imply an urgent need to build and coordinate the required science, technological and policy capacity at all regional and global levels.

## Global-scale science challenges

Initial global-scale modelling results highlight substantive science challenges that have a bearing on the effectiveness of CDR.[2,3,6,16-18] The re-balancing of the land and ocean carbon and energy reservoirs under CDR are poorly constrained at the global and regional levels, and the dynamics of the regional ocean and land systems feedback contributions to the mean global response are uncertain.[3,6] Perhaps the biggest challenges are in the asymmetry of the contrasting quasi-linearity and non-linearity of the relationship between temperature change and cumulative emissions under positive and negative emissions, respectively (Figure 2a).[17,18] The response of surface air temperature to CDR-driven negative emissions does not mirror changes from positive emissions.[17,18]

Under positive emissions, the relationship between cumulative $CO_2$ emissions and temperature change, the Transient Climate Response to cumulative carbon Emissions (TCRE), is quasi-linear and path independent (Figure 2a).[3] This enables TCRE to be used to calculate one of the most policy relevant and socially transforming planetary metrics in support of mitigation policy: the remaining carbon budget.[3] In sharp contrast, the non-linearity of TCRE, and its hysteresis under negative emissions (n-TCRE), arises mainly from the lags in the response of ocean fluxes of $CO_2$ and heat across the base of the ocean mixed layer as well as the divergent responses of net ocean heat and $CO_2$ fluxes to negative emissions arising from the orthogonal profiles of temperature and $CO_2$ in the ocean.[3,17,18] On land, the non-linearities may arise from the slow rates of ecological adjustments, such as from shifts in dominant plant photosynthetic and structural types (biomes) that will occur in response to decreasing atmospheric $CO_2$.[3,16] One of the main practical outcomes from this non-linearity is that CDR is likely not as effective at cooling as positive $CO_2$ emissions are at warming[3] – a potentially significant policy and societal planning and trust challenge.

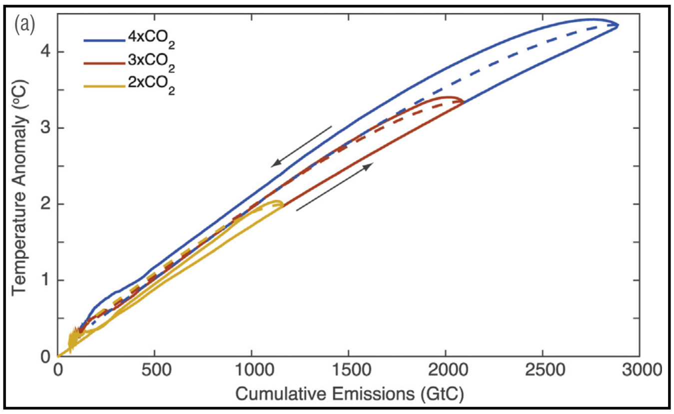
The second possible outcome impacting the effectiveness of CDR is the re-balancing of the anthropogenic $CO_2$ stored in ocean and land sinks during the historical positive emissions period (Figure 2b).[3] Idealised model experiments suggest that about 40–60% of the total $CO_2$ removed from the atmosphere by CDR would be counterbalanced by outgassing from both the contemporary ocean and land sinks (Figure 2b). The question arises: how sensitive are these rates and magnitudes of re-balancing to the scales at which the models capture regional specificities in the physics and biogeochemistry? These could include heterogeneous carbon sinks in soils, dissolved terrestrially fixed carbon exported into the ocean, carbon sinks accumulated on the ocean floor, ocean stratification, upwelling and lateral ocean current transport. There is currently almost no work underway to project and experimentally constrain these processes.

The projected confidence levels for the biogeochemical impacts of CDR are synthesised and assessed in Figure 3, which links the projected temporal effectiveness and scalability of the main intervention types to the confidence and direction of the earth system feedbacks, and ecosystem-scale biogeochemical and biophysical effects and co-benefits.[3] The key point of this global assessment is the widespread low confidence in the impacts and their direction (Figure 3).

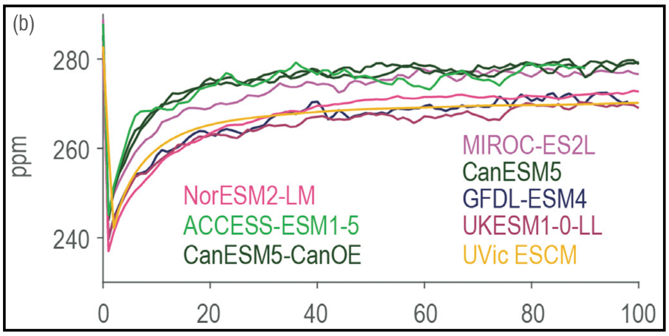
The question then arises: could a regional-scale approach help address these global sources of uncertainty in respect of land and ocean CDR? How coupled regional land–ocean–atmosphere processes contribute to the global impact of CDR through both the carbon and heat fluxes is one of the most pressing science challenges. Can a regional focus with higher temporal and spatial resolution of the process variability strengthen confidence in the assessment of CDR effectiveness? How might natural ocean and terrestrial processes, particularly in the Southern African Regional Earth System (SA-RES), be enhanced to contribute more to the global effectiveness of CDR? This creates an opportunity for South African science to use its comparative regional geographical, climate and ecological advantages to both regional and global benefit.

## Why a regional focus? Addressing the feedback scale challenges

The processes that influence the outcomes of CDR interventions comprise a very wide range of spatial (from one to thousands of kilometres) and temporal (days to decades) scales of variability on

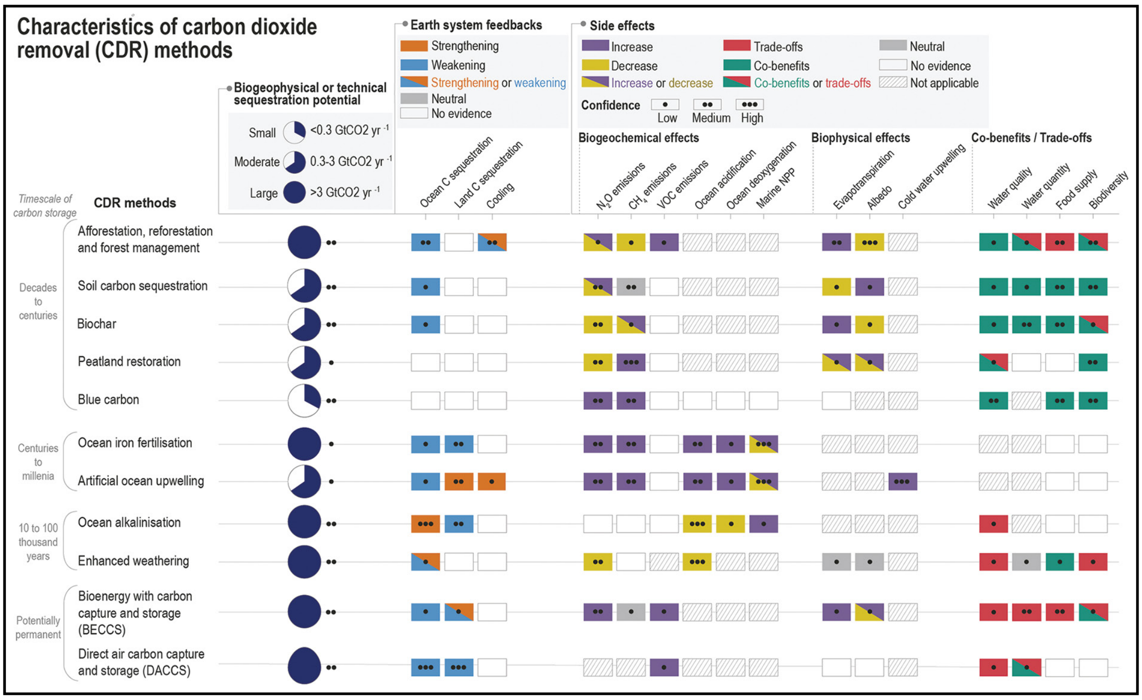



*Source: Zickfeld et al.[17] under licence CC-BY 3.0*



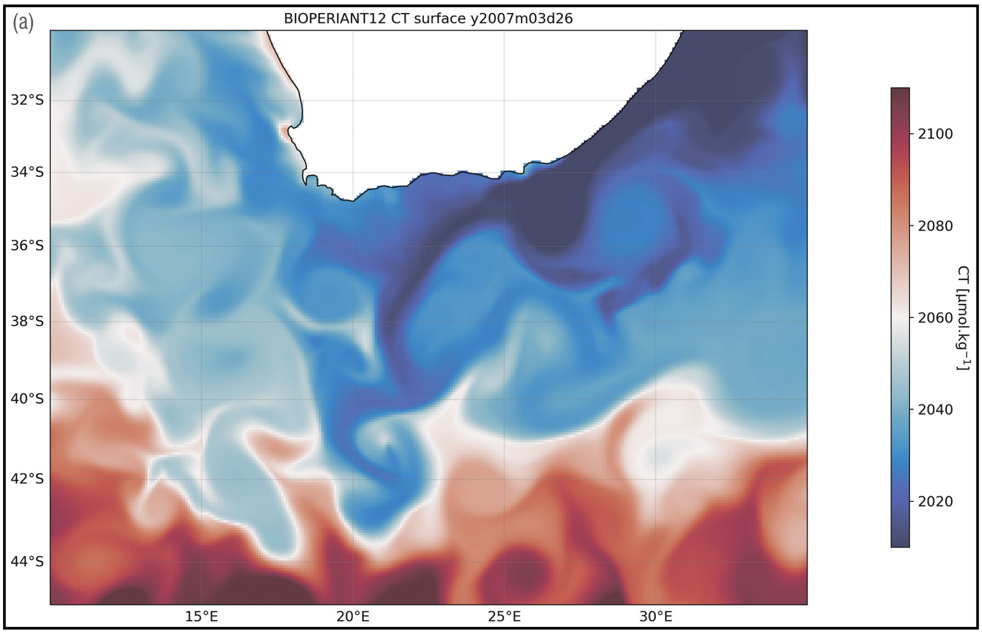

*Source: Canadell et al.[3] with permission (©IPCC)*

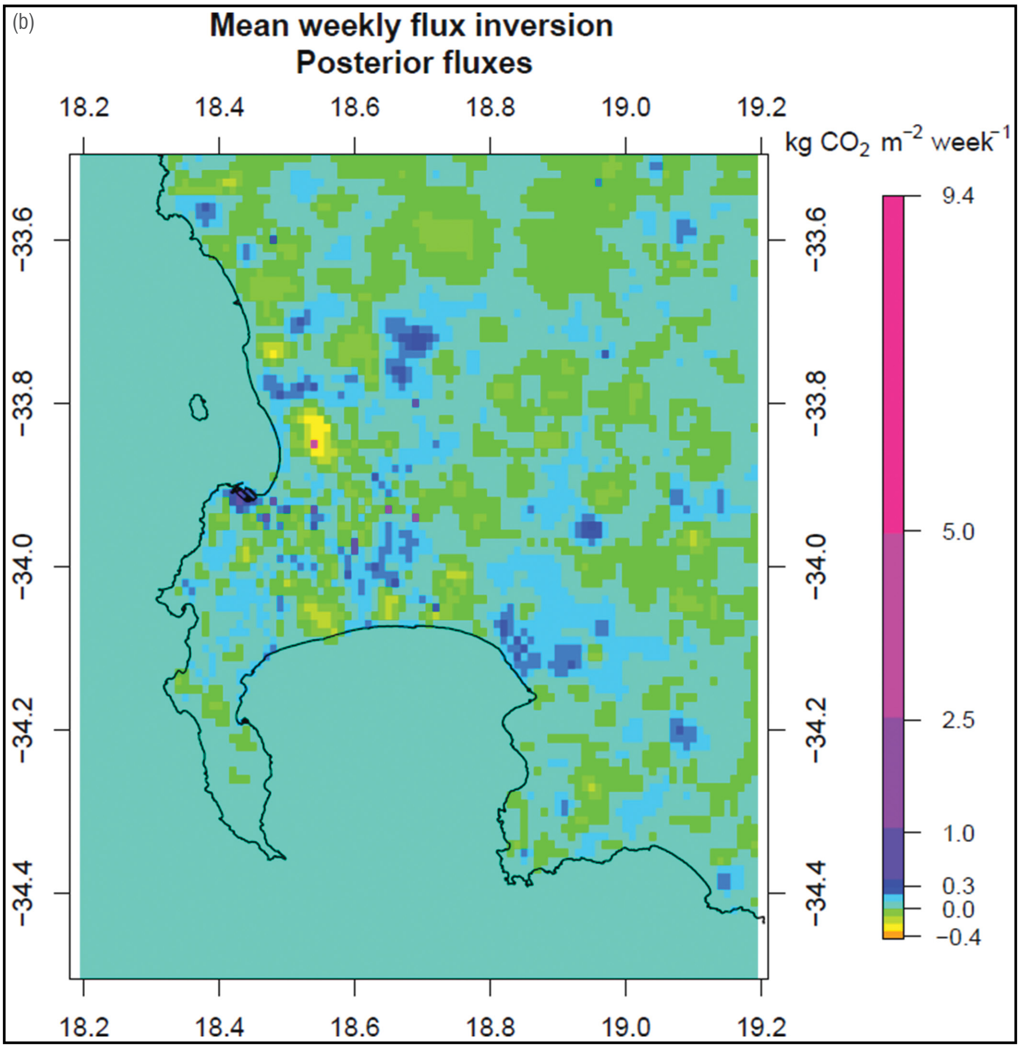
**Figure 2:** Two idealised modelling experiment responses of the global carbon-climate system which have a major influence on the effectiveness of carbon dioxide removal (CDR): (a) the quasi-linear response of warming to increasing cumulative positive emissions contrasted with the non-linear response of cooling to negative cumulative emissions, and (b) the rapid (<20 years) re-balancing of atmospheric $CO_2$ by the land and ocean carbon reservoirs, which is projected to reduce the effectiveness of CDR by 40–60%.[3,17]




*Source: Canadell et al.[3] with permission (©IPCC)*

**Figure 3:** Global synthesis assessment of the scalability, earth system feedbacks, biogeochemical effects and co-benefits of carbon dioxide removal (CDR) interventions. This figure highlights the major science and policy challenges regarding both the scalability as well as the low confidence levels in respect of the magnitude and direction of the responses.[3]




*Source: Nickless et al.[24] under licence CC-BY 4.0*

**Figure 4:** The importance and the challenge of fine spatial scales in SA-RES. (a) A high-resolution model reconstruction of the influence of eddies (±10–100 km) and circulation dynamics on the $CO_2$ gradients in the regional ocean: the warm and low $CO_2$ Agulhas system in the east, the cool and high $CO_2$ Benguela upwelling system in the west and the high $CO_2$ boundary with the Southern Ocean in the south. (b) An inversion model reconstruction of land carbon fluxes in the Western Cape showing the important fine scale (±1–10 km) of the spatial variability in the natural, agricultural, urban and industrial domains of the system[24] (Chang N 2023 January, personal communication).

land and in the ocean. These are critical to understand and project the links and feedbacks between human interventions and the response of natural systems. The integration across these scales presents major modelling and observational challenges, which are well established for positive emissions.[19] In order for global earth system models to run very long projections they need to be set up at medium to low resolutions (25–100 km) that do not capture the spatial and temporal scales critical to the sensitivity to perturbations of the processes.[19] Global model projections at higher resolution are still computationally expensive. The uncoupling of heat and carbon fluxes from the ocean and the slow ecological response of land carbon reservoirs under negative emissions makes it, we suggest, necessary to resolve the finer scales of global earth system models. This, we propose, as a priority towards achieving higher confidence in the projections of non-linear responses that influence the efficacy of CDR. These have profound implications for the resolution choices for models used in projections as well as observations used to evaluate the confidence in those projections. It also presents challenges to integrated assessment models used to understand the sensitivities of the societal–natural systems feedbacks that are critical to the efficacy and scalability and economic outcomes from CDR.

Here we propose that the regional carbon-climate ecosystem science community address this challenge through a dual-linked regional–global observational and modelling approach. There are a number of critical questions: Can a regional focus for the coupled natural–human systems strengthen the global CDR effectiveness and scalability governance policies? Global earth system models implicitly include regional processes and their feedback characteristics on land and in the ocean, so why is there a need for a regional focus?

The proposed Southern African Regional Earth System (SA-RES) with its linked land and ocean ecosystems has unique and highly energetic carbon-climate, biogeochemical and ecological sensitivities and feedbacks that need to be adequately understood.[14,20-23] Studies have highlighted the spatially heterogeneous nature of land and ocean systems, which could influence the effectiveness of CDR interventions (Figure 4a,b).[21,23,24] This is well reflected in the gradients of dissolved $CO_2$ in the ocean (Figure 4a) and the heterogeneous fine-scale gradients of carbon fluxes in the SW Cape (Figure 4b).[24]

Two further aspects of SA-RES that remain particularly weakly constrained are the land–ocean coupling of carbon fluxes from the river basin scale through the estuaries to the coastal and regional ocean as well as the coupling of SA-RES to the Southern Ocean. Recent work on land and the ocean has highlighted that system-scale reconstructions of variability from models and observations are sensitive to small scales of variability in space and time.[20,23,25] These then influence both their suitability for specific interventions and the scaled-up trade-offs, and collectively they define the regional feedbacks, which ultimately set the magnitude of the effectiveness and their net contribution to global negative emissions.

## Final comments

The high-level framing question for this Commentary was: is the South African science–policy community ready and capable of assessing the scalability and effectiveness of regional and/or global CDR interventions? IPCC-AR6 assessments highlighted that CDR is now recognised as the critical global carbon lever to achieve a soft landing for climate within this century. CDR is needed to assist in achieving net-zero by mid-century and then, beyond that, to address the zero emissions commitment from embedded warming from historical emissions. While CDR acts on global warming and climate change through its impact on the global airborne fraction of anthropogenic $CO_2$ (Figure 1), the technological interventions and the resulting feedbacks are likely to be very scale sensitive and regionally differentiated in character. Understanding and projecting the global-scale feedbacks is considered critical to reliably projecting the scalability and effectiveness of CDR and its developmental co-benefits.

Even if ecological and technological interventions were ready and operational at scale, the science is still weak (low confidence) for the land and ocean feedbacks that are likely to have a first-order impact on the evaluation of the effectiveness of regional and global CDR (Figure 1). A regional focus, that builds on South Africa's comparative geographical advantage and catalyses Africa's climate science–mitigation–adaptation nexus is thus proposed as necessary in order to overcome this CDR challenge (Figure 5). This requires an (as-yet non-existent) regional integration of observational and modelling capabilities (Figure 5). The hub science-to-society integration is proposed to be accomplished through a regionally adapted integrated assessment model (SA-IAM). This aims to enable a quantitative examination of the assumptions and trade-offs across the science–society boundary for the coupled human–earth system. The individual capabilities for this already exist and are increasingly better understood by South African scientists and their broader networks (Figure 5). However, these functionalities need to be further developed and coordinated across most national science and policy institutions in support of increasing development choices through skills and avoiding costly errors.

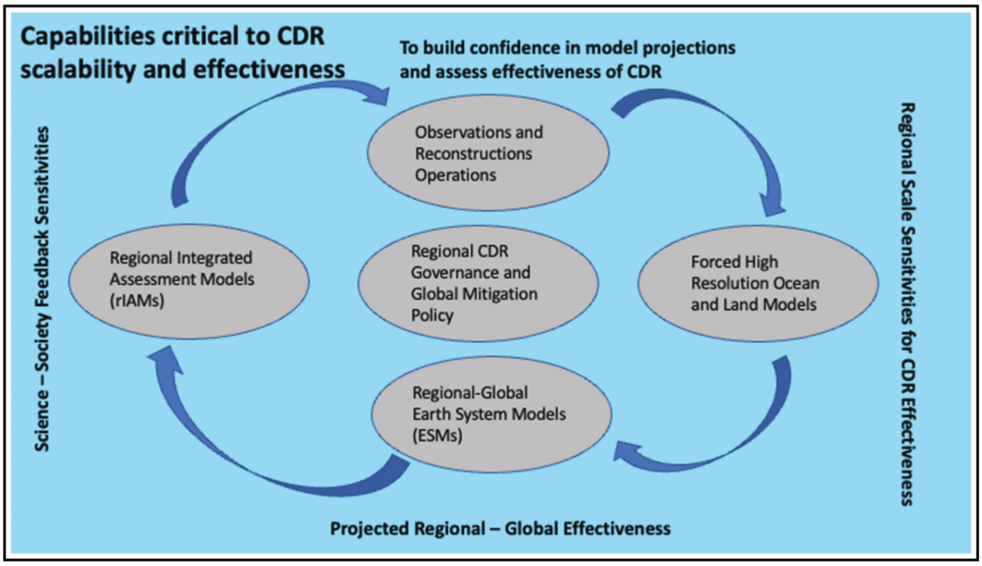



**Figure 5:** A schematic that sets out the proposed value chain of core science–society capabilities critical to a regional carbon dioxide removal (CDR) Science–Policy Hub. A regionally adapted integrated assessment model (IAM) is the platform that links the science to societal needs and investment requirements. The IAM is supported by optimised observing systems to evaluate the effectiveness of regional and global CDR, high-resolution models to evaluate the local trade-offs of interventions, and the earth system models (ESMs) to provide the projections for the effectiveness of CDR.

## Acknowledgements

We acknowledge the constructive comments from the two anonymous readers. P.M.S.M. acknowledges support from the Strategic Fund through the School for Climate Studies at Stellenbosch University.

## Competing interests

We have no competing interests to declare.

## References

1. Rogelj J, Popp A, Calvin KV, Luderer G, Emmerling J, Gernaat D, et al. Scenarios towards limiting global mean temperature increase below 1.5 °C. Nat Clim Chang. 2018;8:325–332. https://doi.org/10.1038/s41558-018-0091-3
2. Lee J-Y, Marotzke J, Bala G, Cao L, Corti S, Dunne JP, et al. Future global climate: Scenario-based projections and near-term information. In: Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, et al., editors. Climate change 2021: The physical science basis. Contribution of Working Group I to the sixth assessment report of the Intergovernmental Panel on Climate Change: Chapter 4. Cambridge: Cambridge University Press; 2023. p. 553–672. https://doi.org/10.1017/9781009157896.006
3. Canadell JG, Monteiro PMS, Costa MH, Cotrim da Cunha L, Cox PM, Eliseev AV, et al. Global carbon and other biogeochemical cycles and feedbacks In: Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, et al., editors. Climate change 2021: The physical science basis. Contribution of Working Group I to the sixth assessment report of the Intergovernmental Panel on Climate Change: Chapter 4. Cambridge: Cambridge University Press; 2023. p. 673–816. https://doi.org/10.1017/9781009157896.007
4. Warszawski L, Kriegler E, Lenton TM, Gaffney O, Jacob D, Klingenfeld D, et al. All options, not silver bullets, needed to limit global warming to 1.5 °C: A scenario appraisal. Environ Res Lett. 2018;16(6), Art. #064037. https://doi.org/10.1088/1748-9326/abfeec
5. Pongratz JA. AIMES Towards climate neutrality: A turning point in land-use history [document on the Internet]. c2023 [cited 2023 May 08]. Available from: https://aimesproject.org/wp-content/uploads/2023/02/AIMES-Bulletin-2023_February_04_Pongratz.pdf
6. Keller DP, Lenton A, Littleton EW, Oschlies A, Scott V, Vaughan NE, et al. The effects of carbon dioxide removal on the carbon cycle. Curr Clim Change Rep. 2018;4:250–265. https://doi.org/10.1007/s40641-018-0104-3
7. OECD. Pricing greenhouse gas emissions: Carbon pricing in South Africa [document on the Internet]. c2021 [cited 2023 May 08]. Available from: https://www.oecd.org/tax/tax-policy/carbon-pricing-south-africa.pdf
8. Marquard A, Ahjum F, Bergh C, Von Blottnitz H, Burton J, Cohen B, et al. Exploring net zero pathways for South Africa – An initial study. Report. Cape Town: University of Cape Town; 2023. https://doi.org/10.25375/uct.22189150.v2
9. Lecocq FH, Winkler JP, Daka S, Fu JS, Gerber S, Kartha V, et al. Mitigation and development pathways in the near- to mid-term. In: Intergovernmental Panel on Climate Change (IPCC), editor. Climate change 2022 – Mitigation of climate change: Working Group III Contribution to the sixth assessment report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press; 2023. p. 409–502. https://doi.org/10.1017/9781009157926.006
10. Bellamy R, Geden O, Fridahl M, Cox E, Palmer J. Editorial: Governing carbon dioxide removal. Front Clim. 2021;3, Art. #816346. https://doi.org/10.3389/fclim.2021.816346
11. Schipper ELF, Revi A, Preston BL, Carr ER, Eriksen SH, Fernandez-Carril LR, et al. Climate resilient development pathways. In: Pörtner H-O, Roberts DC, Tignor M, Poloczanska ES, Mintenbeck K, Alegría A, et al., editors. Climate change 2022: Impacts, adaptation and vulnerability. Contribution of Working Group II to the sixth assessment report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press; 2023. p. 2655–2807. https://doi.org/10.1017/9781009325844.027
12. Ciais P, Tan J, Wang X, Roedenbeck C, Chevallier F, Piao S-L, et al. Five decades of northern land carbon uptake revealed by the interhemispheric $CO_2$ gradient. Nature. 2019;568:221–225. https://doi.org/10.1038/s41586-019-1078-6
13. Keller DP, Brent K, Bach LT, Rickels W. Editorial: The role of ocean-based negative emission technologies for climate mitigation. Front Clim. 2021;3, Art. #743816. https://doi.org/10.3389/fclim.2021.743816
14. Bond WJ, Stevens N, Midgley GF, Lehmann CER. The trouble with trees: Afforestation plans for Africa. Trends Ecol Evol. 2019;34(11):963–965. https://doi.org/10.1016/j.tree.2019.08.003
15. Li X, Zickfeld K, Mathesius S, Kohfeld K, Matthews JBR. Irreversibility of marine climate change impacts under carbon dioxide removal. Geophys Res Lett. 2020;47(17), e2020GL088507. https://doi.org/10.1029/2020GL088507
16. Ehlert D, Zickfeld K. What determines the warming commitment after cessation of $CO_2$ emissions? Environ Res Lett. 2017;12, Art. #015002. https://doi.org/10.1088/1748-9326/aa564a
17. Zickfeld K, MacDougall AH, Matthews HD. On the proportionality between global temperature change and cumulative $CO_2$ emissions during periods of net negative $CO_2$ emissions. Environ Res Lett. 2016;11, Art. #055006. https://doi.org/10.1088/1748-9326/11/5/055006
18. Zickfeld K, Azevedo D, Mathesius S, Matthews HD. Asymmetry in the climate-carbon cycle response to positive and negative $CO_2$ emissions. Nat Clim Chang. 2021;11:613–617. https://doi.org/10.1038/s41558-021-01061-2
19. Hewitt HT, Roberts M, Mathiot P, Biastoch A, Blockley E, Chassignet EP, et al. Resolving and parameterising the ocean mesoscale in earth system models. Curr Clim Change Rep. 2020;6:137–152. https://doi.org/10.1007/s40641-020-00164-w
20. Djeutchouang LM, Chang N, Gregor L, Vichi M, Monteiro PMS. The sensitivity of $pCO_2$ reconstructions to sampling scales across a Southern Ocean sub-domain: A semi-idealized ocean sampling simulation approach. Biogeosciences. 2022;19:4171–4195. https://doi.org/10.5194/bg-19-4171-2022
21. Braby L, Backeberg B, Krug M, Reason C. Quantifying the impact of wind-current feedback on mesoscale variability in forced simulation experiments of the Agulhas Current using an eddy-tracking algorithm. J Geophys Res Oceans. 2020;125(1), e2019JC015365. https://doi.org/10.1029/2019JC015365
22. Hutchings, L, Van der Lingen CD, Shannon LJ, Crawford RJM, Verheye HMS, Bartholomae CH, et al. The Benguela Current: An ecosystem of four components. Prog Oceanogr. 2009;83:1–4. https://doi.org/10.1016/j.pocean.2009.07.046
23. Burt DJ, Fröb F, Ilyina T. The sensitivity of the marine carbonate system to regional ocean alkalinity enhancement. Front Clim. 2021;3, Art. #624075. https://doi.org/10.3389/fclim.2021.624075
24. Nickless A, Rayner PJ, Scholes RJ, Engelbrecht F, Erni B. An atmospheric inversion over the city of Cape Town: Sensitivity analyses. Atmos Chem Phys. 2019;19:7789–7816. https://doi.org/10.5194/acp-19-7789-2019
25. Moncrieff GR, Scheiter S, Slingsby JA. Understanding global change impacts on South African biomes using Dynamic Vegetation Models. S Afr J Bot. 2015;101:16–23. https://doi.org/10.1016/j.sajb.2015.02.004

**AUTHORS:**
Marise C. Breet[1]
Johanna E. de Villiers[2]
Ranel Venter[1]

**AFFILIATIONS:**
[1]Division of Sport Science, Department of Exercise, Sport, and Lifestyle Medicine, Faculty of Medicine and Health Sciences, Stellenbosch University, Stellenbosch, South Africa
[2]Division of Movement Science and Exercise Therapy, Department of Exercise, Sport, and Lifestyle Medicine, Faculty of Medicine and Health Sciences, Stellenbosch University, Stellenbosch, South Africa

**CORRESPONDENCE TO:**
Marise Breet

**EMAIL:**
marise@mbio.co.za

**HOW TO CITE:**
Breet MC, de Villiers JE, Venter R. South African school shoes: Urgent changes required for our children's unique feet. S Afr J Sci. 2023;119(9/10), Art. #16295. https://doi.org/10.17159/sajs.2023/16295

**ARTICLE INCLUDES:**
☐ Peer review
☐ Supplementary material

**KEYWORDS:**
shoe fit, barefoot, foot pathologies, school shoes, children

**PUBLISHED:**
28 September 2023



# South African school shoes: Urgent changes required for our children's unique feet

**Significance:**

Ill-fitting shoes can negatively impact the developing foot. Despite the notable differences in foot shape between South African and European populations, shoe design largely adheres to a universal standard. This discrepancy may arise due to several factors, including limited data on the South African foot morphology, economic constraints, fashion trends, and consumer complacency. As such, the development and manufacturing of school shoes in South Africa face the challenge of accommodating the unique foot shape of the local population. Moreover, consumers should be knowledgeable about the potential long-term adverse consequences associated with wearing ill-fitting shoes.

A recent investigation brought attention to the lack of availability of school shoes in South Africa specifically designed to accommodate the distinctive foot shape of habitually barefoot children.[1] The study revealed that a significant majority of children, approximately 98%, wore shoes that were too narrow for their feet, while 59% wore shoes that were not the appropriate length. These findings corroborate earlier research that found that 67% of children wore ill-fitting school shoes, that is, shoes that were not the correct width and length.[2] It is important to note that both studies assessed foot measurements in a static standing position.[1,2] However, it should be acknowledged that dynamic movements lead to further spreading of the foot, resulting in the lowering of the lateral and medial arches.[3] Consequently, these findings regarding the prevalence of ill-fitting school shoes among school-going children raise even greater concerns.

In compliance with the *South African Schools Act*, the inclusion of a school uniform is deemed vital for attaining social and educational objectives, including the reinforcement of discipline, the creation of an enriched learning environment, and the assurance of overall school safety. Within the realm of South African public schools, the school shoe constitutes an integral component of the mandatory uniform. The legislation emphasises that the uniform should facilitate the child's engagement in all activities with comfort and safety.[4] In South Africa, it is customary for children to spend approximately seven hours per day at school, wearing school shoes. Based on the mentioned research,[1,2] it is possible that many of these children spend many hours in ill-fitting shoes that do not conform to their foot morphology.

The association between wearing improperly fitted footwear and foot health and function is intricate. Foot pain, pes planus, and hallux rigidus in the forefoot can be attributed to or aggravated by inadequate fit, restricted toe box flexibility, and shoes that do not conform to the shape of the foot.[3,5,6] Hallux angle deformities are frequently observed in cases where individuals wear shoes that are either too narrow or too short for their feet.[7-9] A study conducted in South Africa reported that 80% of participants experienced foot pathologies as a result of wearing ill-fitting footwear.[10] However, it is important to note that the study had a limited sample size, consisting of only 60 women.

Foot deformities have been identified as a potential cause of increased medio-lateral postural sway during walking and impaired balance, ultimately impacting an individual's psychological well-being and overall quality of life.[9,11] A noteworthy investigation involving 1238 preschool children from Austria revealed a significant correlation between the risk of developing hallux valgus and the usage of ill-fitting footwear.[12] Furthermore, a compelling study involving monozygotic and dizygotic twins indicated that external factors, such as constricted toe boxes and non-genetic influences, played a crucial role in the development of hallux valgus.[13]

Within numerous African countries, there exists a prevalent inclination among adults and children to engage in daily activities barefoot. A comparative study involving regularly shod Maasai women, partially shod Maasai women, and Korean women demonstrated that Korean women exhibited narrower feet in comparison to both Maasai groups. Notably, the hallux valgus angle was significantly higher among Koreans and the regularly shod Maasai group when contrasted with the partially shod Maasai group.[14] Although hallux valgus tends to be irreversible in adults, interventions targeting children's footwear by adjusting the width to provide ample space and comfort in the toe box hold promise in mitigating the adverse effects associated with this condition.[3,9]

This inquiry prompts the question: what is the prevailing scenario within the South African context? In a study conducted by Hollander et al.[15], it was revealed that children and adolescents in South Africa who habitually go barefoot exhibited a noticeably higher degree of foot flexibility but also a larger hallux valgus angle compared to their German counterparts. The increased hallux angle observed in the barefoot population was unexpected. It is plausible that the prescribed school uniform, which includes more restrictive footwear, may have contributed to the relatively constrained conditions experienced by the habitually barefoot group as compared to their shod counterparts.[15,16]

The intrinsic association between footwear and pathology necessitates the adherence of footwear to specific requirements tailored to the population in question. Various factors should be taken into consideration when designing shoes, including the fit, weight, structure, motion control properties, material composition, and cushioning. Previous studies have acknowledged these general components in the context of South African school shoe manufacturers.[6,17,18] However, it has been observed that these manufacturers have not adequately addressed the fundamental foot measurements necessary to accommodate the unique shape of the South African child's foot. It is important to note that this oversight may not be attributable to ignorance but rather a lack of available data on the foot morphology of South African children and adolescents. Fortunately, recent research endeavours have yielded valuable insights into the shape of habitually barefoot South African children and adolescents, which can be instrumental in guiding the design of shoes that promote healthy foot development in this population.[1]

The question arises as to why South African companies should implement changes based on foot measurements and adapt the current South African school shoe. Several international studies have reported variations in foot shape, particularly in terms of forefoot width, among populations from different continents.[19-21] These studies provide valuable insights into the potential differences in foot morphology across diverse ethnicities and regions.

One study compared four different ethnic groups from Japan, Indonesia, France, and Australia, revealing notable variations in foot dimensions.[21] While the French and Australian samples exhibited longer feet, their feet were relatively narrower compared to those of the Japanese and Indonesian populations. Additionally, the French male individuals had wider feet in comparison to their Australian counterparts. Most foot dimension measurements were similar between the French and Australian samples, except for foot width, where differences were observed. These findings underscore the importance of considering foot width variations when designing and manufacturing footwear.

In more recent research, Mauch et al.[20] conducted a comparative study involving children from Germany and Australia. They found that German preschool children had longer feet, larger ball-of-foot circumferences, and higher dorsal arches than their Australian counterparts. German primary school children also exhibited longer feet, greater ball-of-foot circumference, lower dorsal arch height, and a significantly larger ball-of-foot angle compared to the Australian sample. These findings highlight additional differences in foot morphology between different populations and emphasise the need for tailored footwear solutions.

The underlying mechanisms responsible for the observed inter-continental variations in foot morphology remain unclear.[20] However, the findings suggest that population-specific factors, such as genetics and environmental influences, may contribute to these differences.

The design of a shoe relies on primary data derived from the foot's shape and dimensions, with the primary objective of providing foot protection. To achieve this, a comprehensive understanding of foot biomechanics and structure is essential for the design of the shoe last (the form around which the shoe is constructed).[3,22] It has been observed that South African shoe manufacturers typically employ a shoe design based on the British Mondo Point System, which primarily utilises foot length as a key measurement and increases the shoe's girth in standardised increments.[3,10] As a result, most children can find shoes that accommodate their foot length by selecting the appropriate size; however, options for accommodating foot width are limited.

To ensure a proper fit for the shoe, it is crucial to incorporate a flexible toe box and an articulation line that aligns with the base of the joints. Nevertheless, many shoe manufacturers offer a standard width that may be suitable for most populations accustomed to wearing shoes, while individuals from habitually barefoot populations may encounter difficulties in finding suitable shoe widths. In the past, certain shoe companies have produced half-size shoes specifically for wider feet. However, these sizes are currently scarce in South Africa, possibly due to economic constraints. Considering the prevailing situation concerning the appropriate width of school shoes, it is plausible that the South African adult population also faces challenges in finding shoe designs that align with their foot morphology. If the foundational structure of the shoe last does not align with the shape of the foot, achieving a proper fit for the shoe would prove challenging.[23,24]

The *South African Schools Act 84 of 1996* places significant emphasis on adopting a 'Proudly South African' approach when formulating school uniforms within South Africa.[4] Pertinent questions and challenges can be directed towards the school shoe industry, parents, and educators. One key question is whether the focus on South African manufacturers extends to South African school shoes. Despite clear evidence highlighting differences in foot morphology, there remains a concern regarding the perpetuation of British and American foot morphology data dominating the South African shoe manufacturing industry. Another crucial concern is whether fashion and aesthetic considerations will continue to overshadow the fundamental aspects of shoe measurements specific to the South African population. Furthermore, it is necessary to explore whether parents and educators are prepared to acknowledge and embrace the necessity for research-based shoe designs for their growing children.

Existing research consistently supports the notion that shoes, as an external factor, can have a negative impact, particularly when they are too narrow in the toe box. This evidence challenges the South African shoe industry to adapt the prevailing British lasts used in the production of school shoes for children who are habitually barefoot. It is imperative that prospective longitudinal studies in the future examine the repercussions of ill-fitting shoes during childhood on the development of foot malformations in adulthood.[24]

## Acknowledgements

We thank Carel du Toit for his valuable assistance with proofreading and language editing.

## Competing interests

We have no competing interests to declare.

## References

1. Breet MC, Venter R. Are habitually barefoot children compelled to wear ill-fitting school shoes? A cross-sectional study. BMC Pediatr. 2022;22(1):1–9. https://doi.org/10.1186/s12887-022-03263-9
2. De Villiers JE. Growing feet: Foot metrics and shoe fit in South African school-aged children and adolescents [PhD thesis]. Stellenbosch: Stellenbosch University; 2017. https://scholar.sun.ac.za/handle/10019.1/101191
3. Cheng FT, Perng DB. A systematic approach for developing a foot size information system for shoe last design. Int J Ind Ergon. 1999;25(2):171–185. https://doi.org/10.1016/s0169-8141(98)00098-5
4. South African Schools Act No 84 of 1996. Government Gazette. 1996;(84). Available from: https://www.gov.za/documents/south-african-schools-act
5. Hurst B, Branthwaite H, Greenhalgh A, Chockalingam N. Medical-grade footwear: The impact of fit and comfort. J Foot Ankle Res. 2017;10(1):Art. #2. https://doi.org/10.1186/s13047-016-0184-z
6. Buldt AK, Menz HB. Incorrectly fitted footwear, foot pain and foot disorders: A systematic search and narrative review of the literature. J Foot Ankle Res. 2018;11:Art. #43. https://doi.org/10.1186/s13047-018-0284-z
7. Klein C, Groll-Knapp E, Kundi M. Increased hallux angle in children and its association with insufficient length of footwear: A community based cross-sectional study. BMC Musculoskelet Disord. 2009;10(1):1–9. https://doi.org/10.1186/1471-2474-10-159
8. Wolf S, Simon J, Patikas D, Schuster W, Armbrust P, Döderlein L. Foot motion in children shoes – A comparison of barefoot walking with shod walking in conventional and flexible shoes. Gait Posture. 2008;27(1):51–59. https://doi.org/10.1016/j.gaitpost.2007.01.005
9. Javaid Q. Associated factors of hallux valgus and the deformity impact on quality of life with and without surgery – A review. Isra Med J. 2021;13(3):222–226.
10. Thompson ALT, Zipfel B. The unshod child into womanhood – Forefoot morphology in two populations. Foot. 2005;15(1):22–28. https://doi.org/10.1016/j.foot.2004.10.002
11. Hurn SE, Vicenzino B, Smith MD. Functional impairments characterizing mild, moderate, and severe hallux valgus. Arthritis Care Res. 2015;67(1):80–88. https://doi.org/10.1002/acr.22380
12. Kinz W, Groll-Knapp E, Kundi M. Hallux valgus in pre-school-aged children: The effects of too-short shoes on the hallux angle and the effects of going barefoot on podiatric health. Footwear Sci. 2021;13(1):29–42. https://doi.org/10.1080/19424280.2020.1853826
13. Munteanu SE, Menz HB, Wark JD, Christie JJ, Scurrah KJ, Bui M, et al. Hallux valgus, by nature or nurture? A twin study. Arthritis Care Res. 2017;69(9):1421–1428. https://doi.org/10.1002/acr.23154
14. Choi JY, Babu H, Joseph FN, Stephanie S, Suh JS. Effects of wearing shoes on the feet: Radiographic comparison of middle-aged partially shod Maasai women's feet and regularly shod Maasai and Korean women's feet. Foot Ankle Surg. 2018;24(4):330–335. https://doi.org/10.1016/j.fas.2017.03.012

15. Hollander K, De Villiers JE, Sehner S, Wegscheider K, Braumann KM, Venter R, et al. Growing-up (habitually) barefoot influences the development of foot and arch morphology in children and adolescents. Sci Rep. 2017;7(1):Art. #8079. https://doi.org/10.1038/s41598-017-07868-4

16. Wang Y, Jiang, H, Yu, L, Gao, Z, Liu, W, Mei, Q, et al. Understanding the role of children's footwear on children's feet and gait development: A systematic scoping review. Healthcare. 2023;11:Art. 1418. https://doi.org/10.3390/healthcare11101418

17. Yurt Y, Sener G, Yakut Y. Footwear suitability in Turkish preschool-aged children. Prosthet Orthot Int. 2014;38(3):224–231. http://journals.sagepub.com/doi/10.1177/0309364613497047

18. Barton CJ, Bonanno D, Menz HB. Development and evaluation of a tool for the assessment of footwear characteristics. J Foot Ankle Res. 2009;2:Art. #10. https://doi.org/10.1186/1757-1146-2-10

19. Sacco ICN, Onodera AN, Bosch K, Rosenbaum D. Comparisons of foot anthropometry and plantar arch indices between German and Brazilian children. BMC Pediatr. 2015;15(1):Art. #4. https://doi.org/10.1186/s12887-015-0321-z

20. Mauch M, Mickle KJ, Munro BJ, Dowling AM, Grau S, Steele JR. Do the feet of German and Australian children differ in structure? Implications for children's shoe design. Ergonomics. 2008;51(4):527–539. https://doi.org/10.1080/00140130701660520

21. Kouchi M. Foot dimensions and foot shape: Differences due to growth, generation and ethnic origin. Anthropol Sci. 1998;106(suppl.):161–188. https://doi.org/10.1537/ase.106.supplement_161

22. Putti AB, Arnold GP, Abboud RJ. Differences in foot pressures between Caucasians and Indians. Foot Ankle Surg. 2010;16(4):195–198. https://doi.org/10.1016/j.fas.2009.11.004

23. Walther M, Herold D, Sinderhauf A, Morrison R. Children sport shoes – A systematic review of current literature. Foot Ankle Surg. 2008;14:180–189. https://doi.org/10.1016/j.fas.2008.04.001

24. González Elena ML, Córdoba-Fernández A. Footwear fit in schoolchildren of southern Spain: A population study. BMC Musculoskelet Disord. 2019;20(1):Art. #208. https://doi.org/10.1186/s12891-019-2591-3

# Load shedding as a result of failures at the political-technological interface

AUTHOR:
Mike Muller[1]

AFFILIATION:
[1]School of Governance, University of the Witwatersrand, Johannesburg, South Africa

CORRESPONDENCE TO:
Mike Muller

EMAIL:
Mike.Muller@wits.ac.za

HOW TO CITE:
Muller M. Load shedding as a result of failures at the political-technological interface. S Afr J Sci. 2023;119(9/10), Art. #16595. https://doi.org/10.17159/sajs.2023/16595

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
electricity supply, developmental state, technology policy, megaproject management

PUBLISHED:
31 August 2023

**Significance:**

In effective developmental states, technocrats are 'embedded' in the political system with sufficient autonomy to undertake their tasks. South Africa's current electricity crisis is attributed here, in part, to an initial mistrust between the country's new political leadership and its 'old-order' technocrats following the political transition of 1994. This trust deficit led to policy missteps in the development of new electricity generation. The impact of these missteps was compounded by the adoption of a risky, politically driven, project management strategy. The outcome was not just substantial cost increases but the project delays that resulted in the current 'load shedding'.

Electricity 'load shedding' is not new to South Africa. When demand exceeds supply, distributors have little choice but to restrict use. Limited ability to meet the rapid growth in demand from the gold mining industry after World War II saw supplies rationed.[1] Along with other users, the mining industry was again hit hard by rolling power cuts in early 2008. However, South Africa's current inability to generate sufficient electricity to meet the needs of its people and economy can be attributed to a set of poor decisions (and, indecisions) dating from the 1994 political transition.

This is not to say that the current load shedding is the consequence of the transition to non-racial democracy. The argument focuses rather on the nature of that transition. Specifically, it considers the relationship between an existing, inherently technological, public institution and the new political institutions which positioned themselves as the leaders of a developmental state.

In particular, there was a failure to coordinate the parallel transitions of the country's political institutions and of Eskom, its largest public enterprise. This experience also highlights future risks if a coherent national strategy for energy transition is not established.

## Theoretical background – the role of trust in governance networks

Eskom's institutional challenges are common in contemporary democracies in which important, long-term, technical decisions require political support or sanction for their implementation. The problem is that the outlook of the political institutions on which they depend is inherently short term, defined by contested election cycles.

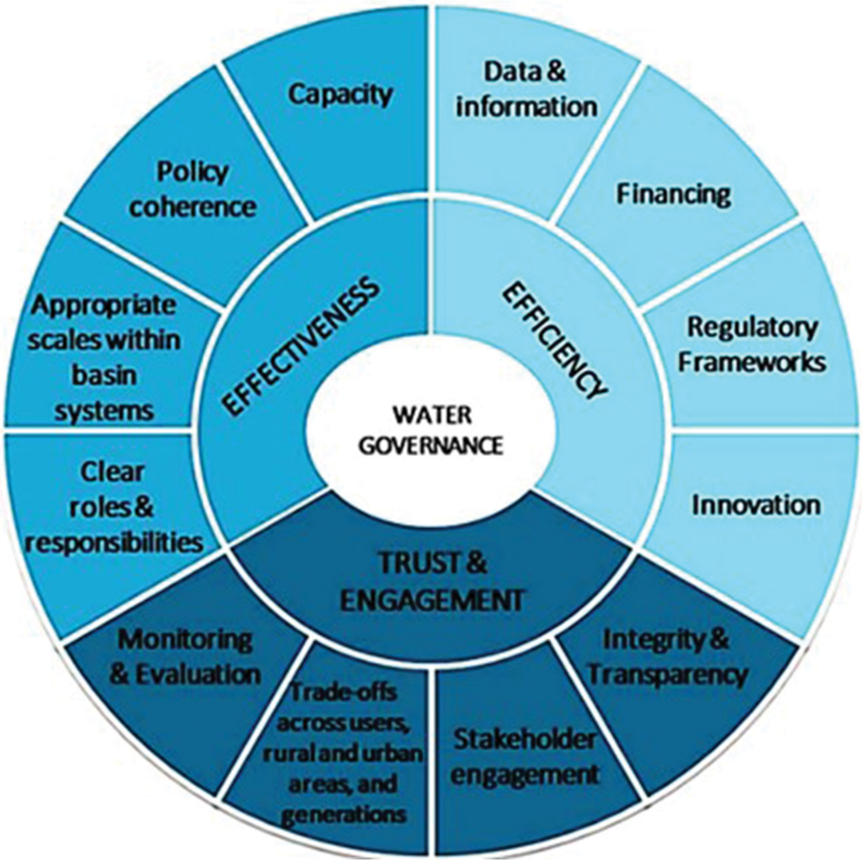
Political support may be provided formally or informally, through a wide range of institutional frameworks that allow the 'technological community' to present their proposals to the 'political community' for evaluation and decision. However, to overcome the 'timing' difference and achieve mutually acceptable outcomes, there must be effective communication and understanding of each other's concerns between the two communities. While formal institutional processes are necessary, they are often not sufficient to allow engagement to proceed to a successful conclusion. To achieve alignment across the diverse and complex domains of technology and politics, an understanding of each other's priorities and constraints underpinned by mutual trust is required.[2] In the related water sector, trust is explicitly identified, alongside efficiency and effectiveness, as a key pillar of successful sector governance.[3]

In successful 'developmental states' and similar contexts, trust relationships based on common histories and backgrounds, play an important role in 'greasing the wheels' of decision-making.[4] Trust between parties is particularly important in the governance of inherently complex operations such as the provision of electricity and water at a national scale. These activities require technical and institutional systems operating across diverse national geographies to supply a diverse community of users, ranging from the poorest individual households to industries of global scope and scale.

Building such systems requires decadal foresight; their financing requires access to large pools of exacting capital backed by credible commitments to pay for it; in operation, the disciplined cooperation of a large, dispersed skilled workforce is needed. But it also requires that the interests of politically powerful actors in the wider society are identified and addressed.

The institutional arrangements needed for Eskom to achieve the desired technical outcome of reliable and affordable electricity supplies had taken decades to establish. There had been substantial missteps, notably a period of substantial over-investment in the 1970s[5], compounded by subsequent economic and political developments. However, by 1994, the organisation was working well[6].

Political transition disrupted the equilibrium because electricity's long-run planning and operational priorities were secondary to the political priorities of the time. Some new political leaders believed that the priorities of the energy sector's leadership were antithetical to their own. And the nature of the transition from minority rule to democratic government was that the technocrats who sought to guide the policymakers had, in large measure, lost their seats at the primary policy tables. These conflicts were aggravated by an international discourse hostile to public enterprise, actively promoting private-sector approaches to the management of traditionally public utility functions.[7]




*Source: OECD*[3]

**Figure 1:** The role of trust in utility governance.

## The initial evolution of the political-technical dialectic

The emergence and growth of Eskom as a public electricity supplier at national scale was integral to South Africa's economic development in the 20th century. The organisation was an early product of the long-running partnership between politics and science policy of Premiers Smuts and Hertzog and Hendrik van der Bijl, their scientific and industrial advisor[8], which has been characterised by some as an early example of developmental state policies[9].

From its creation in 1922, ESCOM (the Electricity Supply Commission as it was initially) expanded rapidly, displacing small municipal and private electricity providers and supplying the state-owned railways and harbours. In 1948, it became the dominant supplier, taking over the highly profitable operations of the Victoria Falls and Transvaal Power Company's coal-fired power stations. The use of a public enterprise to supply the country's electricity was in line with van der Bijl's and Smuts' vision of strategic state participation in the economy.[10] This 'nationalisation' was supported by the private gold mining industry as a non-profit public corporation would provide cheaper electricity.

The public utility was not immune to constraints on growth and suffered an early period of load shedding at the end of World War II in 1945, when electricity demand from gold mines grew rapidly as they went deeper. Unable to increase supply fast enough, due to post-war shortages, ESCOM had to ration supplies to avoid unplanned 'load shedding' interruptions.

Thereafter, ESCOM's generation capacity was increased rapidly, more than doubling by the end of the 1950s and continuing over the following decade, supported by an expanded transmission network. But demand grew erratically. In 1975, when reserve margins fell to just 11% (well below the 15% target), the response was to order new capacity. But, in 1982, with over 20 000 MW on order, demand began to slow and, by 1992, the reserve margin had risen to nearly 40%.

Some critics attributed this excess capacity to over-enthusiastic investment rather than economic stagnation due to political pressures on the apartheid government which promoted Eskom's privatisation.[11] But the critique also led to policy missteps by the new government.[12]

## Seeking a common agenda: 1994–1998

Cheap, plentiful electricity had been a strategic priority for the authoritarian security state whose investment decisions were not accountable to democratic process. The 1994 political settlement changed that dynamic. For many actors in the new government, the electricity sector's strategic decisions were of secondary importance. Their priority was to assert political control over key national institutions. To this end, the *Eskom Amendment Act 1998* gave government formal control over Eskom as sole shareholder, including decision-making powers about future energy investments.

Significant ambiguity remained in government's attitudes. Conservative characterisations of Eskom as an unwieldy and inefficient enterprise that should be privatised[11] informed the Reconstruction and Development Programme's (RDP) call for the establishment of a

> *powerful, independent national electricity regulator … to enforce public policy, ensure long-term financial viability, assure environmental sustainability and act as an ombuds in the event of conflicts between consumers, government and the electricity industry.*[13]

The option of allowing greater private participation in the system also responded to global policy pressures for developing country governments to open their utility markets to private providers.[7]

The more pragmatic position was that while Eskom had been a pillar of the apartheid private-sector economy, its substantial capabilities could be redirected. In addition to accelerating the "electricity for all" programme, the RDP explicitly said: "the benefits of cheap electricity presently enjoyed by large corporations must be extended to all parts of the economy." Electricity could "increase the level of mineral beneficiation ….. (and) employment and add more value to our natural resources before export".[12]

Eskom's technocratic leadership had anticipated the demands of the new democratic dispensation as well as some of the criticisms. The generation surplus was propitious, allowing Eskom to concentrate on expanding electrification to formerly unserved black areas. It enabled Eskom to support minerals beneficiation in energy-intensive industries such as aluminium and chrome smelting.[14] And, to further demonstrate its commitment to the new national interests, Eskom proposed to foster regional economic integration, initially selling cheap coal-fired electricity to southern African countries and later importing hydropower from them.

In retrospect, these initiatives reinforced the view that Eskom had no need for new generation capacity and could focus on expanding household services. Eskom's leadership failed to communicate effectively to the politicians that the inherited surplus capacity was a temporary windfall and that a long timeframe was required to plan and build new generation capacity. So, the political leadership failed to identify and address the critical interventions required to sustain a reliable supply of electricity.

## 1998–2004: Continued political distraction, growing technical concern

Alongside the legislative change to Eskom's mandate, government's 1998 Energy White Paper suggested that the conservative view had prevailed. It promised discussion about public sector investments but also asserted that competition in the sector would move the country towards a competitive electricity market.

Two major decisions were taken. The electricity distribution function would be shifted from Eskom and municipalities to six new regional electricity distributors (REDs). On the supply side, 30% of Eskom's generating capacity would be sold to private investors, a separate transmission company established, and an electricity market introduced to ensure competition between different electricity producers. These reforms were to begin in 2003.[15]

These proposals for liberalisation met predictable political opposition. Over the next decade, the establishment of regional electricity distributors was abandoned due to constitutional objections from both local and provincial governments, unhappy about the loss of functions, funds and influence. "The entire plan appears to be based on the assumption that municipalities can ultimately be compelled to transfer their electricity distribution to a RED" wrote one commentator.[16] Meanwhile, the proposed restructuring of Eskom was not helped by its association with increased private sector involvement and attempts to develop a new type of nuclear reactor.

These issues diverted attention from the growing challenge of ensuring adequate generation capacity. Belatedly, in the face of the new democracy's first major instance of load shedding, Minister of Public Enterprises Alec Erwin acknowledged that:

> *…. In 1998, energy supply was not a major issue, and government thought it was important to focus on energy distribution. Eskom should have at the time been asked to build a base load station, in order to avert the crisis…. The Ministry and Department acknowledge that they incorrectly predicted the short to medium term issues.*[17]

## 2004 – political funding and football aggravate the failures

The electricity policy environment was changed dramatically in May 2004 with South Africa's selection to host the football world cup in 2010. This put a spotlight onto the practical and reputational costs of electricity supply shortfalls, not least because the South African government had to report on its readiness to external agencies.

Football catalysed a national infrastructure programme involving major transport and stadium investments. Secure electricity supply was, however, a prerequisite and it was suddenly recognised that reserve margins were declining rapidly.

Although government had "prohibited Eskom from adding new generation capacity in the expectation that the private sector would do so," Eskom had continued to prepare "Project Alpha" (later Medupi). It was thus able to respond when, by 2004, "it became apparent that the anticipated private sector response was not forthcoming and power deficits would occur around 2007"[18].

A more immediate political priority was that the ruling party needed funds to contest the 2006 local government elections. It is well documented[19] that, shortly after Mohamed Valli Moosa, a member of the ANC's National Executive and finance committees, was appointed as Eskom chairman in August 2005, the Japanese firm Hitachi won the largest contract in Eskom's history to supply boilers for Medupi. Controversially, Chancellor House, an investment arm of the governing ANC, partnered with Hitachi and benefitted from this contract.

A Public Protector's investigation found that there "was a conflict between the personal interest of Mr Moosa in the ANC and his duty towards Eskom" when it awarded the contract to the Hitachi Consortium, in which the ANC had an interest and that "Mr Moosa failed to manage his said conflict of interests …. and therefore acted improperly" but that the award was "not in any way affected by Mr Moosa's improper conduct"[19].

However, there were serious consequences for Eskom and the security of South Africa's electricity supply. Although Eskom had not built a power station for 20 years, it took direct responsibility for managing the project rather than giving 'turnkey' responsibility to a single EPC (engineering, procurement and construction) contractor to deliver the project on time and on budget. This sub-contracting strategy, which allowed the ANC associated company to share the profits in return for political connections, was found by the US Securities and Exchange Commission to be a corrupt practice for which Hitachi paid a USD19 million settlement.[20]

The cost to Eskom (and South Africa) was much higher. Eskom awarded and supervised a multiplicity of separate sub-contracts and was liable for cost increases when, for instance, a change in equipment specifications delayed other contractors. Poor coordination, a 5-year delay and a series of defects more than doubled the costs. But it also reduced Eskom's performance because maintenance of other stations was delayed. Energy availability declined from 81.9% in 2011/2012 to a 2022/2023 low of 58%.[21]

## Post hoc: A case of the megaproject challenge?

It has been suggested that the decision to build Medupi and Kusile was an expensive error and that renewables would have been a better solution.[22] This proposition is ahistorical. In 2005, it was expected that some coal-fired generation would continue for another 40 years. Introduction of new, more efficient generators (in terms of $CO_2$ emission per MWh) would allow closure of older, dirtier generators, making an immediate contribution to $CO_2$ mitigation by enabling more electricity to be generated without increasing emissions.[23] Meanwhile, cost comparisons between renewables and coal often failed to take account of the pace at which renewables could be integrated into the national grid.[24]

In 2014, this long-term perspective still guided global strategies. On the sidelines of a global meeting, Chinese and US climate negotiators argued amicably about "whose coal-fired power stations were most efficient and who was going to reduce their emissions by how much and how quickly"[25]. The Chinese delegate acknowledged that China had planned on reaching 'peak coal' in 2035 but was considering bringing it forward to 2030.[25] South Africa's own peak-plateau–decline strategy, as presented in both the National Climate Change Response Policy (2011) and the National Development Plan (2012), envisaged that emissions would only start to decline in 2035.[26] The choice of supercritical technology for Medupi and Kusile was thus consistent with agreed national policy.

With hindsight, should the risks of completion delays and cost overruns typical of large infrastructure projects[27] have been recognised and addressed? There are project management methodologies to mitigate these challenges, notably the appointment of an overarching EPC contractor. The evidence is that this more robust contracting methodology was not adopted because it would have conflicted with the parallel objective of channelling a portion of the project proceeds to the ruling political party.

The counterfactual suggestion that a *programme* to build renewable generators would have been more reliable and less risky than two large capital projects may also be misleading.[28] While the *project* challenges of procurement and on-site coordination may be addressed by such an alternative, coordination of the different elements of a huge *multi-project programme* between the various loci of decision and control poses similar risks.

South Africa's renewables programme requires the engagement and coherence of approach of many different parties. These include generation project developers, transmission grid managers, system operators responsible for storage and technical regulation as well as payment arrangements together with economic regulators, financiers and spatial planning authorities. Current experience with the rollout of renewables[29] is that these pose equally difficult problems of coordination and impose similar delays and costs.

## Conclusion

While subsequent corruption contributed to Eskom's declining performance, earlier failures of governance and management at the interface between political and technical spheres played a major role. Periods of load shedding due to inadequate generation capacity had occurred both before and after 1994, but their extent is now more persistent and systemic. The 'root causes' include the initial failure of contemporary political leaders to respond timeously to technical advice to expand generating capacity.

Interventions to avoid electricity shortages before and during the 2010 football world cup, accelerated the performance decline, reducing Eskom's planned maintenance and causing a vicious cycle of higher demands on working plants and more unplanned maintenance. Delays in the completion of major new generation projects and their subsequent poor performance compounded pressures on the system. These project management failures were, in part, the consequence of the earlier decision that Eskom would manage the projects directly rather than appointing a 'turnkey' contractor to ensure coordination.

The subsequent deterioration in management control together with personal and political corruption further undermined Eskom's performance. The resulting organisational climate is not conducive to remedial action which has, at times, been actively subverted.

This evidence suggests that the early failure of political and technical leadership to develop a coherent common understanding of South Africa's electricity challenges contributed substantially to the current crisis. In the early years of South Africa's democracy, the leaderships of old technical institutions and new political institutions came from social and political communities with different cultures and values and a history of conflict, resulting in failures of communication at that interface.

Complex technical systems like electricity require competent and empowered technical leadership to achieve effective performance and trust and active cooperation between technical and political leadership is essential where political decisions supervene. The load shedding experience thus highlights broader challenges and opportunities. Effective relationships between technical and political spheres help to manage tensions between the political short-term and longer-term technical priorities and to negotiate and implement policy when there are contests over strategy between different societal interest groups.

Debilitating conflicts are aggravating the electricity sector's current dysfunction as it grapples with the long-term challenges inherent in the energy transition. Urgent interventions are blocked by sectoral interest group 'lawfare' and countervailing political action, increasing the likelihood of supply curtailment and failure.[30] While the 'old' load shedding resulted from political mistrust between historically opposed forces, the 'new' load shedding may be the result of leaderships' failures to develop and embrace a robust, collective strategy for South Africa's energy transition.

## Competing interests

I have no competing interests to declare.

## References

1. Steyn G. Governance, finance and investment: Decision making and risk in the electric power sector [doctoral dissertation]. Brighton: University of Sussex; 2001.
2. Klijn EH, Edelenbos J, Steijn B. Trust in governance networks: Its impacts on outcomes. Admin Soc. 2010;42(2):193–221. https://doi.org/10.1177/0095399710362716
3. Organisation for Economic Co-operation and Development (OECD). OECD principles on water governance [webpage on the Internet]. c2015 [cited 2023 Apr 14]. Available from: www.oecd.org/governance/oecd-principles-on-water-governance.htm
4. Muller M. Decolonising engineering in South Africa – Experience to date and some emerging challenges. S Afr J Sci. 2018;114(5/6), Art. #a0270. https://doi.org/10.17159/sajs.2018/a0270
5. Eberhard A. The political economy of power sector reform in South Africa. In: Victor D, Heller T, editors. The political economy of power sector reform: The experiences of five major developing countries. Cambridge: Cambridge University Press; 2007. p. 215–253. https://doi.org/10.1017/CBO9780511493287.007
6. Kessides I, Bogetic Z, Maurer L. Current and forthcoming issues in the South African electricity sector. World Bank Policy Research Working Paper No. 4197. SSRN. 2007 April 1. Available from: https://ssrn.com/abstract=979669
7. Muller M. The regulation of network infrastructure beyond the Washington consensus. Dev South Africa. 2013;30(4–5):674–686. https://doi.org/10.1080/0376835X.2013.830558
8. Schonland BFJ. Hendrik Johannes van Der Bijl: 1887–1948. Obituary notices of Fellows of the Royal Society. 1950;7(19):27–34. https://doi.org/10.1098/rsbm.1950.0003
9. Freund B. A ghost from the past: The South African developmental state of the 1940s. Transformation: Critical Perspectives on Southern Africa. 2013;81:86–114. https://doi.org/10.1353/trn.2013.0007
10. Christie R. Electricity, industry and class in South Africa. London: Palgrave Macmillan; 1984. https://doi.org/10.1007/978-1-349-07030-5
11. De Villers WJ. Report of the commission of inquiry into the supply of electricity in the Republic of South Africa. Pretoria: Government Printer; 1984.
12. Horwitz RB. The uneasy relation between political and economic reform in South Africa: The case of telecommunications. Afr Aff. 1994;93(372):361–386. https://doi.org/10.1093/oxfordjournals.afraf.a098725
13. African National Congress (ANC). Reconstruction and Development Programme. Johannesburg: ANC; 1993.
14. Kessides JK. The decline and fall of Eskom: A South African tragedy. Report 45. London: The Global Warming Policy Foundation; 2020.
15. National Treasury. Intergovernmental fiscal review. Republic of South Africa; 2003.
16. Johnson V. Electricity restructuring short circuiting municipalities. Local Government Bulletin volume 4 number 2. Cape Town: Dullah Omar Institute; 2002.
17. Parliamentary Monitoring Group. Power outages: Parliamentary briefing by minister and Eskom [webpage on the Internet]. c2008 [cited 2023 Apr 14]. Available from: https://pmg.org.za/committee-meeting/8809/
18. World Bank. Eskom investment support project – Implementation completion and results report. Washington DC: World Bank; 2022.
19. Public Protector. Report on an investigation into an allegation of improper conduct by the former chairperson of the board of directors of Eskom. Report no: 30 of 2008/9. Pretoria; 2008.

20. US Securities and Exchange Commission (SEC). SEC charges Hitachi with FCPA violations [media release on the Internet]. c2015 [cited 2023 Apr 14]. Available from: https://www.sec.gov/news/press-release/2015-212#

21. Mathe ZT. Focus on improving the energy availability factor as a quick win to stop load-shedding. Mail & Guardian. 2023 January 16. Available from: https://mg.co.za/opinion/2023-01-16-focus-on-improving-the-energy-availability-factor-as-a-quick-win-to-stop-load-shedding/

22. Blignaut J. Climate change: The opportunity cost of Medupi and Kusile power stations. J Energy South Afr. 2012;23(4):67–75. https://doi.org/10.17159/2413-3051/2012/v23i4a3179

23. Fouilloux J, Otto M. Medupi and Kusile: Supercritical giants of South Africa. Modern Power Systems. 2009 April 01. Available from: https://www.modernpowersystems.com/features/featuremedupi-and-kusile-supercritical-giants-of-south-africa/

24. Bohlmann J, Bohlmann H, Inglesi-Lotz R, Van Heerden J. An economy-wide evaluation of new power generation in South Africa: The case of Medupi and Kusile. Energy Policy. 2016;97:450–460. https://doi.org/10.1016/j.enpol.2016.07.020

25. Muller M. SA needs to make decisions on energy and act on them. Business Day. 2014 November 25. Available from: https://www.businesslive.co.za/archive/2014-11-25-sa-needs-to-make-decisions-on-energy-and-act-on-them/

26. DFFE. South Africa's low emission development strategy 2050. c2020 [cited 2023 Apr 14]. Available from: https://www.dffe.gov.za/sites/default/files/docs/2020lowemission_developmentstrategy.pdf

27. Flyvbjerg B. What you should know about megaprojects and why: An overview. Proj Manag J. 2014;45(2):6–19. https://doi.org/10.1002/pmj.21409

28. Gregory J. Governance, scale, scope: A review of six South African electricity generation infrastructure megaprojects. Utilities Policy. 2020;66:101103. https://doi.org/10.1016/j.jup.2020.101103

29. Parker D. Delays, inconsistency and lack of clarity blocking South African renewables rollout. Engineering News. 2022 February 11. Available from: https://www.engineeringnews.co.za/article/delays-inconsistency-and-lack-of-clarity-blocking-south-african-renewables-rollout-2022-02-11

30. Muller M. Fracking lawfare may obstruct the just transition. Business Day. 2022 August 25. Available from: https://www.businesslive.co.za/bd/opinion/2022-08-25-mike-muller-fracking-lawfare-may-obstruct-the-just-transition/

# Load shedding in South Africa: Another nail in income inequality?

AUTHOR:
Roula Inglesi-Lotz[1]

AFFILIATION:
[1]Department of Economics, University of Pretoria, Pretoria, South Africa

CORRESPONDENCE TO:
Roula Inglesi-Lotz

EMAIL:
Roula.inglesi-lotz@up.ac.za

HOW TO CITE:
Inglesi-Lotz R. Load shedding in South Africa: Another nail in income inequality? S Afr J Sci. 2023;119(9/10), Art. #16597. https://doi.org/10.17159/sajs.2023/16597

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
load shedding, income inequality, energy ladder, South Africa

PUBLISHED:
31 August 2023

**Significance:**

South African households have been affected by load shedding for over a decade. Low-income households are the most heavily impacted by unreliable electricity supply, rising electricity prices and lack of financial means to absorb such shocks, subject to their living conditions. Marginalised communities struggle to access the advantages of urban areas, deepening the country's income inequalities. Policymaking needs to address the uneven distribution of the impact with policies and programmes that will improve access to finance and technologies for sustainable future solutions. However, there is a catch in the implementation of such policies, as, potentially, measures such as subsidies may exacerbate inequalities and create more problems in the system. Innovative financial programmes are essential to support low-income households and ensure fairness in dealing with load shedding effects while promoting socio-economic development and improving living standards.

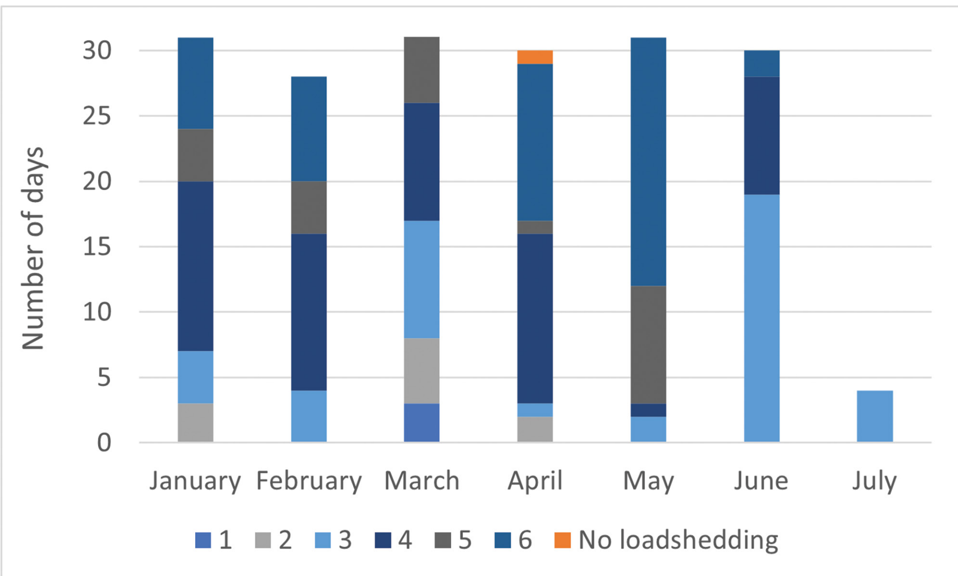
The South African energy crisis is ongoing, with the country experiencing widespread rolling blackouts (load shedding) as supply falls behind demand, threatening to destabilise the national grid.[1] Load shedding started in the late months of 2007 and is ongoing. Eskom, the government-owned national power provider and principal generator, has blamed the rolling blackouts on insufficient generation capacity. These rolling blackouts, or load shedding, are defined as the action to reduce the load on something, especially the interruption of an electricity supply, to avoid excessive load on the generating plant. Such 'load shedding' is conducted at any time that generating units are taken offline for maintenance, repairs, or refuelling (in the case of nuclear plants) with a reserve margin of 8% or less.

Even though load shedding is not the crisis but the *response* to it – a way to mitigate it – persistent load shedding is, understandably, causing much frustration for South African households and businesses, which experience frequent power interruptions. Whether frequent or prolonged, power outages are assumed to limit the economic well-being of households and enterprises by lowering the output of existing electrical equipment and discouraging investments in new welfare-improving and income-generating ones.

Figure 1 demonstrates the number of days of load shedding per month of 2023.[2] The country has barely experienced a day without some stage of load shedding implemented since September 2022, with the average stage in a day continuously rising.

More reliable energy is needed to lower operating expenses and raise productivity and profitability in enterprises. Power outages in South Africa have resulted in sales losses for many businesses, from the retail and service sectors to manufacturing and industry. The expected loss to South Africa's businesses and industries from scheduled power outages is ZAR1 billion per stage daily. As a result, many small and medium-sized enterprises have struggled and eventually closed down, with a loss of thousands of employment positions.

The consequences and impacts of power cuts are not evenly distributed across regions and population groups, which potentially worsens income inequalities that are historically high in South Africa. In this Commentary, the




*Data: News24*[2]

**Figure 1:** Number of days of load shedding by stage per month, 2023.

uneven impact of load shedding on South African electricity consumers is discussed, as well as how it has and will further exacerbate income inequalities in the country. The discussion is informed by theoretical literature and data on how different income groups respond to load shedding incidents. The analysis also takes into consideration suggested and implemented policies for the residential sector within the energy sector and how these can assist or exacerbate existing income inequalities. The ultimate aim is to highlight the uneven impact of load shedding and stimulate discussion on how such disparities can be addressed.

Looking deeper into the socio-economic context of load shedding and its long-term effects on underprivileged groups is critical. Load shedding, a long-standing problem in the country, has intensified pre-existing disparities. Vulnerable households, which account for a sizable part of the population, are disproportionately affected by power outages due to restricted access to backup power sources, insufficient financial resources to deal with protracted outages and dependency on energy for a living. These already-disadvantaged households are further disadvantaged as load shedding disrupts income-generating activities, restricts access to education and healthcare facilities, and reduces overall economic production. The inability to sustain persistent economic activities and acquire a steady income exacerbates the wealth disparity between disadvantaged households and more privileged parts of society. As a result, load shedding adds to the cycle of poverty and impedes upward social mobility, eventually leading to income inequality on a larger scale. Load shedding exacerbates income inequality in the country by maintaining differences in economic opportunity and inhibiting socio-economic advancement. As a result, it is critical to acknowledge the vital role that load shedding plays in perpetuating and growing the gap between different segments of society, needing a thorough grasp of its consequences for income distribution.

## How do households choose where and how to consume energy?

Households in high-income countries have a different energy use profile from those in middle- and low-income countries. Such differences are observed among households with varying income levels, even within the same country. The commonly used theoretical concept of the 'energy ladder' explains the preferred household energy sources at various income levels and how that evolves (Figure 2).[3]

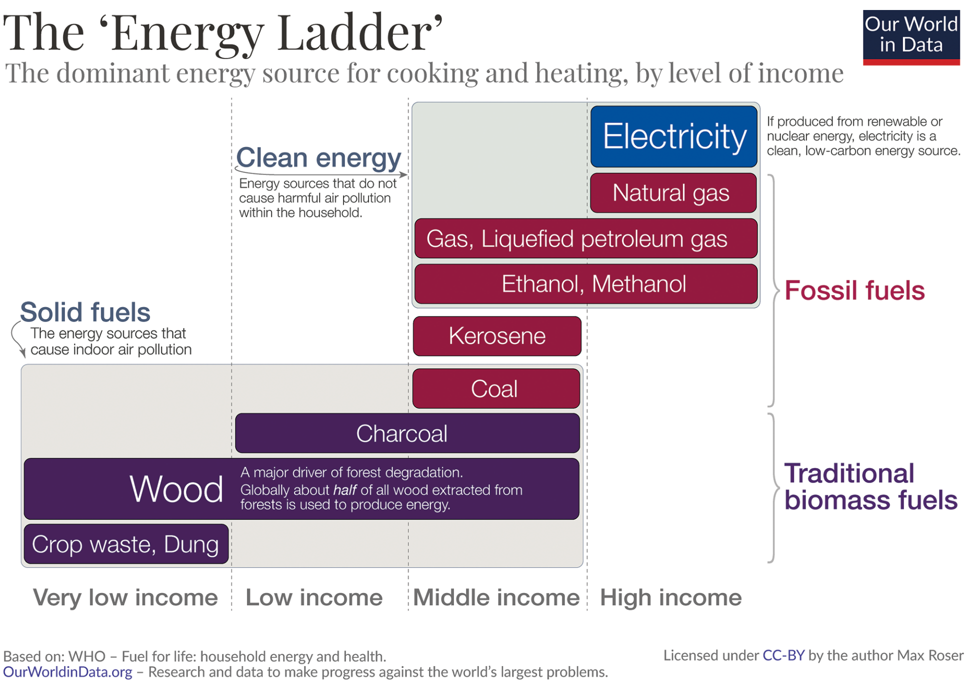
Very low-income households prefer wood and biomass (crop waste and dried dung) as a fuel for cooking, but sometimes also use coal and charcoal. This preference is attributed to a lack of access to the national electricity grid, affordability issues and easiness of use. Fossil fuels are usually burnt on open stoves which results in significant indoor air pollution and exposes household members to pollution and affects their health. According to the 'energy ladder' hypothesis, households evolve from solid fuels to cleaner forms of energy and particularly electricity.

The theoretical concept of the 'energy ladder' is criticised in the literature, mainly for the argument that the only underlying factor for the shift between fuels is the household income level. Although income is indeed an important factor, other conditions and factors play a role in these households' decisions, including access and affordability to technology, availability of resources, urbanisation and living standards. Additionally, fuel-switching depends on price changes, supply reliability, habits and culture, education and demographics of the households.[4]

The fuel preference path is not a series of unconnected steps but a more dynamic and connected system. Also, in practice, fuel-switching only occurs in one direction, from dirty to clean energy alternatives. The concept of 'fuel stacking' advocates that households choose a mix of fuels that change proportions between clean and dirty fuels as their income increases.

With that in mind, Bohlmann and Inglesi-Lotz[5] gave a more thorough picture of how South African households of different income levels control their electricity consumption due to income changes. They showed that low-income households are more sensitive than high-income ones. With the same increase in their disposable income (in percentage terms), low-income households will increase their electricity consumption proportionally higher than high-income ones.

South Africa exhibits higher electrification rates than the rest of the continent (more than 90%). Even though many rural households are considered energy poor, they cannot afford to pay their electricity bills to cover their basic needs. South African households show a behaviour typical of an 'energy ladder' where "households progressively move away from low-quality energy sources such as wood and paraffin




*Source: CC-BY Roser*[3]

**Figure 2:** The energy ladder.

towards convenient and versatile modern sources of energy such as electricity and gas as income rises"[6]. Households with access to electricity use it as their main lighting source, while those without use mainly candles and paraffin. The great majority of electrified households use electricity for cooking purposes. For heating domestic space during winter, households prefer alternative sources to electricity or wearing warmer clothing.[6]

## Unequal impacts of load shedding

Power interruptions lead to disruptions in daily activities, food spoilage and other challenges. Hence, load shedding forces households to make decisions differently as their electricity demand cannot be met during these power cuts. Since 2008 with the first wave of power cuts in South Africa, households have constantly re-evaluated their strategies for responding to the lack of electricity. Primarily, they aim to find alternative resources to substitute electricity. Such options vary from diesel-powered generators to solar panels or wood and charcoal. Households also sometimes decide to postpone activities until electricity has been restored. The latter looks like the least costly option, but such delays put pressure on the grid later and on the everyday life of households.

Load shedding affects all types of households but creates higher risks and poses threats to low-income households that are the most vulnerable in the South African economy. The heterogeneity of households in South Africa in factors such as income, behaviour and preferences, means they are affected differently by load shedding in South Africa.

The reasons for the differences might be internal, due to their different nature, or external, due to their ***geographical location***. Frequently, low-income households are located in areas with ageing and less-reliable infrastructure. Such conditions make them prone to more frequent breakdowns that take more time to be resolved, adding to the scheduled load shedding that higher-income households and businesses, particularly in urban areas, experience.

The issue of load shedding disproportionately affects South Africa's urban poor, who are especially vulnerable to the effects of power outages. Low-income households and communities frequently require additional financial resources to finance ***alternative energy sources*** or other energy expenditures during load shedding. As a result, they cannot obtain fundamental amenities like health care, education, and social services, further restricting specific populations from the advantages of city living.[7]

The energy crisis contributes significantly to exclusivity in South African cities by widening existing inequalities and creating considerable impediments to accessing the benefits of urbanisation for vulnerable groups. Household financial position, in general, is constrained even more for rural low-income households. Low-income households have ***limited access to finance*** to fund backup power sources, such as generators or uninterruptible power supplies (UPS), which can mitigate the impact of load shedding.

The considerable ***fiscal transfers*** targeted at facilitating access to power through initiatives such as ***'Free Basic Electricity'*** for low-income households and the significant resources granted through the 'local equitable share' of around ZAR100 billion certainly play a critical role in minimising the impact of rising power rates on vulnerable areas. Recognising the positive effect of these grants in reducing the electricity expenses burden for low-income households is essential. Given these fiscal transfers, it is reasonable to believe that the direct impact of increased power prices on low-income households may be alleviated to some degree. The supply of 'free basic electricity' and other similar support measures can provide some relief by guaranteeing that financially challenged people can meet their necessary electrical demands. It is crucial to highlight, however, that the effectiveness and reach of these programmes may differ across regions and communities. Challenges such as widespread non-payment in many low-income communities could undercut the intended benefits of these fiscal transfers and prevent them from having the full impact on alleviating the burden of growing power prices. While fiscal transfers are important in tackling the affordability issue, they should not be used to dominate a broader conversation about the implications of increasing energy prices and their possible contribution to income disparity among disadvantaged communities.

Load shedding ***damages equipment*** and makes it difficult for businesses and households to plan accordingly.[8] It is more than evident that low-income households cannot afford to replace the damaged equipment and appliances. The majority of these households cannot afford appropriate insurance that will cover them in such cases.

Low-income households depend more on electrically powered utilities for necessities like heating, cooking, and refrigeration. Load shedding can cause significant disruptions to their daily lives, increasing their vulnerability to food spoilage, exposure to extreme temperatures, and other hazards. Their choice of alternatives, such as **wood and charcoal** for cooking, may worsen their health vulnerabilities. Stoves and other wood-generated appliances used indoors create hazardous air conditions for the members of low-income households. This backward direction in the 'energy ladder' hypothesis is not a South African phenomenon. Internationally, even in developed countries, households that experience or expect shortages from baseload electricity turn to traditional fuels with negative consequences for air quality.[9]

Low-income households need help reducing their energy consumption during load shedding. One such barrier is limited access to **energy-efficient appliances**, which can help to reduce electricity usage. Retrofitting homes with energy-saving measures can also be costly, making it challenging for low-income households to make the necessary changes. As a result, there is a need for targeted efforts to provide low-income households with access to energy-efficient appliances and support to retrofit their homes. Such efforts could help to reduce their energy consumption during peak periods, mitigate the impact of load shedding, and ultimately improve their quality of life.

***Limited access to information*** deepens the inequalities between low- and high-income households. Low-income households need complete information on the timing and duration of load shedding, being challenged to make provisions and plan accordingly – frequent power disruptions in higher stages of load shedding lead to potential health hazards and increased distress.

## Conclusion

Over a decade has passed since the first significant wave of power cuts. South Africa's load shedding has been a persistent phenomenon, causing frustrations to electricity consumers daily. Indeed, load shedding is the reaction to deeper problems in the electricity sector, such as an ageing fleet, lack of proactive maintenance, inefficient management and corruption. It is the response to the mismatches between demand and supply to avoid more extensive national blackouts. The consequences and impact of the power outages are disproportionally distributed across households, potentially worsening income inequality. Here, this uneven effect on consumers and the behaviour of low-income households was discussed. The discussion was informed by theoretical literature and data on how different income groups react to load shedding incidents. The analysis also considered suggested and implemented policies for the residential sector within the energy sector and how they can assist or exacerbate existing income inequalities. Ultimately, the aim was to highlight the uneven impact of load shedding and stimulate discussion on how such disparities can be addressed.

In conclusion, load shedding in South Africa significantly impacts households, particularly low-income households. It has the potential to exacerbate existing income inequalities, which is a concern given South Africa's high level of inequality. Policymakers need to address the uneven distribution of the impact with policies and programmes that will improve access to finance and technologies for sustainable future solutions. Innovative financial programmes are essential to support low-income households and ensure fairness in dealing with load shedding effects while promoting socio-economic development and improving living standards. However, policymakers need to be cautious in implementing such policies to avoid further exacerbating inequalities. By addressing the issue of load shedding and its impact on income inequalities, South

Africa can make strides towards promoting sustainable and equitable development for all its citizens.

The timing and scheduling of load shedding do not consider differences among households, even though there is a clear focus by energy policymakers to assist financially with promoting investment towards renewable, off-grid solutions. In the budget speech of February 2023, the Finance Minister announced new incentives for rooftop solar technologies to address challenges faced by small and medium-sized enterprises and low-income households. However, tax incentives such as this can create more significant problems than they aim to solve as only middle- and high-income households can use the subsidy which does not cover a complete energy-generating solution but only a portion thereof). In contrast, low-income households need help to finance the remaining amounts post-subsidy.[10] Slowly but surely, middle- and high-income households will be off-grid and demand less from Eskom's generated electricity. Lower demand will affect Eskom's viability and profitability, with the possible passing of the burden to the remaining consumers of grid electricity. Addressing power instability is crucial for a fair and sustainable future, promoting socio-economic development and improved living standards.

## Competing interests

I have no competing interests to declare.

## References

1. Mabugu T, Inglesi-Lotz R. The effect of mismatched supply and demand of electricity on economic growth in South Africa. Energy Sources B: Econ Plan. 2022;17(1):Art. #2038731. https://doi.org/10.1080/15567249.2022.2038731
2. Omarjee L. We've had more load shedding so far this year than in the whole of 2022. News24. 2023 May 11 [updated 2023 Jul 04; cited 2023 Aug 17]. Available from: https://www.news24.com/fin24/climate_future/energy/weve-had-more-load-shedding-so-far-this-year-than-in-the-whole-of-2022-20230511
3. Roser M. The 'energy ladder': What energy sources do people on different incomes rely on? Our World in Data. 2021 December 01. Available from: https://ourworldindata.org/energy-ladder#:~:text=The%20so%2Dcalled%20'Energy%20Ladder,heat%20with%20charcoal%20or%20coal
4. Kroon BVD, Brouwer R, Van Beukering PJH. The energy ladder: Theoretical myth or empirical truth? Results from a meta-analysis. Renew Sustain Energy Rev. 2013;20:504–513. https://doi.org/10.1016/j.rser.2012.11.045
5. Bohlmann JA, Inglesi-Lotz R. Examining the determinants of electricity demand by South African households per income level. Energy Policy. 2021;148:Art. #111901. https://doi.org/10.1016/j.enpol.2020.111901
6. Bohlmann JA, Inglesi-Lotz R. Analysing the South African residential sector's energy profile. Renew Sustain Energy Rev. 2018;96:240–252. https://doi.org/10.1016/j.rser.2018.07.052
7. Mangara F, Mutsaa E. Tri-fold Impact of load-shedding in South African cities: Power outage, destruction, and regression [webpage on the Internet]. c2023 [cited 2023 Apr 19]. Available from: https://inclusivecities.ukzn.ac.za/blog/tri-fold-impact-of-load-shedding-in-south-african-cities-power-outage-destruction-and-regression/#:~:text=The%20issue%20of%20load%2Dshedding,energy%20costs%20during%20load%2Dshedding
8. Manyaapelo OC. Load shedding is affecting the economy. IOL. 2023 January 25. Available from: https://www.iol.co.za/news/politics/opinion/load-shedding-is-affecting-the-economy-6196acbf-ee40-4325-8581-ff0a5e7cb21d
9. How Europe is spluttering its way to better air quality. The Economist. 2023 April 13. Available from: https://www.economist.com/europe/2023/04/13/how-europe-is-spluttering-its-way-to-better-air-quality
10. Inglesi-Lotz R, Chitiga-Mabugu M, Bohlmann JA, Bohlmann HR. UP expert opinion: #BudgetSpeech2023 – Yes, indeed, but… [webpage on the Internet]. c2023 [cited 2023 Apr 19]. Available from: https://www.up.ac.za/news/post_3130614-up-expert-opinion-budgetspeech2023-yes-indeed-but

**AUTHORS:**
Stephen R. Clark[1]
Craig McGregor[1]

**AFFILIATION:**
[1]Department of Mechanical and Mechatronic Engineering, Stellenbosch University, Stellenbosch, South Africa

**CORRESPONDENCE TO:**
Stephen Clark

**EMAIL:**
sclark@sun.ac.za

**HOW TO CITE:**
Clark SR, McGregor C. Liquified petroleum gas provides a technically viable and financially feasible means to reduce Eskom's diesel cost burden by 30% to 40%. S Afr J Sci. 2023;119(9/10), Art. #16594. https://doi.org/10.17159/sajs.2023/16594

**ARTICLE INCLUDES:**
☐ Peer review
☐ Supplementary material

**KEYWORDS:**
LPG, energy generation, diesel, Eskom, load shedding

**PUBLISHED:**
31 August 2023

7 AFFORDABLE AND CLEAN ENERGY
8 DECENT WORK AND ECONOMIC GROWTH
9 INDUSTRY, INNOVATION AND INFRASTRUCTURE
11 SUSTAINABLE CITIES AND COMMUNITIES



# Liquified petroleum gas provides a technically viable and financially feasible means to reduce Eskom's diesel cost burden by 30% to 40%

**Significance:**

In early 2022, we analysed the benefits of switching the fuel for the Ankerlig peaking plant from diesel to LPG (liquified petroleum gas or 'bottled' gas) to reduce cost, improve greenhouse gas emissions and provide for the long-term needs of the South African electricity grid. That analysis was published in May 2022 (Clark et al. J Energy South Afr. 2022;33(2):15–23). With increased load shedding and the higher cost of diesel fuel, this recommended fuel change has become even more relevant. In this commentary, we summarise the conclusions from that paper and related studies and update the results to reflect the latest situation within the country.

## Background

Nearly everyone in South Africa knows that load shedding is a significant and growing problem for the country. A recent review in *Daily Maverick* reported that the country suffered over 200 days of load shedding in 2022, with indications that 2023 would be worse.[1] The first day of this year without any load shedding was 19 March 2023. Load shedding up to Stage 6 has become the norm, where Eskom implements rolling blackouts to remove 6 GW of demand from the grid. This major problem must be resolved as quickly as reasonably possible. The cost to the South African economy has been estimated to be more than ZAR500 billion per year.[2] However, long-term planning must be considered to ensure that temporary solutions do not have long-term negative impacts on the system. These include shifting to expensive mid-merit or baseload natural gas fuelled generation, such as offshore gas-fired plants.

In the 2018 iteration of the government's Integrated Resource Plan (IRP), it was assumed that the short-term needs of the system were well provided for, and that no new generation capacity was needed until sometime after 2025. Eskom believed it would improve the energy availability factor of its baseload coal-fuelled plants and was busy commissioning two new plants, Medupi and Kusile, which would add approximately 20% new generation capacity. Almost before the 2018 IRP was finalised, Eskom found significant problems with the new coal plants and performance challenges with the existing plants. The short- to medium-term problems noted in the 2018 update of the IRP led to another IRP update in 2019. These problems at Eskom led to load shedding in late 2018. Unfortunately, the problems with unplanned outages increased in 2019, and so did load shedding. In December 2019, this resulted in Stage 6 load shedding for the first time. These problems have only intensified in subsequent years. This generation shortfall has been mitigated by extensive use of the dispatchable generation provided by the diesel-fuelled gas turbine peaking plants, which has led to high diesel consumption.

As detailed in the 2019 IRP, the short-term problems will be challenging to resolve. Load shedding will be around for several years, with new generation capacity taking years to implement. In 2022, according to a CSIR analysis, the additions to the Eskom generation fleet were 720 MW of coal from the commissioning of the last train of the Kusile plant, plus 419 MW of wind and 75 MW of solar photovoltaic (PV). These additions were more than offset by the continued decrease in the availability of the existing coal fleet, as reported by CSIR: "Eskom fleet average EAF [was] 58.1% for 2022 (relative to 2021 of 61.7%, 2020 of 65%, 2019 of 66.9%)."[3] The decrease in the energy availability factor (EAF) has effectively reduced the system generation capacity by 3.4 GW from 2019 through to 2022. In the meantime, the South African Department of Mineral Resources and Energy (DMRE) is making slow progress with the Renewable Independent Power Producer Programme, awarding 1 GW of solar PV projects in October 2022 for bid round 6 after awarding 1600 MW of wind and 975 MW of solar PV projects in round 5 in October 2021.[4]

The cheapest, most flexible solution to the load shedding problem is the rapid development of renewable generation (solar and wind) at all system levels – whether residential rooftop systems, commercial installations, industrial applications, or utility-scale projects. These projects have been highlighted in the government's report on 'Actions to End Load Shedding and Achieve Energy Security'.[5] The most significant elements of this plan are to expedite the review and approval of the queue of 6 GW of proposed private renewable generation projects and the commitment to encourage and enable commercial, industrial and residential rooftop solar projects. Once approvals are in place, these projects can be implemented quickly at any appropriate size.

One element that is often overlooked in the move towards rapid increases in solar and wind generation is the continued need for firm dispatchable power to provide generation when wind and solar do not meet the demand.[6] Firm capacity is generating capacity that is available at all times, whilst dispatchable capacity can be adjusted or turned on or off as the grid requires. In an interview on Eskom's diesel usage, a South African energy expert, Anton Eberhard, described these firm dispatchable (or peaking) generators as essential in maintaining the balance of the power system.[7] The need for firm dispatchable generation is not just a South African issue; it is true of all power systems and becomes more critical with the increased use of variable generation sources.[8,9]

Because of the significant load shedding due to the shortage of baseload generation capacity, Eskom has been using its peaking generation facilities much more than planned. These plants are fuelled with diesel. For 2022, Eskom forecasted that, with the increased diesel price, the expenditure for diesel fuel for its peaking plants would increase to over

ZAR6.5 billion for the year.[10] The peaking generation in 2022 and associated diesel costs were twice this forecast. Eskom's latest information (12 January 2023) was that it had expended over ZAR14.7 billion for diesel fuel in the last year.[7] This expenditure did not meet all the needs as the operator was budget-constrained. These budget constraints have led to Eskom using these plants at a lower rate than previously, even with the increased load shedding, as shown in Figure 1.[11]

Two plants, Ankerlig (1340 MW) and Gourikwa (740 MW), provide the bulk of the Eskom peaking capacity. Both plants were commissioned in 2007 and are approximately halfway through their 30-year expected lifespan.[12,13] It is technically possible to convert these plants to gas-fired if a gas source were available. As natural gas has traditionally been significantly less expensive on an energy basis than diesel (per GJ), Eskom made plans to make this change. However, they do not have access to gas. The assumption throughout the development of the various IRPs was that gas would be provided by importing liquefied *natural* gas (LNG). However, importing LNG has proven economically challenging due to the high upfront capital cost and low expected utilisation rates of importation facilities. While natural gas prices have traditionally been approximately half the price of diesel fuel, the last year has seen international natural gas prices at twice the price of diesel, as shown in Figure 2.[14] The prices of all fossil fuels worldwide have been impacted by the limitations placed on importing natural gas from Russia into Europe. In particular, the high demand for gas in Europe has changed the price balance between oil and gas.[15] The International Energy Agency (IEA) indicates that gas prices will remain challenged for some years. In their 2022 World Energy Outlook, the IEA predicted that relatively high natural gas prices would remain until the mid-2020s when new gas sources could bring the prices more in line with their long-term pricing relationship. They predicted that these high prices would discourage the construction of new natural gas fuelled power plants.

Liquefied *petroleum* gas (LPG), or 'bottled' gas, is another option for diesel replacement for peaking plants. It is a hydrocarbon gas and, much like natural gas, is a suitable fuel for gas turbine usage and would provide many of the same benefits as LNG. As noted by the World LPG Association:

> *As governments are looking to replace polluting coal, diesel or heavy oil plants, LPG offers a cost-efficient and low emissions alternative for areas that are not connected to the natural gas grid.*[16]

Hitachi, a manufacturer of gas turbine generation plants, has an entire line of systems designed for use with LPG and has installed over 120 plants using LPG as their prime fuel. These are smaller systems designed for local use where natural gas is not readily available.[17] Hitachi reports that running these plants on LPG slightly increases their output compared to diesel fuel and reduces maintenance by 20–30%.

In 2019, as part of an experiment related to solar thermal energy, students at Stellenbosch University converted a Rover 1S/60 gas turbine that was set up to run on kerosene fuel to propane.[18] This is a direct comparison to the conversion of gas turbines fuelled with diesel being converted to LPG fuel. In the experiment, the turbine ran successfully with propane fuel. While in the experiment there was a decrease in the output of turbine from approximately 88% of the rated power using kerosene to 77% using propane, the analysis of the experimental results concluded that this decrease was due to the experimental setup rather than anything to do with the fuel change. The experiment's conclusion was that propane fuel was quite suitable for this gas turbine.[18] As the manufacturers of the Ankerlig gas turbines have indicated that those turbines can be run on LPG fuel, there is no reason to believe that the results from this experimental setup cannot be achieved at Ankerlig.[19]

In their product brochures, Siemens confirms that their V94.2 turbines (model SGT5-2000E), used in Ankerlig and Gourikwa, can utilise LPG as fuel. At Ankerlig and Gourikwa, LPG is currently utilised as a start-up gas for diesel operations as it is easier to start up the turbines with LPG than diesel.[12] However, LPG is currently not available in sufficient volumes onsite to be considered for full-time use in these plants. Full-time operation using LPG as a fuel should be feasible after installing an appropriate LPG storage and delivery system.

In 2017, Sunrise Energy commissioned an LPG importation facility in Saldanha Bay.[20] LPG has the advantage of easy storage and transport (much like diesel) and is much less expensive than diesel. In addition to LPG being cheaper than diesel, it also has lower carbon dioxide emissions. Unlike natural gas (predominantly methane), LPG is not classified as a greenhouse gas.

## Change to LPG

Clark et al.[11] reviewed the relationship between diesel and LPG prices over the past 15 years. That analysis showed that the average LPG price was approximately 60% of the diesel price (on an energy basis). Since this analysis, significant movements in the energy market have occurred due to COVID, the European Energy Crisis and the war in Ukraine. The international energy prices for LNG, diesel and LPG (updated through 2022) are shown in Figure 2 (based on prices published by the US Energy Information Agency for prices at the US Gulf Coast export point).[14] While LNG had a considerable increase in 2021 and 2022, diesel did not change nearly as much, whilst the LPG price hardly changed.

The DMRE-stated March 2023 wholesale price for diesel in South Africa was ZAR20.94 per litre, and the wholesale price for LPG was ZAR9.46 per litre.[21] Converting these to an energy equivalent basis (USD/GJ) gives USD31.84 /GJ for diesel and USD22.19 /GJ for LPG. Therefore, LPG in South Africa is currently priced at 69% of the price of diesel on an energy basis. These numbers are reasonably consistent with the longer-term relationship between these fuels. At ZAR20.94 per litre for diesel, the fuel input cost for electricity is over ZAR5.5 per kWh. Eskom pays approximately 80% of the DMRE-stated retail price rather than the wholesale price. However, it can be expected that the price relationship between the two fuels should be approximately the same.

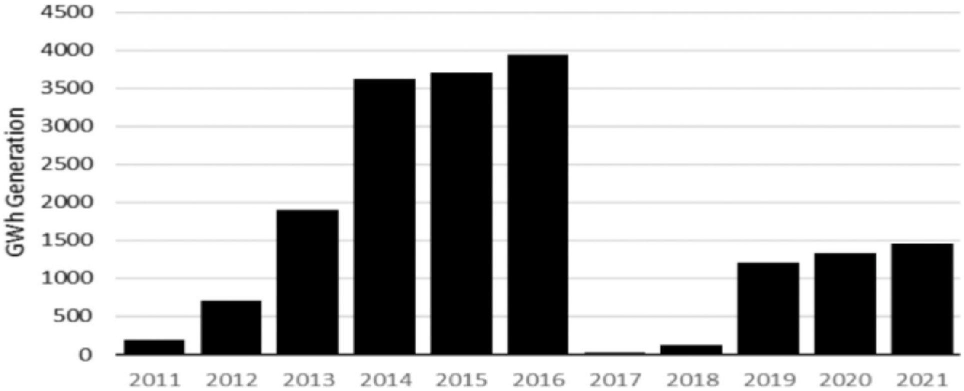



**Figure 1:** Eskom-reported peaking generation, 2011–2021.[11]

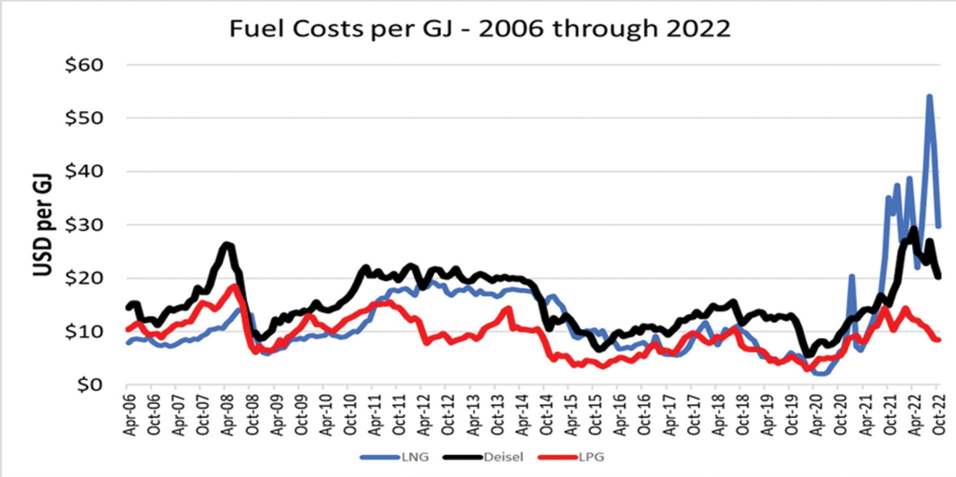



**Figure 2:** Fuel cost comparison (USD/GJ), 2006–2022.[14]

## LPG importation and storage

LPG is currently imported into Saldanha Bay; no new importation facilities would be needed to bring the required volumes of LPG into the area. The Sunrise Energy importation facility currently contains 5500 tonnes of LPG storage. Avedia Energy also has an import facility in Saldanha Bay, commissioned in 2019, with 2200 tonnes of LPG storage. As for all fuels, the challenge remains in the cost-effective storage and delivery of the fuel at high rates for dispatchable usage. Besides being locally available, one advantage of LPG is that it is in a liquid phase at ambient temperature and slightly above atmospheric pressure. Liquid fuels require significantly lower volumes to store.

There is some question regarding the throughput capacity of the Sunrise and Avedia LPG import facilities in Saldanha Bay and their ability to meet the need for the Ankerlig demand. The issue to be determined is the amount of additional LPG storage that might be required. It would likely be necessary for the Saldanha Bay facilities to be expanded to provide additional LPG storage. In addition, LPG fuel storage at the Ankerlig plant itself would also be necessary. In Richards Bay, Bidvest has recently constructed an LPG storage and delivery project. The project has four LPG tanks, each storing 10 000 $m^3$ ($4 \times 5500$ tonnes). The cost of the plant was listed as less than ZAR1 billion. Bidvest completed the project within 2 years from project commencement. A duplicate of this plant would provide 40 000 $m^3$ of LPG storage or 1 PJ of fuel. The current diesel storage at Ankerlig is 32 million litres or 1.2 PJ energy. The 40 000 $m^3$ would be enough to run the 1.3 GW Ankerlig plant for 76 hours, or slightly more than 3 days.

## Fuel delivery to Ankerlig

The Ankerlig generation plant is approximately 100 km from the Saldanha Bay LPG importation facilities. A pipeline would be the lowest-cost means for natural gas and LPG delivery. For diesel delivery to the Ankerlig plant, as the plant is designed for low-capacity peaking use, Eskom has chosen to deliver the fuel by truck, avoiding the capital cost of a pipeline. The 1.3 GW Ankerlig plant requires seven truckloads of diesel per hour of generation. The onsite storage for diesel holds the equivalent of 640 truckloads. As can be deduced from Figure 1, during the maximum usage period for the peaking plants by Eskom, 2014 through 2016, the Ankerlig plant produced over 2000 GWh per year to the grid. As Ankerlig provides approximately 64% of the indicated Eskom peaking capacity, this energy output indicates a capacity factor of over 20% for the plant for these years. In November 2014, it was reported that Eskom operated Ankerlig for over 200 hours at a capacity above 30%. While questions were raised in the press about this operation, Eskom did not report any problems in meeting this need using fuel delivery by truck.[22]

It can be assumed that this same option would likely be considered for LPG delivery to the plant. Compared to seven truckloads of diesel, ten truckloads of LPG would be required to operate the plant per hour of generation. Duplicating the energy storage of diesel would require an LPG storage facility equivalent to the Bidvest Richards Bay LPG plant to be constructed at Ankerlig. Even with the high-capacity operation under periods of load shedding, the number of deliveries of LPG would be only slightly higher than that currently done for diesel. A pipeline delivery system from Saldanha to Ankerlig remains an alternative to truck delivery. However, it is expected that Eskom would not decide to make this investment due to the intermittent usage of the plant.

## Economics of change out to LPG

Assuming the current South African relationship of LPG at 69% of the price of diesel, it is possible to make an economic comparison of the annual cost savings from the fuel changeout. As importation facilities are in place and no pipeline or alternative delivery system investment is required, developing a Bidvest-equivalent LPG storage and delivery system should be the only investment required to allow LPG delivery to Ankerlig. With the ZAR14.7 billion budget for diesel fuel for the year and Ankerlig providing 64% of the dispatchable power, a change to LPG – at 69% of the diesel price – would result in a 1-year savings of ZAR 2.9 billion. As laid out in the original analysis, the cost of making the change should be around ZAR1 billion. Thus, the payout to make the fuel change would be a matter of a few months, with an expected savings of over ZAR1.7 billion in the first year of usage and significant continued savings in subsequent years. Eskom might also have some minor costs to modify the fuel intake and control system for the turbines for full-time LPG usage, but this is not likely to materially change the economics of the fuel change.

## Business structure

Given the probability of load shedding by Eskom continuing for the next several years, this concept should be followed up immediately with all interested parties, including potential investors, Eskom as the customer, and the government as the coordinating party. Companies can bid on the supply of LPG fuel to the Eskom Ankerlig plant, which could be done on a capacity plus usage payment or a simple competitive supply arrangement. With the expected remaining lifespan for the Ankerlig plant of over 15 years and the quick payout time for the expected conversion costs, Eskom should be able to present an attractive contractual arrangement for this fuel change. Outside investors would likely be interested in building the required infrastructure with a suitable fuel supply agreement for Ankerlig.

Even with the planned fuel change, there is no reason for Eskom to abandon the existing diesel storage system, as these units could be converted back to diesel use as necessary. Assuming the diesel fuel system is maintained, Eskom would have complete flexibility to use either LPG or diesel, increasing the fuel security of this plant.

## Effect on load shedding

The change of fuel from diesel to LPG will not solve load shedding, either in the short or long term. However, it will lower the cost of providing the firm dispatchable power that must be provided to keep the system in balance. It will also make it easier for Eskom to use this generation source as needed with less concern about their fuel budget. With the lower fuel cost, should Eskom choose to use these plants at the level it did in 2014 through 2016, it could add over 2000 GWh of generation into the system.

Successful conversion of the Ankerlig peaking plant to readily available lower-cost LPG could also become a model for implementing firm dispatchable generation plants. These facilities will be needed to maintain system balance as the base generation transitions from large coal-fuelled power plants to a system based on variable renewable energy sources from solar and wind.[6]

## Conclusions

The conclusions reached in the 2022 analysis are still valid today and will have increased relevance for the current situation.

- The current load-shedding challenges with the South African grid indicate that the diesel peaking plants will continue to be used at very high levels for the coming years.
- Diesel fuel costs will remain a significant burden on Eskom's finances.
- Switching out the diesel at Ankerlig to LPG fuel can significantly reduce this cost via a low-risk, low-capital project.
- This fuel change would be feasible, with minimal cost to Eskom, as it can be done as a service provided by a private company.
- The fuel changeout should have a quick payback with long-lasting benefits, even when load shedding is resolved. This change should be pursued urgently to reduce the burden on Eskom from the current cost of diesel.
- The fuel changeout will also reduce greenhouse gas emissions.

## Competing interests

We have no competing interests to declare.

## References

1. O'Regan V. Eskom's ability to provide power is steadily worsening, CSIR stats show. Daily Maverick. 2023 February 22. Available from: https://www.dailymaverick.co.za/article/2023-02-22-eskoms-ability-to-provide-power-is-steadily-worsening-csir-stats-show
2. Daily Investor. South Africa lost R560 billion in 2022 because of load-shedding. Daily Investor. 2022 December 14. Available from: https://dailyinvestor.com/south-africa/6860/south-africa-lost-r560-billion-in-2022-because-of-load-shedding/
3. Pierce W, Le Roux M. Statistics of utility-scale power generation in South Africa 2022 [document on the Internet]. c2023 [cited 2023 Mar 25]. Available from: https://www.csir.co.za/sites/default/files/Documents/Statistics%20of%20power%20in%20SA%202022-CSIR-%5BFINAL%5D.pdf
4. South African Department of Mineral Resources and Energy (DMRE). Independent Power Producer Procurement Programme. Pretoria: DMRE; 2022.
5. Government of South Africa. Confronting the energy crisis: Actions to end load shedding and achieve energy security [document on the Internet]. c2023 [cited 2023 Mar 25]. Available from: https://www.stateofthenation.gov.za/assets/downloads/confronting_the_energy_crisis_to_end_load_shedding.pdf
6. Clark S, Van Niekerk J, Petrie J, McGregor C. The role of natural gas in facilitating the transition to renewable electricity generation in South Africa. J Energy South Afr. 2022;33(3):22–35. https://doi.org/10.17159/2413-3051/2022/v33i3a12742
7. Eskom's massive diesel conundrum. BusinessTech. 2023 January 12. Available from: https://businesstech.co.za/news/energy/655567/eskoms-massive-diesel-conundrum/
8. IEA-ETSAP. Renewable energy integration in power grids [document on the Internet]. c2015 [cited 2023 Mar 25]. Available from: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2015/IRENA-ETSAP_Tech_Brief_Power_Grid_Integration_2015.pdf
9. Sepulveda NA, Jenkins JD, De Sisternes FJ, Lester RK. The role of firm low-carbon electricity resources in deep decarbonization of power generation. Joule. 2018;2(11):2403–2420. https://doi.org/10.1016/j.joule.2018.08.006
10. Creamer T. Eskom seeks to show prudency of running diesel plants as its pegs cost of load-shedding at R9.53/kWh. Engineering News. 2022 January 20. Available from: https://www.engineeringnews.co.za/article/eskom-seeks-to-show-prudency-of-running-diesel-plants-as-its-pegs-cost-of-load-shedding-at-r953kwh-2022-01-20
11. Clark S, McGregor C, Van Niekerk JL. Using liquefied petroleum gas to reduce the operating cost of the Ankerlig peaking power plant in South Africa. J Energy South Afr. 2022;33(2):15–23. https://doi.org/10.17159/2413-3051/2022/v33i2a13453
12. Eskom. Ankerlig and Gourikwa gas turbine power stations. 2014 [2020, September 20]. Available from: https://www.eskom.co.za/sites/heritage/Pages/ANKERLIG-AND-GOURIKWA.aspx
13. Hulisani. Avon and Dedisa peaking power [webpage on the Internet]. c2021 [cited 2023 Apr 07]. Available from: https://hulisani.co.za/project/avon-and-dedisa-peaking-power
14. US Energy Information Administration (EIA). Petroleum & other liquids – Spot prices [database on the Internet]. c2022 [cited 2022 Jan 23]. Available from: https://www.eia.gov/dnav/pet/PET_PRI_SPT_S1_M.htm
15. International Energy Agency (IEA). World energy outlook 2022 [webpage on the Internet]. c2022 [cited 2023 Mar 25]. Available from: https://www.iea.org/reports/world-energy-outlook-2022
16. World LPG Association. The role of LPG in shaping the energy transition: 2018 Edition [document on the Internet]. c2018 [cited 2023 Mar 25]. Available from: https://www.wlpga.org/wp-content/uploads/2018/10/The-role-of-LPG-in-shaping-the-energy-transition-2018.pdf
17. Hitachi. LPG burning gas turbine for power generation [document on the Internet]. c2014 [cited 2023 Apr 07]. Available from: https://lpg-apps.org/uploads/Modules/Library/power-generation-lpg-gas-turbines-kuba_hitachi_mitsubishi_japan.pdf
18. Marsh DA. Conversion of a kerosene-fuelled gas turbine to run on propane [Meng thesis]. Stellenbosch: Stellenbosch University; 2019. Available from: https://scholar.sun.ac.za/bitstream/handle/10019.1/106094/marsh_conversion_2019.pdf?sequence=1&isAllowed=y
19. Siemens. Siemens gas turbine portfolio [document on the Internet]. c2020 [cited 2020 Sep 20]. Available from: https://assets.new.siemens.com/siemens/assets/api/uuid:10f4860b140b2456f05d32629d8d758dc00bcc30/gas-turbines-siemens-interactive.pdf
20. South Africa opens Sunrise Energy LPG terminal. Riviera Newsletters. 2017 September 14. Available from: https://www.rivieramm.com/news-content-hub/news-content-hub/south-africa-opens-sunrise-energy-lpg-terminal-27227
21. South African Department of Minerals Resources and Energy (DMRE). Breakdown of petrol, diesel and paraffin prices as on 05 March 2023 [document on the Internet]. c2023 [cited 2023 Mar 25]. Available from: http://www.energy.gov.za/files/esources/petroleum/February2020/Breakdown-of-Prices.pdf
22. Mackenzie-Hoy T. Logistical questions hang over Ankerlig's Nov 2014 diesel burn. Engineering News. 2015 March 24. Available from: https://www.engineeringnews.co.za/article/logistical-questions-hang-over-ankerligs-nov-2014-diesel-burn-2015-03-24/rep_id:4136

Check for updates

# The impact of the increasing residential battery backup systems on load shedding

**AUTHORS:**
Michael J. Ritchie[1]
Jacobus A.A. Engelbrecht[1]
Marthinus J. Booysen[1,2]

**AFFILIATIONS:**
[1]Department of Electrical and Electronic Engineering, Stellenbosch University, Stellenbosch, South Africa
[2]Department of Industrial Engineering, Stellenbosch University, Stellenbosch, South Africa

**CORRESPONDENCE TO:**
MJ Booysen

**EMAIL:**
mjbooysen@sun.ac.za

**HOW TO CITE:**
Ritchie MJ, Engelbrecht JAA, Booysen MJ. The impact of the increasing residential battery backup systems on load shedding. S Afr J Sci. 2023;119(9/10), Art. #16602. https://doi.org/10.17159/sajs.2023/16602

**ARTICLE INCLUDES:**
☐ Peer review
☐ Supplementary material

**KEYWORDS:**
load shedding, battery backups, charging rate, inverter penetration level

**PUBLISHED:**
31 August 2023



**Significance:**

Load shedding has become the norm in South Africa. These rolling blackouts currently range from Stage 1 to Stage 6. Households are disempowered for 2–4 hours and for an average of 1.5–9 hours per day. In financially unequal South Africa, heavy users can afford battery backup solutions to keep the lights on. However, installing these at scale, without solar generation, eventually neuters the utility's ability to stabilise the grid and avert a blackout with shedding. Here we assess and quantify the impact of these interventions using an electricity data set of 12 000 households.

## Background

The lack of energy security in South Africa has resulted in the occurrence of rolling blackouts for, as of 2023, 15 years. This is due to the unreliability of Eskom, the country's main supplier of electricity to the grid, to prevent the energy crisis caused by a lack of supply from its fleet of coal-fuelled power stations. Some of the most fundamental factors that have led to the current state of electricity generation are the lack of investment in new generation capacity since 1998, a revenue shortfall from not having permittance for cost-reflective tariffs, a backlog of older Eskom plants, mismanagement, and operating at a higher-than-benchmarked energy utilisation factor.[1,2] More information regarding the rolling blackout and Eskom's energy crisis can be found in Styan[3].

If the power demand exceeds supply, the generators struggle to keep up and may be required to be disconnected, resulting in a further deficit of supply and increasing the risk of grid collapse and a total blackout for an extended period of potentially 2 weeks.

There are two solutions to the problem: demand management and increased supply. While the latter is time-consuming and capital-intensive, planned power cuts, termed *load shedding*, are the preferred option to reduce demand. Each stage of load shedding reduces the total energy demand by 1 Gigawatt (GW).[2]

The installation of home power backup systems, i.e. inverters and batteries or generators, is an increasingly popular solution to reduce the impact of power cuts for South African users. Domestic users tend to install hybrid inverters with batteries to keep the lights on during load shedding. Although a lithium battery with a capacity of 5 kWh can theoretically supply 5 kW for an hour, reducing the state of charge below 20% diminishes its capacity and causes it to age faster. Solar panels can augment the electricity supply by generating electricity directly to the household and storing excess energy in batteries, while exporting excess electricity to the grid, if the installation allows it.

Because of the additional cost of solar supply and getting approval for it, many users may opt to install only battery backups and inverters. However, installing inverters without solar panels increases the load on Eskom, outside of loadshedding periods, as the batteries charge when power from the grid is restored. This counteracts the utility's (Eskom's) attempt to manage the grid's stability through the demand management technique of load shedding.[4]

Increasingly, those who can afford to do so will revert to power backup solutions. Figure 1 shows the extent to which load shedding has escalated since September 2022. Also shown in the figure is the search appearances of the terms "solar," "battery," and "inverter" on Google Trends. Moreover, searching the term "solar" occurred significantly less than that for "battery" and "inverter," suggesting that a large portion of users consider inverter installations without solar panels.

Inverter prices are decreasing as the market grows, and can predominantly be purchased with a 3 kW, 5 kW or 10 kW power rating. Using a 5 kW inverter as an example, the inverter cost is typically ZAR35 000 and the lithium battery cost is approximately ZAR25 000. Solar panels tend to range from ZAR8000 to ZAR10 000 per kWp (which also pays for the system over time). A certified electrician is required to install the inverter, with installation costs of between ZAR10 000 and ZAR20 000.[4] This amounts to an estimated total of ZAR120 000. The consequence is that many people may opt to install a backup solution without installing a solar power supplement.

In this Commentary, we explore the extent to which penetration of such solar-free backup solutions restricts Eskom's ability to use load shedding to manage grid stability.

## Experimental setup

To answer this question, we present a simulation of a large group of residential households in South Africa to determine the aggregated electricity usage. We evaluate the impact of load shedding on the grid and the effect of users installing inverters.

### *Source data*

The household electricity usage data consists of hourly measured time-series data with a resolution of 0.5 A for 12 000 households measured between 1994 and 2014 and spanning at least one year.[5] Each sample is calculated as the average current used during that hour, measured in amps, for the given household. We determine the *electricity* used for each hour, measured in kilowatt-hours, by assuming a constant operating voltage. The data set

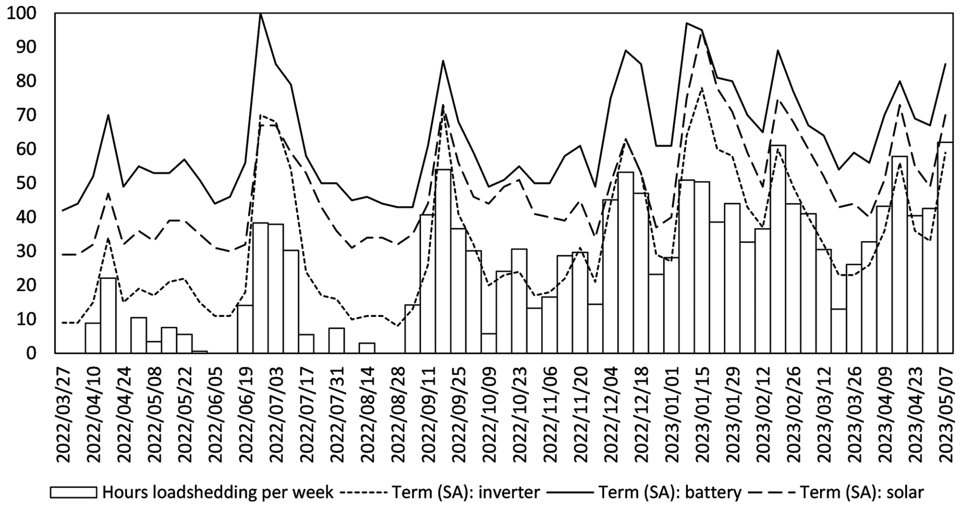



**Figure 1:** The extent to which load shedding has escalated since September 2022. Also shown in the figure is the search appearances of the terms "solar," "battery," and "inverter" on Google Trends.

may include small businesses, but the proportion of small businesses is estimated to be less than 6% of the data set[6], and small businesses are also likely to install inverters to keep the lights on.

The databases, the model and the simulator code are available at https://bit.ly/loadsheddingmodel.

## *System design and limitations*

Here we discuss how load shedding can be mitigated by inverters and batteries, followed by the limitations of the aforementioned components.

### Emulating load shedding

We used two 1-week electricity usage profiles (1 week for summer and 1 week for winter) for 12 000 residential households to obtain the results. The data are presented with hourly time steps but are simulated by the minute to accurately capture loads and battery charging behaviour. The households are separated into 20 groups, each with its own unique load shedding schedule. After each load shedding period, the inverters begin to charge for households belonging to the group whose power has just returned. The duration of load shedding is 2 hours and load shedding begins for each consecutive group in 1-hour periods, resulting in an overlap of load shedding across two groups. An example of the load shedding schedule is represented as follows:

$$\begin{Bmatrix} Zone\,1{:}0011111111 \\ Zone\,2{:}1001111111 \\ Zone\,3{:}1100111111 \\ \vdots \\ Zone\,9{:}1111111100 \end{Bmatrix}$$

where 1 and 0 refer to an hour slot where power is available and unavailable, respectively.

### Emulating inverters and batteries

Inverters with various battery sizes and charging rates are distributed to users whose level of usage is above the penetration threshold, i.e. a penetration level of 15% indicates that inverters are assigned to users within the top 15% usage levels. The inverters are assigned to these users based on their usage intensity, as follows:

1. The top 25% of these users are assigned 10 kWh batteries with a 5 kW charge rate.
2. The middle 50% of these users are assigned 5 kWh batteries with a 2.5 kW charge rate.
3. The bottom 25% of these users are assigned 3 kWh batteries with a 1.5 kW charge rate.

### Inverter and battery limitations

The following limitations are applied *during* load shedding:

1. If the household is drawing more than the inverter capacity, we ignore the household's load that exceeds the inverter rate (the household can use only as much as the inverter allows). We, therefore, assume that a household would adapt its usage to utilise the inverter capability to close its capacity. Admittedly, this is a crude assumption, which could be refined in future work.
2. If the battery's state of charge is reduced to or below 20%, the battery will stop providing power and the household load is zero.

The following limitations are applied *after* load shedding:

1. The battery's charge rate exponentially decreases when its state of charge is above 85%.

## *Simulation setup*

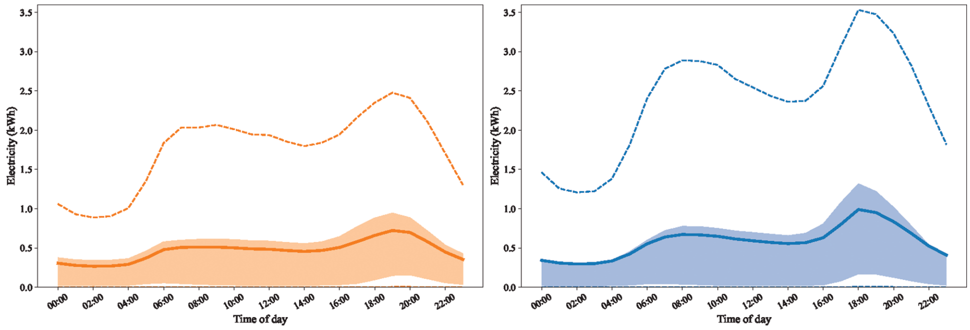
We simulate the electricity usage for all of the households and determine the aggregated load across all 7 days of the week. The data are separated into seasons to compare the effect of seasonality. We determine a measure of how much of the intended load shedding stage the inverters and batteries have curtailed, which is calculated as follows:

$$\%L = \frac{Load(inverters) - Load(loadshedding)}{Load(normal) - Load(loadshedding)} \times 100\%$$

# Results

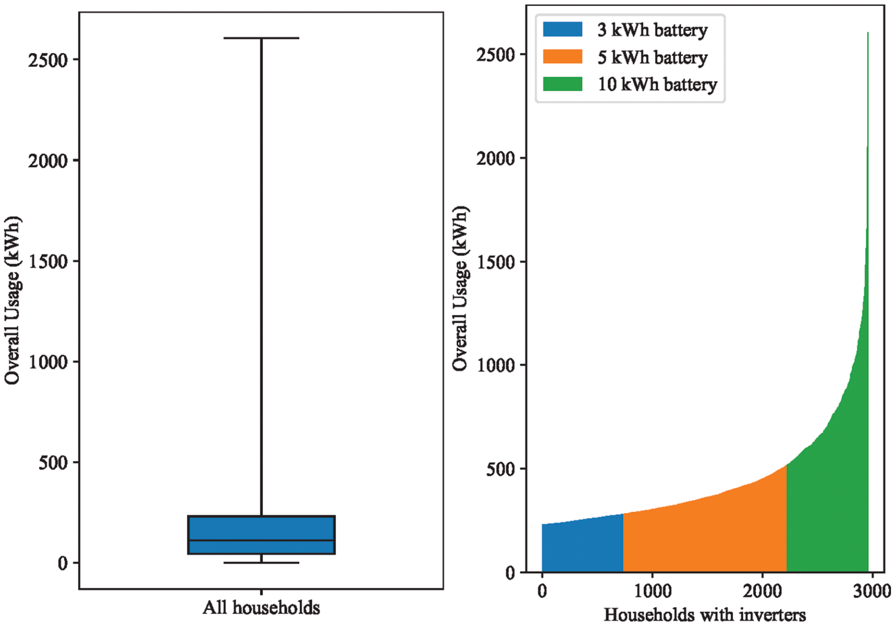
Figure 2 shows the normalised (per household) grid daily energy usage for 12 000 households for (a) summer and (b) winter. The total mean usage for summer and winter is 11.2 kWh and 13.9 kWh, respectively. The upper extremity (dashed line) profile is significantly higher than the interquartile range and indicates that the top 25% of households, which will also be the ones that opt to install the battery backup solutions, have a substantially larger level of usage.

Figure 3a shows a box plot of the overall usage for the summer and winter weeks per household. As seen from the previous figure, there is a large usage bias that is observed in the top 25% of households. Figure 3b shows a bar graph of the distribution of 3 kWh, 5 kWh and 10 kWh inverters to households with a usage level above a penetration level of 25%.



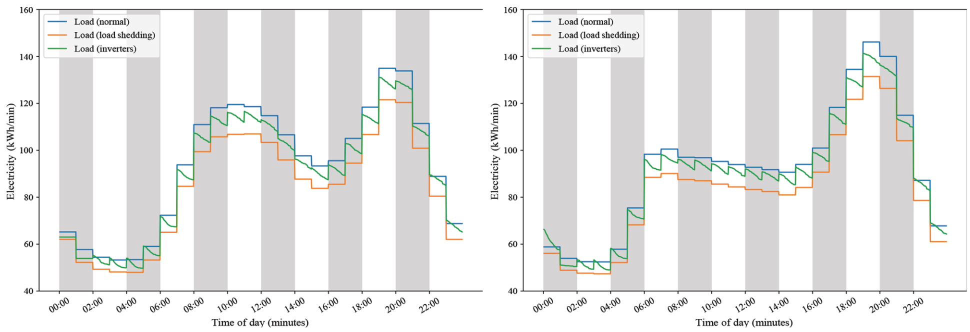

(a) (b)

**Figure 2:** Normalised grid daily energy usage for 12 000 households for (a) summer and (b) winter, where the mean (solid line), interquartile range (shaded area) and extremities (dashed line) are shown for all of the daily usage profiles. Units are kWh per hour, and the total mean daily usage for summer and winter is 11.2 kWh and 13.9 kWh, respectively.




(a) (b)

**Figure 3:** (a) Box plot of the overall usage for the summer and winter weeks per household and (b) a bar graph of the distribution of 3 kWh, 5 kWh and 10 kWh inverters to the top 25% of households.




(a) (b)

**Figure 4:** A plot of *Load*(normal) in blue, *Load*(loadshedding) at 10% of the load in orange, and *Load*(inverters) in green of all 20 zones for a Monday in (a) summer and (b) winter. The results are shown for a penetration level of 15%.

Figure 4 shows a plot of $Load_{normal}$ in blue, $Load_{loadshedding}$ in orange, and $Load_{inverters}$ in green of all of the 20 zones for a Monday in (a) summer and (b) winter. The results are shown for a penetration level of 15%.

As expected, the impact of implementing a notional two stages (10%) of load shedding decreases the overall electricity usage by 10% of the normal load, as shown by the orange usage profile. However, the

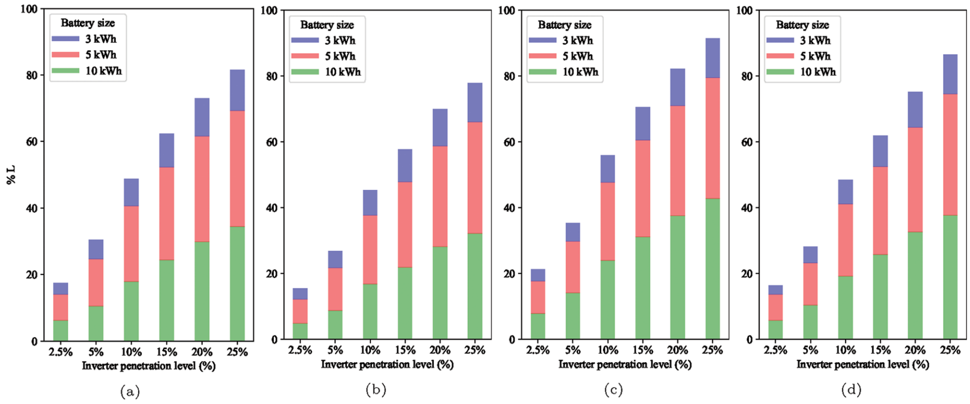



**Figure 5:** Bar graph of the %*L* for penetration levels of 2.5%, 5%, 10%, 15%, 20%, and 25% for (a, b) summer and (c,d) winter. Results are shown for (a,c) the overall daily usage profile and (b,d) the peak hours of the day.

installation of inverters increases the amount of energy demand on the grid and negates the impact of load shedding, as shown by the green usage profile. This is caused by the fact that when the power is returned for a zone that just experienced load shedding, all the inverters will begin to charge at the same time, which, consequently, pulls additional electricity from the grid. Most importantly, a peak is formed at the start of the charging period. With a 15% penetration level, the load peak is closely matched to the normal load, meaning that load shedding is negated. This peak can range from matching the load shedding load to surpassing the normal load by adjusting the inverter penetration level and varying the charging rate.

Figure 5 shows a bar graph of %*L* for a penetration level of 2.5%, 5%, 10%, 15%, 20% and 25% for (a,b) summer and (c,d) winter. Moreover, the contribution of the results for the three battery sizes is indicated. Figure 5a and Figure 5c show the results for the overall daily usage profile, and Figure 5b and Figure 5d show the results obtained during only the peak hours of the day, i.e. between 18:00 and 20:00.

By comparing the results of $Load_{inverters}$ for the various penetration levels, we see that battery charging can undo approximately 85% and 90% of load shedding in summer and winter, respectively, at a PV-free inverter penetration level of 25%. Even with a penetration level of 15%, these values are still as high as approximately 70%.

## Conclusion and recommendations

Electricity has become an essential component of modern society, and load shedding, which causes interruptions in power supply, has a significant impact on individuals and communities. The electricity grid connects people across multiple divides and affects all members of society to varying degrees.

Given the impact of load shedding on everyday life, many of those who can afford it, are opting to install battery backup solutions with inverters. Those who can afford to do so also happen to be the more affluent, who also tend to be the heaviest electricity users. However, given the additional cost and effort of augmenting a battery backup solution with solar power, many may be opting to not install additional generation capacity. The tax refunds announced in the 2023/2024 financial year may not mitigate this effect.

We assessed the impact of domestic users installing battery backup solutions without solar and evaluated the impact of doing so for different penetration levels of such systems into the domestic market. The results show that with a mere 15% penetration, the intended effect of load shedding will be curtailed by just over 60% in summer and just over 70% in winter when considering the whole day's curtailment.

We conclude that the impact of allowing users to charge at 0.5 C (half the battery capacity equivalent), without solar augmentation, will have a dramatic impact on the domestic load, even with only 15% penetration. It is therefore imperative that charging batteries from the grid is restricted to protect the potency of load shedding as a grid-balancing tool. The default charge rate on inverters can be as high as 1 C (e.g. 5 kW for a 5 kWh battery). We recommend that the charging rate of battery backup solutions is restricted to 0.15 C (for example 0.75 kW for a 5 kWh battery) to prevent the high curtailment after a zone is switched on. This lower charge rate should be ample to recharge the battery between bouts of load shedding.

## Data availability

The model and synthesiser source code and data are available at: https://bit.ly/loadsheddingmodel.

## Competing interests

We have no competing interests to declare.

## Acknowledgements

We thank MTN South Africa for their financial support, and EskomSePush for making the load shedding data publicly available.

## References

1. Van Diemen E. Eskom in the dock: 25 years in the making – the real reasons we have rolling blackouts according to de Ruyter. Daily Maverick. 2023 February 26. Available from: https://www.dailymaverick.co.za/article/2023-02-26-25-years-in-the-making-the-real-reasons-we-have-rolling-blackouts-according-to-de-ruyter/

2. Booysen MJ, Rix AJ. South Africa's power grid is under pressure: The how and the why. The Conversation. 2021 November 02. Available from: https://theconversation.com/south-africas-power-grid-is-under-pressure-the-how-and-the-why-170897

3. Styan JB. Blackout: The Eskom crisis. Johannesburg: Jonathan Ball; 2015.

4. Booysen MJ, Rix AJ. Home power backup systems – electrical engineers answer your questions. The Conversation. 2023 February 17. Available from: https://theconversation.com/home-power-backup-systems-electrical-engineers-answer-your-questions-199808

5. Toussaint W. Domestic electrical load metering, hourly data 1994-2014. [data set on the Internet]. c2019 [updated 2020 Apr 15; cited 2023 Apr 24]. Available from: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/759

6. Avenant J. Residential load modelling to predict household consumption for design of photovoltaic systems [Meng thesis]. Stellenbosch: Stellenbosch University; 2022. http://hdl.handle.net/10019.1/126052

Check for updates

# Some real but mostly unconsidered costs hiding in the dark corners of load shedding

**AUTHORS:**
Marthinus J. Booysen[1,2]
Servaas van der Berg[3]
Pieter W. van der Walt[1]

**AFFILIATIONS:**
[1]Department of Electrical and Electronic Engineering, Stellenbosch University, Stellenbosch, South Africa
[2]Department of Industrial Engineering, Stellenbosch University, Stellenbosch, South Africa
[3]Department of Economics, Stellenbosch University, Stellenbosch, South Africa

**CORRESPONDENCE TO:**
MJ Booysen

**EMAIL:**
mjbooysen@sun.ac.za

**HOW TO CITE:**
Booysen MJ, van der Berg S, van der Walt PW. Some real but mostly unconsidered costs hiding in the dark corners of load shedding. S Afr J Sci. 2023;119(9/10), Art. #16596. https://doi.org/10.17159/sajs.2023/16596

**ARTICLE INCLUDES:**
☐ Peer review
☐ Supplementary material

**KEYWORDS:**
load shedding, South Africa, economic growth, social cohesion, demand management

**PUBLISHED:**
31 August 2023



**Significance:**

Culling electrical demand to save a fickle supply costs South Africa more than you may think. Load shedding is South Africa's new norm, affecting each of us, with individual impacts ranging from annoyance to ruin. Much of our vital and collective electrical lifeline is sacrificed on the altars of convenience, inefficiency, and oversight. This Commentary sheds light on some of the hidden costs of load shedding. We call for custodianship of this crucial and limited shared source of life, and call on government to incentivise electricity generation, preferably the renewable and distributed type.

## Context

Our way of life, and often life itself, requires a stable supply of electricity. Establishing the infrastructure to generate this vital energy source is a costly and time-consuming undertaking. Generating electricity also costs money: operational costs include labour costs, maintenance costs, and the cost of fuel – coal in South Africa's case. In South Africa, a parastatal behemoth called Eskom is responsible for the generation, transmission, and a large part of the distribution and retail of electricity. Municipalities do the remaining distribution and sales. Unlike the norm in developed countries, South Africa is especially reliant on electricity for domestic and commercial energy needs. For example, electricity predominantly drives climate control, cooking, water heating, and smelting.

At face value, South Africa's unsubsidised electricity is amongst the cheapest in the world.[1] However, this pricing ignores the numerous government bailouts, which effectively act as hidden subsidies.

Managing electrical grid stability requires a fine balance between supply (generation) and demand (load). Load shedding, a colloquial term for scheduled rolling blackouts, has become the norm in South Africa, with the country currently experiencing load shedding ranging from Stages 1 to 6. Load shedding has resulted in a reduction in the weekly electricity availability factor (hours per week with electricity) from 100% to as low as 63% in the year from March 2022, as shown in Figure 1.[2] Load shedding, by definition, implies that there is an unsatisfied demand for electric energy, and, by design, demand is forcefully reduced to less than available supply, with a safety margin. This forced reduction in demand is needed to ensure that the grid does not collapse, which would result in a complete blackout. Such a blackout could happen in a matter of minutes, and examples exist of such events elsewhere, including in developed countries. A blackout could take days or even weeks to resolve, because the generation units require incremental restarting for synchronisation.[3]

Eskom has a supply crisis, but it cannot be observed in isolation from the demand and potential for demand management. From the demand perspective, the loss of electricity could result in a mere inconvenience (e.g. darkness before bedtime or an inability to cook food at a chosen time), a hindrance (e.g. preventing learners from studying after hours), a health risk (occupants reverting to burning wood for cooking or heating, or broken cold chains for medicines or foodstuffs). Crucially, these energy amputations also lead to lost commerce, production, and business operations – all of which have an adverse impact on economic activity in a country desperate for economic growth.

## Restricted growth and risk to social cohesion

### *Short-term effect on growth*

By affecting production, load shedding curtails economic growth. Modelling by the South African Reserve Bank (SARB) indicates that load shedding may have reduced 2022 economic growth by 2.2–2.5 percentage points per annum.[4] Sustained over several years (and currently, prospects for rapid improvements seem poor), such growth foregone by not utilising existing productive capacity on an annual basis translates into thousands of jobs not created, slower growth of wages, particularly for public servants, and constraints on public sector resources that will affect service delivery and the creation and maintenance of infrastructure.

### *Long-term effect on growth*

Growth foregone in the short term does not include the effect of uncertainty – to which load shedding contributes abundantly – on investor confidence and inflows of international capital. This reduces the growth of long-term productive capacity. Such investment is extremely sensitive to changes in perceptions of the growth prospects of an economy and of the likelihood of social and political stability. Foreign exchange is a particularly valuable resource for a smallish country such as South Africa, where capital goods are mostly imported. Investors require a positive investment climate, where there are perceived favourable returns on investment and relatively low risk of social and political conflict. Positive investor sentiment is difficult to achieve while load shedding is so central in our daily lives.

### *Three phases of post-transition growth*

There have been three phases in South Africa's economic growth since the transition to democracy in 1994, shown in Figure 2. Growth was modest until the turn of the century, largely because investors were uncertain whether the new government would implement fiscally prudent policies. Income grew at only around 1.1% per capita per year.

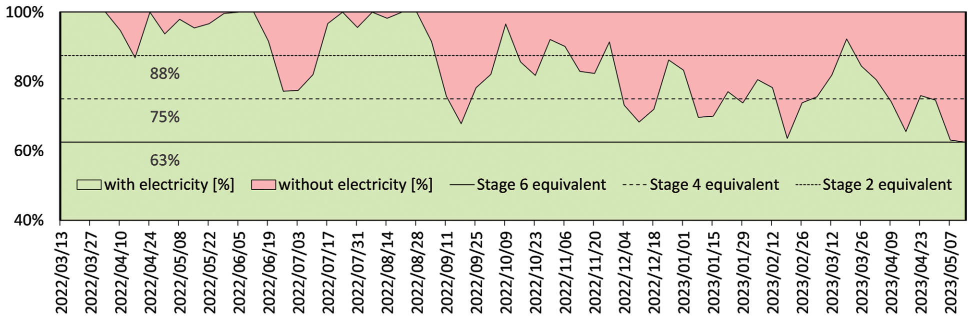



**Figure 1:** The electricity availability factor, expressed as a percentage of the hours per week, has reduced substantially since March 2022. Derived from Eskom Se Push data, 2023.[2]

Once investors became convinced that the new government would indeed implement sensible macroeconomic policies, per capita growth accelerated to a much more respectable 3.7% between 2000 and the 2008 financial crisis. At this rate of growth, average incomes could double in 19 years. Although negative domestic factors might have put an end to this growth spurt soon enough, its end was precipitated by the so-called "great recession" – a global event. While the rest of the world soon recovered and returned to almost similar levels of growth in the subsequent decade, South Africa's homegrown chickens came home to roost. To the litany of corruption, political decay, uncertainty of ownership, employment proscriptions, failed social delivery – at national, provincial, and increasingly also local level, massive unemployment, crime, and violence was added load shedding. We got stuck in a low growth equilibrium, with per capita growth dropping to a meagre 0.8% per year, before the COVID-19 pandemic worsened an already dire situation. At such a growth rate, a doubling of incomes would not take the 19 years mentioned for the 2000–2008 period, but three generations – 90 years. This while world growth almost fully recovered from the financial crisis, and many African countries showed solid growth.

### *A zero-sum game strains social cohesion*

The failure to return to a more acceptable growth after 2008 should be seen against the backdrop of great uncertainty. Load shedding is not the only impediment to strong economic growth – logistical issues relating to export capacity also contributed. But load shedding considerably exacerbated the already fragile business and consumer confidence.

Poor growth has fuelled social frustrations, and social frustrations that boil over in turn create conditions that constrain growth. South Africa thus now faces something close to a zero-sum game in the economic sphere, with per capita resources almost stagnant. This will make it much more difficult to escape from this vicious cycle of conflict and low growth. Living standards are directly affected by economic growth and rising unemployment. Moreover, we require economic growth to increase the social resources available for social spending, in particular spending on those things that our society still needs – school infrastructure, housing, water and sanitation, a strong early childhood development sector, to name just a few. Where growth is stagnant, those who gain can largely do so at the cost of others, heightening the already high levels of underlying conflict in our society.

## Projecting our shed onto GDP

During bouts of shedding, households and businesses are disempowered for 2 to 4 hours at a time, and for an average of 1.5 hours (Stage 2) to 9 hours (Stage 6) per day. The unsatisfied and 'shed' demand does result in lost opportunities for generating revenue from the sale of electricity. However, as Eskom reportedly sells electricity effectively at a loss, we do not consider this cost here.

Although load shedding achieves the goal of severing demand and thereby stabilising the grid, it remains a blunt tool. Consider the electrical loads on the grid. Much of it will be essential and/or could contribute to GDP, but a large part will be non-essential. Separately, much of it will be inefficient, for example, the use of inefficient lighting, the uncontrolled scheduling of water heaters, or the poor thermal insulation of buildings. It therefore stands to reason that a substantial part of the load on the grid, which contributes to the need for load shedding, is discretionary, non-essential or wasteful.

Quantifying the cost of load shedding is a daunting task, with estimates varying. Using SARB's estimation of ZAR899 million per day, and that Stage 6 loadshedding requires 6 GW of load being shed, we can calculate the cost per kilowatt-hour (or colloquially known as a 'unit') of non-essential loads that contributed to the need for load shedding. For a 24-hour cycle, this is the equivalent of 144 GWh (gigawatt-hours) or 144 million kWh of energy shed per day. The cost to the economy of these loads can therefore be estimated as ZAR899/144 kWh = ZAR6/kWh. We will use this crude estimation to try to quantify the cost of some of our discretionary uses.

### *Water heating*

Although hygiene is important, most usages of hot water, especially those involving large volumes, are discretionary. To heat 80 L of water (the volume for a typical bath) during winter requires 5 kWh. This means that the economy subsidises each (discretionary) bath by ZAR30 *if the water is heated from the grid in a period of load shedding*. Additionally, and maybe more importantly, there are more than 5 million electric water heaters in South Africa, and each one consumes an estimated 10 kWh per day, and adds an average of 1 kW to the national evening peak load. Therefore water heating contributes approximately 50 GWh to the daily demand (~10% of the daily grid supply), and 5 GW to peak-time demand. Research has shown that smart, centralised, schedule-control and energy-conscious plumbing can reduce this by up to 22%.[6] An option that can substantially reduce energy required for water heating is heat pumps, but they are expensive, require servicing, and are prone to breakdowns. More importantly, in our abundantly sunny climate, solar thermal and solar photovoltaic (PV) solutions exist that can heat water with limited or no dependence on the grid.

### *Electric vehicles*

Drivers of electric vehicles claim to save a lot of money, because the cost per kilometre of fuel could be four times as much as that of electricity.[7] Although this is beneficial to the owner (a realistic gain of approximately ZAR1.50/km), the economy will effectively be subsidising those with electric vehicles at approximately ZAR6/kWh, *if they charge from the grid when Eskom must shed load*. Given a nominal efficiency of 0.2 kWh/km, that is a subsidy of ZAR1.20/km. In effect, those who can afford to buy electric vehicles and charge from the grid may unwittingly be undermining the economy for their own financial gain. Those who can afford to buy electric vehicles should be encouraged to install their own distributed generation.

### *Poorly insulated dwellings*

South African houses are notoriously poorly insulated. The exact figures are heavily dependent on the specifics of the house in question. However, even a small fin oil heater in a poorly insulated house consumes 2 kWh every hour. The economy would therefore be subsidising poorly insulated houses by ZAR12 per room per hour while the grid is load shedding. Homeowners should be incentivised to improve insulation.

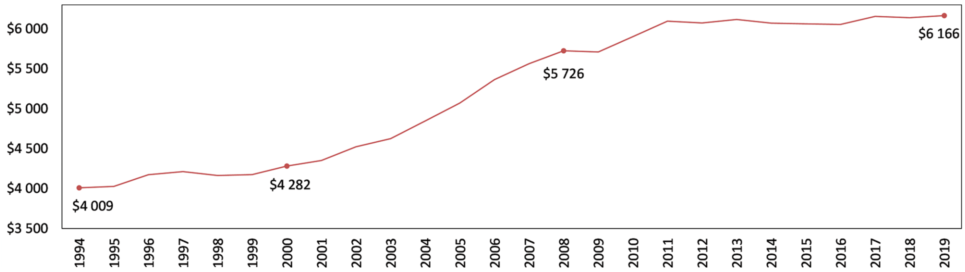



**Figure 2:** Gross national income per capita in 2015 USD, 1994 to 2021.[5]

## Cost to households, schools, and universities

### *Household inverter/battery and solar solutions*

In a desperate attempt to keep the lights on, many households and businesses that can afford to do so, are installing battery-backup solutions. The cost of doing so could range from approximately ZAR25 000 to ZAR100 000 when including only an inverter and batteries, and not including a solar plant. This price increases by approximately 50% when adding solar power and the requisite municipal approvals. Comparing this to our GDP per capita (ZAR110 000), and considering that half of employed South Africans earn less than ZAR60 000 per year from their job, investing in a battery backup with solar is clearly a costly and substantial investment. It is therefore tempting for homeowners to install a battery backup without the solar component.

Our simulations have shown that as much as 70% of peak-time household load shedding is curtailed when 15% of households – i.e. the affluent users that can afford it, who are also the heavy users – install battery-backup solutions without solar supplement. Doing this reduces load shedding to mere shifting the load in time, which can come at a large collective cost because it will seriously limit Eskom's ability to stabilise the grid. Adding solar generation to such systems must be strongly encouraged.

### *Schools*

South Africa's education system is in a dire state, with poor performance in international tests in mathematics and reading. Although a large part of teaching in most schools is not heavily dependent on having electricity, losing valuable teaching time due to load shedding impacts the overstressed system even further and makes it even harder to overcome some of the deficits with subjects such as computer-based technology in many of our poorest schools. In this context, it is crucial to future development that teaching, and school administration, are not interrupted by load shedding and that school expenses are managed well. Schools have two options to manage load shedding: installing a generator and installing battery-backup/inverter systems. It is important to note that the capital and operating costs of these solutions could be reduced substantially if essential loads are split from the rest. Unfortunately, the time, skills, and resources to manage this are rare. The cost to install a generator depends on which loads are hooked up, but could be as much as ZAR250 000 for a small school to ZAR1 million for a large school with a hostel. What breaks the bank is the running cost of the diesel. At one admittedly large school we are supporting, the unbudgeted diesel cost is ZAR72 000 per month. The alternative, a battery-backup/inverter solution, for a school could range from ZAR100 000 for a small school to ZAR2 million for a large school. Again, adding solar provision increases the cost. Schools should be provided with timely guidelines on what is needed and what is affordable to overcome the worst of load shedding.

### *Universities*

Our tertiary institutions are by no means spared from load shedding. As lectures and research cannot wait and sit idle, most universities have installed diesel generators. These generators entail a substantial capital expenditure but the operational cost of running them is frightening. Data provided by the Minister of Higher Education, Science and Innovation in 2023 reported the *daily* cost of diesel at Stage 6 to be ZAR197 019 for the University of Cape Town; ZAR892 473 for the University of Johannesburg; ZAR492 000 for North-West University; ZAR2 201 711 for the University of Pretoria; and ZAR342 939 for Stellenbosch University. The total for the 13 universities reported on was ZAR2 041 909 per day at Stage 3 and ZAR4 469 618 per day at Stage 6. This does not consider the cost of maintenance for the generators. In comparison, the NSFAS (National Student Financial Aid Scheme) budget for 2023 was ZAR47.6 billion, which supports 900 000 students at up to ZAR65 000 per year.[8] This means that for every *day* of Stage 6 load shedding, an additional 69 students could have been supported for a *year's* studies from the savings of the universities alone.

## Conclusion

The pervasiveness of electricity in our lives, and our modern society's dependence on it, is mirrored in the inescapable and extensive impact of its interruption through load shedding. The intricate network that constitutes the so-called grid inextricably links us and our common destiny across its numerous divides. As was the case with Cape Town's #DayZero drought in 2017/2018, it does and will affect us all to some extent.

Although our fragile grid has a severe generation shortfall, its cost and implications must be observed in balance with the extant demand. This is especially true given the flexibility and inefficiencies and discretion inherent in the antipodal demand.

This Commentary has highlighted some hidden costs. An example of an obfuscated cost is that of electric vehicles being charged from the grid during load shedding, which results in a cross-subsidy of its mobility from the fiscus. The cost of electric vehicles means that the owners tend to be wealthy. They should be encouraged to provide their own renewable supply. Another example is that of the discretionary use of large volumes of heated water, which again results in cross-subsidisation. All users, but especially those who can afford to heat large volumes of water, should be incentivised to use the abundant solar resources for heating.

Crucially, many of the hidden costs of load shedding are a consequence of powering non-essential or inefficient loads. Demand management efforts should be redoubled, and participation in such schemes should be incentivised. Moreover, efforts to improve efficiency of existing loads should be rolled out and awareness campaigns initiated. In addition to rolling out centralised large-scale generation and battery storage systems, conditions are right for the speedy deployment of decentralised renewable energy generation and trading thereof.

What is considered expensive, essential, and wasteful is certainly fuzzy and subjective and will be different across the various spectra of our diverse society. But when it comes to balancing the grid to prevent a blackout of unknown duration, especially given our fragile political and economic landscape and societal frustrations, the sums on the left and right of the power scale are prime, and the need for grid stability unambiguous. Demand must be clamped to less than supply, preferably without impacting on essential services and economic powertrains.

Like a nation at war, South Africans are in a tight spot with limited options and restricted room for manoeuvre. This is expected to be the case for years to come. We all need to carefully re-assess our self-indulgences and decisions for the common good. This is not a time to look out for the self, but to keep our collective eyes on the front and focus our sustained and shared survival.

## Competing interests

We have no competing interests to declare.

## References

1. De Ruyter A. Eskom. Eskom-de Ruyter answering affidavit [document on the Internet]. c2023 [cited 2023 Apr 21]. Available from: https://www.scribd.com/document/628038380/Eskom-De-Ruyter-Answering-Affidavit-1
2. EskomSePush [app on the Internet]. Available from: https://esp.info/
3. Booysen MJ, Rix A. South Africa's power grid is under pressure: The how and the why. The Conversation. 2021 November 02. Available from: https://theconversation.com/south-africas-power-grid-is-under-pressure-the-how-and-the-why-170897
4. South African Reserve Bank. Quarterly Bulletin. 2022 December. Box 1. Available from: https://www.resbank.co.za/content/dam/sarb/publications/quarterly-bulletins/quarterly-bulletin-publications/2022/december/01Full%20Quarterly%20Bulletin.pdf
5. World Bank. World development indicators [webpage on the Internet]. No date [cited 2023 Apr 21]. Available from: https://databank.worldbank.org/source/world-development-indicators
6. Engelbrecht JAA, Ritchie MJ, Booysen MJ. Optimal schedule and temperature control of stratified water heaters. Energy Sust Dev. 2021;62:67–81. https://doi.org/10.1016/j.esd.2021.03.009
7. Electric vs petrol vs diesel vehicles in South Africa: Cost per kilometre. BusinessTech. 2022 March 03. Available from: https://businesstech.co.za/news/motoring/564438/electric-vs-petrol-vs-diesel-vehicles-in-south-africa-cost-per-kilometre/
8. National assembly statement by the Minister of Higher Education, Science and Innovation, Dr Blade Nzimande on the PSET readiness for the academic year 2023 and related matter [document on the Internet]. No date [cited 2023 Apr 21]. Available from: https://www.dhet.gov.za/SiteAssets/Media%20Advisory%20and%20Statements%202023/media%20Statement.pdf

# Data gaps will leave scientists 'in the dark': How load shedding is obscuring our understanding of air quality

**AUTHORS:**
Caradee Y. Wright[1,2]
Matthew Benyon[3,4]
Nomfundo Mahlangeni[5]
Thandi Kapwata[6,7]
Tracey Laban[1]
Rebecca M. Garland[2]

**AFFILIATIONS:**
[1]Environment and Health Research Unit, South African Medical Research Council, Pretoria, South Africa
[2]Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa
[3]Centre for Environmental Health and Sustainability, University of Leicester, Leicester, United Kingdom
[4]Institute for Environmental Futures, University of Leicester, Leicester, United Kingdom
[5]Environment and Health Research Unit, South African Medical Research Council, Cape Town, South Africa
[6]Environment and Health Research Unit, South African Medical Research Council, Johannesburg, South Africa
[7]Department of Environmental Health, Faculty of Health Sciences, University of Johannesburg, Johannesburg, South Africa

**CORRESPONDENCE TO:**
Caradee Wright

**EMAIL:**
caradee.wright@mrc.ac.za

**HOW TO CITE:**
Wright CY, Benyon M, Mahlangeni N, Kapwata T, Laban T, Garland RM. Data gaps will leave scientists 'in the dark': How load shedding is obscuring our understanding of air quality. S Afr J Sci. 2023;119(9/10), Art. #16009. https://doi.org/10.17159/sajs.2023/16009

**ARTICLE INCLUDES:**
☐ Peer review
☐ Supplementary material

**KEYWORDS:**
air pollution, air quality management, environmental health, rolling blackouts, South Africa

**FUNDING:**
University of Leicester, South African Medical Research Council, South African National Treasury

**PUBLISHED:**
31 August 2023



**Significance:**

South Africa's scheduled power outages, commonly known as load shedding, are increasing each year due to instability and poor performance of the existing fleet of power stations supplying electricity. The power provider projects that there will be load shedding every week for the next year. Data availability from the existing air quality monitoring stations infrastructure is already sparse over South Africa. Increased load shedding exacerbates this issue as power outages disrupt equipment operation. The collection of long-term and continuous ambient air quality data is needed for air quality-related research, policy and strategy development, and air quality management. The introduction of air quality monitors that are reliable and climate-friendly, such as passive samples, rechargeable battery-powered sensors and renewable energy powered sensors, might be interim interventions to ensure continuous data collection.

Every year, 6.7 million lives are lost prematurely due to the combined impact of outdoor and indoor air pollution.[1] In South Africa in 2019, there were ~30 000 deaths (6% of all deaths) associated with air pollution exposure.[2] However, there is a significant gap in the accurate reporting of air pollutant concentrations in South Africa.[3] Providing measured evidence of both ambient and household air pollution is important because accurate reporting of the significant health consequences of air pollution supports policymakers, decision-makers, and affected communities in their mitigation efforts.

Air quality monitoring and management of ambient air pollution are essential tools to ensure that concentrations of criteria pollutants, such as particulate matter (PM), meet South Africa's National Ambient Air Quality Standards (NAAQSs).[4] Data from air quality monitoring networks are also used extensively in research projects, such as burden of disease estimations and environmental impact assessments. However, disruptions in air quality monitoring have led to incomplete data sets. As a general guideline, it is recommended that data with <75% missing values be used.[5] Obtaining high-quality air pollution data is essential to South Africa's efforts to manage and reduce air pollution related health impacts.

Load shedding poses a significant obstacle to acquiring high-quality air pollution monitoring data. Higher stages of load shedding have a considerable impact on air quality monitoring as equipment cannot operate during power outages. Eskom implements various stages of load shedding, ranging from Stage 1 to Stage 8, which translates to between 2 and 10 hours without electricity per day.[6] Load shedding has become a permanent reality for South Africans, causing interruptions in communication, security, health, and emergency services.[7,8] In 2022, load shedding reached a total of 3773 hours, accounting for 43% of the year. This represents the highest level of load shedding experienced since its implementation in 2007.[8] Despite the termination of the National State of Disaster on electricity supply constraints[9], Eskom continues to implement load shedding, and higher stages of load shedding are projected to be more frequent in the coming months[10].

It is important to measure air pollution to inform air pollution mitigation strategies and to provide data for epidemiological studies to protect public health, particularly in vulnerable populations.[11] Enhancing the assessment and understanding of air quality in low- and middle-income countries is crucial. One notable example is the Global Environment Monitoring System for Air (GEMS Air) established by the United Nations Environment Programme (UNEP).[12] This initiative integrates data from satellites and ground-based air quality reference monitors, while also incorporating data from low-cost sensors, to achieve comprehensive spatial and temporal coverage worldwide.

During extended periods of load shedding, concerns arise regarding whether our air quality monitoring system will be capable of detecting changes in air quality. As the air quality monitoring system feeds data into the South African Air Quality Information System (SAAQIS)[13], the effects on our long-term understanding of air quality and its impacts need to be considered. For example, data gaps lead to estimates with high levels of uncertainty, which is a challenge for assessing the burden of disease and mortality that can be attributed to air pollution. This Commentary focuses on the quality of data captured by SAAQIS on days with load shedding and discusses the implications for air quality monitoring/management and research in South Africa.

## Air quality monitoring in South Africa

There are over 130 air quality monitoring stations that contribute to the measurement of ambient air quality for the SAAQIS (Figure 1). These stations measure PM, sulfur dioxide ($SO_2$), nitrogen dioxide ($NO_2$), ozone ($O_3$), and carbon monoxide (CO), among other air pollutants. These pollutants are evaluated against the NAAQS to assess whether concentrations of pollutants are within acceptable levels for human health. For example, the 24-hour average standard for $PM_{2.5}$ (particulate matter with an aerodynamic diameter smaller than 2.5 $\mu$g/m³) in South Africa is 20 $\mu$g/m³ – this may be compared to the more stringent World Health Organization guideline of 5 $\mu$g/m³.[14] Data available on SAAQIS go through data validation processes and are checked for zero drift, which, if identified, is corrected based on values from the most recent in-situ calibration. In accordance with prescribed standard operating procedures, suspicious data spikes, negative values and other questionable data points are removed

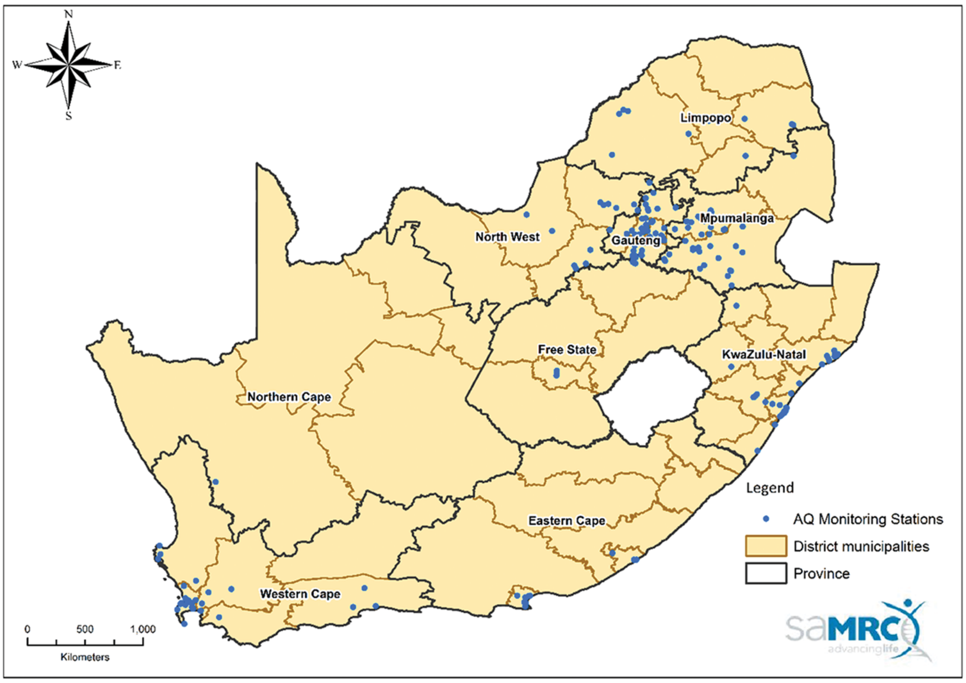



**Figure 1:** Location of the air quality monitoring stations in South Africa that report to the South African Air Quality Information System (SAAQIS) as of January 2023.

from the validated data using default data processing algorithms in SAAQIS. Once the data are downloaded, researchers often repeat quality control procedures.

## Load shedding impacts air quality monitoring data collection

Given that air quality monitoring stations rely on electricity provided by Eskom via the national grid, we expected that load shedding would impact the continuity of air quality data collection. In a recent meeting with the South African Department of Forestry, Fisheries, and the Environment (DFFE) and South African Weather Service (SAWS) it was reported that most air quality monitoring stations were frequently offline due to load shedding.[15] Additionally, there were reports of instrument faults caused by damage to the electronic components of air quality monitoring instruments due to power surges when electricity supply was restored.

For this Commentary, we conducted a preliminary analysis of air quality data from the Diepkloof monitoring station (as an illustrative example) in Soweto, Gauteng, spanning from January 2018 to February 2023. The data were acquired from SAAQIS and processed using Python. We carried out preliminary checks of the data set to assess data quality and completeness prior to conducting the time series in relation to prescribed load shedding schedules. To determine the scheduled load shedding periods for Diepkloof, we used data from a load shedding application.[16] The schedule was validated with the monthly schedule on the Eskom website specifically for the suburb of Diepkloof. An iterative algorithm was then employed to convert the load shedding schedules into a time series format which was then merged with the air quality data, enabling an examination of the association between the quality of the data and load shedding periods. These steps and analyses are deemed preliminary as we are presently carrying out a larger, more extensive analysis with additional pollutants and air quality monitoring stations, together with air quality data from satellites, to fully explore these relationships.

The preliminary analysis showed an association between the timing of load shedding periods and the absence of air quality data at Diepkloof monitoring station during the investigation period (i.e. 22 to 27 February 2023) (Figure 2a). To assess the relationship between the periods of load shedding and missing data, a chi-square test of independence was performed. To meet the assumptions of the test, 10% of the data was randomly sampled, resulting in 452 data points. The null hypothesis posited that the occurrence of missing data was independent of the scheduled load shedding hours and the critical value was 0.05. Table 1 shows the results of the test demonstrating that for all parameters tested the *p*-values were lower than the critical value, thus the null hypothesis was rejected. This indicates that there is a statistically significant relationship between the occurrence of NaN values (missing data) and the occurrence of scheduled load shedding at the Diepkloof SAAQIS monitoring station. No data were collected for the concentrations of $PM_{2.5}$, $NO_2$ and $SO_2$ pollutants during load shedding periods due to the cessation of the air quality monitoring equipment. Following the conclusion of load shedding periods, data collection resumed and pollutant concentrations gradually returned to pre-loadshedding levels. Consequently, the continuity of air quality monitoring was disrupted, resulting in incomplete daily measurements. To consider the difference in data capturing interruptions between a week with load shedding and a week when there was no load shedding, Figure 2b shows concentrations of the three pollutants of interest for Diepkloof at typical concentrations during the month of February. Thus, further investigation is warranted to understand the time-delay observed, especially for $PM_{2.5}$, before the instrument resumes normal monitoring operations after load shedding resumes.

Considering the escalating frequency of load shedding over the years and that the frequency and duration of daily occurrences of load shedding are expected to increase in 2023, the persistence of missing values will remain a major challenge. Load shedding between 1 January 2018 and 31 August 2022 amounted to 746 hours while a shorter period between 1 September 2022 and 27 February 2023 of approximately 6 months had 1075 hours of load shedding. Stage 1 load shedding was instituted between 1 January 2018 and 31 August 2022 and Stage 3 load shedding came into effect between 1 September 2022 and 27 February 2023. The proportion of missing data for air pollutants, excluding ozone, likely due to instrumentation problems was 7–10% between February and August 2022. This proportion of missing data increased to 27–33% in the period September 2022 to February 2023 when load shedding was more frequent (Figure 3).

It is important to note that the load shedding schedules that we used are the estimated on-and-off times for power outages. It is common that load shedding begins and ends at times different from the published on-and-off times. Moreover, there may be other reasons why there are missing data for the air pollutants. These need to be fully explored

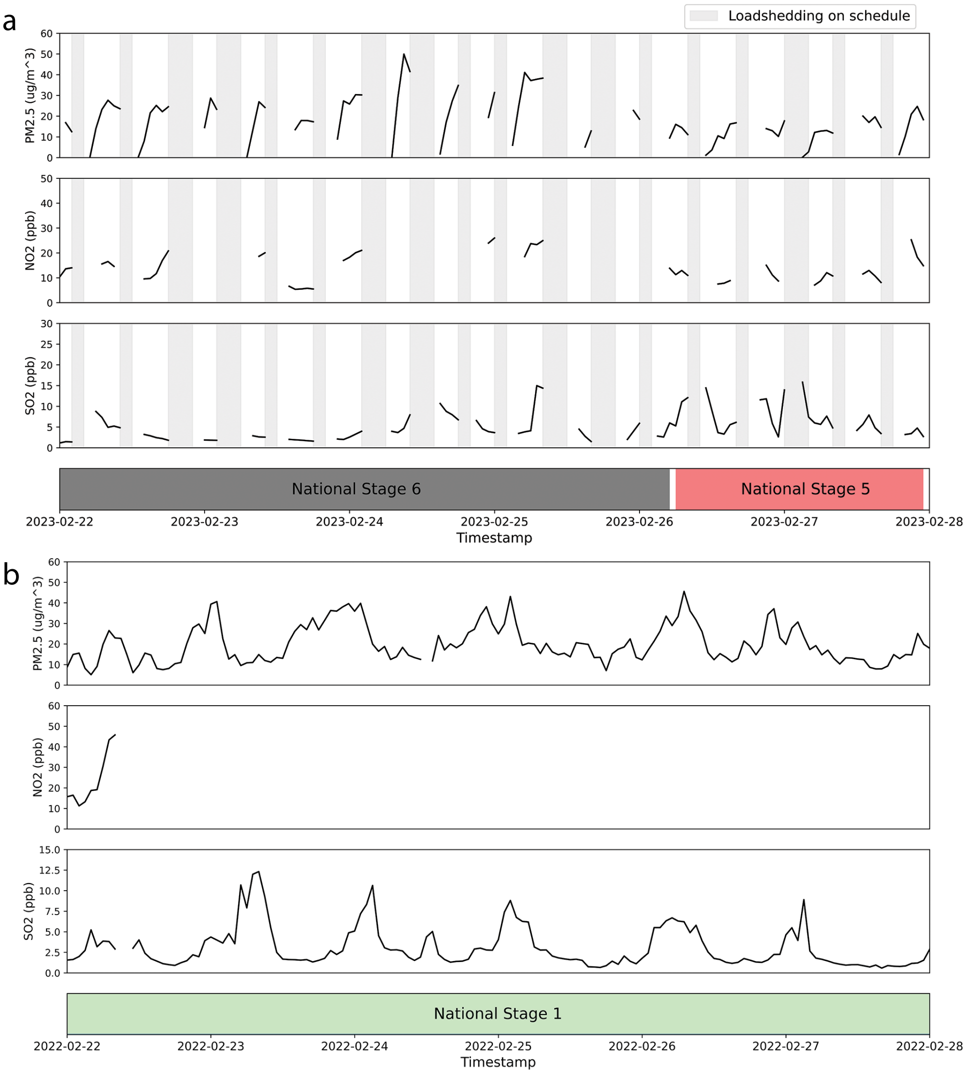



**Figure 2:** Concentrations of $PM_{2.5}$, $NO_2$ and $SO_2$ at the Diepkloof monitoring station between 22 and 27 February 2023 during: (a) load shedding Stages 6 and 5 (the width of the grey column is indicative of the time when load shedding was occurring) and (b) no load shedding (there were no data for $NO_2$ for the majority of this week for reason(s) unknown).

**Table 1:** $X^2$ and *p*-values resulting from the chi-square test of independence. Categories were the occurrence of scheduled load shedding and the occurrence of a NaN (missing data) value for the different parameters.

| Parameter | $X^2$ | *p*-value |
|---|---|---|
| Ambient temperature | 28.4 | < 0.001 |
| Ambient wind direction | 11.6 | < 0.001 |
| Ambient wind speed | 12.5 | < 0.001 |
| CO | 15.6 | < 0.001 |
| $NO_2$ | 12.3 | < 0.001 |
| $NO_x$ | 12.3 | < 0.001 |
| NO | 11.7 | < 0.001 |
| $O_3$ | 19.5 | < 0.001 |
| $PM_{10}$ | 7.23 | 0.007 |
| $PM_{2.5}$ | 3.92 | 0.048 |
| $SO_2$ | 9.47 | 0.002 |

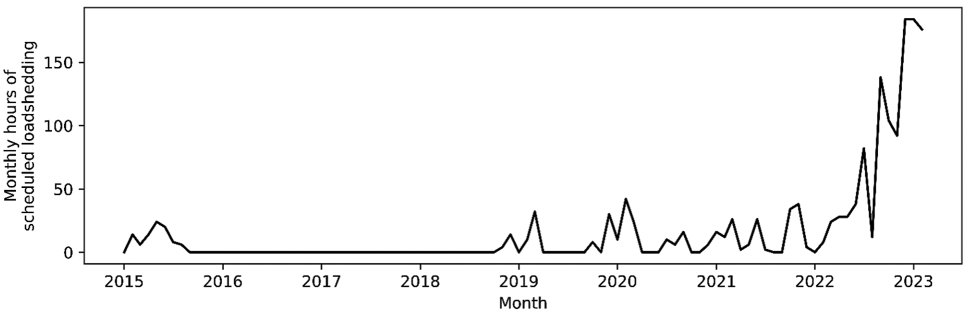



**Figure 3:** Monthly hours of scheduled load shedding between 2015 and 2023.

but may include such things as instrument errors or a problem with electricity being restored, e.g. an electrical short.

The data collected from air quality monitoring stations provide information about air quality at different spatiotemporal scales. Data can be used to assess compliance with ambient air quality standards and review policy measures designed to reduce emissions of pollutants and improve air quality in the long term. Considerable data gaps that occur when the stations are not functioning limit one's ability to measure compliance performance, especially for pollutants such as ozone with 8-hour (running) averaging periods. In addition, data from these monitoring stations are also used to calculate the Air Quality Index (AQI), which is a common tool employed to communicate the air quality in a particular location to the general public. The AQI requires high-resolution data as it is calculated for every hour and communicated to the general public in order to take protective action.[17] If there are hours with no monitoring data, air quality warning systems will fail, short-term peaks will not be reported to the wider public, and citizens will not be alerted to dangerous levels of air pollution.

## Challenges, opportunities and way forward

Given that air pollution is a major global, African and South African human health problem, the need for accurate air quality data is paramount. SAAQIS data are used by researchers in different fields of science including epidemiology[18], chemistry[19], and atmospheric science[20], among others. These data sets feed into air quality management processes and are used to calculate the National Air Quality Indicator[21] reported by the National Ambient Air Quality Officer (NAAQO) in the State of Air in South Africa and Environment Outlook reports[22]. SAAQIS data are also used to consider whether air pollution concentrations are below the NAAQS, and if not, what interventions should be put in place.

Data gaps reflected by the number of 'blank' readings (shown as 'NaN') and likely incorrect recordings after load shedding need to be communicated to all users of SAAQIS via the website. These gaps will affect the collection of continuous data for time series and other methodological uses, affecting the accuracy of the data. It is likely that load shedding is affecting other environmental monitoring systems across the country, and it is important for us to learn from each other how to ensure high-quality, continuous environmental data sets.

As the custodian of SAAQIS, the NAAQO has reported that they are aware of the problem of missing values in the air quality data sets (Gwaze P 2023, oral communication, March 16). They plan to install solar voltaic panels at each air quality monitoring station to ensure a continuous supply of energy to the air quality monitoring instruments as well as to install battery-powered low-cost sensors. However, this process is likely to be rolled out slowly in comparison to the urgency of the problem faced by a lack of continuous data. There is a need for the South African government to update existing policies and regulations and to restructure the electricity supply industry to ensure an increase in the contribution of (ideally) renewable energy sources to the grid.[23,24] Diversifying our source of electricity will ensure that air quality monitoring stations are operational and the collection of data is not disrupted.

## Conclusions

Our resilience to large-scale power outages experienced with load shedding requires a shift in reliance on coal-based electricity generation to renewable energy as much as possible. We must have air quality data during load shedding to understand potential risks to human health from all air pollution sources across the country. For air quality management, as well as for air pollution and associated epidemiological research, we need to act urgently. These air quality data sets are necessary to inform us whether we are placing people and communities at risk from polluted air, thus breaching the constitutional right to an environment that is not harmful to the health of South Africans. Continuous data sets with high capture rates are critical to assess long-term trends in air pollution levels, ensuring an understanding of the impact of implemented policies. Presently, we are 'in the dark' regarding the quality of air where air quality monitoring is occurring sporadically.

## Acknowledgements

The founders of ESP are thanked for the provision of the load shedding time series data and the South African Weather Service / South African Air Quality Information System is thanked for the air quality data. M.B. receives research funding through a PhD scholarship from the Institute for Environmental Futures at the University of Leicester. N.M. was partially supported as a postdoctoral fellow by funding from the South African Medical Research Council (SAMRC) through its Division of Research Capacity Development under the Intra-Mural Postdoctoral Fellow Programme from funding received from the South African National Treasury. The content hereof is the sole responsibility of the authors and does not necessarily represent the official views of the SAMRC or the funders.

## Competing interests

We have no competing interests to declare.

## References

1. World Health Organization. Household air pollution [webpage on the Internet]. c2022 [cited 2023 Jun 05]. Available from: https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health#:~:text=The%20combined%20effects%20of%20ambient,(COPD)%20and%20lung%20cancer

2. Health Effects Institute. The state of air quality and health impacts in Africa [webpage on the Internet]. c2022 [cited 2023 Jun 05]. Available from: https://www.healtheffects.org/announcements/new-state-global-air-special-report-air-quality-and-health-africa

3. Garland R. Clean air journal's 45th anniversary: CAJ archive highlights the importance of continuous air quality measurements. Clean Air J. 2016;26(1):2–3. https://doi.org/10.17159/2410-972X/2016/v26n1a1

4. South African Department of Forestry, Fisheries and the Environment. National Environmental Management: Air Quality Act, 2004 (Act No. 39 of 2004). 2009 National Ambient Air Quality Standards. Available from: https://www.gov.za/sites/default/files/gcis_document/201409/328161210.pdf

5. New Zealand Ministry for the Environment. Good practice guide for air quality monitoring and data management 2009 [webpage on the Internet]. c2009 [cited 2023 Apr 18]. Available from: https://environment.govt.nz/publications/good-practice-guide-for-air-quality-monitoring-and-data-management-2009/

6. Eskom. Interpreting load-shedding stages in South Africa [webpage on the Internet]. No date [cited 2023 Jun 05]. Available from: https://loadshedding.eskom.co.za/LoadShedding/ScheduleInterpretation
7. Mbomvu L, Hlongwane IT, Nxazonke NP, Qayi Z, Bruwer JP. Load shedding and its influence on South African small, medium and micro enterprise profitability, liquidity, efficiency and solvency (2021 April 21). Business Re-Solution Working paper BRS/2021/001. SSRN [cited 2023 Apr 18]. Available from: https://doi.org/10.2139/ssrn.3831513
8. Gehringer C, Rode H, Schomaker M. The effect of electrical load shedding on pediatric hospital admissions in South Africa. Epidemiology. 2018;29(6):841–847. https://doi.org/10.1097/EDE.0000000000000905
9. Government of South Africa. Government terminates National State of Disaster on electricity supply constraints [media release on the Internet]. c2023. [cited 2023 Apr 18]. Available from: https://www.gov.za/speeches/government-terminates-national-state-disaster-%C2%A0electricity-supply-constraints-5-apr-2023
10. South Africa warned of load shedding 'beyond stage 6' as winter approaches. BusinessTech. 2023 April 12. Available from: https://businesstech.co.za/news/energy/679511/south-africa-warned-of-load-shedding-beyond-stage-6-as-winter-approaches/
11. Arowosegbe OO, Röösli M, Künzli N, Saucy A, Adebayo-Ojo TC, Schwartz J, et al. Ensemble averaging using remote sensing data to model spatiotemporal $PM_{10}$ concentrations in sparsely monitored South Africa. Environ Pollut. 2022;310, Art. #119883. https://doi.org/10.1016/j.envpol.2022.119883
12. World Health Organization. Ambient (outdoor) air pollution fact sheet [webpage on the Internet]. c2022 [cited 2023 May 24]. Available from: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health
13. South African Air Quality Information System [homepage on the Internet]. No date [cited 2023 Jun 05]. Available from: https://saaqis.environment.gov.za/
14. World Health Organization. Global air quality guidelines: Particulate matter ($PM_{2.5}$ and $PM_{10}$), ozone, nitrogen dioxide, sulphur dioxide and carbon. Geneva: World Health Organization; 2021. https://apps.who.int/iris/handle/10665/345329
15. Parliamentary Monitoring Group. DFFE & SAWS Q1 & 2 2022/23 Performance with Minister [webpage on the Internet]. c2023 [cited 2023 Apr 18]. Available from: https://pmg.org.za/committee-meeting/36426/
16. ESP [load shedding application]. c2023 [cited 2023 Apr 18]. Available from: https://esp.info/
17. Gwaze P, Mashele SH. South African Air Quality Information System (SAAQIS) mobile application tool: Bringing real time state of air quality to South Africans. Clean Air J. 2018;28(1):3. https://doi.org/10.17159/2410-972X/2018/v28n1a1
18. Altieri KE, Keen SL. Public health benefits of reducing exposure to ambient fine particulate matter in South Africa. Sci Total Environ. 2019;684:610–620. https://doi.org/10.1016/j.scitotenv.2019.05.355
19. Hersey SP, Garland RM, Crosbie E, Shingler T, Sorooshian A, Piketh S, et al. An overview of regional and local characteristics of aerosols in South Africa using satellite, ground, and modeling data. Atmos Chem Phys. 2015:15:4259–4278. https://doi.org/10.5194/acp-15-4259-2015
20. Matandirotya NR, Moletsane SD, Matandirotya E, Burger RP. State of ambient air quality in a low-income urban settlement of South Africa. Sci Afr. 2022;16:e01201. https://doi.org/10.1016/j.sciaf.2022.e01201
21. South African Department of Forestry, Fisheries, and the Environment (DFFE). Update on the air quality monitoring strategy and AAQI stations project [document on the Internet]. c2017 [cited 2023 Jun 05]. Available from: https://saaqis.environment.gov.za/Lekgotla%20Proceedings/2017/2017_2.2-update-on-aqm_strategy---naqi-station-project.pdf
22. South African Department of Forestry, Fisheries, and the Environment (DFFE). 2022 State of Air Report and AQM highlights [document on the Internet]. c2022 [cited 2023 Jun 05]. Available from: https://www.dffe.gov.za/sites/default/files/docs/2022airqualitylekgotlapresentations_stateofair.pdf
23. De Jongh D, Ghoorah D, Makina A. South African renewable energy investment barriers: An investor perspective. J Energy South Afr. 2014;25(2):15–27. https://doi.org/10.17159/2413-3051/2014/v25i2a2665
24. Murombo T. Legal and policy barriers to renewable and sustainable energy sources in South Africa. J World Energy Law Bus. 2016;9:142–165. https://doi.org/10.1093/jwelb/jww001

Check for updates

**AUTHOR:**
Maria C. Marchetti-Mercer[1]

**AFFILIATION:**
[1]Department of Psychology, University of the Witwatersrand, Johannesburg, South Africa

**CORRESPONDENCE TO:**
Maria Marchetti-Mercer

**EMAIL:**
maria.marchetti-mercer@wits.ac.za

**HOW TO CITE:**
Marchetti-Mercer MC. Resilience is not enough: The mental health impact of the ongoing energy crisis in South Africa. S Afr J Sci. 2023;119(9/10), Art. #16608. https://doi.org/10.17159/sajs.2023/16608

**ARTICLE INCLUDES:**
☐ Peer review
☐ Supplementary material

**KEYWORDS:**
load shedding, resilience, mental well-being, psychological distress

**PUBLISHED:**
31 August 2023

3 GOOD HEALTH AND WELL-BEING



# Resilience is not enough: The mental health impact of the ongoing energy crisis in South Africa

**Significance:**

Load shedding has become an intrinsic part of South Africans' daily experience. The devastating impact of load shedding on the economy is often foregrounded, but limited attention has been paid to its effects on the population's mental health. South Africans are often depicted as inherently resilient, able to withstand additional disruptions to their lives. However, this reasoning underestimates the ecosystemic nature of load shedding and its severe psychological impact on people's physical and mental well-being. Fostering the mental health of the population is ultimately the responsibility of the government, which urgently needs to address the energy crisis.

In 2008, Eskom, South Africa's public electricity company, first implemented load shedding. 'Planned rolling blackouts' were deemed necessary to protect the integrity of the national electricity grid. The general public did not expect this decision, which created immediate chaos in the lives of all South Africans. At the time, most people did not understand the terms used in the media to explain this sudden disruption to our daily lives – South Africans were faced with a totally new and alien experience and vocabulary.

Some South Africans have never had electricity at home, but for all who do (whether they are legally or illegally connected), and for all who have come to depend on it in any daily activity – at home, at work, for transport, for communication, and in financial and commercial transactions – the disruption is enormous. Now in 2023 – 15 years later – load shedding is an integral part of almost every day of our lives. South Africans have become electricity experts, and technical jargon has become their daily parlance, as some discuss different types of inverters, generators and the solar panels available in the market in ordinary social conversations.

South Africans are often praised for being resilient. In his most recent State of the Nation Address, on 9 February 2023, President Ramaphosa stated: "The energy crisis is an existential threat to our… social fabric. But… we are, at our most essential, a nation defined by hope and resilience."[1] From a psychological perspective, "[r]esilience is the process and outcome of successfully adapting to difficult or challenging life experiences, especially through mental, emotional, and behavioural flexibility and adjustment to external and internal demands"[2]. Generally, the ability to be resilient is often regarded as a positive personal quality, as it allows people to face adversity and cope with it more effectively. However, one could argue that an overemphasis on the 'resilience' of South Africans can trivialise and underestimate the role that the prevailing socio-political and economic forces in the country are playing in creating a daily living experience where 'resilience' is continually needed in order to survive and is constantly tested by the ongoing energy crisis. From an ecosystemic perspective, the daily experiences of people can be understood in terms of multiple interconnecting systems. Most South Africans already face massive structural challenges, such as poverty, unemployment, and high levels of crime. A collapsing national energy system that disrupts economic activities, as well as all citizens' educational, social and family pursuits, may be the psychological tipping point for many. Not surprisingly, there has been a rise in emigration figures: migration agencies reported increased interest in 2022, especially from younger people.[3]

Increasingly, the prevailing complaints from South Africans relate not only to the practical disruptions caused by load shedding, but also to the psychological distress that accompanies them. The mental health impact of load shedding is particularly critical because the country is just emerging from the COVID-19 pandemic, which had a significant impact on the economy[4], and on people's social and mental well-being[5,6]. The aftermath of the pandemic, especially for those who suffered severe economic losses and who lost their jobs, is severe.

Since mid-2022, load shedding has become progressively worse. In 2022, 200 days of load shedding were recorded, escalating to Stage 6. At the time of writing, at the beginning of June 2023, the country was moving between Stages 5 and 6, with rumours that Stages 8 to 10 were likely as the heart of winter approached.

Given the enormous economic disparity in South African society, load shedding gravely affects most South Africans' daily living activities, with accompanying psychological distress. Only a privileged few are able to cushion themselves from the worst aspects of load shedding and some of the daily inconvenience of limited electricity supply by installing inverters, generators and solar panels in their homes and workplaces. The constant disruptions brought about by blackouts have led to an overwhelming sense of frustration and uncertainty, as people feel they have no control of their lives and environment. Breakages, cable theft, and sabotage exacerbate the situation by adding unscheduled outages, some for as long as 10 days. Outages are also increasingly leading to water shortages as reservoirs cannot refill in the short periods between outages, and to telecommunication problems. Significantly, such disruptions in work and educational activities, as well as in the domestic domain, have been shown internationally to increase psychological distress.[7]

Disruption in electricity supply potentially further affects people's physical health, which can in turn increase people's risk of developing mental health problems. At the most basic level, food safety is affected as constant refrigeration cannot be guaranteed, placing people at risk of food poisoning and diarrhoeal diseases. Water purification systems also require electricity to function optimally, so load shedding leaves people without water or facing the threat of contaminated water.[7] South Africa has recently experienced a cholera outbreak and the contamination of rivers and sea water in different provinces, because sewage plants did not have electricity.[8,9]

International research has shown that 'energy poverty', where people have inadequate access to electricity, for example, in order to keep warm during the winter months, can lead to adverse health outcomes, with concomitant depression and anxiety.[10] Exposure to household air pollution from solid fuel combustion, often the only source of energy available in poorer communities, has also been linked to strong depression outcomes.[11]

The absence of uninterrupted electricity adds to stress in some occupational sectors, especially in the public health sector, where workers have already suffered significant stress because of the COVID-19 pandemic. Hospitals have called to be exempted from load shedding, but at the time of writing, only a few hospitals countrywide are exempted. The rest rely mostly on generators to keep their operations going. For example, the Chris Hani Baragwanath Hospital in Soweto, which claims to be the third largest hospital in the world (with more than 400 buildings, 3200 beds and more than 6500 staff), has to work on very large generators. The effects of power outages on health care are diverse and far-reaching. Hospital equipment and other services rely heavily on power.[7] In hospitals that are not exempt from load shedding, and in clinics in rural areas which have no access to generators, patients and health workers are at risk, and experience frustration and uncertainty. People with disabilities and other health concerns may indeed face severe risk, as they rely heavily on medical devices, such as life-support equipment, that require an electricity supply. This situation places additional stress on their caregivers. The QuadPara Association of South Africa has also drawn attention to the plight of many disabled people as a result of load shedding.[12]

The above examples illustrate how the effects of load shedding on the larger social and work environment impact on mental health. The World Health Organization defines mental health as "a state of mental well-being that enables people to cope with the stresses of life, realize their abilities, learn well and work well and contribute to their community"[13]. Mental health is best understood as a state of wellness that goes beyond the absence of mental disorders, but falls along a continuum of well-being that is experienced differently from person to person, affected by the social, economic and political systems in which people operate.[13]

Mental health concerns are not new in South African society, especially in the economic aftermath of the COVID-19 pandemic.[5,6] Structural concerns such as rural poverty and a lack of infrastructure, poor education and unemployment have long been linked to poorer psychological well-being.[14] Many South Africans – already traumatised by discrimination, inequality and interpersonal violence – are vulnerable to mental health challenges.[13,15]

Against this background, the overwhelming impact of load shedding on individuals' mental well-being must be considered from an ecosystemic perspective. Such an approach highlights the important role that different systems and contexts play in individuals' functioning and well-being. Although Bronfenbrenner's ecological systems theory[16,17] was not developed specifically to address mental health, it offers important insights into how the ongoing energy crisis influences individuals' mental health and well-being at several levels.

In his original theory, Bronfenbrenner[16] proposes that the ecological environment is made up of four different systems: the micro-, meso-, exo- and macrosystems. In the microsystem, which consists of the relationship and connection between an individual and that individual's close environment such as the home, workplace, and school[17], load shedding disrupts daily work, home routine and activities, including education. For instance, if someone relies on electricity for medical equipment or to work and/or study from home, unreliability of the electricity supply can create stress and anxiety. The mesosystem refers to interrelations between a person's immediate environment or key settings, for example, between home and the workplace, or home and school.[17] Load shedding regularly disrupts such connections, as people struggle to attend school or work regularly, leading to increased stress and anxiety levels as people struggle to maintain their responsibilities and relationships. The exosystem – social structures such as a person's world of work, public agencies, and the mass media[17] – is also affected. Exosystem factors, such as a lack of supportive social policies, community resources, or access to health care, can pose barriers to mental health services, promotion and interventions. Load shedding can reduce the availability of basic resources such as food and water, increase the risk of crime and violence in some areas, as well as restrict access to proper health care, adding to stress and anxiety.[5,6] Lastly, the macrosystem refers to the broader cultural, social, and economic values, beliefs, and norms that influence people's development, consisting of the blueprints of society, including laws and regulations.[17] Ongoing load shedding can be directly attributed to such larger systemic socio-political factors, resulting in political instability, inadequate investment in infrastructure and corruption.

In later work, Bronfenbrenner included the concept of the chronosystem to address changes over time, both in an individual and in the larger environment.[18] The current political and economic context of post-apartheid South Africa, and specifically 'state capture', has weakened Eskom's ability to provide stable energy, increasing the occurrence of load shedding. Stressing the interconnectedness of the different systems within which people function thus provides us with a useful theoretical lens to understand the far-reaching impact of load shedding on the mental health of South Africans.

In this context of psychological vulnerability, it is concerning that current information regarding the psychological impact of load shedding is largely anecdotal, rather than based on rigorous academic research, and it is mainly media articles that point out that load shedding adds to South Africans' levels of anxiety and depression. Recently, psychologists interviewed on television and radio have also raised awareness around increasing mental distress, and some have suggested strategies for dealing with the psychological distress caused by load shedding. However, thus far, there is a lack of comprehensive South African academic research examining the specific impact of load shedding on mental health so that we can address its consequences more effectively. One exception is a recent online survey by the South African Depression and Anxiety Group (SADAG), which collected data from a sample of 1831 participants. Not surprisingly, it found increased symptoms of depression and anxiety amongst participants arising from constant exposure to load shedding.[19] Research in other countries that also experience regular electricity outages, such as Zambia, shows that load shedding affects daily routines and school performance, reducing time for studying, social and family activities, and leading to economic and psychological problems.[11]

Academic psychological research to explore the micro- and macro-elements associated with load shedding must thus be urgently initiated to illuminate the complex relationship between load shedding and mental well-being. Three main issues should be foregrounded. Firstly, it is vital to identify, examine and assess the common stressors related to load shedding. Secondly, the coping strategies currently employed by people and their perceived effectiveness need to be investigated. Lastly, the relationship between the perceived stressors linked to load shedding, and mental health in general, needs to be determined. Data gathering should be targeted at providing a strong evidence-based foundation for developing targeted psycho-social interventions to support mental health in the current South African context.

Mental health professionals must also push back against the larger socio-political system when it ignores its responsibility to stop creating a living situation where 'resilience' is continuously required in order to survive, and where the responsibility for mental health rests on the shoulders of the dedicated few. South Africans may not reach a state of 'learned helplessness'[20] where people come to believe that no matter what they do, nothing will change their situation, and where they passively accept their circumstances. Such a sense of resignation and long-suffering forbearance – even if it is portrayed and praised as 'resilience' – must be resisted at all costs.

The mental health impact of load shedding on the lived experiences of all South Africans must move to the forefront of current debates. It is not a secondary issue. It has been argued that it is crucial to conduct comprehensive academic research on the psychological impact of these constant disruptions to our daily lives, exploring its effects on both individual and familial functioning, but this is not enough. Mental health professionals need to become activists: they need to speak out and draw attention to the extreme psychological distress that the South

African population is experiencing because of continued and worsening load shedding – the larger socio-political system has a fundamental responsibility to promote the well-being of all South Africa's people.

## Competing interests

I have no competing interests to declare.

## References

1. Crowe J. South Africans are defined by our resilience. IOL. 2023 February 18. Available from: https://www.iol.co.za/opinion/letters/south-africans-are-defined-by-our-resilience-e352ba15-0289-4efc-b6ab-129d4a4caf21
2. American Psychological Association dictionary of psychology [homepage on the Internet]. No date [cited 2023 Apr 14]. Available from: https://dictionary.apa.org/
3. Spike in South Africans looking to emigrate – with load shedding the 'final nail in the coffin'. BusinessTech. 2022 July 16. Available from: https://businesstech.co.za/news/wealth/606290/spike-in-south-africans-looking-to-emigrate-with-load-shedding-the-final-nail-in-the-coffin/
4. OECD. South Africa economic snapshot: Economic forecast summary (November 2022) [webpage on the Internet]. c2022 [cited 2023 Apr 10]. Available from: https://www.oecd.org/economy/south-africa-economic-snapshot/
5. Lentoor AG. Stress in times of SARS-COV-2 related lockdown in South Africa: A community based cross-sectional survey. Glob J Med Public Health. 2022;11(2):1–8. https://www.gjmedph.com/Uploads/O6_COVID-19%20Stress.pdf
6. Nguse S, Wassenaar D. Mental health and COVID-19 in South Africa. S Afr J Psychol. 2021;51(2):304–313. https://doi.org/10.1177/00812463211001543
7. Phiri C, Mpundu G, Chama E, Ngandu M, Mwanza EL, Aluckal E, et al. An assessment of the impact of electricity power outages on university students' life, Lusaka, Zambia: A university perspective study [document on the Internet]. c2021 [cited 2023 Apr 20]. Available from: https://www.researchgate.net/publication/350524976_An_assessment_of_the_impact_of_Electricity_Power_Outages_on_University_Students'_life_Lusaka_Zambia_A_University_perspective_study
8. Chepape L. Ageing infrastructure and load-shedding cause sewage spills and beach closures in Cape Town. Mail & Guardian. 2023 January 04 [cited 2023 Apr 10]. Available from: https://mg.co.za/environment/2023-01-04-ageing-infrastructure-and-load-shedding-cause-sewage-spills-and-beach-closures-in-cape-town/
9. Gounden T. Rand Water assures residents the water it distributes is cholera-free. Eye Witness News. 2023 March 11 [cited 2023 Apr 12]. Available from: https://ewn.co.za/0001/01/01/rand-water-says-the-water-distributed-to-residents-is-cholera-free
10. Bentley R, Daniel L, Li Y, Baker E, Li A. The effect of energy poverty on mental health, cardiovascular disease and respiratory health: A longitudinal analysis. Lancet Reg Health West Pac. 2023;35:Art. #100734. https://doi.org/10.1016/j.lanwpc.2023.100734
11. Liu Y, Chen X, Yan Z. Depression in the house: The effects of household air pollution from solid fuel use among the middle-aged and older population in China. Sci Total Environ. 2020;703:Art. #134706. https://doi.org/10.1016/j.scitotenv.2019.134706
12. Loadshedding life-threatening to people with disabilities. Rolling Inspiration. 2022 July 07. Available from: https://www.rollinginspiration.co.za/loadshedding-life-threatening-to-people-with-disabilities/
13. World Health Organization (WHO). Mental health: Strengthening our response [webpage on the Internet]. c2022 [cited 2023 Apr 20]. Available from: https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response
14. Khumalo IP, Temane QM, Wissing MP. Socio-demographic variables, general psychological well-being and the mental health continuum in an African context. Soc Indic Res. 2012;105:419–442. https://doi.org/10.1007/s11205-010-9777-2
15. Jack H, Wagner RG, Petersen I, Thom R, Newton CR, Stein A, et al. Closing the mental health treatment gap in South Africa: A review of costs and cost-effectiveness. Glob Health Action. 2024;7(1):Art. #23431. https://doi.org/10.3402/gha.v7.23431
16. Bronfenbrenner U. Reality and research in the ecology of human development. Proc Am Philos Soc. 1975;119(6):439–469.
17. Bronfenbrenner U. Toward an experimental ecology of human development. Am Psychol. 1977;32(7):513–531. https://doi.org/10.1037/0003-066X.32.7.513
18. Bronfenbrenner U. Ecology of the family as a context for human development: Research perspectives. Dev Psychol. 1986;22(6):723–742. https://doi.org/10.1037/0012-1649.22.6.723
19. South African Depression and Anxiety Group (SADAG). SADAG's new load shedding survey results sheds some light on the impact on mental health in South Africa [media release on the Internet]. 2023 March 05 [cited 2023 Apr 20]. Available from: https://www.sadag.org/index.php?option=com_content&view=article&id=3282:sadag-s-new-load-shedding-survey-results-sheds-some-light-on-the-impact-on-mental-health-in-south-africa&catid=149:press-releases&Itemid=101
20. Seligman ME. Learned helplessness. Annu Rev Med. 1972;23(1):407–412. https://doi.org/10.1146/annurev.me.23.020172.002203

# Load shedding and mental health in South Africa: Methodological challenges of establishing causal links

AUTHORS:
Jason Bantjes[1,2]
Leslie Swartz[3]

AFFILIATIONS:
[1]Mental Health, Alcohol, Substance Use and Tobacco (MAST) Research Unit, South African Medical Research Council, Cape Town, South Africa
[2]Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa
[3]Department of Psychology, Stellenbosch University, Stellenbosch, South Africa

CORRESPONDENCE TO:
Jason Bantjes

EMAIL:
Jason.Bantjes@mrc.ac.za

HOW TO CITE:
Bantjes J, Swartz L. Load shedding and mental health in South Africa: Methodological challenges of establishing causal links. S Afr J Sci. 2023;119(9/10), Art. #16661. https://doi.org/10.17159/sajs.2023/16661

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
load shedding, mental health, psychiatrisation, South Africa, depression, anxiety

PUBLISHED:
31 August 2023



**Significance:**

We critically interrogate assertions that load shedding has a deleterious impact on mental health and explore the methodological challenges of establishing causal links empirically. We highlight the lack of empirical data to support a causal link and the problem of conflating psychological distress with psychopathology. In addition, we set out the methodological problems associated with collecting data to show that load shedding impacts the prevalence of mental disorders. While it may make superficial strategic sense for activists to link load shedding to mental health as a political strategy to raise awareness, this approach could have long-term negative consequences.

Load shedding in South Africa has undoubtedly resulted in social and economic disruptions and made people's lives more complicated, but it is not clear whether load shedding has resulted in significant changes in rates of mental illness. Claims have been made in the media that load shedding has had a marked deleterious impact on people's mental health, although the evidence offered in support of these claims is far from rigorous. In this Commentary, we critically interrogate assertions that load shedding has led to an increase in the prevalence of mental disorders and the data that have been offered to support these claims. We discuss the ideological and methodological challenges of trying to collect the data needed to prove such assertions. We highlight how linking mental health to load shedding might further political ends (such as normalising mental illness), while also arguing that claiming such links in the absence of sound empirical data can also have unintended harmful consequences, including misleading the public and trivialising serious mental illness.

## Claims that load shedding adversely affects mental health

There have been several recent media reports about the impact of load shedding on mental health. The *Financial Mail*, in an article entitled 'By the numbers: What load-shedding does to your mental health' (12 March 2023), published the bold claim that "…long-term projected feelings of hopelessness are having a negative impact on people's mental health"[1]. Under the sensational headline 'Load shedding leading to anxiety and depression which can be fatal, say psychologists', *The Citizen* (23 February 2023) stated that "load shedding can even affect those who never had mental issues before"[2]. *The Daily Maverick* (6 July 2022) reported: "Load shedding is adding to the anxiety, depression and mental health toll among South Africans," adding that load shedding is "causing a wave of stress and anxiety which, for many, could lead to depression"[3]. Similar headlines include 'Load-shedding bound to lead to depression and anxiety, says psychologist' (*Times Live*, 16 January 2023)[4], 'Anxiety and stress exacerbated by load shedding' (IOL, 19 January 2023)[5], and 'Negative impact of rolling blackouts on mental health' (*SABC News*, 7 March 2023)[6].

These media reports share three distinctive characteristics, which we discuss in more detail below. First, they use rhetoric to construct a crisis narrative. Second, they lack sound empirical evidence to support claims about deteriorating mental health outcomes due to load shedding. Third, they conflate psychological distress with psychopathology.

## Crisis narratives

Crisis narratives are stories which construct events as disastrous and call people to action in order to avert almost certain devastation. Crisis narratives are part of a dystopian genre and can be found in some accounts of climate change and environmental threats[7], infectious disease epidemics[8], child hunger[9], and other public health emergencies[10-13]. Crisis narratives function to frame collective understandings of risks, warnings, and possible harms, thus potentially mobilising society into action to 'combat' the 'foe'.[10] Crisis narratives may have important positive consequences. For example, they may create cohesion among of group of people who come together to avert the crisis. These narratives may, however, also be divisive in that they can set up an 'us versus them' dialectic which can be unhelpful in that those 'in the know' are presented as experts while alternative views of others are dismissed.[13] At their worst, crisis narratives sow panic and create confusion through the use of hyperbolic rhetoric and the distortion of evidence to fit a particular narrative for political purposes.[11]

Media reports of the psychological impact of load shedding seldom talk about resilience, adaptation, and creativity. Instead, they use the language of psychiatric disease (depression and anxiety) and words like "fatal" *(The Citizen)*[2] and "toll" (*The Daily Maverick*)[3] which imply that load shedding will lead to mental health casualties, while also framing mental illness as an inevitable and certain consequence of load shedding (*Times Live*)[4]. Crucially, some reports distort the available evidence to fit their account of a crisis by overreaching on what can be concluded from empirical data, as we discuss below. Even more worrying is the potential conflation of something which is clearly disruptive and bad for society (i.e. load shedding) with a mental health crisis. At worst, if one of the bases on which we claim load shedding is bad is that of poor mental health, then some may assume that if mental health outcomes are not rigorously shown to be poor, this makes load shedding less serious, which is not the case. Constructing a crisis narrative about load shedding and mental health may appear to be a useful short-term political strategy for mental health activism, but in the long run it could backfire if empirical data do not support the claims.

## Lack of sound empirical evidence

Typically, media reports justify their claims that load shedding negatively impacts mental health either by quoting experts (usually psychologists or psychiatrists) who speculate from a position of authority without offering any empirical evidence, or by quoting data from a cross-sectional survey conducted by the South African Depression and Anxiety Group (SADAG).

The mental health experts quoted in media reports use, it appears, common-sense reasoning to argue that depression and anxiety are precipitated by uncomfortable emotions that accompany load shedding, including feelings of impotence, uncertainty, loss of control, frustration, anger, and fear about crime. From the way they are presented in media reports, it seems as if the experts assume as a self-evident 'fact' that load shedding causes a significant number of people to have uncomfortable feelings and that people are likely to be unable to regulate or tolerate these uncomfortable emotions without developing a depressive or anxiety disorder. By appealing to the authority of experts, the media reports frame the human psyche as fragile and position people as highly susceptible to emotional discomfort, while discrediting South Africans' capacity to adapt to adverse circumstances. It is notable that reports quoting experts start with the axiomatic statement that load shedding is difficult and precipitates uncomfortable feelings such as powerlessness and anger. But then the reports slip into asserting that uncomfortable feelings will naturally lead to psychopathology (i.e. mood and anxiety disorders). In slipping between these two ideas, the reports use a logical fallacy and conflate psychological distress with psychological disorders (a theme we return to later). It is important to stress here that the reporting on what experts say may itself be highly selective and sensational – as consumers of such reports we generally do not have access to the full text and context of what experts in fact said. We also do not know how many (if any) experts may have refused to engage with the kinds of questions asked in this kind of reporting.

The second main source of evidence in media reports about load shedding and mental health is survey data collected in early 2023 by SADAG, a South African non-profit organisation established to provide support to people living with mental health problems and to serve as a patient advocacy group.[14] SADAG has several partners, including universities, government departments, drug companies, and for-profit professional mental health services, as listed on their website.[14] According to various news reports and a press release posted on the SADAG website[15], a cross-sectional survey was conducted with a self-selected (i.e. non-probability) sample of 1836 respondents (out of 30 000 "members of the SADAG community" who were invited to complete the survey, i.e. a 6.1% response rate). It is not clear how or what data were collected, but it appears from the press release that a combination of quantitative and qualitative data were collected. The press release on the SADAG website states that "4 in 10 people reported depression, and 62% of people struggled with anxiety and panic." The press release further states that "1 in 10 have contemplated suicide or had thoughts of suicide. …. (and) 31% reported problematic family relationships, and feelings of isolation." The press release includes a description of the strategies that "members of the SADAG community" report using to cope with load shedding. This press release is not intended to be a scientific report and it is unfair to subject it to the usual standards expected of scholarly writing. Nonetheless, the press release has been widely cited and is being used to support a crisis narrative that links load shedding to increasing rates of depression, anxiety and fatal mental health outcomes. As such, we believe we are justified (if not required) as mental health professionals and scientists to interrogate the results of the survey and the basis on which truth claims are made.

There are several serious methodological problems with the survey (as it is reported in the SADAG press release), chief among these is that the survey relies on a non-probability sample drawn from the "SADAG community," uses a cross-sectional research design with self-reported data, and provides no information about the validity or reliability of the survey instrument.

Non-probability samples cannot be used to generalise to the whole population, especially if the sample is drawn from a community that has been set up as a support group for "people living with mental health problems" (and is thus explicitly not representative of all South Africans). Drawing a sample from a particular delineated subset of the population introduces sampling bias, which makes generalising the findings of the survey to the whole population invalid. The very low response rate (6.1%) creates problems even with generalising the survey findings to the SADAG community. To say anything valid and reliable about the prevalence of mental disorders, one would need to draw a large-enough representative sample from the population using random sampling (i.e. attempting to ensure survey respondents are recruited at random from the whole population to ensure that everyone in the country has an equal chance of participating in the survey). The question of the sample size (i.e. what is a big enough sample?) is also important, especially when one is trying to measure phenomena that naturally have a low base rate, as might be the case for severe depressive illnesses.[16,17]

Cross-sectional research designs, like that used in the SADAG survey, can at best identify correlations and associations between variables but cannot be used to infer causality, making it impossible to conclude that load shedding is *the* cause, or even *a* cause, of the depression, anxiety and suicidal thoughts reported by survey respondents. To make any valid inference about causality, at the very least one would require longitudinal data or interrupted time-series data with measures for mental health taken before, during, and after load shedding. Robert Koch, writing in the 19th century in the context of microbiological organisms, originally postulated four criteria necessary to infer that an organism causes a disease, namely: (1) the organism must be found in diseased individuals but not in healthy individuals; (2) the organism must be found in the diseased individual; (3) inoculating a healthy individual with the organism must precipitate the disease; and (4) the organism must be re-isolated from the inoculated diseased individual.[18] Koch's postulates have been the subject of controversy and debate in infectious disease medicine, and have subsequently been revised (even Koch revised his own postulates by eventually dropping the first criterion), but nonetheless have served as an invaluable guide to the discovery of the specific causes of various infectious diseases.[19,20] Importantly, by proposing these criteria, Koch forced medical scientists to think carefully about (and justify) the necessary and sufficient conditions to infer that an agent causes a disease. In psychiatry it is much harder to propose a set of necessary and sufficient conditions for claiming that any mental illness is caused by X.[21] This is so for several reasons, including the multifactorial nature and causes of mental disorders in general, and the obvious question of relating the onset of a disorder to its purported causes. Simply put, for X to cause Y in epidemiological terms, it makes sense that X must predate Y. In psychiatric epidemiology in general, researchers are often dealing with disorders of slow or unclear onset – one does not move from being asymptomatic one day to having a clear case of a disorder the next, and prodromes for disorders may be diffuse and difficult to assess. In the case of the SADAG study, if it is the case that the participant pool were people already experiencing symptoms of anxiety and depression, for example, and then experienced a subjective exacerbation of symptoms during load shedding, one cannot conclude that load shedding caused the anxiety and depression. A further difficulty with cross-sectional research of this kind is that of recall bias – it is not unusual for people experiencing difficulties to attribute these to proximal stressful events which may, in fact, have post-dated the difficulties.

Finally, without any information about the reliability and validity of the survey instrument used in the SADAG study, one cannot even be sure that the survey assesses depressive or anxiety disorders. Diagnosing mental disorders is a specialised task that psychologists and psychiatrists spend years learning and requires more than counting the number of symptoms a person reports or asking someone if they are depressed. Screening instruments can be used to identify people who are likely to meet diagnostic criteria for a mental disorder, but before any screening instrument can be used in a survey it needs to be carefully validated to ensure that it is both reliable and valid.[22] Screening instruments usually consist of a list of symptoms which a respondent is asked to endorse, thus yielding a symptom count. Researchers who use screening instruments determine a cut-off point (i.e. a symptom score) that differentiates respondents who are likely to meet diagnostic criteria for a disorder from those who are not. This cut-off point must be statistically established for a particular population, along with the sensitivity and specificity of the instrument.[22] Of course, it is possible that the SADAG study used reliable instruments that have been validated for their survey population (i.e. "the SADAG community"), but without knowing this information we cannot assess the accuracy of any prevalence rates quoted for depression or anxiety disorders.

Indeed, it is possible that SADAG attended to most or even all the concerns we have raised above, but in the absence of clear reporting of methods used, it is a mistake to rely on the findings to generalise about load shedding causing mental health problems in South Africa. Just as the science community should not rely on media reports for information on the seriousness of climate change, full and informed reporting of methods is necessary for assessment of the accuracy of claims made in relation to mental health issues.

## Conflating psychological distress and psychopathology

Media reports about load shedding and mental health (and indeed the SADAG research) seem to conflate psychological distress with psychopathology. Psychological distress is a transient state of emotional discomfort and is a common reaction to the day-to-day vicissitudes of life.[23] It is normal to feel psychological distress and it is unrealistic to expect that we will never experience hardship, struggle, or uncomfortable feelings. Psychological distress does not require treatment from a mental health professional and usually resolves with time and appropriate support from family and friends. By comparison, psychopathology (i.e. mental disorders) are severe persistent disturbances of thinking and feeling, which cause marked impairments in social, interpersonal and occupational functioning.[24] Many mental disorders require treatment by a mental health professional and are considered to be serious mind-brain illnesses. Of course, mental illness cannot be understood as a binary phenomenon and there is a growing trend in psychiatry towards understanding psychological functioning on a dynamic continuum; as such it is not always easy to draw a line between psychological distress and psychopathology. However, if we use these constructs interchangeably (as appears to be the case in media reports and the SADAG study on loadshedding and mental health), we run the risk of trivialising serious mental illnesses which are debilitating by conflating them with the uncomfortable feelings of everyday struggles. In part, the problem here is that the term 'mental health' is increasingly being used as an elastic construct and as a catch-all phrase to denote everything from severe serious mental illness to normal responses to stressful situations. Using the discourse of 'mental health' is a political act which can help to normalise and destigmatise mental illness. But the discourse of 'mental health'" (which does not differentiate between psychological distress and mental disorders) can trivialise severe mental illness and obscure the needs of people with severe debilitating mind-brain illnesses.

Load shedding probably causes psychological distress and discomfort for many people, and may even make some people angry, irritable, and uncertain some of the time. And load shedding may present challenges for people with pre-existing mental health conditions. But having feelings about load-shedding (even strong feelings) does not mean that one has a mental disorder or that one requires psychiatric treatment, as some media reports and the SADAG press release seem to imply. This is not to say that it is impossible that some South Africans without pre-existing mental health vulnerabilities could experience severe, clinical levels of anxiety or distress due to load shedding (for example, those who depend on regular power supply to run life-saving home medical equipment, those who are at significantly increased risk of victimisation during load shedding, or those living in communities where power outages last several days at a time). It is easy for the implicit conflation between psychological distress and psychopathology in the SADAG press release to slip past unnoticed because of the growing trend towards using psychiatric speak (i.e. terms like depression, anxiety, panic attack) in everyday situations to describe everyday experiences. Even if there is a continuum between everyday distress and psychopathology, different points on the continuum have different meanings and, crucially, different implications for services. Psychiatrisation, is part of a larger growing tendency in contemporary society to medicalise everyday life by turning the daily problems of living into illnesses that require treatment from a medical professional.[25,26] Psychiatrisation leads to over-diagnosis and over-treatment of mental disorders, and reflects the expansion of psychiatry's reach into everyday life and the growth of the medical-industrial complex.[27]

## Conclusion

It is important and helpful to raise awareness of mental disorders, to challenge stigma by normalising people's experience of living with a mental illness, and to promote access to treatment and support services for people who are struggling, as SADAG and others are committed to doing. A possible unintended consequence of mental health activism, however, may be the conflation of psychological distress with psychopathology or using inappropriate data to further a political cause, however just this larger cause may be. Load shedding is a serious issue, likely to have far-reaching consequences for South Africa. Similarly, it is well established from sound research that resources for appropriate evidence-based mental health interventions are not sufficient in South Africa[28], nor further afield[29,30]. It may make superficial strategic sense for activists to link the topic of load shedding to mental illness through the (re)production of a crisis narrative as a political strategy to raise awareness, reduce stigma and mobilise resources. But, at this stage, there are not robust data available to make such links, and falsely asserting that there are, could undermine the credibility of mental health activists in the long run, especially if the data are subsequently shown not to support the activists' claims. If data linking load shedding and mental health are collected and fully reported scientifically, these results will be important for mental health resource planning. The current state of evidence suggests, however, that arguments about load shedding as a factor in the development of mental disorders are premature. Clearly, more, and better, research is needed.

## Competing interests

We have no competing interests to declare.

## References

1. Malan M. By the numbers: What load-shedding does to your mental health. Financial Mail. 2023 April 12 [cited 2023 Apr 28]. Available from: https://www.businesslive.co.za/fm/fm-fox/2023-04-12-by-the-numbers-what-load-shedding-does-to-your-mental-health/
2. Patel F. Load shedding leading to anxiety which can be fatal, say psychologists. The Citizen. 2023 February 23 [cited 2023 Apr 28]. Available from: https://www.citizen.co.za/news/south-africa/load-shedding-anxiety-depression-fatal-psychologists/
3. Pollitt R. Load shedding is adding to the anxiety, depression and.... Daily Maverick. 2022 July 06 [cited 2023 Apr 28]. Available from: https://www.dailymaverick.co.za/article/2022-07-06-load-shedding-is-adding-to-the-anxiety-depression-and-mental-health-toll-among-south-africans/
4. Masweneng K. Load-shedding bound to lead to depression and anxiety, says psychologist. Times Live. 2023 January 16 [cited 2023 May 03]. Available from: https://www.timeslive.co.za/news/south-africa/2023-01-16-load-shedding-bound-to-lead-to-depression-and-anxiety-says-psychologist/
5. Makgatho L. Anxiety and stress exacerbated by load shedding. IOL. 2023 January 19 [cited 2023 May 01]. Available from: https://www.iol.co.za/sundayindependent/news/anxiety-and-stress-exacerbated-by-load-shedding-f739e3dd-6d32-4516-b132-dae495845210
6. Myeza Z. Negative impact of rolling blackouts on mental health. SABC News. 2023 March 07 [cited 2023 May 01]. Available from: https://www.sabcnews.com/sabcnews/negative-impact-of-rolling-blackouts-on-mental-health/
7. Warner J, García AK, Haller T. Environmental crisis narratives in drylands. In: Kronenburg García A, Haller T, Van Dijk H, Samimi C, Warner J, editors. Drylands facing change: Interventions, investments and identities. London: Routledge; 2022. p. 32–50. https://doi.org/10.4324/9781003174486-4
8. Lees S, Enria L, James M. Contesting the crisis narrative: Epidemic accounts in Sierra Leone, Tanzania, and Democratic Republic of the Congo. Disasters. 2023;47(1):78–98. https://doi.org/10.1111/disa.12535

9. Fram MS, Frongillo EA. Backpack programs and the crisis narrative of child hunger – a critical review of the rationale, targeting, and potential benefits and harms of an expanding but untested model of practice. Adv Nutr. 2018;9(1):1–8. https://doi.org/10.1093/advances/nmx008

10. Seeger MW, Sellnow TL. Narratives of crisis: Telling stories of ruin and renewal. Redwood City, CA: Stanford University Press; 2016.

11. Glowacki EM, Taylor MA. Health hyperbolism: A study in health crisis rhetoric. Qual Health Res. 2020;30(12):1953–1964. https://doi.org/10.1177/1049732320916466

12. Liu BF, Austin L, Lee YI, Jin Y, Kim S. Telling the tale: The role of narratives in helping people respond to crises. J Appl Commun Res. 2020;48(3):328–349. https://doi.org/10.1080/00909882.2020.1756377

13. Sellnow TL, Seeger MW. Theorizing crisis communication. 2nd ed. Hoboken, NJ: Wiley-Blackwell; 2020. Available from: https://www.wiley.com/en-us/Theorizing+Crisis+Communication%2C+2nd+Edition-p-9781119615989

14. South African Depression and Anxiety Group (SADAG) [homepage on the Internet]. No date [cited 2023 Apr 28]. Available from: https://www.sadag.org/

15. Mashigo L. SADAG conducts research on the psychological affects of load shedding. IOL. 2023 March 05 [cited 2023 Apr 29]. Available from: https://www.iol.co.za/the-star/news/sadag-conducts-research-on-the-psychological-affects-of-load-shedding-7cf17b12-9206-43d3-8cfc-336c2f0c8067

16. Jablensky A. Research methods in psychiatric epidemiology: An overview. Aus N Z J Psychiatry. 2002;36(3):298–310. https://doi.org/10.1046/j.1440-1614.2002.01035.x

17. Shoukri MM, Asyali MH, Donner A. Sample size requirements for the design of reliability study: Review and new results. Stat Methods Med Res. 2004;13(4):251–271. https://doi.org/10.1191/0962280204sm365ra

18. Segre JA. What does it take to satisfy Koch's postulates two centuries later? Microbial genomics and *Propionibacteria acnes*. J Invest Dermatol. 2013;133(9):2141–2142. https://doi.org/10.1038/jid.2013.260

19. Singh S, Segal LN. A lung pathobiont story: Thinking outside the Koch's postulate box. Cell Host Microbe. 2022;30(9):1196–1198. https://doi.org/10.1016/j.chom.2022.08.012

20. Byrd AL, Segre JA. Adapting Koch's postulates. Science. 2016;351(6270):224–226. https://doi.org/10.1126/science.aad6753

21. Dalén P. Causal explanations in psychiatry: A critique of some current concepts. Br J Psychiatry. 1969;115(519):129–137. https://doi.org/10.1192/bjp.115.519.129

22. Eaton WW, Hall ALF, Macdonald R, Mckibben J. Case identification in psychiatric epidemiology: A review. Int Rev Psychiatry. 2009;19(5):497–507. https://doi.org/10.1080/09540260701564906

23. Payton AR. Mental health, mental illness, and psychological distress: Same continuum or distinct phenomena? (with correction). J Health Soc Behav. 2009;50(2):213–227. https://doi.org/10.1177/002214650905000207

24. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. Washington DC: American Psychiatric Association; 2013. https://doi.org/10.1176/appi.books.9780890425596

25. Haslam N, Tse JSY, De Deyne S. Concept creep and psychiatrization. Front Sociol. 2021;6:226. https://doi.org/10.3389/fsoc.2021.806147

26. Rapley M, Moncrieff J, Dillon J. Carving nature at its joints? DSM and the medicalization of everyday life. In: Rapley M, Moncrieff J, Dillon J, editors. De-medicalizing misery. London: Palgrave Macmillan; 2011. p. 1–9. https://doi.org/10.1057/9780230342507_1

27. Graham SS. Dis-ease or disease? Ontological rarefaction in the medical-industrial complex. J Med Humanities. 2011;32(3):167–186. https://doi.org/10.1007/s10912-011-9137-5

28. Docrat S, Besada D, Cleary S, Daviaud E, Lund C. Mental health system costs, resources and constraints in South Africa: A national survey. Health Policy Plan. 2019;34(9):706–719. https://doi.org/10.1093/heapol/czz085

29. Troup J, Fuhr DC, Woodward A, Sondorp E, Roberts B. Barriers and facilitators for scaling up mental health and psychosocial support interventions in low- and middle-income countries for populations affected by humanitarian crises: A systematic review. Int J Ment Health Syst. 2021;15(1):Art. #5. https://doi.org/10.1186/s13033-020-00431-1

30. MacDougall AG, Lysaght R, Casey R, Janzen Le Ber M, Mutiso V, Ndetei D, et al. Heeding the call of an innovative paradigm for mental health interventions in low- and middle-income countries. Psychiatr Serv. 2021;73(7):812–814. https://doi.org/10.1176/appi.ps.202000125

# The value of multibeam bathymetry in marine spatial planning in South Africa: A review

**AUTHORS:**
Thamsanqa F. Wanda[1,2]
Errol A. Wiles[1,3]
Hayley C. Cawthra[4,5]
Anton de Wit[2]

**AFFILIATIONS:**
[1]South African Institute for Aquatic Biodiversity, Makhanda, South Africa
[2]Department of Geosciences, Nelson Mandela University, Gqeberha, South Africa
[3]Institute for Coastal and Marine Research, Nelson Mandela University, Gqeberha, South Africa
[4]Minerals and Energy Unit, Council for Geoscience, Cape Town, South Africa
[5]African Centre for Coastal Palaeoscience, Nelson Mandela University, Gqeberha, South Africa

**CORRESPONDENCE TO:**
Thamsanqa Wanda

**EMAIL:**
t.wanda@saiab.nrf.ac.za

**DATES:**
**Received:** 11 July 2022
**Revised:** 06 May 2023
**Accepted:** 08 May 2023
**Published:** 28 Sep. 2023

**HOW TO CITE:**
Wanda TF, Wiles EA, Cawthra HC, de Wit A. The value of multibeam bathymetry in marine spatial planning in South Africa: A review. S Afr J Sci. 2023;119(9/10), Art. #14320. https://doi.org/10.17159/sajs.2023/14320

**ARTICLE INCLUDES:**
☒ Peer review
☐ Supplementary material

**DATA AVAILABILITY:**
☐ Open data set
☐ All data included
☐ On request from author(s)
☒ Not available
☐ Not applicable

**EDITORS:**
Jennifer Fitchett
Adriaan van der Walt

**KEYWORDS:**
hydroacoustic surveys, benthic habitat, biodiversity, climate change, seafloor geomorphology

**FUNDING:**
South African National Research Foundation (MND210629616959), South African Institute for Aquatic Biodiversity



Given a growing global population and shift to embrace the blue economy, a need for marine spatial planning (MSP) has emerged in South Africa to sustainably resolve the rising conflicts over the use of marine and seabed resources and services. A well-developed marine spatial plan yields numerous ecological, social and economic benefits. These are achieved through mediating between spatially conflicting economic drivers' interests (e.g. commercial fishing, tourism, mining), preventing their activities from compromising thresholds of an environment's sustainability. Within the MSP framework, high-resolution geospatial datasets are required to document and describe the seabed in the highest possible detail. At any scale, integrated analysis of seabed geomorphology and habitats is anticipated to greatly improve the understanding of ecosystem functioning from a multidisciplinary perspective, whilst improving MSP procedures and management of marine space. South Africa is the first of few African countries to have an approved and implemented MSP framework, but is still somewhat behind globally in implementing large-scale regional hydroacoustic surveys to cover the country's vast offshore territory. The deficiency of hydroacoustic surveys is perhaps due to a relative lack of funds and poor communication about the value of multibeam echo-sounder (MBES) derived data, whilst marine geoscience remains a scarce skill in the country. This review paper presents a geological perspective of MSP and explores (1) the value that seabed mapping offers MSP specifically and (2) the need to increase seabed mapping with MBES, using a recently initiated project from the South African east coast as a case study.

**Significance:**

The collected MBES data (our case study) provides unprecedented seabed detail of the complex reef habitat and adjacent areas within specific management zones of the uThukela Banks Marine Protected Area. We reveal seabed features and their spatial distribution at a scale not possible using earlier (singlebeam) seabed mapping techniques. These high-resolution data will enable a better understanding of east coast marine habitats whilst contributing to improved spatial management of areas within and around the uThukela Banks Marine Protected Area.

## Introduction

Marine spatial planning (MSP) is a process currently being employed on a global scale towards efficient management of the ocean space, and therefore the blue economy[1,2] (i.e. sustainable exploitation, preservation and regeneration of the marine environment) and addressing the needs of growing global populations (e.g. Operation Phakisa[2-4]). Globally[5-12] and locally, the introduction of an ecosystem-based MSP process towards management of the marine space is well supported.[4,13,14] The decision-making process is guided by the quality of data on which it is based. At present, about 75 countries that have marine borders on all major oceans have commenced with MSP initiatives.[15] Some of these countries (such as China[16], Canada[17,18], the United States of America's California mapping programme[19], Ireland[20] and Australia[21]) have actively applied improved technology[22,23] (e.g. multibeam echo-sounders/MBES) in benthic habitat mapping, which has led to effective MSP results. A case study from the long-established Irish programme INFOMAR (integrated mapping for the sustainable development of Ireland's marine resource) showed investment benefits of seabed mapping initiative across all marine sectors to be 4–6-fold compared to the initial capital investment.[24] In South Africa, the development of an MSP framework began in 2015[9], identifying spatial plans encompassed by the country's exclusive economic zone (EEZ). This enabled South Africa to take the lead over other African countries by being the first to have the Department of Environmental Affairs (now known as the Department of Forestry, Fisheries and the Environment) draft an MSP legislation in 2017.[4,13,25] The draft was later approved by the government as *Act No. 16 of 2018*.[14]

Globally, there is a drive to protect the ocean's biodiversity and/or ecosystems[26-28]; however, only 9% of the world's seabed has been mapped to resolutions at an appropriate scale for MSP and management[29,30]. Marine protected areas (MPAs) are a significant component of the South African MSP process, and receive an elevated status and restricted use based on their valuable environmental products and services.[31] There are 41 MPAs (total area 5.4%) within South Africa's EEZ[32], defined on ecosystem conservation and/or socio-economic objectives, but only a fraction of these MPAs are associated with the high-resolution geospatial context provided by MBES data (e.g. Protea Banks MPA[33], the uThukela MPA[34] and Cape Canyon MPA[35]). The management, protection and monitoring of the oceanic resources and services that are not holistically understood pose significant challenges (e.g. lacking the contribution of the abiotic factors, such as the seabed geomorphology, in an attempt to understand the declining continental shelf biodiversity).

Marine habitat and resource mapping has become a global prerequisite for spatial management of the ocean's resources and services.[36,37] The availability of high-resolution bathymetric data has increased modelling capabilities in marine science.[18,30,38-42] However, MBES data are only one component (although an important one) of benthic habitat mapping, which requires various inputs (e.g. seabed imagery, sediment samples, backscatter and bathymetry).

MBESs have rapidly become the preferred hydroacoustic surveying technique in marine habitat mapping for its capability to simultaneously and reliably obtain bathymetry and backscatter data.[30,37,42,43] MBESs form the foundation for the Seabed 2030 initiative[44] (which aims at completely assimilating data and mapping the world's seabed by 2030). This technique uses sound to collect hundreds of thousands to millions of georeferenced depth soundings which are processed to create a 3D digital terrain model of the surveyed area. Hence, MBES data capture high spatial resolution imagery of the seabed's geomorphology[45] which is used to characterise marine habitats in terms of their georeferenced location, spatial extent and geomorphological characteristics[38]. In South Africa, the most recent of these interdisciplinary research projects was carried out in the Western Cape by Pillay et al.[42], who applied machine learning and algorithms to seabed classification.

Benthic habitat mapping is recognised as a need to be expanded upon in South Africa[41], because the continental shelf bathymetry has been previously recognised to be poorly resolved[40]. One of the benthic habitats recognised globally for the multitude of coastal goods and services it offers is reefs[46] (i.e. outcropping rock on the seabed), which form the primary focus presented in the case study here. Reefs host diverse marine ecosystems[47] and surveying reefs with MBES in South Africa is anticipated to improve the understanding of their functioning, besides supporting multiple resource management objectives (e.g. delineating MPAs)[41,45]. Seabed characteristics (e.g. depth, rugosity and substrate) and their interaction with the local oceanographic conditions form the building blocks that contribute to the functioning of marine ecosystems and structure of benthic guild (habitat provision) at regional spatial scales.[30,42,48] In this paper, we aim to (1) highlight the gaps in allocating spatially restrictive boundaries without high-resolution hydroacoustic mapping within South Africa's EEZ, especially in areas demarcated as MPAs, (2) provide a link between geology, biological and oceanographic space utilising high-resolution bathymetry data, and (3) present initial results from a case study of a selected site from the uThukela MPA ACEP (African Coelacanth Ecosystem Programme) Smart Zones Project.

## Multibeam echo-sounders and benthic habitat mapping

Within the South African context, there is increased interest in developing enhanced high-resolution seabed models.[42] Recent marine geological mapping and marine research[41,42] have presented a strong case for the possible role of MBES data in the realm of benthic habitat mapping.[18,21,38] MBES provides the baseline data (bathymetry, backscatter, slope, etc.) from which the seabed and habitat maps can be derived and interpreted in conjunction with ground-truth data[20], thus providing a detailed bathymetric surface which can in turn be used for planning and decision-making. In addition, backscatter intensity data provide an indication of seabed character (cf. Montereale-Gavazzi et al.[49]), discerning between sediment classes and consolidated seabed. The system frequencies are optimised for specific depth ranges and must be employed accordingly.[36,37,43,50,51] MBES data can be processed to produce dataset derivatives (e.g. rugosity, slope, bathymetric position index and aspect maps), which offer enhanced and detailed properties of the substrate known to influence benthic diversity. Marine mapping efforts have begun to match in data quality and resolution those of the terrestrial realm[18], and therefore are as informative as terrestrial topographic observations.

## Methods

For this study, sites between Durban and Richards Bay were selected for MBES mapping based on (1) features identified on low-resolution regional datasets[40,52] (and references therein) (2) being known recreational fishing areas (likely reefs) and (3) their location relative to the uThukela MPA management zones. Hydroacoustic surveys were carried out on the reefs and their adjacent areas in the 40–100 m depth zones, depending on site location. The Geophysics and Mapping Platform (GeMaP) provided by ACEP was used to acquire MBES data. The platform's Seabat Reson 7101 MBES and SBG systems (Navsight Apogee inertial navigation system) were used during data collection, whilst maintaining on average a ±10% overlap between the adjacent survey lines to increase data integrity at the swath edges. Sound velocity profiles were collected periodically throughout the day, monitored at <3 m/s difference between the profile velocity from the live sound velocity. Raw data were processed using HYPACK (2022) and Qinertia (v. 3.0.5966) to produce gridded (3 m cells) digital terrain models of the seabed. Golden Software Surfer (v. 23.3.202) and ESRI ArcMap 10.1 were used to visualise and interpret data and derivatives.

## Case study: The uThukela Banks Marine Protected Area

The uThukela Banks MPA (5666 km$^2$; Figure 1) proclaimed in 2019[32] is situated on South Africa's wave-dominated[53] east coast continental shelf. This MPA is subdivided into zones with varying degrees of accessibility and/or protection from human pressure[54] (Figure 1) to regulate the declining continental shelf biodiversity[55]. The uThukela River terrestrial sediments and associated offshore unconsolidated material deposits have actively shaped the eastern KwaZulu-Natal (KZN) continental shelf since the break-up of Gondwana.[56] Further details of the shelf stratigraphy of the area have been documented by Hicks and Green[57] (and references therein). This MPA has not received much attention in terms of high-resolution MBES mapping despite encompassing an extensive subaqueous delta[58,59] and multiple reef complexes noted for high biodiversity[52]. Green et al.[34] recently carried out an extensive MBES survey, providing new insights into higher-resolved geomorphology within the uThukela Banks MPA. This extensive survey showed the value of incorporating geological data into informing MPA management (and therefore, MSP). Prior to this, low-resolution bathymetry and regional sedimentary facies had been described from the region in 2007.[58] A study in 2012[52] developed a systematic framework for assessment of biodiversity and marine biodiversity protection for KZN by recognising spatial priorities for sustainable conservation efforts (formerly known as the SeaPLAN project). However, this was based on relatively low-resolution bathymetry data in which the geomorphology is poorly resolved. The singlebeam echosounder (SBES) mapping efforts of De Wet and Compton[40] (published in 2021) do span the uThukela MPA, but only contribute low-resolution and less-detailed bathymetry. To date, through the ACEP Smart Zones Project (unpublished data), 13 of 14 selected sites have been surveyed with MBES at an average coverage of 17 km$^2$ per site, providing ca. 230 km$^2$ of new MBES data over key localities within and adjacent to the uThukela MPA (Figure 1).

## Results

With this case study, we focus on one of the surveyed ACEP Smart Zones MPA Project sites (at 40–70 m depth range; Figure 2) which shows a locally steeply inclined (13°) seaward slope. Compared to the SBES dataset (Inset A; Figure 2)[40], we note that the new (MBES) data allow a precise distinction between sediment-starved areas from those adjacent zones (i.e. either rich in sediments and/or featureless seabed), and between different types of outcrop morphology. Areas adjacent to the outcrop likely represent unconsolidated sediments described by flat-lying sediments and bedforms. These large superimposed subaqueous dunes form discontinuous fields along the mid-shelf (Inset B; Figure 2). The sediment-starved areas are characterised by exposed reef. Reef geomorphology is variable with rugged reef pinnacles and ridges bordered by low-relief reefs fringed by adjacent relatively flat and/or featureless seabed. The reefs are commonly oriented approximately coast-parallel on the inner- to mid-shelf, representing submerged shorelines.[33,59] Abrupt changes in depth of the seabed, as observed on the outer edges of the reefs (Figure 2), are recognised as prominent zones of overstepped submerged shorelines.[33,59]

## Discussion

### *Application of MBES to substrate analysis*

The data example provided here underscores the distribution of reef and adjacent sediments (and/or featureless seabed) at this survey site. Reefs play an integral role in marine habitats as they are biodiversity hotspots on the continental shelf (e.g. Caribbean nations reefs[60]) and are vulnerable to both anthropogenic and natural impacts.[61] Thus reefs and organisms that inhabit them benefit from proclamation

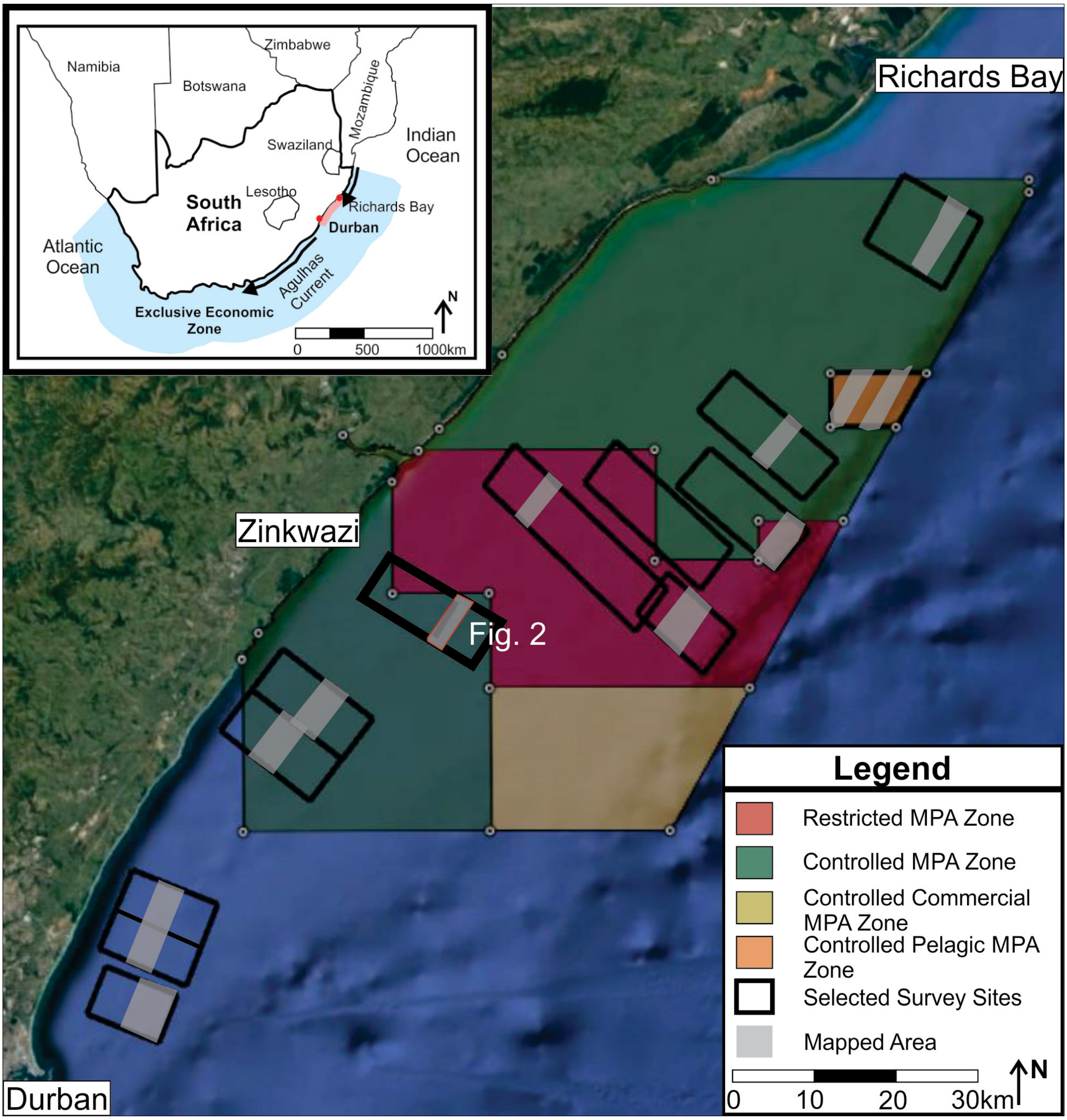



**Figure 1:** Selected survey sites and their location relative to the uThukela Marine Protected Area (MPA) zones. Note: the mapped (unpublished) areas within the selected site blocks are highlighted.

of these localities as MPAs. Duran et al.[62] previously reported that understanding the temporal and spatial variation of reef communities is of critical importance to monitoring stressors on health and ecosystem functioning (e.g. fishing[63]) and global stressors (e.g. climate change[64]). The high-relief reef (less vulnerable to inundation by sediments) could provide stable habitat for long-term inhabitants, whereas the low-relief reef (including the high-relief reef edges) preferentially hosts short-term inhabitants as it is more vulnerable to burial by the dynamic sediments (cf. Harman et al.[65]). The interaction of the high-relief reef with the localised dynamic system induces turbulence, promoting the growth and density of substrate attached organisms[66], and therefore, creating accommodation for biodiversity abundance[65]. This, therefore, emphasises how the seabed geomorphology is the fundamental building block to the systems that exist on and/or above its surface.[42,48] The uThukela MPA case study demonstrates improved resolution detail of the reef in the dataset (Figure 2), which was previously not well represented from the lower-resolution datasets.[40] Within the sediment-rich regions, the presence of bedforms marks a clear indication of current-driven sediment transport on the seabed, which will be further investigated in more detail at a later stage. Bedforms and their sediment dynamics have been studied from northern to southern KZN[67] and are mostly considered the result of the poleward-flowing geostrophic Agulhas Current[33]. These adjacent sediment-rich regions play a role in the broader ecosystem and habitat provision.[68]

### *MBES contribution towards MSP*

The improvements made to raw and primary data quality (reliability) impact the country's MSP process and decision-making outcomes positively.[69] MBES can be used for both short- and long-term monitoring of dynamic seabed sediment processes.[39] This is vital in monitoring the spatial and temporal habitat-use compatibility, a potentially valuable contribution towards decision-making in regions of conflicting spatial interests between economic drivers. Through its use in monitoring the low-relief reef and its inhabitants, MBES data will prove critical,

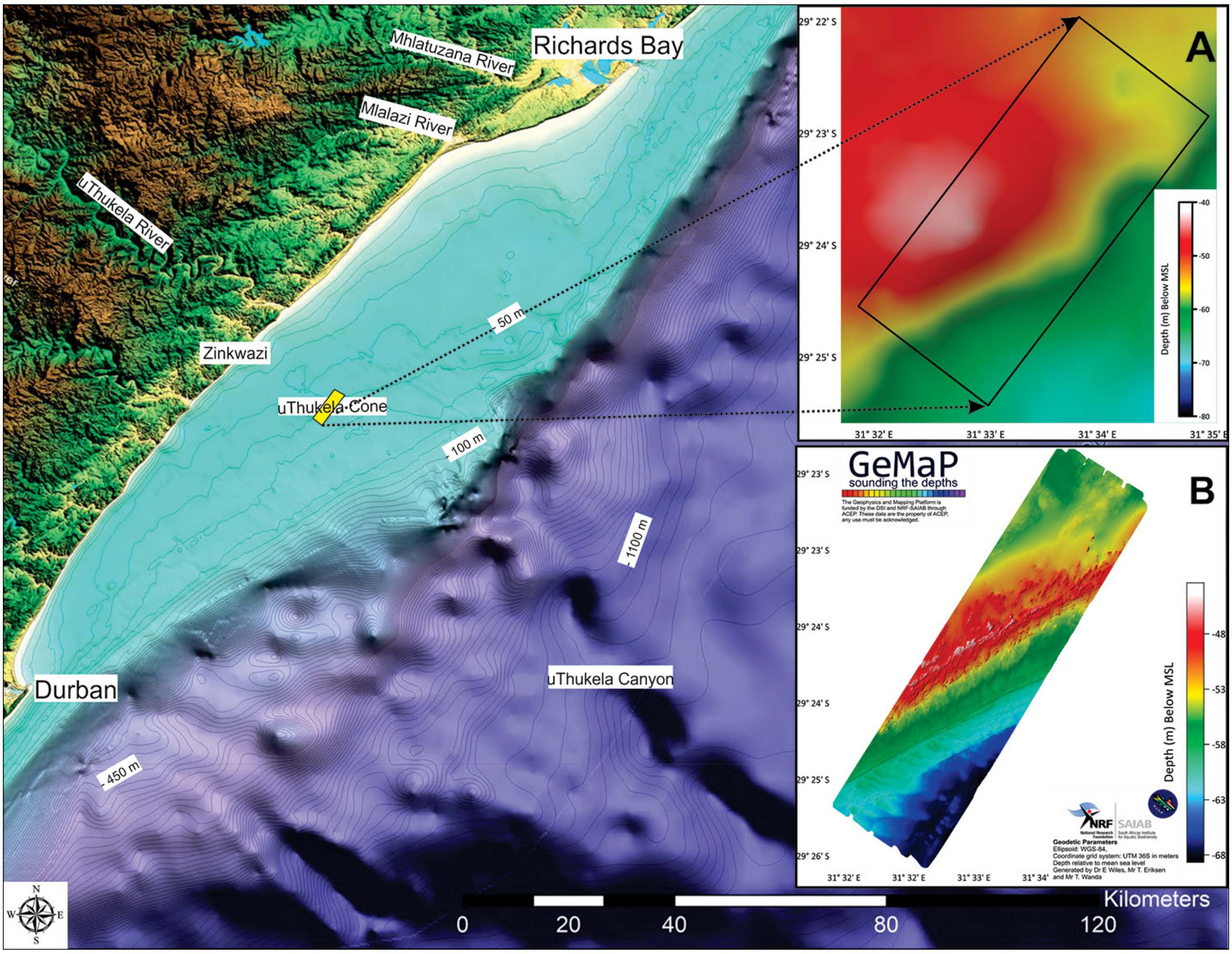



**Figure 2:** Hydroacoustic surveys of a target area within the uThukela Banks Marine Protected Area. Inset A: Previously available low-resolution bathymetry data[37] (modified). Inset B: Higher-resolved seabed bathymetry from the on-going ACEP Smart Zones Project, offshore Zinkwazi.

as reef resilience requires withstanding press and/or pulse types of sedimentation (including in situ sedimentation) disturbances.[70] The MBES data and its derivatives enable detailed seabed geomorphology to be modelled. The rugged geomorphological character of reefs could likely provide several biological niches for varying ecosystems.[38]

## *Gaps in current South African MSP in relation to geoscientific research*

Recent global initiatives (e.g. the National Development Plan 2030 goals[44]) have encouraged a dedicated focus to elevate the significance of seabed mapping. Although detailed high-resolution data do exist, much of these data remain locked under commercial and military embargo, without public access. Restricted access to existing commercial data is acknowledged globally as a significant challenge.[71] High-resolution mapping within South African MSP and therefore the ability to quantify the contribution of seabed type and characteristics in ocean system functions is scarce. Of the 41 MPAs within South Africa's EEZ, none is covered entirely by MBES data, with relatively limited (or focused) coverage where data do exist (e.g. uThukela MPA[34]). However, demarcation based on rich biodiversity, even without marine geological context, is testament to the work of researchers from marine sub-disciplines (biological and physical oceanography), with qualitatively described marine habitats in the literature (cf. trait-based assessment by Ortega-Cineros et al.[72]). To date, our knowledge (from MBES) of the structural properties of the uThukela Banks MPA seabed is limited (e.g. focused multibeam bathymetry[34]). Given that South Africa's MSP management framework is still in its early years of implementation, the current framework is likely subject to revision and fine-tuning for application in future marine spatial plans. The globally endorsed ecosystem-based MSP for management of the ocean space itself has often been problematic to translate into operational management and further enhance work already done in MSP.[73]

## *Reflections on the future direction of MSP from a geoscience perspective in South Africa*

The growing world population, if left unchecked, threatens the replenishment and sustainability of marine natural resources.[2] There is no better time to initiate planning for probable future conflicts than the present.[74] It is critically important for South Africa to aim for growth towards effective marine spatial plans. Tasks suggested for future consideration include an expansion of mapping the seabed and defining habitat through this mapping to cover (most broadly) South Africa's EEZ. Closer collaboration between government departments, research organisations and the private sector could be fostered to achieve an integrated goal (of multi-functional marine data), from the available national budget. The MSP process is rapid, output-oriented, and in many instances authority-driven (by various sectors), on a set budget and reliant on the quality of data input.[12,75] These factors are anticipated to significantly contribute to heterogeneous approved, implemented and developed MSP initiatives. Improvements to the baseline primary data will ensure that these heterogeneous initiatives stem from detailed maps for ocean use and management (through the allocation of meaningful spatial boundaries).

Ehler[76] proposed that by the year 2030, a third of the world's EEZs will be covered by government-approved marine spatial plans. The mapping of South African marine geomorphology and habitats by MBES would contribute substantially to the goals of the National Development Plan

2030[44] and global alignment (e.g. United Nations Sustainable Development Goals). In addition, it will contribute foundational knowledge as well as health monitoring systems within MSP[77] for interested sectors. There is no best single method to perform strategic spatial planning[78], and thus a variety of techniques and data sources are required. Hydroacoustic surveys are not without their challenges; equipment, including a suitable vessel, and software are costly, complex and require specialist installation and operation. Surveys are time consuming and are subject to suitable weather conditions. Hence, projects need to have appropriate budgets, personnel (scarce skills in South Africa) and time to generate bathymetric products (Phase 1) before any complementary fieldwork (Phase 2). Phase 2 efforts (baited remote underwater video, remotely operated vehicles, sediment grabs, benthic sleds, etc.) would not only benefit in terms of site selection and sampling design using a digital terrain model and derivatives (rugosity) but also from the geospatial context of the larger study area. The ACEP-supported SMART ZONES MPA Project has been initiated to achieve this on a small scale, with mapping activities taking place across strategic reef sites within and adjacent to the uThukela MPA, followed by biological, oceanographic and remote-imagery sampling campaigns. For this project, the GeMaP based in KZN will provide access to high-resolution bathymetric mapping tools and vessels to collect essential bathymetry data, upon which biological and oceanographic sampling and modelling will be based.

## Conclusion

Seabed composition and substrate structure have a significant impact on marine biological and oceanographic systems[42,67], ranging from the role of a specific ecological niche to the general marine habitat describing a particular biome. This contribution demonstrates the effectiveness of using MBES in hydroacoustic surveys, where seabed features are resolved in much greater detail and accuracy whilst revealing new features and seabed interactions in higher detail than previously available techniques could achieve. Therefore, South African researchers, MSP practitioners and the government will greatly benefit in making better decisions when planning, monitoring and protecting such MPAs. Our case study shows the level of detail that can be achieved by mapping reef habitat and adjacent areas. Technological improvements of MBES are anticipated to greatly benefit South Africa's marine management sector.

Hydroacoustic and bathymetric surveys are well known in marine geosciences; however, increased exposure to broader elements of marine science in general is encouraged to allow meaningful integration and holistic knowledge generation. Such integration is essential for MSP, providing multi-functionality and data integrity.[79] The surveys carried out in the uThukela MPA and neighbouring sites serve to highlight the insufficient detail in our knowledge of the present-day seabed bathymetry. These surveys provide a preview of the value MBES will add (including cross-discipline collaboration future developments) for the South African government and marine research practitioners within the EEZ. This will assist in the further progression of work already done in MSP, provided the required skillset and funds are available.

## Competing interests

We have no competing interests to declare.

## Authors' contributions

T.F.W.: Conceptualisation, data collection, data analysis, writing – the initial draft, writing – revisions. E.A.W.: Supervision, data collection, funding acquisition, writing – review and editing. H.C.C.: Supervision, writing – review and editing. A.d.W.: Supervision, validation.

## References

1. Harris J, Lombard AT. Socio-ecological resilience in South Africa's blue economy: The role of marine protected areas. Policy Briefing No 176. Johannesburg: South African Institute of International Affairs; 2018.
2. Ehler CN, Zaucha J, Gee K. Maritime/marine spatial planning at the interface of research and practice. In: Zaucha J, Gee K, editors. Maritime spatial planning: Past, present, future. London: Palgrave; 2019. p. 1–21. https://doi.org/10.1007/978-3-319-98696-8_1
3. South African Department of Planning, Monitoring and Evaluation (DPME). Operation Phakisa: Oceans Economy Lab [webpage on the Internet]. c2015 [cited 2022 Jul 11]. Available from: https://www.operationphakisa.gov.za/operations/oel/pages/default.aspx
4. Dorrington RA, Lombard AT, Bornman TG, Adams JB, Cawthra HC, Deyzel S, et al. Working together for our oceans: A marine spatial plan for Algoa Bay, South Africa. S Afr J Sci. 2018;114(3/4), Art. #a0247. https://doi.org/10.17159/sajs.2018/a0247
5. Ehler CN, Douvere F. Marine spatial planning: A step-by-step approach toward ecosystem-based management. IOC Manual and Guides No. 53 (Intergovernmental Oceanographic Commission and Man and the Biosphere Programme). Paris: Unesco; 2009.
6. Agardy T, Di Sciara GN, Christie P. Mind the gap: Addressing the shortcomings of marine protected areas through large scale marine spatial planning. Mar Policy. 2011;35:226–232. https://doi.org/10.1016/j.marpol.2010.10.006
7. Ehler CN. Pan-Arctic marine spatial planning: An idea whose time has come. In: Tedsen E, Cavalieri S, Kraemer RA, editors. Arctic marine governance. Heidelberg: Springer; 2014. p. 199–213. https://doi.org/10.1007/978-3-642-38595-7_9
8. Manea E, Di Carlo D, Depellegrin D, Agardy T, Gissi E. Multidimensional assessment of supporting ecosystem services for marine spatial planning of the Adriatic Sea. Ecol Indic. 2019;101:821–837. https://doi.org/10.1016/j.ecolind.2018.12.017
9. Santos CF, Ehler CN, Agardy T, Andrad F, Orbach MK, Crowder LB. Marine spatial planning. In: Sheppard C, editor. World seas: An environmental evaluation. Volume three: Ecological issues and environmental impacts. 2nd ed. London: Academic Press; 2019. p. 571–592. https://doi.org/10.1016/B978-0-12-805052-1.00033-4
10. Gilliland PM, Laffoley D. Key elements and steps in the process of developing ecosystem-based marine spatial planning. Mar Policy. 2008;32(5):787–796. https://doi.org/10.1016/j.marpol.2008.03.022
11. Foley MM, Halpern BS, Micheli F, Armsby MH, Caldwell MR, Crain CM, et al. Guiding ecological principles for marine spatial planning. Mar Policy. 2010;34:955–966. https://doi.org/10.1016/j.marpol.2010.02.001
12. Pınarbaşı K, Galparsoro I, Borja Á, Stelzenmüller V, Ehler CN, Gimpel A. Decision support tools in marine spatial planning: Present applications, gaps and future perspectives. Mar Policy. 2017;83:83–91. https://doi.org/10.1016/j.marpol.2017.05.031
13. Lombard AT, Dorrington RA, Reed JR, Ortega-Cisneros K, Penry GS, Pichegru L, et al. Key challenges in advancing an ecosystem-based approach to marine spatial planning under economic growth imperatives. Front Mar Sci. 2019;6:146. https://doi.org/10.3389/fmars.2019.00146
14. Republic of South Africa. Act No. 16 of 2018: Marine Spatial Planning Act, 2018. Government Gazette. 2019. Available from: https://www.dffe.gov.za/sites/default/files/gazetted_notices/msp_developmentofmarinespatialplans_g42444gon647.pdf
15. Ehler CN. Two decades of progress in marine spatial planning. Mar Policy. 2020. https://doi.org/10.1016/j.marpol.2020.104134
16. Feng R, Chen X, Li P, Zhou L, Yu J. Development of China's marine functional zoning: A preliminary analysis. Ocean Coast Manag. 2016;131:39–44. https://doi.org/10.1016/j.ocecoaman.2016.08.011
17. Kostylev VE, Todd BJ, Shaw J. Benthic habitat mapping in Canada – a perspective. J Ocean Technol. 2008;3:7–12.
18. Brown CJ, Sameoto JA, Smith SJ. Multiple methods, maps, and management applications: Purpose made seafloor maps in support of ocean management. J Sea Res. 2012;72:1–13. https://doi.org/10.1016/j.seares.2012.04.009
19. Johnson SY, Cochrane GR, Golden NE, Dartnell P, Hartwell SR, Cochran SA, et al. The California seafloor and coastal mapping program – providing science and geospatial data for California's state waters. Ocean Coast Manag. 2017;140:88–104. https://doi.org/10.1016/j.ocecoaman.2017.02.004
20. Brown CJ, Blondel P. Developments in the application of multibeam sonar backscatter for seafloor habitat mapping. Appl Acoust. 2009;70(10):1242–1247. https://doi.org/10.1016/j.apacoust.2008.08.004
21. Baker EK, Harris PT. Habitat mapping and marine management. In: Harris PT, Baker E, editors.Seafloor geomorphology as benthic habitat. Amsterdam: Elsevier; 2020. p. 17–33. https://doi.org/10.1016/B978-0-12-814960-7.00002-6

22. Kenny AJ, Cato I, Desprez M, Fader G, Schuttenhelm RTE, Side J. An overview of seabed-mapping technologies in the context of marine habitat classification. ICES J Mar Sci. 2003;60(2):411–418. https://doi.org/10.1016/S1054-3139(03)00006-7

23. Weatherall P, Marks KM, Jakobsson M, Schmitt T, Tani S, Arndt JE, et al. A new digital bathymetric model of the world's oceans. Earth Space Sci. 2015;2:331–345. https://doi.org/10.1002/2015EA000107

24. O'Toole R, Judge M, Sachetti F, Furey T, Mac Craith E, Sheehan K, et al. Mapping Ireland's coastal, shelf and deep water environments using illustrative case studies to highlight the impact of seabed mapping on the generation of blue knowledge. London: Geological Society; 2020. https://doi.org/10.1144/SP505-2019-207

25. South African Department of Environmental Affairs. Marine Spatial Planning Bill [B9-2017]. Government Gazette no. 40726 of 28 March 2017. 2017. Available from: https://pmg.org.za/files/170413B9-2017_Marine_Spatial_Planning.pdf

26. Roberts CM, Page RW, O'Leary BC, Allen HL, Yates KL, Tudhope AW, et al. 30×30: A blueprint for ocean protection. University of Oxford, University of York: Greenpeace; 2019.

27. Visalli ME, Best BD, Cabral RB, Cheung WWL, Clark NA, Garilao C, et al. Data-driven approach for highlighting priority areas for protection in marine areas beyond national jurisdiction. Mar Policy. 2020;122:103927. https://doi.org/10.1016/j.marpol.2020.103927

28. Sala E, Mayorga J, Bradley D, Cabral RB, Atwood TB, Auber A, et al. Protecting the global ocean for biodiversity, food and climate. Nature. 2021;592(7854):397–402. https://doi.org/10.1038/s41586-021-03371-z

29. Mayer L, Jakobsson M, Allen G, Dorschel B, Falconer R, Ferrini V, et al. The Nippon Foundation-GEBCO Seabed 2030 project: The quest to see the world's oceans completely mapped by 2030. Geosciences. 2018;8:63. https://doi.org/10.3390/geosciences8020063

30. Ierodiaconou D, Schimel AC, Kennedy D, Monk J, Gaylard G, Young M, et al. Combining pixel and object based image analysis of ultra-high resolution multibeam bathymetry and backscatter for habitat mapping in shallow marine waters. Mar Geophys Res. 2018;39:271–288. https://doi.org/10.1007/s11001-017-9338-z

31. Edgar GJ, Stuart-Smith RD, Willis TJ, Kininmonth S, Baker SC, Banks S, et al. Global conservation outcomes depend on marine protected areas with five key features. Nature. 2014;506:216–220. https://doi.org/10.1038/nature13022

32. Mann-Lang JB, Branch GM, Mann BQ, Sink KJ, Kirkman SP, Adams R. Social and economic effects of marine protected areas in South Africa, with recommendations for future assessments. Afr J Mar Sci. 2021;43(3):367–387. https://doi.org/10.2989/1814232X.2021.1961166

33. Pretorius L, Green AN, Cooper JAG, Hahn A, Zabel M. Outer- to inner-shelf response to stepped sea-level rise: Insights from incised valleys and submerged shorelines. Mar Geol. 2019;416:105979. https://doi.org/10.1016/j.margeo.2019.105979

34. Green AN, Engelbrecht LD, Sink K, Mackay CF, Fennessy ST, Michalsen K. Lifting back the waters: Marine geophysics provides new insights into the Uthukela Banks Marine Protected Area. S Afr J Sci. 2022;118(11/12), Art. #14020. https://doi.org/10.17159/sajs.2022/14020

35. Palan K, Green AN, Andrews B, Sink K, Wiles EA. A morphometric analysis of the fluid flow features of the southern Orange Basin, South Africa. Mar Geol. 2020;423:106145. https://doi.org/10.1016/j.margeo.2020.106145

36. Baker EK, Harris PT. Habitat mapping and marine management. Seafloor Geomorphol Benthic Habitat. 2021;21:23–38. https://doi.org/10.1016/B978-0-12-385140-6.00002-5

37. Brown CJ, Beaudoin J, Brissette M, Gazzola V. Multispectral multibeam echo sounder backscatter as a tool for improved seafloor characterization. Geosciences. 2019;9(3):126. https://doi.org/10.3390/geosciences9030126

38. Brown CJ, Smith SJ, Lawton P, Anderson JT. Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. Estuar Coast Shelf Sci. 2011;92(3):502–520. https://doi.org/10.1016/j.ecss.2011.02.007

39. Bosman A, Romagnoli C, Madricardo F, Correggiari A, Remia A, Zubalich R, et al. Short-term evolution of Po della Pila delta lobe from time lapse high-resolution multibeam bathymetry (2013–2016). Estuar Coast Shelf Sci. 2020;233:106533. https://doi.org/10.1016/j.ecss.2019.106533

40. De Wet WM, Compton JS. Bathymetry of the South African continental shelf. Geo-Mar Lett. 2021;41(3). https://doi.org/10.1007/s00367-021-00701-y

41. Pillay T, Cawthra HC, Lombard AT. Characterisation of seafloor substrate using advanced processing of multibeam bathymetry, backscatter, and sidescan sonar in Table Bay, South Africa. Mar Geol. 2020;429:106332. https://doi.org/10.1016/j.margeo.2020.106332

42. Pillay T, Cawthra HC, Lombard AT. Integration of machine learning using hydroacoustic techniques and sediment sampling to refine substrate description in the Western Cape, South Africa. Mar Geol. 2021;440:106599. https://doi.org/10.1016/j.margeo.2021.106599

43. Gaida TC, Mohammadloo TH, Snellen M, Simons DG. Mapping the seabed and shallow subsurface with multi-frequency multibeam echosounders. Remote Sens. 2019;12(1):52. https://doi.org/10.3390/rs12010052

44. National Development Plan 2030: Our future - make it work [document on the Internet]. c2012 [cited 2022 Jul 11]. Available from: https://www.gov.za/sites/default/files/gcis_document/201409/ndp-2030-our-future-make-it-workr.pdf

45. Costa BM, Battista TA, Pittman SJ. Comparative evaluation of airborne LiDAR and ship-based multibeam SoNAR bathymetry and intensity for mapping coral reef ecosystems. Remote Sens Environ. 2009;113(5):1082–1100. https://doi.org/10.1016/j.rse.2009.01.015

46. Laing SCS, Schleyer MH, Turpie JK. Ecosystem service values of sediment generation and entrapment by marginal coral reefs at Sodwana Bay, South Africa. Afr J Mar Sci. 2020;42(2):199–207. https://doi.org/10.2989/1814232X.2020.1771415

47. Berg H, Öhman MC, Troëng S, Lindèn O. Environmental economics of coral reef destruction in Sri Lanka. Ambio. 1998;27:627–634.

48. Wilson MFJ, O'Connell B, Brown C, Guinan JC, Grehan AJ. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. Mar Geod. 2007;30:3–35. https://doi.org/10.1080/01490410701295962

49. Montereale-Gavazzi G, Madricardo F, Janowski L, Kruss A, Blondel P, Sigovini M, et al. Evaluation of seabed mapping methods for fine-scale classification of extremely shallow benthic habitats-application to the Venice Lagoon, Italy. Estuar Coast Shelf Sci. 2016;170:45–60. https://doi.org/10.1016/j.ecss.2015.12.014

50. Williams KL, Jackson DR, Tang D, Briggs KB, Thorsos EI. Acoustic backscattering from a sand and a sand/mud environment: Experiments and data/model comparisons. IEEE J Ocean Eng. 2009;34:388–398. https://doi.org/10.1109/JOE.2009.2018335

51. Hughes Clarke JE. Multispectral acoustic backscatter from multibeam, improved classification potential. In: Proceedings of the United States Hydrographic Conference; 2015 March 16–19; National Harbor, MD, USA. p. 15–19 .

52. Harris JM, Livingstone T, Lombard AT, Lagabrielle E, Haupt P, Sink K, et al. Coastal and marine biodiversity plan for KwaZulu-Natal. Spatial priorities for the conservation of coastal and marine biodiversity in KwaZulu-Natal. Ezemvelo KZN Wildlife Scientific Services Technical Report. 2012.

53. Cooper JAG. Gemorphological variability among microtidal estuaries from the wave-dominated South African coast. Geomorphology. 2001;40:99–122. https://doi.org/10.1016/S0169-555X(01)00039-3

54. Kirkman SP, Mann BQ, Sink KJ, Adams R, Livingstone T-C, Mann-Lang JB, et al. Evaluating the evidence for ecological effectiveness of South Africa's marine protected areas. Afr J Mar Sci. 2021;43(3):389–412. https://doi.org/10.2989/1814232X.2021.1962975

55. McCauley DJ, Pinsky ML, Palumbi SR, Estes JA, Joyce FH, Warner RR. Marine defaunation: Animal loss in the global ocean. Science. 2015;347, Art. #1255641. https://doi.org/10.1126/science.1255641

56. Dingle RV. Continental margin subsidence: A comparison between the east and the west coasts of Africa. In: Scruton RA, editor. Dynamics of passive margins. Washington DC: American Geophysical Union; 1982. p. 59–71.

57. Hicks N, Green AN. Sedimentology and depositional architecture of a submarine delta-fan complex in the Durban Basin, South Africa. Mar Pet Geol. 2016;78:390–404. https://doi.org/10.1016/j.marpetgeo.2016.09.032

58. Bosman C, Uken R, Leuci R, Smith AM, Sinclair D. Shelf sediments off the Thukela River mouth: Complex interaction between fluvial and oceanographic processes. S Afr J Sci. 2007;103:490–492.

59. Engelbrecht LD, Green AN, Cooper JAG, Mackay CF. Antecedent geological control on transgressive delta and shoreline preservation: Examples from the SE African shelf. Mar Geol. 2022;454:106934. https://doi.org/10.1016/j.margeo.2022.106934

60. Miloslavich P, Díaz JM, Klein E, Alvarado JJ, Díaz C, Gobin J, et al. Marine Biodiversity in the Caribbean: Regional Estimates and Distribution Patterns. PLoS ONE. 2010;5(8),e11916. https://doi.org/10.1371/journal.pone.0011916

61. Hereher ME. Assessment of Egypt's Red Sea coastal sensitivity to climate change. Environ Earth Sci. 2015;74(4):2831–2843. https://doi.org/10.1007/s12665-015-4304-z

62. Duran A, Shantz AA, Burkepile DE, Collado-Vides L, Ferrer VM, Palma L, et al. Fishing, pollution, climate change, and the long-term decline of coral reefs off Havana, Cuba. Bull Mar Sci. 2017;94(2). https://doi.org/10.5343/bms.2017.1061

63. Cinner JE, Hunchery C, MacNeil MA, Graham AJ, McClanahan TR, Maina J, et al. Bright spots among the world's coral reefs. Nature. 2016;535:416–419. https://doi.org/10.1038/nature18607

64. Hughes TP, Kerry JT, Alvarez-Noriega M, Alvarez-Romero JG, Anderson KD, Baird AH, et al. Global warming and recurrent mass bleaching of corals. Nature. 2017;543:373–377. https://doi.org/10.1038/nature21707

65. Harman N, Harvey SE, Kendrick GA. Differences in fish assemblages from different reef habitats at Hamelin Bay, south-western Australia. Mar Freshw Res. 2003;54(2):177–184. https://doi.org/10.1071/MF02040

66. Guichard F, Bourget E. Topographic heterogeneity, hydrodynamics, and benthic community structure: A scale-dependent cascade. Mar Ecol Progress Series. 1998;171:59–70. https://doi.org/10.3354/meps171059

67. Green AN, Flemming BW, Cooper JAG, Wanda TF. Bedform evolution and dynamics of a geostrophic current-swept shelf, northern KwaZulu-Natal, South Africa. Geo-Mar Lett. 2022;42(1):1–22. https://doi.org/10.1007/s00367-021-00722-7

68. MacKay C, Untiedt C, Hein L. Local habitat drivers of macrobenthos in the northern, central and southern KwaZulu-Natal Bight, South Africa. Afr J Mar Sci. 2016;38(suppl1):S105–S121. https://doi.org/10.2989/1814232X.2016.1146631

69. Flynn S, Meaney W, Leadbetter AM, Fisher JP, Nic Aonghusa C. Lessons from a marine spatial planning data management process for Ireland. Int J Digit Earth. 2021;14(2):139–157. https://doi.org/10.1080/17538947.2020.1808720

70. Harris PT. Shelf and deep-sea sedimentary environments and physical benthic disturbance regimes: A review and synthesis. Mar Geol. 2014;353:169–184. https://doi.org/10.1016/j.margeo.2014.03.023

71. Woelfl A, Snaith H, Amirebrahimi S, Devey C, Dorschel B, Ferrini V, et al. Seafloor mapping – the challenge of a truly global ocean bathymetry. Front Mar Sci. 2018;6:283. https://doi.org/10.3389/fmars.2019.00283

72. Ortega-Cisneros K, Yokwana S, Sauer W, Cochrane K, Cockcroft A, James NC, et al. Assessment of the likely sensitivity to climate change for the key marine species in the southern Benguela system. Afr J Mar Sci. 2018;40(3):279–292. https://doi.org/10.2989/1814232X.2018.1512526

73. Young OR, Osherenko G, Ogden J, Crowder LB, Wilson JA, Day JC, et al. Solving the crisis in ocean governance: Place-based management of marine ecosystems. Environment. 2007;49(4):20–32. https://doi.org/10.3200/ENVT.49.4.20-33

74. Ehler CN. 13 Myths of marine spatial planning. Mar Ecosyst Manag. 2012;5:5–7.

75. Collie JS, (Vic) Adamowicz WL, Beck MW, Craig B, Essington TE, Fluharty D, et al. Marine spatial planning in practice. Estuar Coast Shelf Sci. 2013;117:1–11. https://doi.org/10.1016/j.ecss.2012.11.010

76. Ehler CN. World-wide status and trends of maritime/marine spatial planning. Presented at: The 2nd International Conference on Marine/Maritime Spatial Planning; 2017 March 15–17; Paris, France.

77. Lucieer V, Roche M, Degrendele K, Malik M, Dolan M, Lamarche G. User expectations for multibeam echo-sounders backscatter strength data-looking back into the future. Mar Geophys Res. 2018;39:23–40. https://doi.org/10.1007/s11001-017-9316-5

78. Albrechts L. Strategic (spatial) planning reexamined. Environ Plann B Plann Des. 2004;31(5):743–758. https://doi.org/10.1068/b3065

79. Pfaff MC, Logston RC, Raemakers SJPN, Hermes JC, Blamey LK, Cawthra HC, et al. A synthesis of three decades of socio-ecological change in False Bay, South Africa: Setting the scene for multidisciplinary research and management. Elementa. 2019;7(32):1–49. https://doi.org/10.1525/elementa.367

Check for updates

**AUTHORS:**
Ahmed C. Bawa[1*]
Anastassios Pouris[2]

**AFFILIATIONS:**
[1]Universities South Africa, Pretoria, South Africa
[2]Institute for Technological Innovation, University of Pretoria, Pretoria, South Africa
*Current: Johannesburg Business School, University of Johannesburg, Johannesburg, South Africa; Tokyo College, University of Tokyo, Tokyo, Japan

**CORRESPONDENCE TO:**
Ahmed Bawa

**EMAIL:**
ahmed.bawa@jbs.ac.za

**DATES:**
**Received:** 25 Sep. 2022
**Revised:** 16 Mar. 2023
**Accepted:** 20 May 2023
**Published:** 28 Sep. 2023

**HOW TO CITE:**
Bawa AC, Pouris A. An assessment of the economic impact of South Africa's public universities. S Afr J Sci. 2023;119(9/10), Art. #14851. https://doi.org/10.17159/sajs.2023/14851

**ARTICLE INCLUDES:**
☒ Peer review
☐ Supplementary material

**DATA AVAILABILITY:**
☐ Open data set
☐ All data included
☒ On request from author(s)
☐ Not available
☐ Not applicable

**EDITORS:**
Chrissie Boughey
Nkosinathi Madondo

**KEYWORDS:**
universities, economic impact, higher education, gross-value added

**FUNDING:**
Universities South Africa



# An assessment of the economic impact of South Africa's public universities

With the understanding that universities play multiple social purposes, we aimed to provide an estimation of the economic impact of the public universities in South Africa. Using models described in the literature, we estimated economic benefits from four university activities – university exports, research at universities, the production of graduates and universities as business entities. Comparative analysis shows that, as an economic sector, Higher Education contributes more to South Africa's gross value added than other economic sectors such as Wood and Wood Products, Textiles, Clothing and Leather Goods, or Paper and Paper Products. It is comparable to sectors such as Gold Mining, and Beverages and Tobacco. Taking into account a number of assumptions, which are explained in the text, for 2018 the total economic impact was estimated at about ZAR513 billion. Governmental expenditure on higher education in that year was ZAR66 billion. These figures produce a cost–benefit ratio for the sector of 1 : 7.7, considering only these four university activities.

**Significance:**

The South African public higher education institutions form a very significant economic sector within the national economy. In terms of the gross value added, it is very similar in size to the gold mining industry. Taking into account four activities of universities, this sector contributes about ZAR500 billion annually to the economy, which is likely to be an underestimate of the actual contribution. This finding opens the way for policymakers to understand the importance of the sector as an area of investment. Recognising the potential limitations of the use of the modelling developed for other economies, our study indicates the importance of further work to indigenise the economic modelling for local conditions.

## Introduction

Universities are knowledge-intensive social institutions. They are created by societies with the understanding and expectation that they play vital, complex roles in multi-layered democracies such as South Africa. Universities are special as knowledge-intensive institutions in the sense that they have students. They are expected to produce new generations of socially engaged professionals, experts, and intellectuals. In societies like South Africa, which are deeply unequal, they are expected to generate social mobility and a society that is more socially just; this, notwithstanding the fact that they are deeply embedded in the political economies in which they are located. They are expected to generate new knowledge that enhances the growth of society's understanding of itself, of nature, and of the universe. Universities are necessary for building social cohesion. They preserve, transmit and recreate culture. These are all non-tangible outcomes which are measured qualitatively. There are many other activities which bring benefits which are currently unmeasurable.[1]

On the other hand, there are tangible, quantitatively measurable outcomes of the work done by universities. The international literature identifies a number of university activities that may be characterised as bringing economic benefits to a society. They produce knowledge, data and information that are taken up by the economy for the purposes of product development which may be industrial, service-related, or of social relevance. They produce graduates who become business developers or are employed in higher paying jobs and contribute to the national personal income tax base. Furthermore, universities are substantial business entities in the contexts in which they find themselves. Universities South Africa (USAf), an umbrella body that represents public universities in South Africa, engaged in this study to develop a broad understanding of the contribution of its member universities to the national economy.

Internationally, economic impact assessments of universities are regularly undertaken in order to inform governments and society at large of the economic importance of the sector and to advocate for suitable public and private investments in higher education. When performed on longitudinal timescales, these analyses also help inform universities and university systems of the ways in which they impact society and how these may be further shaped by strategic interventions. Examples of these studies include those by Oxford Economics[2] for the UK, Biggar Economics[3] for the League of European Research Universities (LERU) and KPMG Econtech[4] for Australian universities.

The economic impact of higher education is defined by Beck et al.[5] as "the difference between existing economic activity in a region given the presence of the institution and the level that would have been present if the institution did not exist". In an editorial of *Nature*[6], it was stated that "every government and organization that funds research wants to support science that makes a difference". Siegfried et al.[7] argue that the main purpose of the universities' impact studies is to express the value of an institution or the whole higher education sector, often to assist in sourcing funding, obtaining a subvention, or answering to criticism.

There are few published investigations related to the impact of the higher education sector in South Africa.[8] Furthermore, a number of investigations focus on one particular institution only. For example, Dyason[9] and Coetzee[10] investigate North-West University and Orr[11] Stellenbosch University.

Here we report the impact of South African public universities using methodologies developed for other systems of universities, all in the Global North. There are limitations related to this approach and where practical alternative approaches are suggested. Further, this analysis is based on 2018 data and therefore the impact of the COVID-19 pandemic is missed.

## Literature review

There is a multitude of investigations related to various aspects of the impact on economies of education in general and higher education in particular. Hanushek[12] identifies that cognitive skills can explain most of the differences in growth rates across countries. Psacharopoulos et al.[13] review investigations estimating the returns on primary, secondary and higher education. Bloom et al.[14] argue that, despite the encouragement of the international community that African countries should not neglect higher educationthe international community of the African governments, their research has identified that tertiary education has an important role in promoting economic growth and alleviating poverty. Tilak[15] found that different levels of education affect development outcomes differently; for some development outcomes, primary and secondary education may be more important than tertiary education, while for others such as income growth rate, tertiary education may be more important.

The above literature, although informative, does not focus on quantifying the impact of the various activities of higher education on the economy. The estimation of impact is not a straightforward process. As Bowen[16] mentions: "For individuals, the outcomes of higher education are harvested over their lifetimes averaging fifty to sixty years after graduation from college. For society the impacts may persist through centuries." For example, individuals benefit through the 'graduation premium' in the form of higher wages, while society at large benefits through enhanced economic growth, the provision of services of all kinds, and the benefits that accrue to particular local, provincial or national government structures or other social institutions.

Recent studies[2,3,17,18] have focused on only four activities amongst a plethora of university activities. These four activities are:

- The economic impacts of university exports (international students).
- The economic impacts of research at universities.
- The economic impacts of the production of graduates with enhanced knowledge and skills.
- The economic impacts of universities as business entities.

It is emphasised that researchers have chosen to focus on these particular activities of universities. There are others. The UK Department for Business, Innovation & Skills[1] identified more than 20 different concepts ranging from greater social cohesion, trust and tolerance, to increased entrepreneurial activity and productivity. What this means is that any estimation of economic benefits should be considered to be conservative.

## Methodology and results

Impact assessments of universities and university systems are developed through a range of different approaches and, generally, one would expect a composite picture to emerge. These assessments range from econometric approaches, qualitative analyses, national and international comparisons and the input–output modelling of the various domains in which higher education makes a contribution. Each approach is dependent on the underlying assumptions. For example, one of the major weaknesses of input–output models is that the relationships between sectors are assumed to be constant. This affects the estimation of multipliers used in these analyses.[19] The question that arises is how the uncertainty may be reduced.

Computable general equilibrium (CGE) modelling approaches allow for more elastic relationships between sectors and factors of production. However, because of the complexity of these CGE models, industrial detail is lost and modellers must at times make heroic assumptions to operationalise the model.[20]

There are two parts to this investigation. The first is the positioning of the university sector in the economy. For this we provide a number of comparative graphs which position the higher education sector alongside other industry sectors. The second, as has been pointed out above, is an estimation of the economic impact of four distinct industry operations. For this we utilise the modelling approaches used internationally, which we describe below. The four areas of university operation focused on are: university exports (as in international students); research and innovation activities; the production of graduates; and universities as business entities. As was pointed out earlier, there are indeed other activities that universities engage in which were not considered in this study. Examples of these are consultancy services, student residences, and service provision as in hospital services. There is limited methodological literature for those activities. If one considers the differentiation within South Africa's higher education system, the four areas that have been chosen reflect substantial areas of activity for all 26 public universities.

### *South African universities in terms of gross value added*

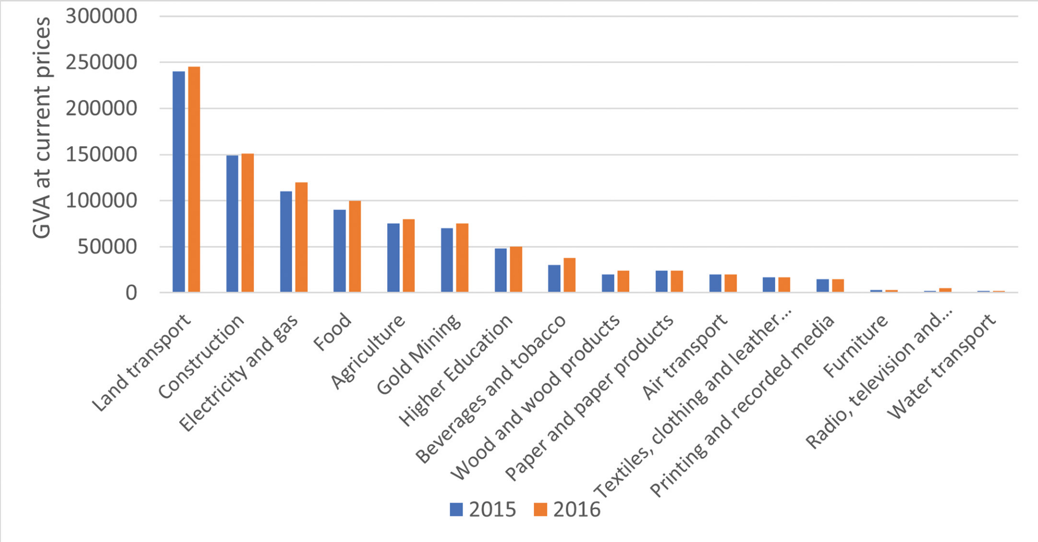
There are more than one (1) million students at the 26 public universities in South Africa. This represents a participation rate of about 21% of 18 to 24 year olds and the system produces more than 200 000 graduates a year. The unemployment rate of graduates is considerably lower than the general unemployment rate. In 2019, the university sector employed 162 865 individuals at all levels, of which 64 921 were permanent.[21] The university sector produces just less than 1% of the total research output of the world. It is a significant industry sector in the South African economy.

Even though the higher education sector in South Africa includes the Council on Higher Education (CHE), a number of other statutory and non-statutory bodies and a number of private institutions that offer higher education qualifications, for the purpose of these comparisons, the higher education sector refers just to the 26 public universities. To illustrate the position of the higher education sector in the economy relative to other economic sectors, we used data provided by the South African Reserve Bank to compare the gross value added (GVA) of the sector in Figure 1. GVA is one way to measure the contribution of a sector to the economy. The number represents "a quantitative assessment of the value of goods and services produced minus the cost of inputs and materials used in the production process"[22]. Figure 1 shows the GVA of a number of sectors for 2015 and 2016.[23] From this comparison, we can see that the contribution of the Higher Education sector to the country's GVA exceeds that of Wood and Wood Products; Textiles Clothing and Leather Goods; and Paper and Paper Products. It is comparable with sectors such as Gold Mining; and Beverages and Tobacco.

Figure 2 provides a measure of state spending on higher education as a percentage of GDP. This may not be the most effective measure, but it does allow a comparison with governmental expenditure in other national systems. As the graph indicates, as a percentage of GDP, the value of the higher education sector in South Africa in 2014 was 0.74%. This fell to 0.68% in 2018, and then increased somewhat following implementation of the recommendations of the Commission of Inquiry into Higher Education and Training, established by then President Jacob Zuma and chaired by Justice JA Heher. One of these recommendations was the need to raise the spend on higher education from 0.74% to 1% of GDP. For the 2019/2020 financial year, this figure was at about 0.9%. For the 2022/2023 financial year, the higher education block grant rose by just 0.9% on the previous year, significantly below CPI, and so we may expect another decline in this percentage.

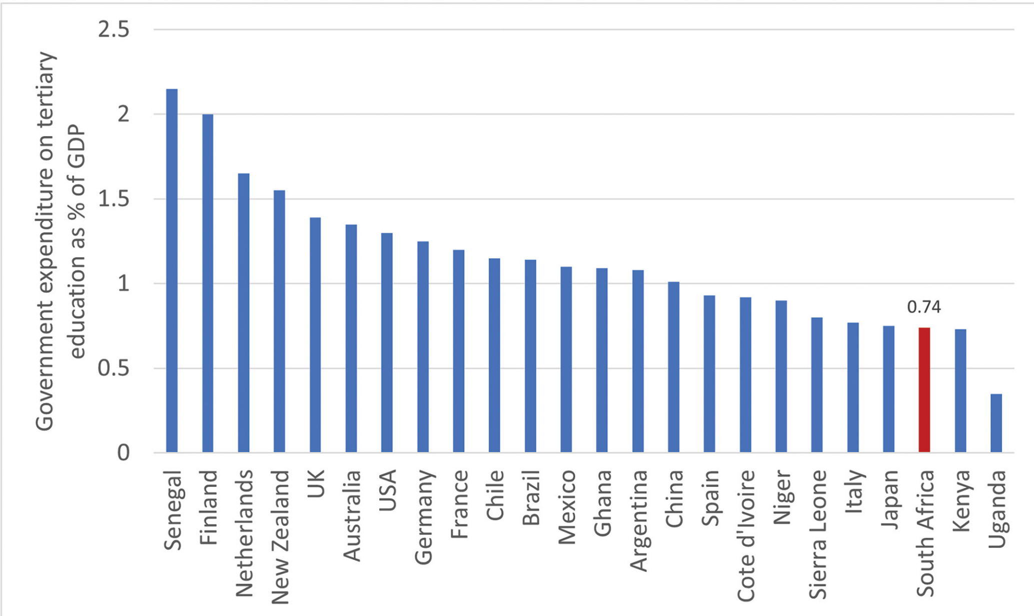
Figure 2 shows that South Africa ranks in a low position in terms of government expenditure to higher education, compared with not only developed countries but also some of the BRICS countries and other emerging and African countries. In this comparison, the countries in the sample, except for Uganda, show similar or higher shares of government expenditure to tertiary education in 2014. South Africa's low contribution is contrary to the National Development Plan which calls for significant growth in the sector.

Figure 3 provides another measure of the state of government spending on higher education. While in real rand terms, there has been a steady increase in subsidies to higher education, as a percentage of the overall national budget, this spend has been completely steady in a range




*Source: Stats SA*[23]

**Figure 1:** Gross value added (GVA) at current prices for various economic sectors, 2015 and 2016.




*Source: World Bank*[24] *under licence CC-BY 4.0*

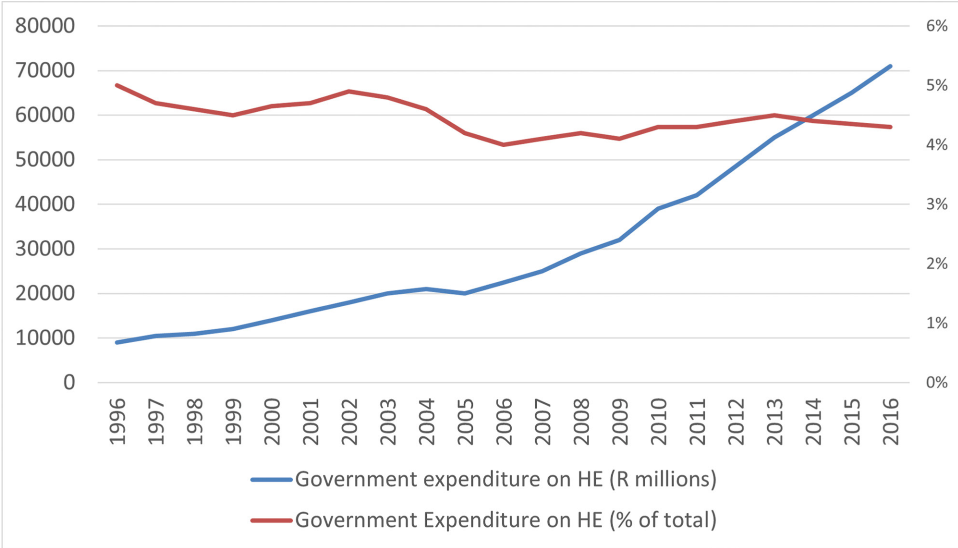
**Figure 2:** Government expenditure on tertiary education as a percentage of GDP (%), 2014.

between 4% and 5%. This finding illustrates that national policymakers continue to see higher education as an area of expenditure rather than one of investment.

The rapid growth in actual funding is attributable to the very significant increase in national spending on the new student funding system (via the National Student Financial Aid Scheme (NSFAS)) introduced at the ANC Electoral Conference in December 2017. This figure grew from ZAR11.8 billion in 2017/2018 to ZAR37.0 billion in 2019/2020. In that year, higher education subsidies reached ZAR42.3 billion, indicating that we are fast approaching the point at which government spending on student funding for students from poor and working-class backgrounds will overtake government spending on universities via the block subsidy grant and earmarked grants. Furthermore, the annual shortfall in funding from the national Department of Higher Education and Training (DHET) to NSFAS has been compensated for partly out of the governmental subsidy and earmarked grant allocation to universities, placing the medium- to long-term sustainability of the sector at great risk.

If one considers that the gross domestic expenditure on research and development (GERD) has continued its precipitous decline as a fraction of GDP and that, more specifically, private sector spend on research and development (R&D) has declined to the extent that government spending on R&D has now overtaken private sector spending, this flat line in state investment in higher education (as a fraction of national budget spend) is of deep concern.[24]

There will have been some shift in the position of the Higher Education sector in the national economy in recent years, especially with the impact of COVID-19 and the subsequent impact on the funding of higher education, admission statistics and cuts in funding for research and innovation. Future analyses will detect these.




*Source: Stats SA*[23]

**Figure 3:** Government expenditure on the Higher Education (HE) sector (millions of rands and % share of total).

## *The economic impact of universities*

In this section, as was pointed out above, we estimate the impact of four activities of universities on the national economy. It has to be emphasised again that there are other operations and activities that universities engage in which may require different estimation approaches. Here we focus on university exports, the impacts through the research done at universities, the impacts of human capital development and the impacts of universities acting as business entities.

### Impacts through university exports

The assessment of educational exports refers to money spent in the country by international students at South African universities. Approximately 7% of the students at the 26 public universities are not South African, and are predominantly from other southern African countries.[25] While the proportion of international students in other higher education systems in the Global North is much higher, it still makes sense for us to include this analysis. The Southern Africa Development Community (SADC)[26] education protocols (1997) commit universities to charge local fees for students from member states on the basis that they are subsidised by the South African state to the same extent as local students.

The literature shows that there are two approaches to this analysis, depending on the availability of data. Combining the individual fee income of international students with the number of students enrolled in a particular year provides an approach to calculating the total tuition fee income. Making assumptions on average study duration we can calculate the tuition fee income per overseas student from start to completion of studies. Subsequently, the values can be used to estimate the present value. This approach is used when we would like to estimate the export value of a particular type of education over its total time of duration. Alternatively, we can estimate the educational exports for a particular year. Students should be distinguished as undergraduates or postgraduates as the tuition fees and duration of study might differ. For the purpose of this analysis we used the second approach.

In addition to the tuition fee income that foreign students generate, they incur expenditures on non-tuition related items whilst studying, including general expenses (e.g. on mortgage/rent, food and household supplies, utilities, transport costs, medical and health costs, communication costs) as well as study-related expenses (e.g. on textbooks, stationery, non-tuition university fees). Similarly, visits from relatives and friends contribute to exports by the universities.

Our estimates are based on data derived from DHET[27] which show that in 2018 there were 62 326 undergraduate, just under 10 000 doctoral, and a similar number of other postgraduate non-South African students registered at South African universities.

The estimated average tuition fee for undergraduates was ZAR41 911 during 2018. Average postgraduate fees were estimated at ZAR40 000 per year. Apart from the tuition fees, students face non-tuition-related fees. The average cost for a year of study was estimated at ZAR122 545.[27] On the basis of these assumptions, it was estimated that the total income from international students is ZAR10 billion per annum, while undergraduate and postgraduate international students generate ZAR7.8 billion and ZAR2.2 billion per annum, respectively. The expenditure of family and friends who visit international students in South Africa must then be taken into account. It is assumed that these expenditures are due to the presence of international students. For this investigation we assumed that each undergraduate student received three friends/family members per year who in turn spent three (3) days in South Africa. It was estimated that this expenditure amounts to just over ZAR1 billion.[28] In summary, the direct impact associated with the expenditure of international students and their families was estimated to be ZAR11.0 billion pa.

In relation to the indirect and induced effect (or knock-on effect) of the international students on the South African economy, we utilised the approach of benefit-transfer as described by London Economics[17]. The approach is used to estimate economic values for ecosystem services by transferring available information from studies already completed in another location and/or context. Estimates of the economic multipliers relating to the expenditures of universities in other countries (e.g. Australia) show that the multiplier is 1 : 3 for resources spent within the university sector. The multiplier for off-campus expenditure is 1 : 3.14. That is, for ZAR1 million of expenditure at the universities, a total of ZAR3 million output is generated through the rest of the economy.[17] Transferring these multipliers for South Africa, the total impact (direct + indirect + induced) for expenses within universities is ZAR33 billion. This assumes that the multipliers developed for the Australian case hold for the South African case.

Notwithstanding the pressure for places in the South African higher education sector, this estimation of the impact of international students on the South African economy opens the way for a more rigorous analysis of whether the number of students from other countries should be expanded without jeopardising the participation of South Africans in higher education.

### Impact of research activities

The economic impact of research activities at South African universities was estimated by combining information on the research-related income accrued by the universities available in a particular year, with estimates from the wider economic literature on the extent to which public investment in research activity results in additional productivity (i.e. positive 'productivity spillovers').

Assuming that the direct economic impact of research generated by South African universities is equal to the funding that these universities receive for the purposes of research-related activities each year, the direct effect of the research activities of South African universities can be derived from the reports of the Department of Science and Innovation (DSI) and the Human Sciences Research Council (HSRC).

The total university sector research and development (R&D) expenditure during 2017 was ZAR12.6 billion.[29] In addition to the direct impact of research activities in terms of the income derived by universities (and subsequent expenditure), the wider academic literature indicates that investments in intangible assets, such as R&D, may induce positive externalities in the broader economy. Strong evidence of the existence of market sector productivity spillovers from public R&D expenditure is provided in the international literature.[17,30] These reports estimate that the elasticity of market sector productivity with respect to public spending on higher education R&D stands at 0.175. In other words, at the margin, previous findings suggest that a 1% increase in public spending on university research is associated with an increase of 0.175% in the rest of the economy.

Using this approach, others[17] have inferred an average spillover multiplier of 9.76 associated with certain Australian universities. Using the value transfer approach, we assumed that for every ZAR1 invested in university research, an additional economic output of ZAR9.76 is generated across the rest of the South African economy. In order to test the order of magnitude of the multiplier for the South African economy we utilised the assessment of the Technology and Human Resources for Industry Programme (THRIP).[31] THRIP provided government incentives to the private sector to generate research and innovation partnerships between industry and the higher education sector. Among the approaches used in the assessment was a survey of business stakeholders who had invested in the programme. On the question "how much revenue is your company expected to earn from selling goods or services incorporating THRIP technology?", the median respondent (50%) answered that 5 years after the completion of the project the expected revenue would be ZAR5 million. The average of the responses was R24 million. Looking at the 10-year horizon after completion of the projects, the predicted median revenue increases to ZAR40 million and the average to R224 million. The present average value of ZAR24 million of 5 years in the future, with an interest rate of 5%, is ZAR18.8 million. Taking into account that the average THRIP investment per project was ZAR1.5 million, the multiplier is 12.5, which is close to the value used according to the value transfer mentioned above of 9.76. It must be emphasised that these were grants made to support the strengthening of the research–innovation nexus and, therefore, to support a special category of projects.

To summarise, the direct impact of research activities at universities is ZAR12.6 billion. Using the multiplier developed for Australian universities, the spillover impact plus the direct impact is about ZAR123 billion.

### Impact on human capital

To measure the economic benefits that accrue from the production of graduates by higher education institutions, we utilised the Wilkins report.[32] This approach uses as 'treatment' group those individuals in possession of the qualification of interest (as their highest qualification), and the 'counterfactual' group consists of individuals with comparable personal and socio-economic characteristics but with the next lower (adjacent) level of qualification.

The central feature of this calculation is that the private return on education, as documented by PricewaterhouseCoopers and Wilkins[32,33], is that graduates can expect higher lifetime earnings as a result of their investment in their education, and the tax authorities will receive substantially more in taxes from those higher educated. We used a discount rate in order to transform future streams of benefits into current values.

Using the lifelong benefits approach[32] for bachelor's degrees transformed into South African rands with purchasing power parity[34,35], we obtained a value of ZAR1 450 000 with a 7% discount rate (and ZAR3 784 500 with a discount rate of 3.5%) per graduate. This approach considers that, as graduates go up the academic ladder, their income increases and their contribution to the fiscus via taxation and via direct spending increases.

In 2018 there were 100 740 students who graduated with a bachelor's degree or equivalent, producing an economic impact of ZAR146 billion using a 7% discount rate (ZAR381 billion for a discount rate of 3.5%) over a period of 30 years. The South African higher education system produces in excess of 200 000 graduates per annum, far in excess of the 100 740 considered here but the rest are qualifications which are not easily comparable to those used in the Bergstrom model. What this means is that this figure is a significant underestimate of the economic impact of the human resource outputs of universities.

Using the same approach for the 6801 master's and 1810 doctoral students who graduated during 2018 and who are South African, we estimated the economic impact of the master's graduates to be ZAR2.6 billion (ZAR6.8 billion for a discount rate of 3.5%) and that of the doctoral graduates to be ZAR0.5 billion (ZAR1.3 billion for a discount rate of 3.5%). Hence, the aggregate economic benefit is about ZAR149 billion (and ZAR387 billion for a discount rate of 3.5%).

The above estimates must be adjusted to consider the high mobility of the educated population and the potential for them to immigrate before their retirement. Emigration statistics in South Africa are not reliable. However, a number usually quoted is that 20 000 professionals at different stages of their careers leave the country every year. This is about 21% of the cohort of 2018 if it is assumed that the emigration patterns pertain to graduates with bachelor's degrees or equivalent. The impact of emigrations is catered for by reducing the aggregate economic benefit by 10.5%. With the 7% discount rate, this leaves the aggregate output as approximately ZAR134 billion (ZAR347 billion with a discount rate of 3.5%).

### Universities as economic entities

Traditionally, the estimated economic impact of universities has almost exclusively been based on the direct, indirect and induced impact of universities on their local, regional or national economies. These approaches consider a university as an economic entity creating output within the local economy by purchasing products and services from different industries/suppliers and hiring employees.

The sum of the direct, indirect and induced effects constitutes the *gross* economic impact on the local economy. Using data for 2018[23] we identified that universities had an expense cash flow of ZAR66 billion of which ZAR40 billion was compensation for employees. Based on this, the total direct impact associated with the universities' expenditures (in terms of monetary output) was estimated at ZAR66 billion. To determine the full economic impact of this expenditure, we used the value transfer method.[17] Multiplying the gross expenditure by 3.0 (as a multiplier) gives ZAR198 billion.

The above estimates must be adjusted in order to avoid double counting. Hence, from the direct impact we must reduce the total research income (ZAR12.67 billion) as this was included in the estimate of research impact. Similar care must be exercised with regard to university exports. Hence, we reduce the tuition-fee income generated from students who are not South African nationals and the on-campus non-tuition-fee income generated from these students; this is ZAR4.4 billion. Hence, the direct impact is closer to ZAR49 billion which produces a total impact of ZAR147 billion.

### Aggregate economic impact of universities

This is a first attempt at understanding the extent to which South Africa's public universities contribute to the national economy through four activities: the education of international students, the extent of performance and production of research, the education of students and the production of graduates, and the extent to which universities contribute as business entities. Care has been taken to avoid double-counting because of obvious overlaps between these activities.

The total impact is estimated at just more than ZAR510 billion. In 2018 the inputs into the sector were of the order of ZAR66 billion, and this produces a cost–benefit ratio of 1 : 7.7.

It must be noted that there are a number of limitations to this study. First is its dependence on modelling done for other economies, all of which are developed economies. This urges the undertaking of research that would provide the basis for the development of multipliers that apply to the South African economy. Second is the use of sometimes disjointed data sets. The development of the National Education Research Database being undertaken by DHET in partnership with the University of the Witwatersrand will help with this. Third is the absence of a counterfactual condition. What if the universities did not exist? Would South Africa have exported its students to other university systems and what would the cost–benefit analysis of that have been?

## Concluding remarks

We report the results of an effort to position the public higher education sector in the South African economy and to identify the cost–benefits of the sector. We identified that the contribution of the public Higher Education sector to the country's value added exceeds that of Wood and Wood Products; Textiles, Clothing and Leather Goods; and Paper and Paper Products. It is comparable with sectors such as Gold Mining and Beverages and Tobacco. Estimation of the cost–benefits of the sector identified a ratio of 1 to 7.7.

We argue that, in addition to its many non-tangible contributions, the university system is an important economic sector that contributes significantly to the national fiscus and should be seen as an area of investment rather than as an area of expenditure. Areas for further research have been identified. This study should be repeated regularly with the provision that research is engaged in to satisfy its indigenisation.

## Acknowledgements

This article reports on a Universities South Africa (USAf) project which is one of two projects focused on the economic impact, funding and sustainability of higher education. This particular project is funded by USAf. We thank the committee and staff of USAf who were involved in managing the project. We also thank Professor Roula Inglesi-Lotz, at the University of Pretoria, for valuable insights and suggestions on an earlier version of this articleand the anonymous external reviewers for their suggestions. We value the commitment of fellow scholars to the processes of peer review.

## Competing interests

We have no competing interests to declare.

## Authors' contributions

A.C.B.: Conceptualisation, data analysis, validation, secondary writer of initial draft, writing – revisions, project leadership, funding acquisition. A.P.: Conceptualisation, methodology, data collection, data analysis, validation, writing – the initial draft, secondary writer of revisions, project management, funding acquisition.

## References

1. UK Department for Business, Innovation & Skills (BIS). The benefits of higher education participation for individuals and society: Key findings and reports. London: BIS; 2013.
2. Oxford Economics. The economic impact of universities in 2014–2015. London: Oxford Economics; 2017.
3. Biggar Economics. Economic contribution of the LERU universities. Roslin: Biggar Economics; 2015.
4. KPMG Econtech. Economic modelling of improved funding and reform arrangements for Australian universities. Canberra: Universities Australia; 2009.
5. Beck R, Elliott D, Meisel J, Wagner M. Economic impact studies of regional public colleges and universities. Growth Change. 1995;26(2):245–260. https://doi.org/10.1111/j.1468-2257.1995.tb00170.x
6. Nature. Impact: The search for the science that matters. Nature Special. 2013 October 17. Available from: https://www.nature.com/collections/jjnybhlpcr
7. Siegfried JJ, Sanderson AR, McHenry P. The economic impact of colleges and universities. Econ Educ Rev. 2007;26(5):546–558. https://doi.org/10.1016/j.econedurev.2006.07.010
8. Pouris A, Inglesi-Lotz R. The contribution of higher education institutions to the South African economy. S Afr J Sci. 2014;110(Art. #a0059. https://doi.org/10.1590/sajs.2014/a0059
9. Dyason D, Rossouw R, Kleynhans EPJ. Economic impact assessment of a South African university campus: A case for promoting on-campus contact learning. S Afr J Econ Man Sci. 2019;22(1):1–11. http://dx.doi.org/10.4102/sajems.v22i1.3153
10. Coetzee H, Nell W. Measuring impact and contributions of South African universities in communities: The case of the North-West University. Dev South Afr. 2018;35(6):774–790. https://doi.org/10.1080/0376835X.2018.1475218
11. Orr D. Higher education strategic partnerships: The impact of Stellenbosch University's community interaction agreements on local development [PhD dissertation]. Stellenbosch: Stellenbosch University; 2016.
12. Hanushek EA. Will more higher education improve economic growth. Oxf Rev Econ Policy. 2016;32(4):538–552. https://doi.org/10.1093/oxrep/grw025
13. Psacharopoulos G, Patrinos H. Returns to investment in education: A further update. Educ Econ. 2004;12(2):111–134. https://doi.org/10.1080/0964529042000239140
14. Bloom DE, Canning D, Chan K, Luca DL. Higher education and economic growth in Africa. Int J Afr High Educ. 2014;1(1):23–57. https://doi.org/10.6017/ijahe.v1i1.5643
15. Tilak JBT. Higher education and development. Paper presented at: International Seminar, University XXI; 2003 November 25–27; Brasilia, Brazil. Available from: https://www.academia.edu/15336783/Higher_Education_and_Development
16. Bowen H. The individual and social value of American higher education. New York: Routledge; 2018. https://doi.org/10.4324/9781351309929-19
17. London Economics. The economic impact of group of eight universities. London: London Economics; 2018.
18. New Zealand Institute for Economic Research (NZIER). Economic impact of universities. Wellington: NZIER; 2016.
19. Van Heerden J, Bohlmann H, Giesecke J, Mackochekanwa A, Roos L. Higher education impact: Universities in the South African economy. Pretoria:Higher Education South Africa (HESA); 2007.
20. Harris T, Deller SC, Goetz, S. Linkages of the agricultural sector: Models and precautions. In: Van Allen KN, editor. Encyclopedia of agriculture and food systems. Cambridge, MA: Academic Press; 2014. p. 148–155. https://doi.org/10.1016/B978-0-444-52512-3.00124-8
21. Council on Higher Education (CHE). Vital statistics. Pretoria: CHE; 2021.
22. O'Farrell L. What is gross value added (GVA)? City-REDI Blog. 2019 May 15. Available from: https://blog.bham.ac.uk/cityredi/what-is-gross-value-added-gva/
23. Inglesi-Lotz R, Pouris A. Examining the economic impact of public universities in South Africa. Pretoria: Universities South Africa (USAf); 2019.
24. UNESCO Institute for Statistics. Expenditure on tertiary education (% of government expenditure on education) [data set on the Internet]. World Bank Data. 2018. Available from: https://data.worldbank.org/indicator/SE.XPD.TERT.ZS
25. South African Department of Science and Innovation (DSI). Draft science, technology and innovation decadal plan. Pretoria: DSI; 2022.
26. Southern African Development Community (SADC). Protocol on education & training 1997. Gaborone: SADC; 2017.
27. South African Department of Higher Education and Training (DHET). Statistics on post-school education and training in South Africa. Pretoria: DHET; 2018.
28. How much it costs to study in South Africa. BusinessTech. 2018 January 14. Available from: https://businesstech.co.za/news/lifestyle/219191/university-fees-2018-how-much-it-costs-to-study-in-south-africa/
29. Universities South Africa (USAf). Economic impact assessment of South African public universities. Pretoria: USAf; 2020.
30. National Advisory Council on Innovationo (NACI). South African science, technology and innovation indicators report. Pretoria: NACI; 2020.

31. Pouris A. Implementation and impact assessment of THRIP. Report to DPME. Pretoria: Business Enterprises at University of Pretoria; 2014.

32. Wilkins R. The household, income and labour dynamics in Australia Survey: Selected findings from waves 1 to 12. Victoria: Melbourne Institute of Applied Economic and Social Research; 2015. Available from: https://melbourneinstitute.unimelb.edu.au/_data/assets/pdf_file0006/2155506/hilda-statreport-2015.pdf

33. PricewaterhouseCoopers (PWC). The economic benefits of higher education qualifications. Pretoria: PWC; 2005.

34. Bergstrom JC, De Civita P. Status of benefit transfer in the United States and Canada: A review. Can J Agr Econ. 1999;47(1):79–87. https://doi.org/10.1111/j.1744-7976.1999.tb00218.x

35. Organisation for Economic Cooperation and Development (OECD). Purchasing power parities [webpage on the Internet]. No date [cited 2020 Jul]. Available from: https://data.oecd.org/conversion/purcasing-power-parities-ppp.htm

# Association of accelerometry-based and self-reported physical activity with cardiovascular risk in South African children

**AUTHORS:**
Patricia Arnaiz[1]
Felix Guntlisbergen[1]
Denis Infanger[1]
Markus Gerber[1]
Larissa Adams[2]
Danielle Dolley[2]
Nandi Joubert[1,3,4]
Madeleine Nienaber[2]
Siphesihle Nqweniso[2]
Rosa du Randt[2]
Peter Steinmann[3,4]
Jürg Utzinger[3,4]
Cheryl Walter[2]
Uwe Pühse[1]
Ivan Müller[1]

**AFFILIATIONS:**
[1]Department of Sport, Exercise and Health, University of Basel, Basel, Switzerland
[2]Department of Human Movement Science, Nelson Mandela University, Gqeberha, South Africa
[3]Swiss Tropical and Public Health Institute, Allschwil, Switzerland
[4]University of Basel, Basel, Switzerland

**CORRESPONDENCE TO:**
Patricia Arnaiz

**EMAIL:**
patricia.arnaiz@unibas.ch

**DATES:**
**Received:** 24 Jan. 2023
**Revised:** 02 July 2023
**Accepted:** 03 July 2023
**Published:** 28 Sep. 2023

**HOW TO CITE:**
Arnaiz P, Guntlisbergen F, Infanger D, Gerber M, Adams L, Dolley D, et al. Association of accelerometry-based and self-reported physical activity with cardiovascular risk in South African children. S Afr J Sci. 2023;119(9/10), Art. #15494. https://doi.org/10.17159/sajs.2023/15494

**ARTICLE INCLUDES:**
☒ Peer review
☐ Supplementary material

**DATA AVAILABILITY:**
☒ Open data set
☐ All data included
☐ On request from author(s)
☐ Not available
☐ Not applicable

**EDITOR:**
Pascal Bessong

**KEYWORDS:**
physical activity, accelerometry, self-report, cardiovascular health, children, South Africa

**FUNDING:**
Novartis Foundation, Swiss National Science Foundation (grant no. 192651)



The burden of non-communicable diseases is increasing, with risk factors emerging early in life. Physical activity reduces cardiovascular risk, but limited evidence exists for children from lower-income countries and mostly relies on self-reported methods that might be inaccurate and biased. We aimed to compare self-reported and accelerometer-measured physical activity in relation to cardiovascular risk markers in children from underserved communities in South Africa. We analysed cross-sectional data from 594 children aged 8 to 13. Physical activity was measured via accelerometry and the Physical Activity Questionnaire for Older Children (PAQ-C). Correlation analyses and linear regression models examined the relationship between accelerometer-measured and self-reported physical activity and their association with cardiovascular risk markers (body mass index, blood pressure, blood lipid profile and glycated haemoglobin). Results show a positive but weak correlation between PAQ-C scores and accelerometer-measured moderate-to-vigorous physical activity (MVPA). MVPA was inversely associated with body mass index, whilst sedentary behaviour correlated positively with lipid levels. PAQ-C scores were inversely associated with systolic blood pressure. The comparison of self-reported and accelerometer-measured physical activity in children from Gqeberha, South Africa, revealed inconsistencies in their correlation and association with cardiovascular risk markers. Accelerometry provided a more accurate cardiovascular risk estimation than PAQ-C, although associations were weak. Further, longitudinal studies should investigate the predictive power of both methodologies. These findings inform researchers and public health practitioners in the choice of method for physical activity appraisal beyond practical considerations, especially when combined with cardiovascular risk and in lower-income settings.

**Significance:**

We explore two widely used methods to assess physical activity levels in children. By comparing both methods, we expose inconsistencies in their correlation and association with cardiovascular risk markers. These data can guide researchers and public health practitioners in the use of one method beyond practical considerations. Whilst this work focuses on children from marginalised areas of South Africa, the issues explored are of relevance to other lower-income settings.

## Introduction

Cardiovascular diseases (CVD) are the leading cause of death worldwide,[1] and their burden is increasing in low- and middle-income countries (LMICs).[2] Meanwhile, physical activity (PA) has substantial positive effects on CVD mortality. In fact, beneficial health outcomes begin with very modest amounts of moderate-to-vigorous physical activity (MVPA).[3] Although the positive effects of PA on cardiovascular health have already been observed in childhood, the relationship between PA and cardiovascular risk markers (CRMs) has predominantly been studied in adults.[4] Furthermore, most studies examining PA behaviour have originated from high-income countries (HICs).

Behaviours associated with PA are complex constructs that differ according to socioeconomic status, region and cultural context.[5] Yet, PA health benefits observed in HICs have been broadly extrapolated to LMICs due to the scarcity of evidence from these regions. An umbrella systematic review found that 3.1% of studies concerned with promoting PA in children and adolescents were from LMICs.[6] Of those, only one originated from Africa. This observation was corroborated by Guthold et al., who reported that sub-Saharan Africa had the least available data on PA amongst adolescents worldwide.[7] Moreover, estimates of PA levels in LMICs are heterogeneous, partly due to an unstandardised use of different measurement methods.

With current physical inactivity estimates arising from self-reported methods, a device-based PA data gap exists, especially in LMICs.[7] A systematic review of PA trends in sub-Saharan Africa found that 72.2% of studies relied on self-report methods to assess PA.[8] Self-reporting provides a convenient way to assess activity patterns across large populations in a short time.[9,10] However, self-reports are prone to inaccuracy and bias originating from recall errors, the social desirability effect and difficulties understanding the questions.[11] Wearable devices such as accelerometry are seen as a more accurate alternative,[12] as they can quantify energy expenditure and estimate the intensity, duration and frequency of PA.[13] Nevertheless, accelerometry technology is costly and time consuming, especially on a large scale, whilst it also involves a range of subjective decisions.[14] As a consequence, the lack of consensus on PA assessment instruments and data management limits the comparability of studies.

Multiple studies have exposed differences and paucity of agreement between self-report methods and device-based assessments. For example, the Scottish Health Survey 2003 found that more than 75% of children reported meeting the recommended 60 minutes of moderate-to-vigorous physical activity (MVPA) per day,[15] whilst other studies from the United Kingdom using accelerometry reported a prevalence of less than 5%.[16-18] In South Africa, estimates for meeting PA recommendations range between 35.8%[8] and 77%[19] and vary both between and within instruments.[20-22] It is therefore not surprising that previous studies have shown a weak-to-moderate correlation between self-reported and device-based PA assessments.[13,23] Filling the PA data gap with more harmonised, device-based and larger sample studies is necessary to promote best PA practices in LMICs.[7]

Given the scarcity of data on PA behaviour from LMICs and the differences in the use and scope of self-report questionnaires versus accelerometry-based PA measurements, the aim of this study was twofold. First, we compared self-reported and accelerometer-measured PA amongst a large sample of school children from South Africa. Secondly, we examined their association with different CRMs. Based on the existing evidence, we hypothesised that self-reporting and accelerometry would vary in their PA estimates and association with CRMs.

## Materials and methods

### *Study design and setting*

This study was part of the KaziBantu project, a school-based intervention programme that attempts to promote sustainable lifestyle changes to achieve better health within disadvantaged communities in Gqeberha, South Africa. The *KaziBantu* programme was designed as a cluster-randomised controlled trial (RCT) and included eight schools that were randomly allocated to an intervention group (four schools) and a control group (four schools).[24] The study was structured such that after the completion of a baseline assessment in January 2019, children from the four intervention schools participated in the *KaziKidz* health promotion intervention for 32 weeks. The trial was registered at ISRCTN on 11 July 2018 under the registration number 18485542.

### *Participants*

Data from 981 children were initially collected during a baseline assessment in early 2019. The children were aged between 8 and 16 years and were attending grades 4–6. Only children between the ages of 8 and 13 were retained in the final study sample, as there was only one child at age 14 and 16, respectively. After excluding children due to lack of consent or dropping out ($n = 14$), reporting an impairment during data collection ($n = 238$), not answering all questions from the PAQ-C ($n = 39$), having no or invalid ActiGraph measurements ($n = 75$), missing information on height and weight ($n = 27$), or being outside the age range ($n = 2$), the final study sample consisted of 586 children (301 boys and 285 girls). The minimal sample size was originally calculated for the cluster RCT, as described in the study protocol.[24] A posteriori power analysis conducted with G*Power 3.1 Software (Heinrich Heine Universität Düsseldorf, Germany) revealed that the sample of 586 participants was sufficiently powered to demonstrate a weak correlation ($r = 0.102$) between self-reported and accelerometry-based PA (assuming an alpha error of 0.05 and a power of 0.80).[25]

## Data collection

### *Accelerometry-based assessment of physical activity*

The ActiGraph accelerometry device (ActiGraph wGT3X-BT, Pensacola, Florida, USA) was used to measure PA. Participants were directed to wear the device for seven consecutive days around the hip. They were allowed to remove the ActiGraph for activities that involved water contact (e.g. swimming or showering). Accelerometers were set up at a sampling rate of 30 Hz and ran on the latest firmware version (version 1.9.2). Analysis was carried out with the ActiLife software (version 6.13.4), using data set up at epochs of 10 s.

To be eligible for the data evaluation, the ActiGraph had to be worn for at least four valid weekdays and at least one weekend day. A day was considered valid if the ActiGraph had been worn for at least eight hours during that day.[26] Sleep time was removed and during waking time, non-wear periods, defined and identified based on the Troiano 2007 algorithm, were excluded from the analysis. The different PA intensities were categorised according to the cut-off points from Evenson 2008 for children.[27]

### *Self-reported physical activity*

A simplified version of the Physical Activity Questionnaire for Older Children (PAQ-C) was used to assess children's PA behaviour over the previous week.[28] Specifically, children ranked their personal PA level by answering questions 2–8, whereas question 10 inquired whether children were fit to perform PA. Later, a summary activity score between 1 and 5 (1 = lowest PA level, 5 = highest PA level) was calculated for questions 2–8. Children who reported sickness in question 10 or failed to answer one or more questions were excluded from the study sample ($n = 277$).

### *Blood pressure*

Resting BP was measured after the children were directed to be seated for 5 min. BP was measured three times with a pause of 1 min between each measurement. A calibrated Omron digital blood pressure monitor (Omron M6 AC model; Hoofddorp, The Netherlands) was used by nurses or biokineticists for the measurements. Only the second and third measurements were used to calculate an average for systolic (SBP) and diastolic BP (DBP). Elevated blood pressure in children was defined as above 120/80 mmHg or the 90th percentile according to sex, age and height reference values.[29]

### *Blood lipid profiles and glycated haemoglobin*

A point-of-care instrument (Alere Afinion AS 100 Analyzer, Abbott Technologies; Abbott Park, United States of America) was used to determine the blood lipid profiles (BLP) and glycated haemoglobin (HbA1c) concentrations. A healthcare worker first cleaned the fingertips with an alcohol swab and then pricked it with a safety lancet. Two drops of blood were carefully squeezed out of the finger, but only the second drop was collected for analysis. The device delivered the results within 8 min. All devices used were tested and calibrated before the procedure.[24]

The BLP included total cholesterol (TC), low-density lipoprotein (LDL), high-density lipoprotein (HDL), triglycerides (TG), non-high-density lipoprotein (non-HDL) and the ratio between TC and HDL (TC/HDL). Acceptable serum lipid and lipoprotein concentrations for children are <4.4 mmol/L for TC, <2.8 mmol/L for LDL, >1.2 mmol/L for HDL, <1 mmol/L for TG, <3.1 mmol/L for non-HDL and <3.7 for TC/HDL.[30] Levels of HbA1c revealed participants' average plasma glucose level over the past 8–12 weeks and were reported as a percentage of the total haemoglobin (%). Individuals with an HbA1c of 6.5% or higher are diagnosed with diabetes.[31]

### *Body mass index*

The body height of the children was measured by a stadiometer with an accuracy of 0.1 cm. Participants were instructed to keep their backs erect, and their shoulders relaxed. The body weight was measured by standing on a digital weighing scale (Tanita MC-580; Tanita, Tokyo, Japan) with an accuracy of 0.1 kg. For each participant, the BMI was calculated by dividing body weight (kg) by the square of body height ($m^2$). Based on sex and age references specified by the World Health Organization (WHO), BMI for age was calculated and subsequently classified as "Thin" if it was below −2 standard deviation (SD), "Normal weight" if between −2 and 1 SD, "Overweight" if above 1 SD, or "Obese" if above 2 SD.[32]

### *Statistical analyses*

Statistical analyses were conducted with the SPSS Statistics program (IBM SPSS Statistics for Mac, Version 27). Descriptive statistics were calculated as frequencies (%) for categorical variables and as medians (with interquartile range) for relatively symmetric, as well as skewed continuous variables. To identify differences between boys and girls, Mann–Whitney *U*-tests were conducted for continuous data, and Pearson's chi-squared tests were used for categorical data. Significance was set at $p \leq 0.05$ for all statistical analyses, and all tests were two-sided.

The relationship between PAQ-C scores and accelerometry-derived PA metrics was assessed using Spearman's rank correlation coefficients ($\rho$). Accelerometer wear time, age and sex were included as control variables. Cohen's correlation guidelines were used to evaluate the effect sizes as follows: $|\rho|=0.1$ small effect size, $|\rho|=0.3$ medium effect size, $|\rho|=0.5$ large effect size.[33] Linear regression models were used to analyse the associations between PA and sedentary behaviour (SB) with CRMs. All models were adjusted for the influence of height, weight, sex, age and accelerometer wear time except BMI which was controlled only for sex, age and wear time. Unstandardised regression coefficients were used to assess effect size and discuss clinical relevance.

### *Ethics approval and consent to participate*

The procedures of the *KaziBantu* study comply with the Declaration of Helsinki and have received ethical approval from the Nelson Mandela University Ethics Committee (reference #H18-HEA-HMS-001; dated 26 March 2018), Eastern Cape Department of Health (reference #EC_201804_007; dated 5 June 2018), and Eastern Cape Department of Education (dated 9 May 2018). The study was also cleared by the ethical review board of the Ethics Committee Northwest and Central Switzerland (reference #R-2018-00047; dated 1 March 2018). Each possible participant was informed about the study's objectives, procedures, risks and benefits. Participation in this study was voluntary and withdrawing was possible at any time with no further consequences. Oral approval (assent) had to be given by the participating children, whilst written informed consent was given by the corresponding parent or guardian.

## Results

### *Descriptive characteristics*

Characteristics of the study participants are presented in Table 1. With a median of 70.6 min of MVPA per day, 64.8% of the children achieved

**Table 1:** Descriptive characteristics of school-age children from Gqeberha, South Africa, in January 2019

| Variable | All ($N$ = 586) | Boys ($n$ = 301) | Girls ($n$ = 285) | $p$-value[a] | $\eta^2$ or Cramer's $V$[b] |
|---|---|---|---|---|---|
| Anthropometric | | | | | |
| Age | 10.0 (10.0–11.0) | 11.0 (10.0–11.0) | 10.0 (9.0–11.0) | **0.002** | 0.02 |
| Height (cm) | 139.5 (133.5–146.0) | 138.6 (132.9–144.3) | 140.8 (134.5–146.9) | **0.013** | 0.01 |
| Weight (kg) | 33.3 (28.9–39.6) | 32.1 (28.3–37.2) | 35.4 (30.0–41.9) | **<0.001** | 0.03 |
| BMI[c] (kg/m$^2$) | 17.1 (15.6–19.2) | 16.7 (15.4–18.2) | 17.8 (15.8–20.3) | **<0.001** | 0.03 |
| Blood pressure | | | | | |
| SBP[d] (mmHg) | 107.8 (99.0–116.3) | 106.5 (98.0–115.5) | 109.0 (99.5–117.0) | **0.032** | 0.01 |
| DBP[e] (mmHg) | 65.5 (60.5–71.5) | 64.5 (59.5–69.8) | 66.5 (61.5–72.5) | **0.019** | 0.01 |
| Blood lipid profile | | | | | |
| TG[f] (mmol/L) | 0.8 (0.6–1.0) | 0.7 (0.6–0.9) | 0.9 (0.7–1.2) | **<0.001** | 0.07 |
| TC[g] (mmol/L) | 3.7 (3.3–4.2) | 3.6 (3.2–4.1) | 3.8 (3.4–4.3) | **0.010** | 0.01 |
| LDL[h] (mmol/L) | 2.0 (1.7–2.4) | 1.9 (1.6–2.4) | 2.0 (1.7–2.4) | 0.063 | 0.01 |
| HDL[i] (mmol/L) | 1.3 (1.1–1.5) | 1.3 (1.1–1.5) | 1.3 (1.1–1.5) | 0.168 | <0.01 |
| Non-HDL[j] (mmol/L) | 2.4 (2.0–2.8) | 2.3 (2.0–2.8) | 2.5 (2.1–2.9) | **<0.001** | 0.03 |
| TC/HDL[k] (ratio) | 2.9 (2.5–3.3) | 2.8 (2.4–3.2) | 3.0 (2.7–3.5) | **<0.001** | 0.03 |
| Blood sugar | | | | | |
| HbA1c[l] (%) | 5.4 (5.3–5.6) | 5.4 (5.3–5.6) | 5.4 (5.3–5.6) | 0.383 | <0.01 |
| PAQ-C[m] | | | | | |
| Median score (1–5) | 2.9 (2.4–3.4) | 3.0 (2.4–3.6) | 2.7 (2.3–3.2) | **<0.001** | 0.02 |
| ActiGraph | | | | | |
| Sedentary time (%) | 64.2 (60.3–67.9) | 62.4 (58.7–66.5) | 65.6 (61.9–69.0) | **<0.001** | 0.06 |
| Light activity time (%) | 28.9 (26.0–31.6) | 29.2 (26.0–31.8) | 28.6 (26.0–31.1) | 0.154 | <0.01 |
| Moderate activity time (%) | 4.8 (3.8–6.1) | 5.5 (4.4–6.7) | 4.1 (3.3–5.1) | **<0.001** | 0.16 |
| Vigorous activity time (%) | 2.0 (1.3–2.9) | 2.5 (1.8–3.6) | 1.6 (1.1–2.1) | **<0.001** | 0.18 |
| MVPA[n] (min/day) | 70.6 (52.6–89.4) | 83.8 (64.4–104.1) | 57.7 (45.5–75.2) | **<0.001** | 0.21 |
| Wear time (min/day) | 1034.0 (1002.1–1053.4) | 1040.6 (1008.7–1059.6) | 1028.5 (997.2–1049.3) | **<0.001** | 0.03 |
| Meets physical activity guidelines[o] (Yes/No; %) | 64.8/35.2 | 81.7/18.3 | 47.0/53.0 | **<0.001** | 0.37 |

*Note: Values in bold indicate statistically significant results. Data are median (IQR) or percentage.*

*[a]Between-sex differences assessed by Mann-Whitney-U-Test or Pearson's chi-square test. [b]Effect size indicated by $\eta^2$ for continuous data and Cramer's V for categorical data. [c]Body mass index; [d]Systolic blood pressure, [e]Diastolic blood pressure, [f]Triglycerides, [g]Total cholesterol, [h]Low-density lipoprotein, [i]High-density lipoprotein, [j]Difference between TC and HDL, [k]Ratio between TC and HDL, [l]Glycated haemoglobin, [m]Physical Activity Questionnaire for Older Children, [n]Moderate- to vigorous-intensity physical activity, [o]More than or equal to (Yes), or less than (No) 60 minutes of MVPA per day.*

the recommended minimum of 60 min of MVPA per day. Overall, study participants spent 64.2% of their daily time in SB, 28.9% in light PA, 4.8% in moderate PA, and 2.0% in vigorous PA. The median wear time of the ActiGraph accelerometer was 17.2 h per day. The median PAQ-C score was 2.9, which lies slightly above half the possible value between 1 (lowest) and 5 (highest).

Significant differences were found between girls and boys. BMI, SBP, DBP, TG, TC, non-HDL, and TC/HDL were higher in girls than in boys (BMI: 17.8 vs. 16.7, $p < 0.001$; SBP: 109.0 mmHg vs. 106.5 mmHg, $p = 0.032$; DBP: 66.5 mmHg vs. 64.5 mmHg, $p = 0.019$; TG: 0.9 mmol/L vs. 0.7 mmol/L, $p < 0.001$; TC: 3.8 mmol/L vs. 3.6 mmol/L, $p = 0.010$; non-HDL: 2.5 mmol/L vs. 2.3 mmol/L, $p < 0.001$; TC/HDL: 3.0 vs. 2.8, $p < 0.001$). Boys were significantly more active than girls, as indicated by both the questionnaire (3.0 vs. 2.7, $p < 0.001$) and accelerometry (MVPA min/day 83.8 vs. 57.7, $p < 0.001$), and spent less time engaging in SB than girls (62.4% vs. 65.6%, $p < 0.001$, respectively). The WHO recommendations for PA were achieved by 81.7% of boys compared with only 47.0% of girls.

### *Correlation analysis between PAQ-C and wGT3X-BT ActiGraphy*

We found significant but weak associations between PAQ-C scores and PA metrics measured by the ActiGraph accelerometry (Table 2). The PAQ-C scores were positively associated with average MVPA minutes per day ($\rho$=0.10, $p = 0.015$). The PAQ-C scores also correlated with the percentage of time spent engaging in moderate- or vigorous-intensity activities ($\rho$=0.09, $p = 0.035$; $\rho$=0.10, $p = 0.013$, respectively). However, we found little evidence for an association between PAQ-C scores and the time spent in SB ($p = 0.72$) or light PA ($p = 0.31$). Association patterns are depicted in Figure 1.

### *Associations between physical activity and sedentary behaviour with cardiovascular risk markers*

Significant associations were observed between selected CRMs and PA metrics measured by PAQ-C and ActiGraph (Table 3). The ActiGraph MVPA was significantly and inversely associated with BMI (MVPA: beta=−0.031; CI = −0.043, −0.020; $p < 0.001$). The PAQ-C scores were inversely associated with SBP (beta=−1.563; CI = −2.926, −0.200; $p = 0.025$). Time spent in SB was positively associated with TC, LDL and non-HDL (TC: beta=0.001, CI = 0.000, 0.002, $p = 0.018$; LDL: beta=0.001, CI = 0.001, 0.002, $p = 0.002$; non-HDL: beta=0.001, CI = 0.000, 0.002, $p = 0.008$). Figure 2 provides a graphical representation of significant association patterns. We found little evidence for an association of DBP, HDL, TG, TC/HDL and HbA1c with self-reported PA levels or with accelerometer-measured MVPA and SB.

## Discussion

This study compared self-reported (PAQ-C) and accelerometry-based (ActiGraphy) PA and their association with selected CRMs in a population of school-aged children from underserved communities in South Africa. It was found that PAQ-C scores were weakly associated with ActiGraph-measured MVPA and that the two PA assessment methods were inconsistent in detecting relationships with CRMs.

PAQ-C scores correlated positively with ActiGraph-measured MVPA levels. Thus, children who rated themselves as being more physically active according to their PAQ-C scores achieved higher levels of MVPA as measured via accelerometry. However, the strength of the relationship between the two methods was weak. These findings agree with the results from other studies questioning the convergent validity of the PAQ-C.[34-36] A recent meta-analysis by Marasso and colleagues identified a moderate pooled correlation coefficient between the PAQ-C

**Table 2:** Spearman rank correlations between PAQ-C scores and ActiGraph measurements of physical activity ($N = 586$)

| Self-reported physical activity | Accelerometer-measured physical activity | | | | |
|---|---|---|---|---|---|
| | *Average MVPA[a] (min/day)* | *Sedentary time (%)* | *Light activity time (%)* | *Moderate activity time (%)* | *Vigorous activity time (%)* |
| PAQ-C[b] score (1–5) | **0.101 ($p = 0.015$)** | −0.015 ($p = 0.719$) | −0.042 ($p = 0.309$) | **0.087 ($p = 0.035$)** | **0.103 ($p = 0.013$)** |

*Note: Values in bold indicate statistically significant results; models are adjusted for sex, age and accelerometer wear time.*

*[a]Moderate- to vigorous-intensity physical activity, [b]Physical Activity Questionnaire for Older Children*

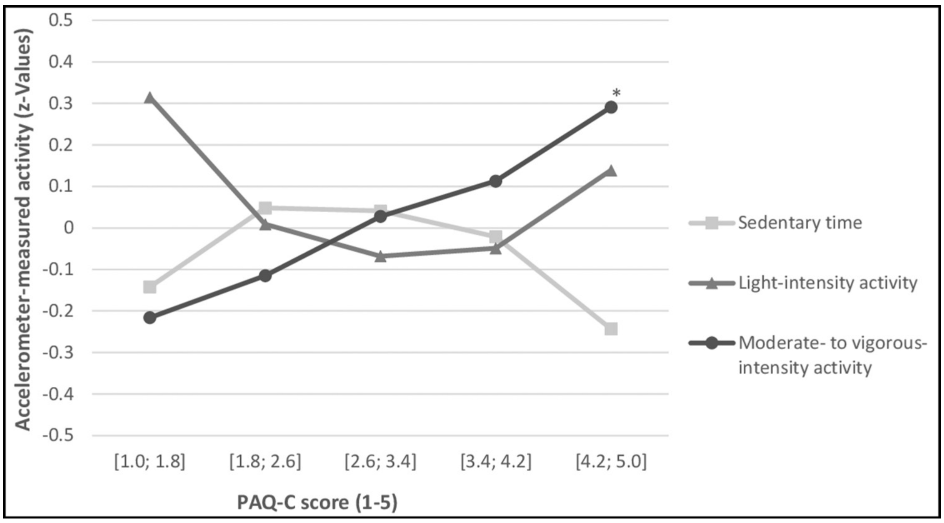



**Figure 1:** Correlation between accelerometry-based and self-reported physical activity. Moderate-to-vigorous physical activity showed a significant positive correlation with Physical Activity Questionnaire for Older Children (PAQ-C) scores (*$p = 0.015$) in school-aged children from Gqeberha, South Africa.

**Table 3:** Associations between cardiovascular risk markers with physical activity and sedentary behaviour

| Dependent variable | | PAQ-C[a] score (1–5) | | | ActiGraph MVPA[b] (min/day) | | | Sedentary time (min/day) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | N | *beta* | 95% CI | *p*-value | *beta* | 95% CI | *p*-value | *beta* | 95% CI | *p*-value |
| BMI[c,d] (kg/m²) | 586 | −0.383 | −0.773, 0.006 | 0.054 | **−0.031** | **−0.043, −0.020** | **<0.001** | <0.001 | −0.005, 0.005 | 0.989 |
| SBP[e] (mmHg) | 577 | **−1.563** | **−2.926, −0.200** | **0.025** | 0.012 | −0.031, 0.055 | 0.584 | −0.002 | −0.021, 0.017 | 0.814 |
| DBP[f] (mmHg) | 577 | −0.911 | −2.077, 0.255 | 0.125 | 0.006 | −0.030, 0.043 | 0.729 | −0.002 | −0.018, 0.015 | 0.852 |
| TC[g] (mmol/L) | 492 | 0.068 | −0.008, 0.143 | 0.079 | <0.001 | −0.002, 0.003 | 0.834 | **0.001** | **0.000, 0.002** | **0.018** |
| LDL[h] (mmol/L) | 486 | 0.033 | −0.029, 0.095 | 0.294 | <0.001 | −0.002, 0.002 | 0.894 | **0.001** | **0.001, 0.002** | **0.002** |
| HDL[i] (mmol/L) | 492 | 0.020 | −0.015, 0.054 | 0.260 | <0.001 | −0.001, 0.001 | 0.627 | <0.001 | 0.000, 0.001 | 0.817 |
| TG[j] (mmol/L) | 492 | 0.062 | −0.004, 0.129 | 0.067 | <0.001 | −0.002, 0.002 | 0.912 | −0.001 | −0.001, 0.000 | 0.254 |
| Non-HDL[k] (mmol/L) | 492 | 0.048 | −0.017, 0.112 | 0.146 | <0.001 | −0.002, 0.002 | 0.988 | **0.001** | **0.000, 0.002** | **0.008** |
| TC/HDL[l] (ratio) | 492 | 0.015 | −0.057, 0.087 | 0.685 | <0.001 | −0.002, 0.002 | 0.937 | 0.001 | 0.000, 0.002 | 0.155 |
| HbA1c[m] (%) | 500 | 0.008 | −0.019, 0.035 | 0.548 | <0.001 | −0.001, 0.001 | 0.569 | <0.001 | 0.000, 0.000 | 0.629 |

*Note: Values in bold indicate statistically significant results. Unstandardised coefficients (beta) from linear regression models were adjusted for sex, age, height, weight and accelerometer wear time. The number of children included in models (N) varies due to missing data for some cardiovascular risk markers.*

*[a]Physical Activity Questionnaire for Older Children; [b]Moderate-to-vigorous physical activity; [c]Body mass index; [d]Model adjusted only for sex, age and wear time; [e]Systolic blood pressure, [f]Diastolic blood pressure, [g]Total cholesterol, [h]Low-density lipoprotein, [i]High-density lipoprotein, [j]Triglycerides, [k]Difference between TC and HDL, [l]Ratio between TC and HDL, [m]Glycated haemoglobin*

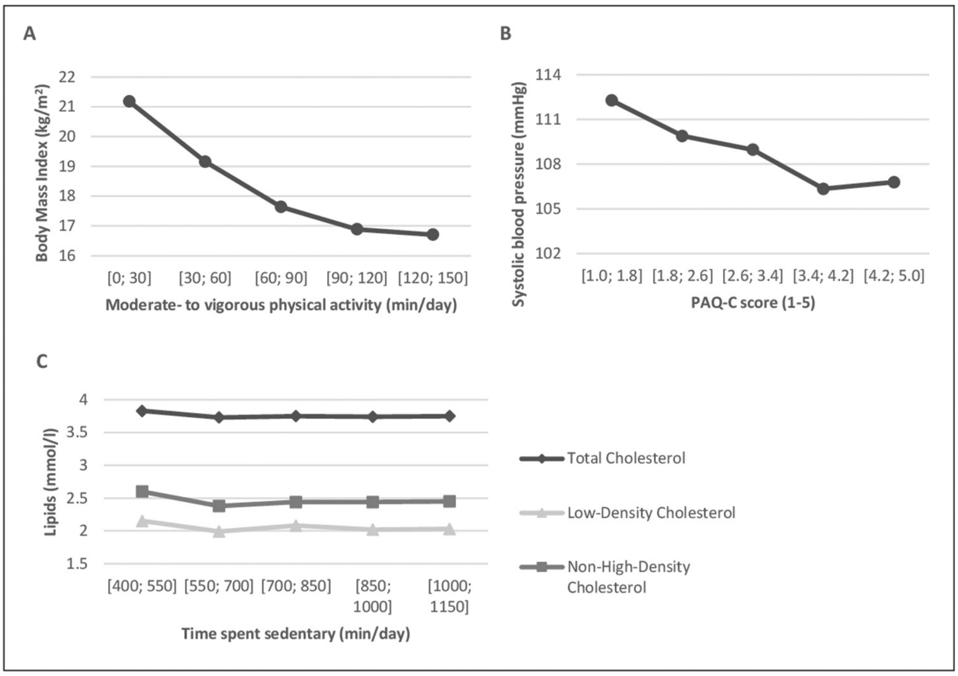



**Figure 2:** Significant associations between physical activity and cardiovascular risk markers. A significant inverse association between (A) accelerometer-measured MVPA and BMI and (B) PAQ-C scores and SBP. (C) A positive association between SB and lipids can be observed for school-aged children from Gqeberha, South Africa.

and accelerometry measurements, whilst also reporting wide correlation variability in the studies included. The low correlation coefficient and the weak explanatory power observed in this study contribute to the notion that PAQ-C and accelerometry have an inconsistent agreement for measuring MVPA. Furthermore, the PAQ-C would not fulfil the standards of a PA questionnaire proposed by Terwee et al., whereby correlation coefficients with accelerometer-measured MVPA must be at least moderate.[37] Because we only included valid accelerometry data in our analyses to ensure a representative mean wear time (17.23 out of a maximum of 18 h), we deem the presented accelerometry-based assessments as accurate. Therefore, we argue that attempting to derive children's objective MVPA levels from their PAQ-C scores, and vice versa, is linked to a high level of uncertainty.

Differences between self-reported and accelerometer-measured PA can be interpreted as children's difficulties in assessing their own PA habits. It has been observed that memory errors play an important role, particularly when dealing with children.[38] Another factor potentially contributing to discrepancies is the social desirability bias, where study participants report higher PA levels to be viewed favourably by others. Moreover, differences might stem from the ActiGraph not being able to accurately measure activities with little upper body movement such as cycling. During cycling, the upper body is not accelerated in any direction as happens when, for example playing soccer, a popular activity amongst children in South Africa.[39] Previous studies have also pointed out that the PAQ-C and accelerometry do not actually measure the same construct.[40] Accelerometry measures the exact duration, frequency and intensity of body movement, whilst the PAQ-C provides information about self-reported PA behaviour, activity types and settings in which PA is performed. Hence, the simultaneous use of PAQ-C and accelerometry could allow for a more comprehensive study of the relationship between self-perceived and actual PA and customised PA recommendations.

Regarding cardiometabolic risk, ActiGraph-measured MVPA was inversely related to BMI. Increasing MVPA by 15 min per day was associated with a BMI reduction of approximately −0.47 points. The significant association between accelerometer-measured MVPA and BMI demonstrates that PA, particularly MVPA, can be a crucial contributor to weight control at a young age.[41] Especially girls should be encouraged to increase their time spent in MVPA, as they spent substantially less time in MVPA and showed higher BMI values compared to boys (57.7 vs. 83.8 min/day and 17.8 vs. 16.7 kg/m$^2$, respectively). In contrast, children's self-reported PA levels were not significantly associated with BMI. Thus, children who considered themselves as more physically active did not exhibit lower BMI values. More consistent associations with BMI for device-based PA than for self-reported PA have previously been reported and were attributed to self-report bias.[42]

Diverging results were also obtained for the association between PA and SBP. In this study, SBP was weakly and inversely associated with self-reported PA but not with ActiGraph MVPA levels. Research has not shown that self-reported PA estimates are more likely to identify a significant association with BP compared with device-based PA measurements. In contrast, some studies have observed a significant inverse association between PA and BP using device-based methods,[43-45] although this relationship has not always been established.[46-48] The inconsistency of findings on the relationship between PA and BP amongst children and adolescents may be explained by varying durations of PA measurement, a lack of consistency in methodology, the influence of childhood adiposity, and wide ranges of age groups included in the studies.[49]

The activity assessment by accelerometry also provides estimates of the time children spend in SB, which allowed for the analysis of the association between CRMs and physical inactivity. Higher levels of SB were related to small increases in TC, LDL and non-HDL concentrations, whilst neither self-reported nor accelerometer-measured PA was associated with any BLP parameter. Increasing daily SB by 15 min was associated with a predicted increase of +0.015 mmol/L in TC, LDL and non-HDL. Because girls showed higher levels of SB compared to boys (65.6 vs. 62.4% per day, respectively) and higher concentrations in TC (3.8 vs. 3.6 mmol/L, respectively) and non-HDL (2.5 vs. 2.3 mmol/L, respectively), they should particularly be encouraged to reduce time spent in SB. It has been hypothesised that PA and SB have independent effects on lipoprotein metabolism, with PA more strongly affecting HDL and TG, and SB being rather related to the 'bad' cholesterol (LDL and non-HDL).[50] Consistent with the findings of this study, it has been claimed that PA does not lower LDL and TC.[51] However, the significant effect of MVPA on HDL and TG that has been previously observed was not found in this study.[52] Our results conform to those of the European Youth Heart Study that revealed non-significant relationships between MVPA and BLP.[53]

The results of this work must be interpreted considering the following limitations. First, this study has a cross-sectional design, which means that causality cannot be inferred because temporality is not known. Second, 8.0% of children with ActiGraph data were excluded from the analyses due to not meeting our wear time requirements. Third, PA was assessed over the period of 1 week, but since PA is highly variable in children, the measure comprised 1 week may not fully reflect children's true PA levels. Fourth, the limitations of the assessment methods must be considered. PAQ-C is prone to recall errors, social desirability effect and difficulties understanding the questions. Also, more than 25% of study participants reported that an impairment had prevented them from engaging in their usual PA and were therefore not included in the analysis. Limitations of the ActiGraph accelerometry include its inability to accurately measure activities with little upper body movement such as cycling or weight training when worn around the hip. Lastly, the impact of children's dietary habits was not accounted for in the associations between PA and SB with CRMs.

## Conclusion

We found a weak relationship between self-reported (PAQ-C) and accelerometer-measured (ActiGraph) MVPA levels in a paediatric population from low-income areas in South Africa. Thus, we advise caution when comparing studies that are based on diverse methodologies for assessing PA. As PA measured via accelerometry and SB showed stronger associations with CRMs, the wearable device ActiGraph allowed for a more accurate CVD risk estimation compared to the PAQ-C questionnaire. However, our results point towards a weak association between PA and CRMs. Therefore, it is of interest to further investigate the association between different PA measurement methods in longitudinal studies, especially under researched paediatric populations from LMICs.

## Acknowledgements

We thank all members of the *KaziBantu* team who contributed to the realisation of the project and who were involved in data collection in Gqeberha, South Africa. We also appreciate all learners of the *KaziBantu* schools in Gqeberha for their participation in the project and the teachers and principals for ensuring smooth collaboration. We thank the reviewers for their inputs and acknowledge the Novartis Foundation (Basel, Switzerland) and the Swiss National Science Foundation (Bern, Switzerland) for funding the research.

## Competing interests

We have no competing interest to declare.

## Authors' contributions

P.A.: Conceptualisation, writing – the initial draft, data compilation, data curation, interpretation of results, student supervision. F.G.: Conceptualisation, writing – the initial draft, data compilation, statistical analyses, interpretation of results. D.I.: Statistical input, writing – revisions. M.G.: Writing – revisions. L.A.: Data collection, writing – revisions. D.D.: Data collection, writing – revisions. N.J.: Data collection, writing – revisions. M.N.: Data collection, writing – revisions. S.N.: Data collection, writing – revisions. R.d.R.: Writing – revisions. P.S.: Writing – revisions. J.U.: Writing – revisions. C.W.: Writing – revisions. U.P.: Writing – revisions. I.M.: Conceptualisation, data collection, writing – revisions.

## References

1. WHO. The top 10 causes of death [webpage on the Internet]. c2020 [cited 2021 Jul 05]. Available from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

2. Roth GA, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: A systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018;392(10159):1736–1788.

3. Kraus WE, Powell KE, Haskell WL, Janz KF, Campbell WW, Jakicic JM, et al. Physical activity, all-cause and cardiovascular mortality, and cardiovascular disease. Med Sci Sports Exerc. 2019;51(6):1270–1281. https://doi.org/10.1249/MSS.0000000000001939

4. Ruiz JR, Ortega FB. Physical activity and cardiovascular disease risk factors in children and adolescents. Curr Cardiovasc Risk Rep. 2009;3(4):281–287. https://doi.org/10.1007/s12170-009-0043-6

5. Siefken K, Varela AR, Waqanivalu T, Schulenkorf N. Physical activity in low- and middle-income countries. 1st ed. London: Routledge; 2022. Available from: https://www.taylorfrancis.com/books/9780429344732

6. Barbosa Filho VC, Minatto G, Mota J, Silva KS, de Campos W, Lopes A da S. Promoting physical activity for children and adolescents in low- and middle-income countries: An umbrella systematic review. Prev Med. 2016;88:115–126. https://doi.org/10.1016/j.ypmed.2016.03.025

7. Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: A pooled analysis of 298 population-based surveys with 1·6 million participants. Lancet Child Adolesc Health. 2020;4(1):23–35. https://doi.org/10.1016/S2352-4642(19)30323-2

8. Muthuri SK, Onywera VO, Tremblay MS, Broyles ST, Chaput JP, Fogelholm M, et al. Relationships between parental education and overweight with childhood overweight and physical activity in 9–11 year old children: Results from a 12-country study. PLoS ONE. 2016;11(8), e0147746. https://doi.org/10.1371/journal.pone.0147746

9. Nigg CR, Fuchs R, Gerber M, Jekauc D, Koch T, Krell-Roesch J, et al. Assessing physical activity through questionnaires – A consensus of best practices and future directions. Psychol Sport Exerc. 2020;50:101715. https://doi.org/10.1016/j.psychsport.2020.101715

10. Benítez-Porres J. Cut-off values for classifying active children and adolescentes using the Physical Activity Questionnaire: PAQ-C and PAQ-A. Nutr Hosp. 2016;33(5). Available from: http://revista.nutricionhospitalaria.net/index.php/nh/article/view/564

11. Sallis JF. Self-report measures of children's physical activity. J Sch Health. 1991;61(5):215–219. https://doi.org/10.1111/j.1746-1561.1991.tb06017.x

12. Burchartz A, Anedda B, Auerswald T, Giurgiu M, Hill H, Ketelhut S, et al. Assessing physical behavior through accelerometry – State of the science, best practices and future directions. Psychol Sport Exerc. 2020;49:101703. https://doi.org/10.1016/j.psychsport.2020.101703

13. Marasso D, Lupo C, Collura S, Rainoldi A, Brustio PR. Subjective versus objective measure of physical activity: A systematic review and meta-analysis of the convergent validity of the physical activity questionnaire for children (PAQ-C). Int J Environ Res Public Health. 2021;18(7):3413. https://doi.org/10.3390/ijerph18073413

14. Prince SA, Adamo KB, Hamel M, Hardt J, Connor Gorber S, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. Int J Behav Nutr Phys Act. 2008;5(1):56. https://doi.org/10.1186/1479-5868-5-56

15. Scottish Executive. Scottish health survey 2003: Summary of key findings. Edinburgh: Scottish Executive The Stationery Office; 2005.

16. Ness AR, Leary SD, Mattocks C, Blair SN, Reilly JJ, Wells J, et al. Objectively measured physical activity and fat mass in a large cohort of children. PLoS Med. 2007 March 20;4(3), e97. https://doi.org/10.1371/journal.pmed.0040097

17. Reilly J, Jackson D, Montgomery C, Kelly L, Slater C, Grant S, et al. Total energy expenditure and physical activity in young Scottish children: Mixed longitudinal study. The Lancet. 2004;363(9404):211–212. https://doi.org/10.1016/S0140-6736(03)15331-7

18. Riddoch CJ, Mattocks C, Deere K, Saunders J, Kirkby J, Tilling K, et al. Objective measurement of levels and patterns of physical activity. Arch Dis Child. 2007;92(11):963–969. https://doi.org/10.1136/adc.2006.112136

19. Sedumedi CM, Janssen X, Reilly JJ, Kruger HS, Monyeki MA. Association between objectively determined physical activity levels and body composition in 6–8-year-old children from a black South African population: BC–IT study. Int J Environ Res Public Health. 2021;18(12):6453. https://doi.org/10.3390/ijerph18126453

20. Nqweniso S, Walter C, du Randt R, Aerts A, Adams L, Degen J, et al. Prevention of overweight and hypertension through cardiorespiratory fitness and extracurricular sport participation among South African schoolchildren. Sustainability. 2020;12(16):6581. https://doi.org/10.3390/su12166581

21. Salvini M, Gall S, Müller I, Walter C, du Randt R, Steinmann P, et al. Physical activity and health-related quality of life among schoolchildren from disadvantaged neighbourhoods in Port Elizabeth, South Africa. Qual Life Res. 2018;27(1):205–216. https://doi.org/10.1007/s11136-017-1707-1

22. Van Biljon A, McKune AJ, DuBose KD, Kolanisi U, Semple SJ. Physical activity levels in urban-based South African learners: A cross-sectional study of 7 348 participants. S Afr Med J. 2018;108(2):126. https://doi.org/10.7196/SAMJ.2017.v108i2.12766

23. Koolhaas C, Van Rooij F, Cepeda M, Tiemeier H, Franco O, Schoufour J. Physical activity derived from questionnaires and wrist-worn accelerometers: Comparability and the role of demographic, lifestyle, and health factors among a population-based sample of older adults. Clin Epidemiol. 2017;10:1–16. https://doi.org/10.2147/CLEP.S147613

24. Müller I, Smith D, Adams L, Aerts A, Damons BP, Degen J, et al. Effects of a school-based health intervention program in marginalized communities of Port Elizabeth, South Africa (the Kazibantu study): Protocol for a randomized controlled trial. JMIR Res Protoc. 2019;8(7), e14097. https://doi.org/10.2196/14097

25. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behav Res Methods. 2009;41(4):1149–1160. https://doi.org/10.3758/BRM.41.4.1149

26. Janz KF, Medema-Johnson HC, Letuchy EM, Burns TL, Gilmore JME, Torner JC, et al. Subjective and objective measures of physical activity in relationship to bone mineral content during late childhood: The Iowa Bone Development Study. Br J Sports Med. 2008;42(8):658–663. https://doi.org/10.1136/bjsm.2008.047779

27. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. J Sports Sci. 2008;26(14):1557–1565. https://doi.org/10.1080/02640410802334196

28. Kowalski KC, Crocker PRE, Donen RM. The physical activity questionnaire for older children (PAQ-C) and adolescents (PAQ-A) manual. Saskatoon: University of Saskatchewan; 2004. Available from: https://www.prismsports.org/UserFiles/file/PAQ_manual_ScoringandPDF.pdf

29. Flynn JT, Kaelber DC, Baker-Smith CM, Blowey D, Carroll AE, Daniels SR, et al. Clinical practice guideline for screening and management of high blood pressure in children and adolescents. Pediatrics. 2017;140(3), e20171904. https://doi.org/10.1542/peds.2017-3035

30. Orimadegun BE. Dyslipidaemia in African children and adolescents. In: Management of Dyslipidemia [Internet]. IntechOpen; 2021 [cited 2022 Jan 16]. Available from: https://doi.org/10.5772/intechopen.96804

31. The International Expert Committee. International expert committee report on the role of the A1C assay in the diagnosis of diabetes. Diabetes Care. 2009;32(7):1327–1334. https://doi.org/10.2337/dc09-9033

32. WHO. BMI-for-age (5-19 years) [webpage on the Internet]. c2007 [cited 2021 Nov 9]. Available from: https://www.who.int/tools/growth-reference-data-for-5to19-years/indicators/bmi-for-age

33. Cohen J. Statistical power analysis for the Behavioral Sciences. 2nd ed. New York: Routledge; 1988.

34. Benítez-Porres J, López-Fernández I, Raya JF, Álvarez Carnero S, Alvero-Cruz JR, Álvarez Carnero E. Reliability and validity of the PAQ-C questionnaire to assess physical activity in children. J Sch Health. 2016;86(9):677–685. https://doi.org/10.1111/josh.12418

35. Ben Jemaa H, Mankaï A, Mahjoub F, Kortobi B, Khlifi S, Draoui J, et al. Physical activity level assessed by accelerometer and PAQ-C in Tunisian children. Ann Nutr Metab. 2018;73(3):234–240. https://doi.org/10.1159/000492673

36. Chan CHS, Ha ASC, Ng JYY, Lubans DR. Associations between fundamental movement skill competence, physical activity and psycho-social determinants in Hong Kong Chinese children. J Sports Sci. 2019;37(2):229–236. https://doi.org/10.1080/02640414.2018.1490055

37. Terwee CB, Mokkink LB, Van Poppel MNM, Chinapaw MJM, Van Mechelen W, De Vet HCW. Qualitative attributes and measurement properties of physical activity questionnaires: A checklist. Sports Med. 2010;40(7):525–537. https://doi.org/10.2165/11531370-000000000-00000

38. Sirard JR, Pate RR. Physical activity assessment in children and adolescents. Sports Med. 2001;31(6):439–454. https://doi.org/10.2165/00007256-200131060-00004

39. Alegi P, Bolsmann C. South Africa and the global game: Football, apartheid and beyond. London: Routledge; 2010.

40. Ham SA, Reis JP, Strath SJ, Dubose KD, Ainsworth BE. Discrepancies between methods of identifying objectively determined physical activity. Med Sci Sports Exerc. 2007;39(1):52–58. https://doi.org/10.1249/01.mss.0000235886.17229.42

41. Janssen I, LeBlanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. Int J Behav Nutr Phys Act. 2010;7(40):16. https://doi.org/10.1186/1479-5868-7-40

42. Wanner M, Richard A, Martin B, Faeh D, Rohrmann S. Associations between self-reported and objectively measured physical activity, sedentary behavior and overweight/obesity in NHANES 2003–2006. Int J Obes. 2017;41(1):186–193. https://doi.org/10.1038/ijo.2016.168

43. Owen CG, Nightingale CM, Rudnicka AR, Sattar N, Cook DG, Ekelund U, et al. Physical activity, obesity and cardiometabolic risk factors in 9- to 10-year-old UK children of white European, South Asian and black African-Caribbean origin: The Child Heart And health Study in England (CHASE). Diabetologia. 2010;53(8):1620–1630. https://doi.org/10.1007/s00125-010-1781-1

44. Leary SD, Ness AR, Smith GD, Mattocks C, Deere K, Blair SN, et al. Physical activity and blood pressure in childhood: Findings from a population-based study. Hypertension. 2008;51(1):92–98. https://doi.org/10.1161/HYPERTENSIONAHA.107.099051

45. Andersen LB, Riddoch C, Kriemler S, Hills A. Physical activity and cardiovascular risk factors in children. Br J Sports Med. 2011;45(11):871–876. https://doi.org/10.1136/bjsports-2011-090333

46. Brage S, Wedderkopp N, Ekelund U, Franks PW, Wareham NJ, Andersen LB, et al. Features of the metabolic syndrome are associated with objectively measured physical activity and fitness in Danish children. Diabetes Care. 2004;27(9):2141–2148. https://doi.org/10.2337/diacare.27.9.2141

47. Craig SB, Bandini LG, Lichtenstein AH, Schaefer EJ, Dietz WH. The impact of physical activity on lipids, lipoproteins, and blood pressure in preadolescent girls. Pediatrics. 1996;98(3):389–395. https://doi.org/10.1542/peds.98.3.389

48. Klesges RC, Keith Haddock C, Eck LH. A multimethod approach to the measurement of childhood physical activity and its relationship to blood pressure and body weight. J Pediatr. 1990;116(6):888–893. https://doi.org/10.1016/S0022-3476(05)80645-6

49. Knowles G, Pallan M, Thomas GN, Ekelund U, Cheng KK, Barrett T, et al. Physical activity and blood pressure in primary school children: A longitudinal study. Hypertension. 2013;61(1):70–75. https://doi.org/10.1161/HYPERTENSIONAHA.112.201277

50. Celis-Morales CA, Perez-Bravo F, Ibañez L, Salas C, Bailey MES, Gill JMR. Objective vs. self-reported physical activity and sedentary time: effects of measurement method on relationships with risk biomarkers. PLoS ONE. 2012;7(5), e36345. https://doi.org/10.1371/journal.pone.0036345

51. Ahmed HM, Blaha MJ, Nasir K, Rivera JJ, Blumenthal RS. Effects of physical activity on cardiovascular disease. Am J Cardiol. 2012;109(2):288–295. https://doi.org/10.1016/j.amjcard.2011.08.042

52. Carroll S, Dudfield M. What is the relationship between exercise and metabolic abnormalities?: A review of the metabolic syndrome. Sports Med. 2004;34(6):371–418. https://doi.org/10.2165/00007256-200434060-00004

53. Ekelund U, Anderssen SA, Froberg K, Sardinha LB, Andersen LB, Brage S, et al. Independent associations of physical activity and cardiorespiratory fitness with metabolic risk factors in children: The European youth heart study. Diabetologia. 2007;50(9):1832–1840. https://doi.org/10.1007/s00125-007-0762-5

Check for updates

**AUTHORS:**
Nkechi V. Enwuru[1]
Solayide A. Adesida[2]
Christian A. Enwuru[3]
Udoma E. Mendie[1]

**AFFILIATIONS:**
[1]Department of Pharmaceutical Microbiology and Biotechnology, Faculty of Pharmacy, College of Medicine, University of Lagos, Lagos, Nigeria
[2]Department of Microbiology, Faculty of Science, University of Lagos, Lagos, Nigeria
[3]Centre for Infectious Disease Research, Microbiology Department, Nigerian Institute of Medical Research, Lagos, Nigeria

**CORRESPONDENCE TO:**
Solayide Adesida

**EMAIL:**
sadesida@unilag.edu.ng

**DATES:**
**Received:** 23 Apr. 2021
**Revised:** 05 July 2023
**Accepted:** 08 July 2023
**Published:** 28 Sep. 2023

**HOW TO CITE:**
Enwuru NV, Adesida SA, Enwuru CA, Mendie UE. Accessory gene regulators and virulence genes associated with the pathogenicity of *Staphylococcus aureus* from clinical and community settings in Lagos, Nigeria. S Afr J Sci. 2023;119(9/10), Art. #10901. https://doi.org/10.17159/sajs.2023/10901

**ARTICLE INCLUDES:**
☒ Peer review
☐ Supplementary material

**DATA AVAILABILITY:**
☐ Open data set
☒ All data included
☐ On request from author(s)
☐ Not available
☐ Not applicable

**EDITOR:**
Pascal Bessong

**KEYWORDS:**
accessory gene regulators, pathogenesis, *Staphylococcus aureus*, superantigens

**FUNDING:**
Tertiary Education Trust Fund (ETF/ES/AST, D/UNIV/LAGOS/VOL.2, 2010)



# Accessory gene regulators and virulence genes associated with the pathogenicity of *Staphylococcus aureus* from clinical and community settings in Lagos, Nigeria

*Staphylococcus aureus* is a prominent pathogen that causes serious community and hospital-acquired infections globally. Its pathogenicity is attributed to a variety of secreted and cell surface associated proteins that are modulated by the quorum-sensing accessory gene regulator (*agr*) system. In this study, we investigated the presence of toxin genes and *agr* involved with *S. aureus* from clinical samples and apparently healthy individuals. Unequivocal identification of the isolates was obtained with the Vitek 2 system. We screened 70 clinical (CL) and 22 community (C) *S. aureus* strains for the methicillin resistance (*mec*A) gene, *agr* and superantigens (SAg) (enterotoxins and toxic shock syndrome toxin-1) using PCR techniques. A total of 12 clinical isolates were classified as methicillin-resistant *S. aureus* (MRSA); 89 isolates belonged to one of the four *agr* groups ($agr^{1-4}$), and 3 isolates were non-typeable. Of the *agr* groups, $agr^{1}$ was the most prominent and mostly consisted of isolates from pus/wounds. The methicillin-susceptible *S. aureus* (MSSA) isolates were distributed within the four *agr* groups while MRSA strains were restricted to $agr^{1}$ and $agr^{3}$. The most common enterotoxin gene, *sei*, was likewise more prevalent in MSSA strains than in MRSA strains, where *sea* predominated. The co-existence of two or more enterotoxins was confirmed in 40% of the isolates. *sea* occurred through all the *agr* groups except $agr^{3}$ and *sei* was not found in $agr^{1}$ and $agr^{4}$. The toxic shock toxin (*tst*) gene was detected in six MSSA. These findings suggest that MSSA may cause more lethal infections than MRSA because of the increased frequency of toxic genotypes seen in MSSA strains.

**Significance:**

- Isolates in the $agr^{1,3}$ groups had more SAg toxin genes, whereas isolates in the $agr^{4}$ groups possessed more *tst* genes.
- The MSSA isolates contained higher proportions of virulence genes than MRSA.
- The clinical implications of this discovery include that MSSA may cause more lethal infections than MRSA due to the greater number of toxigenic genotypes discovered.

## Introduction

*Staphylococcus aureus* is a dynamic Gram-positive pathogen that lives as a harmless commensal bacterium on the skin and mucosal surfaces of humans and other animals.[1] It has the potential to multiply in the blood and other tissues, triggering serious medical conditions.[2] It is widely considered as one of the leading causes of hospital- and community-acquired infections globally.[3] Evidently, the organism features prominently in 8–33% of cases of skin, soft-tissue, and bloodstream infections that can result in significant morbidity and mortality.[4]

Depending on its growth phases, *S. aureus* is able to utilise a wide range of virulence factors to initiate and establish infections in susceptible hosts. Typically, in the lag and early exponential growth phrases, the pathogen releases cell-wall-associated factors that aid in tissue adhesion and immune system evasion.[5] When the bacterial population gets to the late exponential growth phase, it begins to secrete a wide range of exoproteins, including proteases, haemolysins, and superantigens (SAgs) while also down-regulating cell-wall-associated factors, resulting in biofilm dispersion and dissemination of infection.[5]

Staphylococcal superantigens (SAgs) are notable exotoxins which play a critical role in *S. aureus* infections. They have been categorised into staphylococcal enterotoxins (SE), staphylococcal enterotoxin-*like* (SE*l*) proteins and toxic-shock syndrome toxin[6] encoded by the *tst* gene. The adhesion and invasion phases of *S. aureus* development are characterised pre-eminently by its population-density-dependent behaviour. The synchronisation and swift transition between these two phases is achieved through a cell-to-cell communication mechanism known as quorum sensing (QS). The majority of these virulence genes are regulated by the accessory gene regulator (*agr*) system which is divided into four (1-4) *agr* groups.[7,8]

The *agr* region is crucial in pathogenesis and actively controls the expression of virulence factors, heterogeneous resistance in methicillin-resistant *S. aureus* (MRSA), and biofilm development.[9-11] It co-regulates the expression of several exoproteins including $\alpha$-, $\beta$-, $\gamma$-haemolysin as well as lipases, phenol-soluble modulins and TSST-1 while down regulating the synthesis of cell-wall-associated proteins such as protein A, coagulase, and fibronectin binding protein.[12,13] Several studies have been conducted to investigate the association of *agr* groups with certain biological traits in *S. aureus* and the outcomes have indicated that some enterotoxin clusters – namely *seg, sei, sem, sen*, and *seo* – are linked to $agr^{4}$ in a number of strains.[14]

In Tehran, Choopani et al.[15] noticed that MRSA isolates had the *seb* gene but that the TSST-1 precursor antigen was predominantly expressed by *agr*-III and IV strains. A study Some studies in southwest Nigeria indicated that 89% of the isolates screened had at least one SE: *seo* (34%) was the most prominent SE, then *seg* (30%) and *sea* (21%), whereas toxic shock syndrome toxin (TSST), *seb, sec, see, sej, sel, sem, ser* and *seu* were not detected. Similarly, the co-existence of *seo*/*seg* and *sei*/*seg* genes were recognised.[16]

Also, Akinduti et al.[17] reported that the bulk of the enterotoxins detected among clinical methicillin-sensitive *S. aureus* isolates in their study were confined to *agr*[2]. Nonetheless, information on the extent of superantigen genetic heterogeneity in community and clinical *S. aureus* populations in Nigeria is limited. Consequently, in the present study, we focused on determining the presence of staphylococcal superantigen genes in *S. aureus* isolates from two healthcare institutions and apparently healthy volunteers. Co-existence of the toxigenic genes with the *agr* groups and sources of the isolates was also assessed.

## Materials and methods

### *Bacterial isolates*

We analysed 92 previously described non-duplicated *S. aureus* strains collected over a period of 2 years from two tertiary care hospitals (Lagos University Teaching Hospital (LUTH), Idi-araba and National Orthopaedic Hospital, Igbobi).[18] The sample consisted of 22 nasal (NS) isolates obtained from apparently healthy individuals and designated as community (C) strains and 70 isolates obtained from various specimens submitted to the microbiological laboratories of the hospitals, classified as the clinical (CL) strains.

### *Bacterial identification and methicillin resistance determination*

The isolates were identified using Vitek 2 automated systems (BioMérieux, Marcy L'Étoile, France). Antibiotic susceptibility testing of the isolates, also carried out by the Vitek 2 automated system, had been determined previously.[18]

### *Genomic DNA extraction*

Genomic DNA for the evaluation of *mec*A, enterotoxin, toxic shock syndrome toxin-1 genes and *agr* determinants was extracted as previously described.[18] The quality and concentration of DNA were estimated spectrophotometrically.

### *mecA PCR*

Strains that demonstrated phenotypic resistance to methicillin were subjected to *mec*A PCR as previously described.[19] The PCR amplification was carried out in a 25 μL reaction volume containing 1 μL of template DNA, 10 μL of 2 x master mix of 1 x PCR buffer, 1.5 mmol/L $MgCl_2$, 0.15 mmol/L dNTP, and 1.25 IU Taq DNA polymerase, 0.7 μL of 0.8 μmol/L of each primer and 12.6 μL of deionised water. For the amplification of the 533 base pair (bp) fragment, the *mec*A-specific primer pairs employed were MECAP4: 5' - TCCAGATTACAACTTCACCAGG - 3', and MECAP7: 5' - CCACTTCATATCTTGTAACG - 3'. The PCR products were then separated by agarose gel electrophoresis and visualised with ethidium bromide staining. A 100 bp DNA ladder was used as a molecular weight marker.

### *Determination of agr groups*

We used the primers and thermal cycling conditions for *agr* groups differentiation described by Shopsin et al.[20] The *S. aureus* strains RN6390 (*agr* group I), RN6607 (*agr* group II), RN8465 (*agr* group III), and RN4850 (*agr* group IV), graciously provided by the Medical Microbiology Laboratory, Otto-von-Guericke University, Germany, were used as controls.

### *Detection of superantigen genes*

Multiplex PCR was used to determine enterotoxin genes associated with *S. aureus*.[21] We employed the primers for classical staphylococcal enterotoxin genes (*sea, seb, sec* and *sed*) as well as toxic shock syndrome toxin-1 (*tst*) and the SE*l* (*see, seg, seh, sei, sej*). For quality control, *S. aureus* ATCC 19095 was used for *sec, seh, seg* and *sei* while *S. aureus* FR I913 was employed for *sea, see, tst* and *S. aureus* ATCC 14458 for *seb* gene.

### *Statistical analyses*

The data were analysed using GraphPad Prism 7 software (San Diego, CA, USA) and Microsoft Excel. Data for frequencies in percentages and absolute values are shown in charts and tables.

### *Ethical approval*

The study was approved by the Institutional Review Board (IRB) of the College of Medicine, University of Lagos, Nigeria (reference number: CM/COM/8/VOL.XIX).

## Results

### Staphylococcus aureus *distribution and methicillin resistance*

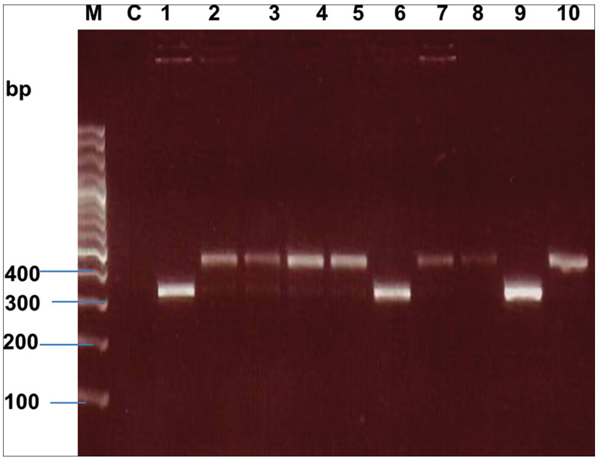
The *S. aureus* isolates investigated in this study originated mainly from three body sites: the genitourinary tract, nasal cavity and dermatological region of the body. Additionally, isolates were obtained from sputum (4.3%), blood (7.6%), and the ear or eye (6, 6.5%). Twelve clinical isolates were identified as MRSA and harboured the *mec*A gene (Table 1).

### *Classification of agr alleles*

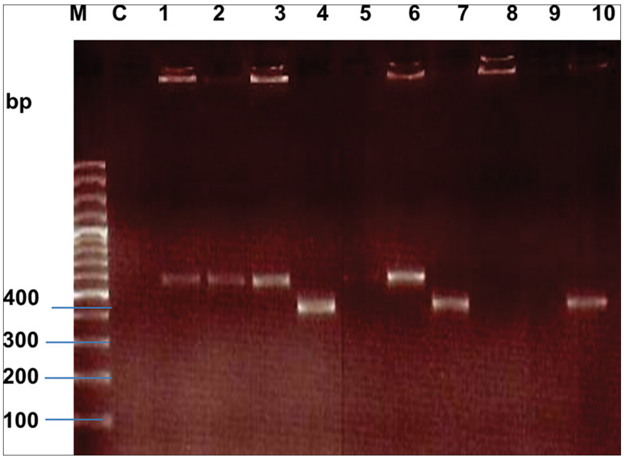
With the exception of three isolates that were non-typeable, the *S. aureus* isolates were classified into four groups using the *agr*-PCR technique (Figure 1). Of the MRSA strains, 67% were classified into *agr*[1] and 33% into *agr*[3]. Among the CL-MSSA strains, Of the CL-MSSA strains23 (42%) belonged to *agr*[1], 14 (25%) to *agr*[2], 11 (20%) to *agr*[3], and 7 (13%) to *agr*[4]. Also, 9 (41%) of the C-MSSA strains belonged to *agr*[1], followed by *agr*[2] (8; 36%), *ag*r[3] (4; 18%) and *agr*[4] (1; 5%) (see Figure 2).

**Table 1:** Distribution of *Staphylococcus aureus* determined as methicillin-resistant strains

| Isolate source | Number of *S. aureus* isolates (%) | Number of MSSA | Number of MRSA |
|---|---|---|---|
| **Genitourinary tract** | | | |
| High vaginal swabs | 3 | 2 | 1 |
| Seminal fluid | 1 | – | 1 |
| Urethral swabs | 2 | 1 | 1 |
| Catheter tips | 3 | 3 | 0 |
| Urine | 8 | 7 | 1 |
| | | | |
| **Nasal cavity** | 22 (23.9) | 22 | – |
| | | | |
| **Dermatological region** (pus or wound swabs) | 33 (35.7) | 27 | 6 |
| | | | |
| **Other** | | | |
| Blood | 7 (7.6) | 6 | 1 |
| Ears or eyes | 5 (5.5) | 5 | 0 |
| Sputum | 6 (6.5) | 5 | 1 |




**(a) Clinical methicillin-resistant *Staphylococcus aureus* (CL-MRSA).** Lane M: molecular weight marker (100 kb). Lane C: negative *S. aureus* control. Lanes 1, 6 and 9 are PCR amplicons from the primer set of *agr*$^{3}$ alleles using DNA from CL-MRSA strains. Lanes 2–5, 7, 8 and 10 are PCR amplicons from the primer set of *agr*$^{1}$ alleles using DNA from MRSA strains.




**(b) Clinical methicillin-susceptible *Staphylococcus aureus* (CL-MSSA).** Lane M: molecular weight marker (100 kb), Lane C: negative *S. aureus* control. Lanes 1–3, 5 and 6 are PCR amplicons from the primer set of *agr*$^{4}$ alleles using DNA from CL-MSSA strains and show *agr4* with 657 kb. Lanes 4, 7, and 10 are PCR amplicons from the primer set of *agr*$^{2}$ alleles using DNA with 572 kb.

**Figure 1:** Agarose gel electrophoresis of PCR products of *agr* alleles of *Staphylococcus aureus*.

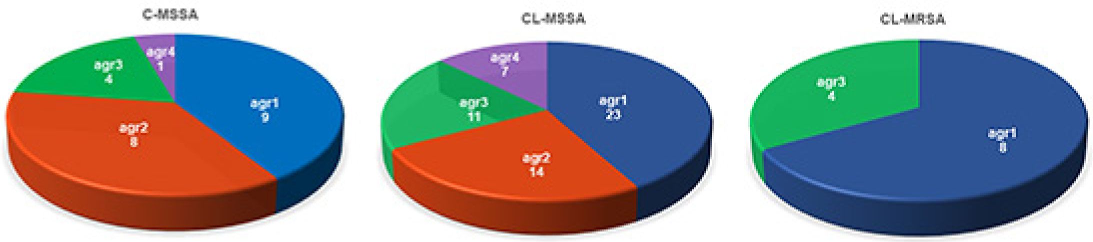



**Figure 2:** Prevalence of *agr* types among the *Staphylococcus aureus* strains.

### *Staphylococcal superantigen (SAg) gene profile among the isolates*

Eight (8) distinct SE genes (*sea, seb, sec, sed, seg, seh, sei* and *sej*) were recognised and none of the isolates was positive for *see* gene. Unlike the MSSA strains, only three types of SAg gene were detected among the MRSA strains, with *sea* accounting for the majority (10, 83%), followed by *seh* (6, 50%), and *sei* (2, 17%) (Figure 3). The 70 MSSA strains had *sei* (47, 67.1%) as the most prominent enterotoxin gene and *sed* was detected only in C-MSSA. CL-MSSA had 40 (60%) *sei*, 16 (22.9%) *sea*, 10 (14.3%) *seb*, 5 (7.1%) *sec*, 16 (22.9%) *seg* and no *sed, seh* nor *sej* were detected. In addition, 4 (5.7%) *sea, seb, seg*, 7 (10%) *sei*, 2 (2.9%) *sed* and 1 (1.4%) *sej* were identified amongst the C-MSSA strains. In all, 8 C-MSSA strains had no detectable SAg genes and the 6 *tst* genes discovered were peculiar to MSSA strains.

### *Enterotoxin genes co-existence and* agr *groups*

In all, 18 CL-MSSA strains had 2 or 3 SAg gene combinations, with the highest number (*sea-sei*) occurring in 7 strains. Other combinations of the enterotoxin genes were also detected among the community strains; 5 MRSA strains co-harboured *sea* and *seh* genes and *sea, seh* and *sei* co-existed in 1 MRSA strain. The distribution of the toxin genes within the *agr* groups is shown in Table 2.

The *sea* and *seh genes* were associated with *agr*$^{1}$ among the MRSA while *sea, seh* and *sei* were associated with *agr*$^{3}$. For the CL-MSSA, *sea, seb, sec, seg, seh* and *sei* genes were connected to *agr*$^{3}$, whereas *sea, seb, sec, seg, sei* and *tst* genes were related to *agr*$^{4}$. Regarding C-MSSA, the *sea, seb, sed* and *sei* genes were associated with *agr*$^{1}$, while *sea* and *sei* genes were found within *agr*$^{2}$. *seb, sed, seg, sei* and *tst* genes were in *agr*$^{3}$ and *sea, seg* and *tst* genes were associated with *agr*$^{4}$. Two *tst* genes were found in *agr*$^{1}$ and *agr*$^{4}$ groups in the CL-MSSA, and one each in the *agr*$^{3}$ and *agr*$^{4}$ groups among the C-MSSA.

### *Distribution of agr genes and SAg in relation to specimen types*

Whilst some specimens fitted into certain *agr* classes, all specimen types were represented in *agr*$^{1}$ group except the ear/eye samples (Table 3). The most prevalent genes (*sea* and *sei*) were predominantly detected in pus (35%) and wound (50%) isolates. Table 4 shows other observed frequencies.

## Discussion

In this study, the occurrence of methicillin resistance among the *S. aureus* isolates examined was 13%. Individuals harbouring MRSA, which is frequently multidrug resistant, are at risk of serious threat to themselves, and also play a role in the dissemination of the pathogen in hospital and community settings. Although the 29/307 MRSA recorded among hospitalised oncology patients examined in Ahvaz, Khuzestan Province, southwest of Iran[22] was much lower, it could not be ruled out that the higher rate observed in this study is unconnected to the antibiotic use pattern in Nigeria.[23] The isolates we examined were mostly from clinical sources and about 24% of the whole collection were from individuals who had no known indications for antibiotic use. The National Orthopaedic Hospital in Lagos has the capacity to accommodate up to 450 trauma

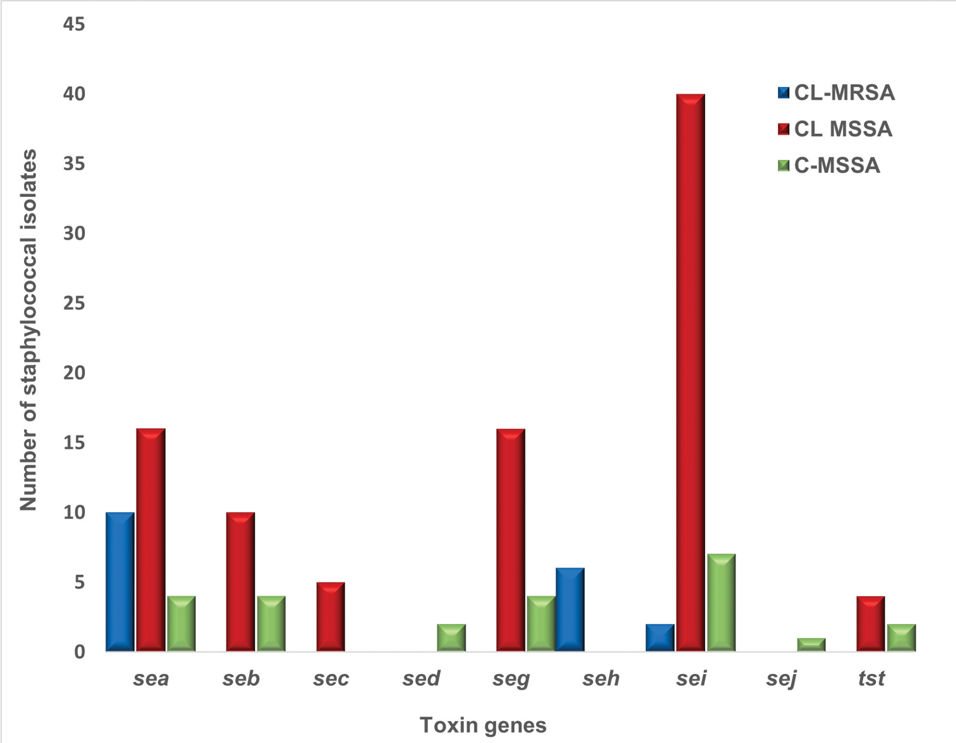



**Figure 3:** Superantigen gene profiles of the *Staphylococcus aureus* strains.

**Table 2:** Co-existence of staphylococcal superantigen genes in relation to *agr* types

| Toxin genes | Number of CL-MRSA (12) | | Number of CL-MSSA (55) | | | | Number of C-MSSA (22) | | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *agr*[1] (8, 67%) | *agr*[3] (4, 33%) | *agr*[1] (23, 42%) | *agr*[2] (14, 25%) | *agr*[3] (11, 20%) | *agr*[4] (7, 13%) | *agr*[1] (9, 41%) | *agr*[2] (8, 36%) | *agr*[3] (4, 18%) | *agr*[4] (1, 5%) | |
| *sea* | 8 | 2 | 6 | 4 | 3 | 2 | 1 | 2 | – | 1 | 29 (36%) |
| *seb* | – | – | 4 | 2 | 1 | 3 | 3 | – | 1 | – | 14 (15%) |
| *sec* | – | – | 1 | – | 3 | 1 | – | – | – | – | 5 (6%) |
| *sed* | – | – | – | – | – | – | 1 | – | 1 | – | 2 (2%) |
| *see* | – | – | – | – | – | – | – | – | – | – | 0 (0%) |
| *seg* | – | – | 3 | 2 | 4 | 4 | – | – | 3 | 1 | 17 (18%) |
| *seh* | 4 | 2 | – | – | 2 | – | – | – | – | – | 8 (9%) |
| *sei* | – | 2 | 9 | 8 | 11 | 6 | 2 | 2 | 3 | – | 43 (47%) |
| *sej* | – | – | – | – | – | – | – | – | – | 1 | 1 (1%) |
| *tst* | – | – | 2 | – | – | 2 | – | – | 1 | 1 | 6 (7%) |

**Table 3:** Distribution of *agr* groups of the *Staphylococcus aureus* isolates by specimen types

| *agr* groups | Number of *S. aureus* isolates from the samples | | | | | |
|---|---|---|---|---|---|---|
| | Pus/wound | Genitourinary tract | Blood | Sputum | Ears/eyes | Nasal swab |
| *agr*[1] | 14 | 11 | 2 | 4 | – | 9 |
| *agr*[2] | 7 | 2 | 3 | – | 2 | 8 |
| *agr*[3] | 7 | 6 | 1 | – | 1 | 4 |
| *agr*[4] | 4 | 2 | – | – | 1 | 1 |
| Total | 32 | 21 | 6 | 4 | 4 | 22 |

**Table 4:** Relative incidence of SAg genes in relation to the various types of specimen

| SAg genes | Type of specimen | | | | | |
|---|---|---|---|---|---|---|
| | Pus/wound ($n = 34$) | Genitourinary tract ($n = 19$) | Blood ($n = 7$) | Sputum ($n = 4$) | Ears/eyes ($n = 6$) | Nasal swab ($n = 22$) |
| *sea* | 12 | 9 | 3 | 1 | 1 | 4 |
| *seb* | 1 | 2 | 2 | – | 1 | 4 |
| *sec* | 5 | 2 | – | – | – | – |
| *sed* | – | – | – | – | – | 2 |
| *see* | – | – | – | – | – | – |
| *seg* | 7 | 6 | 2 | – | 1 | 2 |
| *seh* | 3 | 2 | 2 | – | – | – |
| *sei* | 17 | 13 | 4 | 2 | 3 | 7 |
| *sej* | – | – | – | – | – | 1 |
| *tst* | 2 | – | – | 1 | 1 | 2 |

*(–) denotes negative*

patients, while the Lagos University Teaching Hospital, a leading tertiary hospital with over 761 admission beds, is the facility of last resort and referral for all disease conditions. As tertiary healthcare institutions, both hospitals may have a high prescription culture for antibiotics. Furthermore, antibiotics are widely accessible and can be purchased without a prescription over the counter. This has conferred selective pressure on the majority of bacterial pathogens in our environment.

With the exception of three non-typeable *S. aureus* isolates, all isolates examined in this study were defined into the four *agr* groups. In *S. aureus*, the accessory gene regulator (agr) plays a vital role in the temporal expression of a wide range of bacterial virulence factors. A large proportion of the MRSA strains (58.6%) fitted into the *agr*[1] and *agr*[3] clusters. The CL-MSSA and NS-MSSA strains, on the other hand, belonged to *agr*[1-4]. These outcomes are in contrast to those of a study conducted in Poland[24] in which no *agr*[4] was detected, but are comparable to the findings of Elazhari et al.[21] Likewise, the majority (42%) of the isolates belonging to *agr* group 1 were MRSA, which was in agreement with the findings of other studies.[20,23]

In the present communication, the number of superantigen (SAg) genes detected varied significantly across the population of the *S. aureus* analysed, but was more pronounced among the methicillin-susceptible strains than their methicillin-resistant counterparts. The clinical relevance of this observation is that MSSA acquisition can be more deleterious. Elsewhere, Elazhari and others[21] indicated that 19 of the MSSA strains examined in their study harboured SAg genes varying from 1 to 11, which supports our assumption. From another perspective, Ayeni et al.[16] discovered neither *seb* or *sec* among their collection of *S. aureus* strains in southern Nigeria, which contradicts our findings. Additionally, our findings revealed a significant relationship between *sea, seh* and MRSA: 83% and 50% of MRSA strains had the *sea* and *seh* genes, respectively. In contrast, Ali et al.[14] discovered the presence of the *seb* gene in MRSA strains. This signifies that the enterotoxin gene profiles of *S. aureus* may vary substantially depending on geographical area and population structure.[25]

We also found the simultaneous presence of *sea* and *seh* genes in five MRSA strains, as well as the occurrence of two or more distinct enterotoxin genes in 18 CL-MSSA. Four toxin genes (*seb, seg, sei, tst*) co-existed in one C-MSSA strain. The *tst* gene was exclusively present in CL-MSSA and C-MSSA isolates and was not detected in MRSA. Previously, Danelli and colleagues[26] explained that *tst* carriage is commonly related with MSSA, which is consistent with our findings. Besides the possibility that MSSA isolates may have a lower genetic fitness burden because they do not have the SCC*mec* element[27], Varshney et al.[28] postulated that MSSA strains have an increased potential to secrete toxins than MRSA. *S. aureus* is known to have an outstanding array of virulence characteristics for initiating infections. This event may have major implications for public health.[29] It has also been insinuated that even modest quantities of staphylococcal enterotoxins could elicit T-cell activation[30], resulting in systemic infections including staphylococcal enterotoxin-induced shock and autoimmunity[31]. Similarly, Schmidt et al.[32] speculated that TSST-1 expression was independent of *S. aureus* methicillin sensitivity. This could further explain the carriage of *tst* genes among the MSSA analysed in this study, albeit at low prevalence. Therefore, it is possible that there may be additional, as yet unexplained, variables that affect *tst* carriage. Further studies are needed to better understand TSST-1 expression mechanisms in *S. aureus*.

In our analysis, the a*gr*[1] group was represented in all specimens investigated, except for ear/eye samples, demonstrating some similarity with the results of Javdan and co-workers[33]. However, its dominance was higher in pus/wound and genitourinary tract samples. This is in contrast to the findings of Elazhari et al.[21] These authors observed that *agr*[2] and *agr*[3] were prevalent in isolates from pus/wounds. This disparity between our findings and those of others may not be unconnected to sampling variability and study timeframe. While the *agr*[1]-MRSA had more enterotoxin genes, most of the toxigenic CL-MSSA and NS-MSSA belonged to *agr*[3]. The findings of this study also indicate that the expression of SAg genes was not unconnected to the types of specimens. The capacity to categorise infections based on sites or specimen types may give insight into the extent to which microorganisms play a role in disease initiation and progression. *sei*, the most prominent enterotoxin gene, was identified in all samples anaysed, but no *seb, sec, seg*, and *seh* were found among sputa isolates.

Meanwhile, the *sej* gene was only detected in sputum samples, while the *tst* genes were mainly present in pus/wound and nasal samples. Evidence from our study also shows that isolates that tested positive for *sei* and *sea* were significantly related with pus/wounds ($p<0.05$). This could alter the pathophysiology of wounds and present survival of the pathogen. Gergova et al.[34] noted that more virulence genes were found in invasive *S. aureus* compared to isolates from non-invasive sites (nasopharyngeal secretion, skin lesion, urogenital tract and eye secretion). Their result corroborates our findings. They also reported that several of the genes (*sei, sea*, and *seg*) were found in individuals who had died from staphylococcal bacteraemia, underlining the possible serious clinical consequences of superantigenic *S. aureus* strains.

## Conclusions

Our findings indicate that *agr*[1] expression seems to be important for colonisation and establishment of *S. aureus* infections. Although the superantigen gene content of the pathogen differs significantly between MRSA and MSSA, the presence of the genes in MRSA poses an increased public health risk. The preponderance of MSSA and the connection of superantigenic genes may result in an expansion of strains with higher pathogenicity, which could lead to therapeutic dead ends. Future research may be required to determine the relationships between MSSA and *tst* carriage, as well as whether these associations are restricted to certain *agr* groups.

## Acknowledgements

We are grateful to the staff of the Department of Microbiology, Otto-von-Guericke University, Germany, for technical support. We gratefully recognise Dr Beniam Ghebremedhin's invaluable professional support during the planning and development of the research. Financial aid was granted to N.V.E. by TETFUND (Tertiary Education Trust Fund, Grant No. ETF/ES/AST & D/UNIV/LAGOS/VOL.2, 2010).

## Competing interests

We have no competing interests to declare.

## Authors' contributions

N.V.E., S.A.A. and C.A.E.: Conceptualisation, methodology, data collection and analysis, writing – initial draft and proofreading. N.V.E., C.A.E., S.A.A., U.E.M.: Validation, project leadership, supervision. All authors read and approved the final manuscript.

## References

1. Karsten B, Michael ZD, Robert LS. *Staphylococcus, Micrococcus*, and other catalase-positive cocci. In: Manual of clinical microbiology. 12th ed. Washington DC: ASM Press; 2021. p. 11–13.
2. Jakub MK, Alexander RH. *Staphylococcus aureus* bloodstream infections: Pathogenesis and regulatory mechanisms. Curr Opin Microbiol. 2020;53:51–60. https://doi.org/10.1016/j.mib.2020.02.005
3. Mandeal SM, Ghosh AK, Pati BR. Dissemination of antibiotic resistance in methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant *S. aureus* strains isolated from hospital effluents. Am J Infect Control. 2015;43:87–88. https://doi.org/10.1016/j.ajic.2015.08.015
4. Thammavongsa V, Kim HK, Missiakos D, Schneewind O. Staphylococcal manipulation of host immune response. Nat Rev Microbiol. 2015;13(9):529–543. https://doi.org/10.1038/nrmicro3521
5. Wang B, Muir TW. Regulation of virulence in *Staphylococcus aureus*: Molecular mechanisms and remaining puzzles. Cell Chem Biol. 2016;23(2):214–224. https://doi.org/10.1016/j.chembiol.2016.01.004
6. Jhalka K, Tara CS, Dipendra T. *Staphylococcus aureus* and staphylococcal food-borne disease: An ongoing challenge in public health. BioMed Res Int. 2014;2014:1–9. https://doi.org/10.1155/2014/827965
7. Maleki DT, Ghalavand Z, Laabei M, Nikmanesh B, Houri H, Kodori M, et al. Molecular analysis of accessory gene regulator functionality and virulence genes in *Staphylococcus aureus* derived from paediatric wound infections. Infect Genet Evol. 2019;73:255–260. https://doi.org/10.1016/j.meegid.2019.05.013
8. Tan L, Huang Y, Shang W, Yang Y, Peng H, Hu Z, et al. Accessory gene regulator (*agr*) allelic variants in cognate *Staphylococcus aureus* strain display similar phenotypes. Front Microbiol. 2022;13:700894. https://doi.org/10.3389/fmicb.2022.700894
9. Jarraud S, Mougel C, Thioulouse J, Lina G, Meugnier H, Forey F, et al. Relationships between *Staphylococcus aureus* genetic background, virulence factors, *agr* groups (alleles), and human disease. Infect Immun. 2002;70:631–641. https://doi.org/10.1128/IAI.70.2.631-641.2002
10. Singh R, Ray P. Quorum sensing-mediated of staphylococcal virulence and antibiotic resistance. Future Microbiol. 2014;9:669–681. https://doi.org/10.2217/fmb.14.31
11. Tahmasebi H, Dehbashi S, Arabestani MR. Association between the accessory gene regulator (agr) locus and the presence of superantigen genes in clinical isolates of methicillin-resistant *Staphylococcus aureus*. BMC Res Notes. 2019;12:130. https://doi.org/10.1186/s13104-019-4166-7
12. Christian J, Alexander RH. Regulation of *Staphylococcus aureus* virulence. Microbiol Spectr. 2019;7(2):10. https://doi.org/10.1128/microbiolspec.GPP3-0031-2018
13. Gordon YCC, Justin SB, Michael O. Pathogenicity and virulence of *Staphylococcus aureus*. Virulence. 2021;12(1):547–569. https://doi.org/10.1080/21505594.2021.1878688
14. Yan X, Su-Yun Q, Kai-Hu Y, Fang D, Wen-Qi S, Chen S, et al. Clinical and molecular characteristics of *Staphylococcus aureus* isolated from Chinese children: Association among the *agr* groups and genotypes, virulence genes and disease types. World J Pediatr. 2021;17:180–188. https://doi.org/10.1007/s12519-021-00421-4
15. Choopani A, Heiat M, Amini E, Golpuch M, Aghamollaei H. The relationship between the presence of enterotoxin type b gene and antibiotic resistance in *Staphylococcus aureus*. J Appl Biotechnol Rep. 2015;2(1):203–206.
16. Ayeni FA, Ruppitsch W, Allerberger F. Molecular characterization of clonal lineage and staphylococcal toxin genes from *S. aureus* in Southern Nigeria. Peer J. 2018;6, e5204. https://doi.org/10.7717/peerj.5204
17. Akinduti AP, Osiyemi JA, Banjo TT, Ejilude O, El-Ashker M, Adeyemi AG, et al. Clonal diversity and spatial dissemination of multi-antibiotics resistant *Staphylococcus aureus* pathotypes in Southwest Nigeria. PLoS ONE. 2021;16(2), e0247013. https://doi.org/10.1371/journal.pone.0247013
18. Enwuru NV, Adesida SA, Enwuru CA, Ghebremedhin B, Mendie UE, Coker AO. Genetics of bi-component leukocidin and drug resistance in nasal and clinical *Staphylococcus aureus* in Lagos, Nigeria. Microb Pathog. 2018;115:1–7. https://doi.org/10.1016/j.micpath.2017.12.030
19. Ender M, Berger-Bächi B, McCallumn N. Variability in SCC*mec*N1 spreading among injection drug users in Zurich, Switzerland. BMC Microbiol. 2007;7:62. https://doi.org/10.1186/1471-2180-7-62
20. Shopsin B, Mathema B, Alcabes P, Said-Salim B, Lina G, Matsuka A, et al. Prevalence of *agr* specificity groups among *Staphylococcus aureus* strains colonizing children and their guardians. J Clin Microbiol. 2003;41(1):456–459. https://doi.org/10.1128/JCM.41.1.456-459.2003
21. Elazhari M, Elhabchi D, Zerouali K, Dersi N, Elmalki A, Hassar M, et al. Prevalence and distribution of superantigen toxin genes in clinical community isolates of *Staphylococcus aureus*. J Bacteriol Parasitol. 2011;2:107. https://doi.org/10.4172/2155-9597.1000107
22. Montazeri EA, Khosravi AD, Khazaei S, Sabbagh A. Prevalence of methicillin resistance and superantigenic toxins in *Staphylococcus aureus* strains isolated from patients with cancer. BMC Microbiol. 2021;21:262. https://doi.org/10.1186/s12866-021-02319-7
23. Chukwu EE, Oladele DA, Enwuru CA, Gogwan PL, Abuh D, Audu RA, et al. Antimicrobial resistance awareness and antibiotic prescribing behavior among healthcare workers in Nigeria: A national survey. BMC Infect Dis. 2021;21(1):22. https://doi.org/10.1186/s12879-020-05689-x
24. Grazul M, Balcerczak E, Sienkiewicz M. Analysis of the presence of the virulence and regulation genes from *Staphylococcus aureus* (*S. aureus*) in coagulase negative staphylococci and the influence of the staphylococcal cross-talk on their functions. Int J Environ Res Public Health. 2023;20(6):5155. https://doi.org/10.3390/ijerph20065155
25. Kolawole DO, Adeyanju A, Schaumburg F, Akinyoola AL, Lawal OO, Amusa YB, et al. Characterization of colonizing *Staphylococcus aureus* isolated from surgical wards' patients in a Nigerian university hospital. PLoS ONE. 2013;8(7), e68721. https://doi.org/10.1371/journal.pone.0068721
26. Danelli T, Duarte FC, de Oliveira TA, da Silva RS, Alfieri DF, Gonçalves GB, et al. Nasal carriage by *Staphylococcus aureus* among healthcare workers and students attending a university hospital in southern Brazil: prevalence, phenotypic, and molecular characteristics. Interdiscip Perspect Infect Dis. 2020;2020:3808036. https://doi.org/10.1155/2020/3808036
27. Zhao H, Xu S, Yang H, He C, Xu X, Hu F, et al. Molecular typing and variations in amount of tst gene expression of TSST-1-producing clinical *Staphylococcus aureus* isolates. Front Microbiol. 2019;10(1388). https://doi.org/10.3389/fmicb.2019.01388
28. Varshney AK, Martinez LR, Hamilton SM, Bryant AE, Levi MH, Gialanella P, et al. Augmented production of Panton-Valentine leukocidin toxin in methicillin-resistant and methicillin-susceptible *Staphylococcus aureus* is associated with worse outcome in a murine skin infection model. J Infect Dis. 2010;201:92–96. https://doi.org/10.1086/648613

29. Cheung GYC, Bae JS, Otto M. Pathogenicity and virulence of *Staphylococcus aureus*. Virulence. 2021;12(1):547–569. https://doi.org/10.1080/21505594.2021.1878688

30. Pinchuk IV, Beswick EJ, Reyes VE. Staphylococcal enterotoxins. Toxins. 2010;2(8):2177–2197. https://doi.org/10.3390/toxins2082177

31. Principato M, Qian BF. Staphylococcal enterotoxins in the etiopathogenesis of mucosal autoimmunity within the gastrointestinal tract. Toxins. 2014;6(5):1471–1489. https://doi.org/10.3390/toxins6051471

32. Schmidt KA, Manna AC, Gill S, Cheung AL. SarT, a repressor of alpha-hemolysin in *Staphylococcus aureus*. Infect Immun. 2001;69:4749–4758. https://doi.org/10.1128/IAI.69.8.4749-4758.2001

33. Javdan S, Narimani T, Abadi MSS, Gholipour A. *Agr* typing of *Staphylococcus aureus* species isolated from clinical samples in training hospitals of Isfahan and Shahrekord. BMC Res Notes. 2019;12:363. https://doi.org/10.1186/s13104-019-4396-8

34. Gergova RT, Tsitou VS, Gergova II, Muhtarova AA, Mitov IG. Correlation of methicillin resistance and virulence genes of *Staphylococcus aureus* with infection types and mode of acquisition in Sofia, Bulgaria. Afr J Clin Exper Microbiol. 2019;20(4):280–288. https://doi.org/10.4314/ajcem.v20i4.3

Check for updates

**AUTHORS:**
Lindumusa Myeni[1]
Mokhele E. Moeletsi[1,2]

**AFFILIATIONS:**
[1]Agricultural Research Council – Natural Resources and Engineering, Pretoria, South Africa
[2]Risks and Vulnerability Assessment Centre, University of Limpopo, Sovenga, South Africa

**CORRESPONDENCE TO:**
Lindumusa Myeni

**EMAIL:**
MyeniL@arc.agric.za

**DATES:**
**Received:** 18 Oct. 2022
**Revised:** 08 June 2023
**Accepted:** 09 June 2023
**Published:** 28 Sep. 2023

**HOW TO CITE:**
Myeni L, Moeletsi ME. Assessing the adoption of improved seeds as a coping strategy to climate variability under smallholder farming conditions in South Africa. S Afr J Sci. 2023;119(9/10), Art. #15001. https://doi.org/10.17159/sajs.2023/15001

**ARTICLE INCLUDES:**
☒ Peer review
☒ Supplementary material

**DATA AVAILABILITY:**
☐ Open data set
☐ All data included
☒ On request from author(s)
☐ Not available
☐ Not applicable

**EDITOR:**
Teresa Coutinho

**KEYWORDS:**
climate-related risks, climate-resilient seeds, climate-smart agriculture, farmers' seed systems, maize seed varieties

**FUNDING:**
Horizon 2020 Framework Programme (727201)



# Assessing the adoption of improved seeds as a coping strategy to climate variability under smallholder farming conditions in South Africa

This study was undertaken to examine the adoption rate, constraints and factors determining the uptake of improved seed varieties in smallholder farming conditions of South Africa, using a maize crop. Primary data were collected from 279 smallholder farmers in the Maluti-a-Phofung municipality of the Free State Province using a household questionnaire and were validated through focus group discussions with key local informants. Descriptive statistics, frequency analysis and a binary logistic model were used for data analysis. Results show that the majority of the farmers (>69%) were growing uncertified local seed varieties saved from their previous harvest or acquired from neighbouring farmers. The lack of awareness and knowledge of, inadequate information, unaffordability and unavailability of improved seeds, low income and landlessness were major constraints limiting the adoption rate. Age and income were the main variables that had a positive and significant effect on the adoption of improved seed varieties, whilst ownership of livestock had a significant and negative influence on their adoption. As a result, the promotion of improved seed varieties needs to be supported by conducive and effectual institutional policies that will improve the provision of extension services, capacity-building initiatives, subsidy programmes, financial resources and good agricultural lands, and reduce the transaction costs of improved seed varieties. The recommendations from this study can be used by the government and other developmental organisations to enhance the wider uptake and use of improved seed varieties in smallholder farming conditions not only in the study area but also in other regions with similar challenges.

**Significance:**

- A lack of awareness and knowledge of, inadequate information, unaffordability and unavailability of improved seeds, low income and landlessness are key constraints limiting adoption.
- We recommend the provision of knowledge and awareness-raising of improved seed varieties through enhanced extension services and capacity-building initiatives as key strategies to improve the adoption rate.
- The recommendations of this study can be used to enhance the wider uptake and use of improved seeds and assist smallholder farmers to cope effectively with climate variability and change.

## Introduction

Agriculture, particularly rainfed crop production, is extremely susceptible to the impacts of climate change.[1] The changing rainfall patterns (i.e. rainfall onsets and amounts), elevated air temperatures and increase in the incidents of weather-related risks such as droughts, floods, frost, and crop diseases and pests attributed to climate change are a serious threat to agricultural production and food security.[1-3] These weather-related risks are expected to increase in both frequency and magnitude and may push 122 million more people into extreme poverty by the year 2030.[4] African smallholder farmers are amongst the most vulnerable as a result of their high dependence on rainfed crop production and they have limited adaptation capacity due to the lack of knowledge or skills, limited resources and financial constraints.[1,5,6] There is an urgent need for the development and implementation of effectual adaptation and mitigation strategies to combat the negative effects of climate variability on crop production and contribute towards the achievement of the Sustainable Development Goals.[7,8]

Climate-smart agriculture (CSA) is a viable approach to sustainably increase productivity and improve the climate resilience of smallholder farmers whilst at the same time reducing agriculture's contribution to environmental contamination and climate change.[9,10] The use of improved seeds, particularly climate-resilient varieties, is one of the most important and effective CSA strategies to improve crop production and enhance climate resilience in smallholder farming conditions.[7,11,12] Some of these improved seed varieties possess desirable traits such as drought and heat tolerance, early maturing, high yielding, pest and disease resistance, better milling or cooking quality, better nutritional content and better nitrogen-use efficiency.[13-16] Depending on local context, exposure and vulnerability, these traits can help smallholder farmers in the mitigation of abiotic and biotic stresses that are prevailing in the specific region.[1,17,18]

The planting of locally adapted seed varieties of good quality at the right time whilst using sustainable agricultural practices and the recommended application rates of other inputs provides healthy crop establishment and subsequently good crop yields.[13,19-23] Furthermore, the use of high-quality seeds also ensures the efficient use of other inputs such as labour, fertilisers and pesticides in farming systems.[24] Several other crop management strategies to cope with climate change, such as changing seed densities or seed rates and changing crop varieties and association, all depend on the stable availability and affordability of good-quality seeds. Furthermore, research has shown that improved seed varieties performed significantly better than local seed varieties over a wide range of agro-climatic conditions, under both optimum and stress growing conditions in Africa.[18,23,25-27] Moreover, empirical studies in different African countries have shown that the adoption of improved seed varieties increases yields, yield stability, food security and profitability and reduces poverty in smallholder farming conditions.[12,14,18,28]

Despite substantial efforts in research on the development and promotion of improved seed varieties as part of the Green Revolution initiative[29], the wider uptake and use of improved seed varieties by smallholder farmers is still low and unsatisfactory in Africa[16,17,26,30]. Thus, most African smallholder farmers plant local or indigenous varieties (landraces) that produce lower yields and are often most vulnerable to abiotic and biotic stresses, regardless of numerous improved seed varieties that are commercially available.[14,31] Consequently, crop yields of smallholder farmers remain very low and stagnant despite the proliferation of input subsidy programmes of the national governments and developmental agencies in Africa.[17,27] For example, the Agricultural Research Council (South Africa) has been involved in projects such as the Water Efficient Maize for Africa (WEMA) and Improved Maize for African Soils (IMAS) aimed to develop, test and disseminate improved maize varieties to improve yields and food security under the threat of climate variability in smallholder farming conditions.[32] However, the low adoption of improved seeds by smallholder farmers is a serious concern of the government and other development organisations as climate change is already and anticipated to continue exposing the most vulnerable communities to food and nutrient insecurities, poverty, economic crisis, environmental and ecosystem degradation and health issues.[33,34]

The understanding of the main challenges limiting the wider uptake and use of improved seed varieties by smallholder farmers is critical for the development and implementation of effectual policies, programmes and initiatives aimed at improving crop production, food security and climate resilience in smallholder farming conditions with greater potential in climate mitigation.[5,8,35,36] Previous studies have indicated that restricted availability of improved seed varieties, high prices of seeds, risk aversion, preference for local varieties, low income, lack of knowledge and credits are amongst the major constraints limiting smallholder farmers from adopting improved seed varieties.[1,13,17,19,37] This previous research showed that the key barriers and determinants of adoption of improved seed varieties are local-specific due to the differences in local seed systems, institutions, policies, political and cultural ideologies, availability of resources and infrastructure.[1,35]

Few studies have investigated the factors influencing the adoption of improved seed varieties in South Africa.[21,28,36,38] Thus, the understanding of key barriers and determinants for the adoption of improved seed varieties is limited in South Africa, particularly in smallholder farming conditions. Given that South Africa has one of the most established and competitive seed sectors in Africa[39,40], the understanding of barriers and determinants for the adoption of improved seed varieties could give new and deep insights into factors to be considered when promoting improved seed varieties in smallholder farming conditions. Therefore, the present study aims to identify the key barriers and determinants of the adoption of improved seed varieties in the Maluti-a-Phofung municipality of the Free State Province of South Africa. The present study only focuses on the maize (*Zea mays* L.) crop, one of the main staple foods in the southern African region.[31,41]

In the present study, firstly we determined the adoption rate of improved maize seed varieties; secondly, we identified the constraints associated with the adoption of improved maize seed varieties; and thirdly, we investigated the factors influencing the adoption of improved maize seed varieties. Lastly, we recommend pertinent approaches and pathways to increase the use of improved maize seed varieties in the study area.

## Materials and methods

### *Study site description*

This study was conducted in diverse villages surrounding Harrismith and Phuthaditjhaba towns of the Maluti-a-Phofung municipality in the Thabo Mofutsanyana district of the Free State province of South Africa (Figure 1). The Free State is a summer rainfall region that contributes over 30% of the total maize production of the country, mostly under rainfed cropping systems.[42-45] This region has a subtropical highland climate with mean annual precipitation ranging between 500 and 900 mm and a mean annual air temperature of 17 °C.[44] There is high spatial rainfall variability in this region due to orographic patterns.[42,43,45]

The recent increase in the incidents of droughts, frost, hailstorms, and crop pests and diseases are the major weather-related risks affecting agricultural production in this region.[42-44] Therefore, the wider uptake and use of improved seed varieties, particularly climate-resilient varieties that possess desirable traits such as drought and heat tolerance, early maturing, high yielding, disease and pest resistance is expected to improve crop production and enhance climate resilience in this region.

### *Data collection*

The study relies on survey data gathered by the ARC as part of the InnovAfrica project (www.innovafrica.eu). This 4-year international collaborative project was initiated in 2017 with the aim of improving the food and nutritional security of African smallholders in six countries, including South Africa.[46] Detailed information about the activities of the

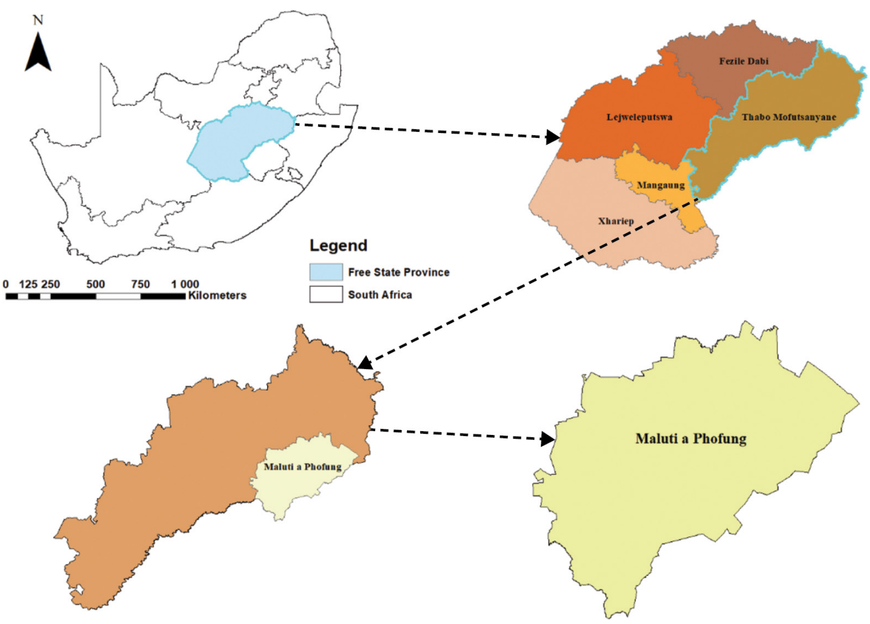



**Figure 1:** Location of the study area.[5]

InnovAfrica project in the South African case study was presented by Myeni et al.[33] Primary data were collected using a pre-tested, structured household questionnaire following a systematic random sampling technique whilst deliberately targeting smallholder farmers. Details of the procedures undertaken during the household survey were presented by Myeni et al.[5,8] Ethical clearance to conduct the study was provided by the Ethical Clearance Committee of the Agricultural Research Council – Natural Resources and Engineering. Data collected from a total of 600 smallholder farmers through the household questionnaire were compiled in a Microsoft Excel spreadsheet and underwent a data quality process to detect and eliminate incomplete, erroneous and suspicious data. Data on the potential pathways to improve the use of improved seed varieties within the context of climate change adaptation were collected through participatory engagements with multiple local stakeholders such as beneficiary farmers, extension officers, farmer associations, agro-dealers and traditional leadership during focus group discussions. Furthermore, these focus group discussions were used for the validation of the data collected through the household survey. After the data quality process, a total of 279 questionnaires that had good-quality data were reserved for further data analysis using the Statistical Package for the Social Sciences (SPSS) version 23.0. Details of the procedures undertaken during data analysis are presented in Section 1 of the supplementary material.

## Results and discussion

The descriptive analysis of the demographic characteristics of household heads in the study area is presented in Section 2 of the supplementary material.

### *Types and sources of seeds grown*

The results indicate that the majority of smallholder farmers grow local seed varieties (76%) whilst only about 14% grow improved seed varieties, thereby suggesting a very low level of adoption and use of improved seed varieties in the study area (Table 1). The results also show that about 10% of the farmers grow unknown seed varieties. These findings indicate that some of the farmers plant any seed varieties that are accessible and available to them without any prior knowledge of their specific characteristics and traits.

The majority of the smallholder farmers planted uncertified seeds either from their own farm (41%) or acquired from other farmers (28%), whilst about 30% purchased their seeds from cooperatives and private traders (Table 2). These findings indicate that the informal farmers' seed system was the main source of seeds grown by smallholder farmers (>69%) despite its low yield potential, as also reported by previous African studies.[14,19,33] Generally, the seeds are old varieties that have been circulating for decades and they vary from poor to fair in quality.

**Table 1:** Types of maize variety grown ($n = 279$)

| Type of maize variety | Frequency | Percentage (%) |
|---|---|---|
| Local seeds | 213 | 76.3 |
| Improved seed varieties | 38 | 13.6 |
| Unknown | 28 | 10.1 |

**Table 2:** Sources of seeds ($n = 279$)

| Source of seeds | Frequency | Percentage (%) |
|---|---|---|
| Government | *3* | 1.1 |
| Cooperatives | 46 | 16.5 |
| Private traders (e.g. agro-dealers) | *37* | 13.3 |
| Own farm | 113 | 40.5 |
| Other farmers | *79* | 28.3 |
| NGO | *1* | 0.4 |

### *Determinants to the adoption of improved seed varieties*

The results show that the adoption of improved seed varieties was driven mostly by factors such as easy and stable access to seeds (32%), high yield (29%), taste (14%) and multi-purpose (12%) (Table 3). These findings indicate that it is important to consider the farmers' preferences when selecting and promoting particular improved seed varieties in the study area. Remarkably, drought tolerance (8%), early maturation (5%) and tolerance to pests and diseases (0.4%) were the least reasons stimulating the farmers to adopt improved varieties. This suggests that most of the farmers lacked knowledge and information regarding seed varieties, their attributes, as well as their critical roles in combating the effects of climate variability. The focus group discussions revealed that most of the farmers plant local seeds mainly because this variety has easy and stable accessibility for free or at a low cost.

### *Constraints to adoption of improved seed varieties*

The results indicate that most of the farmers (76%) perceived no quality problems with their seeds (Table 4). The focus group discussions revealed that most farmers were mostly unaware of the quality indicators used to define good-quality seeds and their estimations of quality were only based on the physical appearance of the seeds. At least 15% of farmers indicated that a low germination rate was one of the quality-related challenges being observed in their seeds whilst only 6% of respondents indicated that unhealthy seeds were the key constraint.

The majority of the farmers (29%) perceived that they had not experienced any non-availability of improved maize seeds in the study area (Table 5). Furthermore, the results indicate that a limited number of suppliers (21%), poor road infrastructure (17%), limited quantity available from the supplier (14%), long distance to the supplier (11%) and unsuitable package size (7%) were the major reasons limiting the availability of improved seeds in the study area. This suggests that although improved seed varieties are readily available from the formal market, the high transaction cost of improved seed varieties attributed to the long distances to the suppliers, the limited number of suppliers, poor road infrastructure and unsuitable package size were the major constraints limiting the adoption of improved seed varieties in the study area. Our findings are in agreement with previous studies that have

**Table 3:** Reasons for the choice of seed variety grown ($n = 279$)

| Reason for choice of the variety grown | Frequency | Percentage (%) |
|---|---|---|
| Easy and stable access to seeds | 88 | 31.5 |
| Drought tolerant | 21 | 7.5 |
| High yielding | 82 | 29.4 |
| Early maturing | 13 | 4.7 |
| Taste | 40 | 14.4 |
| Tolerant to pests and diseases | 1 | 0.4 |
| Multi-purpose e.g. grain and fodder | 34 | 12.2 |

**Table 4:** Seed quality problems ($n = 279$)

| Quality problem | Frequency | Percentage (%) |
|---|---|---|
| No problem | 231 | 75.8 |
| Low germination rate | 42 | 15.1 |
| Unclean seed (low physical seed purity) | 16 | 5.7 |
| Unhealthy seeds (carry pests and diseases) | 3 | 1.1 |
| Not pure variety | 6 | 2.2 |
| Fake seeds on sale | 2 | 0.5 |

**Table 5:** Seed availability problems ($n = 279$)

| Availability problem | Frequency | Percentage (%) |
|---|---|---|
| No problem | 172 | 28.7 |
| A limited number of suppliers | 59 | 21.1 |
| Limited quantity available from the supplier | 39 | 14.0 |
| Unsuitable package size | 20 | 7.2 |
| Long distance to the supplier | 31 | 11.1 |
| Poor road infrastructure | 47 | 16.8 |
| Not available at the right time | 4 | 1.4 |

**Table 6:** Seed access problems ($n = 279$)

| Access problem | Frequency | Percentage (%) |
|---|---|---|
| No problem | 176 | 58.4 |
| Limited purchasing power | 65 | 23.3 |
| Limited access through social networks | 20 | 7.2 |
| Not eligible for support or subsidy | 31 | 11.1 |

reported similar constraints to the utilisation of improved seeds by smallholder farmers.[47,48]

The majority of the farmers (58%) perceived that seeds were accessible without any challenges (Table 6). This is probably because most of them use seeds that are saved from the previous harvest or acquired from neighbouring farmers. Furthermore, the results indicate that limited purchasing power (23%), ineligibility for support or subsidy (11%) and limited access to social networks (7%) were the major challenges faced in accessing improved seed varieties in the study area. This suggests that although improved seed varieties are readily available from the formal market, the farmers cannot afford to purchase them due to financial constraints. Similar findings have been reported in previous studies where the cost of improved seed varieties has been identified as one of the key constraints to the adoption of improved seed varieties by African smallholder farmers.[1,36,49]

## *Factors influencing the adoption of improved seed varieties by smallholder farmers*

### Characteristics of adopters and non-adopters of improved seed varieties

The results show significant mean differences between adopters and non-adopters only regarding gender, income and land, whilst the other differences were marginal (Table 7). The results suggest that male farmers were more likely to adopt improved seed varieties at a 10% significance level due to their higher chances of accessing new information, agricultural resources and financial resources compared to their female counterparts. Our findings are similar to previous studies that postulated that male farmers were more likely to adopt improved seed varieties.[35,36] Off-farm and on-farm income for adopters was significantly larger than that of non-adopters at a 1% level, thereby suggesting that farmers with higher incomes were more likely to adopt improved seed varieties than poor farmers, also reported by.[17,36] The results indicate that the majority of adopters owned more land than the non-adopters, suggesting that farmers who owned land were more likely to adopt improved seeds at a 10% level. The results also reveal that the average farm size for adopters was significantly larger than that of non-adopters at a 5% level, thereby suggesting that farmers with larger farm sizes were more likely to adopt improved seed varieties than their counterparts. This could be due to those farmers having enough land to

**Table 7:** Characteristics of adopters and non-adopters of improved seed varieties

| Independent variable | Means | | |
|---|---|---|---|
| | Non-adopter | Adopter | Difference |
| Gender | 0.49 | 0.61 | –0.12* |
| Age | 57.90 | 60.11 | –2.20 |
| Level of education | 5.09 | 6.11 | –1.02 |
| Occupation | 0.77 | 0.84 | –0.07 |
| Off-farm income | 1778.71 | 2537.90 | –759.19*** |
| On-farm income | 86.51 | 426.32 | –339.80*** |
| Access to media | 0.95 | 0.95 | 0.00 |
| Distance to market | 16.62 | 17.20 | –0.57 |
| Ownership of transport | 0.57 | 0.66 | –0.09 |
| Ownership of land | 0.72 | 0.84 | –0.12* |
| Farm size | 0.36 | 0.56 | –0.20** |
| Ownership of livestock | 0.25 | 0.18 | 0.06 |
| Access to extension services | 0.01 | 0.03 | –0.02 |

**p < 0.1, **p < 0.05, ***p < 0.01*

experiment and test new varieties whilst preserving some space for their preferred local varieties. Our findings are in agreement with previous studies that reported a significant positive relationship between farm size and the probability of adoption of improved seed varieties.[17,26,30,35,38]

### Factors affecting the adoption of improved seed varieties by smallholder farmers

The results estimated using a binary logistic regression model show that age ($p < 0.1$), off-farm income ($p < 0.05$), on-farm income ($p < 0.01$) and ownership of livestock ($p < 0.05$) are the key explanatory variables that had a substantial effect on the adoption of improved seed varieties by smallholder farmers (Table 8). Only the key factors that had a significant influence on the adoption of improved seed varieties in the study area are briefly discussed in this sub-section.

#### Age

The results indicate a significant positive relationship between the age of the farmer and the probability of adoption of improved seed varieties at a 10% level. They reveal that for a unit increase in age, the odds ratio of a smallholder farmer adopting improved seed varieties is 1.033. The possible explanation for this is that older farmers might be more experienced and some might have worked for commercial farmers and thus been exposed to the different types of improved seed varieties. Furthermore, some of the older farmers could have observed over the years that their traditional seed varieties were becoming more unreliable and gave low yields and thus, they were exploring different coping strategies to meet their production needs. Our findings are in line with previous studies that reported a positive relationship between the age of a farmer and the adoption of improved maize varieties.[50,51] However, they differ from other studies that postulate that younger farmers were more likely to adopt improved seed varieties due to their enthusiasm for taking risks and exploring new technologies.[17,30]

#### Income

The results indicate that farmers with higher off-farm income were 1.000 times more likely to plant improved seed varieties at a 5% level whilst farmers with higher on-farm income were 1.001 times more likely to plant improved seed varieties at a 1% level. These findings suggest that

**Table 8:** Estimated results on the factors affecting the adoption of improved seed varieties ($n = 279$)

| Explanatory variable | Coefficient | Odds ratio | SE | $z$ | $p > z$ |
|---|---|---|---|---|---|
| Gender | 0.623 | 1.865 | 0.406 | 2.350 | 0.125 |
| Age | 0.033* | 1.033 | 0.019 | 2.967 | 0.085 |
| Level of education | 0.068 | 1.071 | 0.055 | 1.560 | 0.212 |
| Occupation | 0.247 | 1.280 | 0.559 | 0.196 | 0.658 |
| Off-farm income | 0.002** | 1.000 | 0.000 | 4.628 | 0.031 |
| On-farm income | 0.001*** | 1.001 | 0.000 | 7.522 | 0.006 |
| Access to media | −0.095 | 0.910 | 0.816 | 0.014 | 0.907 |
| Distance to market | 0.002 | 1.002 | 0.034 | 0.003 | 0.954 |
| Ownership of transport | 0.846 | 2.331 | 0.614 | 1.902 | 0.168 |
| Ownership of land | 0.829 | 2.291 | 0.531 | 2.435 | 0.119 |
| Farm size | −0.087 | 0.916 | 0.466 | 0.035 | 0.852 |
| Ownership of livestock | −1.299** | 0.273 | 0.609 | 4.548 | 0.033 |
| Access to extension services | 1.355 | 3.876 | 1.484 | 0.834 | 0.361 |

*SE, standard error*

*$*p < 0.1$, $**p < 0.05$, $***p < 0.01$*

both off-farm and on-farm income were the key determinants of the adoption of improved seed varieties in the study area. This substantiates the view that access to high income provides farmers with greater powers to purchase improved seed varieties, as also reported by previous studies.[36,52]

**Ownership of livestock**

The results indicate that ownership of livestock had a negative and significant effect (−1.299) on the adoption of improved seed varieties at a 5% level. For a unit increase in the ownership of livestock, the odds ratio of a smallholder farmer adopting improved seed varieties is 0.273. This suggests that farmers who owned livestock were more comfortable and likely to feed their livestock with their traditional varieties for which they were more familiar with their characteristics and traits than unknown improved seed varieties with high uncertainties. This indicates that farmers lack information on these new varieties and how different livestock could respond to them. Our findings differ from previous studies that postulate that ownership of livestock had a significant positive effect on the adoption of improved seed varieties.[53,54] This could be due to the fact that those studies often associated livestock ownership with wealth and, hence, the purchasing power of inputs. However, most of the farmers in our study area were not selling their livestock but instead keeping them for their own consumption and traditions.

## Conclusions and recommendations

This study was undertaken to examine the adoption rate, constraints and factors determining the adoption of improved seed varieties in smallholder farming conditions in the Maluti-a-Phofung municipality of the Free State Province of South Africa, using a maize crop. The results show that most of the smallholder farmers grew uncertified local seed varieties saved from their previous harvest or acquired from neighbouring farmers. The adoption of seed varieties was driven mostly by factors such as easy and stable access, high yield, taste and multi-purpose use. The results also reveal that most of the farmers used improved seed varieties without full knowledge of their specific characteristics, traits and ecological requirements, and how to effectively use them. The major barriers limiting smallholder farmers from adopting the improved seed varieties include lack of awareness and knowledge, inadequate information, unaffordability and unavailability of improved seeds, low income and landlessness. Age and income were the key factors that significantly and positively affected the adoption of improved seed varieties, whilst ownership of livestock had a significant negative effect on their adoption.

Based on the findings, the provision of knowledge and awareness-raising of improved seed varieties through enhanced extension services and capacity-building initiatives are highly recommended to overcome knowledge constraints. This study also encourages the government to improve the existing subsidy programmes and enhance the provision and accessibility of financial resources that are inclusive for smallholder farmers, including women and youth. To improve the availability and affordability of good-quality seed varieties, we recommend that improved seed varieties be sold at a local market in small and affordable micro-packs (2 or 5 kg). The government is also encouraged to improve the infrastructure in remote rural areas to reduce transaction costs for transportation inputs and marketing products. Smallholder farmers are encouraged to work in groups and buy inputs such as improved seeds and fertilisers in bulk at bargain purchasing and transportation costs. We also recommend that community-based seed banks are developed and improved to ensure that the required quantities of locally adapted seeds of good quality are available in remote areas during the planting season and at affordable prices. The recommendations of this study can be used by the government and other developmental organisations to enhance the wider uptake and use of improved seed varieties in smallholder farming conditions and assist farmers to cope effectively with climate variability and change.

## Acknowledgements

Knowledge Intelligence Applications GmbH (KIAG) is gratefully acknowledged for the provision of KIPUS software used for data capturing and online storage. We also gratefully acknowledge the Kenya Agricultural & Livestock Research Organisation (KALRO) for the provision of a questionnaire used for primary data collection. Mulalo Thavhana (ARC) is thanked for her technical assistance during the data collection. We acknowledge the Free State Department of Agriculture and Rural Development extension officers and smallholder farmers for their time and their active participation in the household survey as well as the focus group discussions. Dr Thomas Fyfield (ARC) is thanked for his editorial assistance. This research was funded by the European Union's H2020 Research and Innovation Programme, Grant Agreement No. 727201.

## Competing interests

We have no competing interests to declare.

## Authors' contributions

L.M.: Conceptualisation, methodology, software, validation, formal analysis, writing – original draft preparation, writing – review and editing. M.E.M.: Methodology, writing – review and editing. All authors have read and agreed to the accepted version of the manuscript.

## References

1. Acevedo M, Pixley K, Zinyengere N, Meng S, Tufan H, Cichy K, et al. A scoping review of adoption of climate-resilient crops by small-scale producers in low- and middle-income countries. Nat Plants. 2020;6(10):1231–1241. https://doi.org/10.1038/s41477-020-00783-z
2. Tesfahunegn GB, Mekonen K, Tekle A. Farmers' perception on causes, indicators and determinants of climate change in northern Ethiopia: Implication for developing adaptation strategies. Appl Geogr. 2016;73:1–12. https://doi.org/10.1016/j.apgeog.2016.05.009
3. Vernooy R, Sthapit B, Otieno G, Shrestha P, Gupta A. The roles of community seed banks in climate change adaption. Dev Pract. 2017;27(3):316–327. https://doi.org/10.1080/09614524.2017.1294653

4. Jafino BA, Walsh B, Rozenberg J, Hallegatte S. Revised estimates of the impact of climate change on extreme poverty by 2030. Policy Research Working Paper. Washington DC: World Bank; 2020. p. 17. Available from: https://openknowledge.worldbank.org/entities/publication/a549a5ee-71 cd-5ed4-bcf3-3a8cb508b199

5. Myeni L, Moeletsi M, Thavhana M, Randela M, Mokoena L. Barriers affecting sustainable agricultural productivity of smallholder farmers in the eastern free state of South Africa. Sustainability. 2019;11(11):1–18. https://doi.org/10.3390/su11113003

6. De Souza K, Kituyi E, Harvey B, Leone M, Murali KS, Ford JD. Vulnerability to climate change in three hot spots in Africa and Asia: Key issues for policy-relevant adaptation and resilience-building research. Reg Environ Change. 2015;15(1):747–753. https://doi.org/10.1007/s10113-015-0755-8

7. Lunduka RW, Mateva KI, Magorokosho C, Manjeru P. Impact of adoption of drought-tolerant maize varieties on total maize production in south Eastern Zimbabwe. Clim Dev. 2019;11(1):35–46. https://doi.org/10.1080/17565529.2017.1372269

8. Myeni L, Moeletsi ME. Factors determining the adoption of strategies used by smallholder farmers to cope with climate variability in the eastern Free State, South Africa. Agriculture. 2020;10(9):410–426. https://doi.org/10.3390/agriculture10090410

9. Lipper L, Thornton P, Campbell BM, Baedeker T, Braimoh A, Bwalya M, et al. Climate-smart agriculture for food security. Nat Clim Chang. 2014;4(12):1068–1072. https://doi.org/10.1038/nclimate2437

10. Makate C. Effective scaling of climate smart agriculture innovations in African smallholder agriculture: A review of approaches, policy and institutional strategy needs. Environ Sci Policy. 2019;96(1):37–51. https://doi.org/10.1016/j.envsci.2019.01.014

11. Amondo E, Simtowe F, Rahut DB, Erenstein O. Productivity and production risk effects of adopting drought-tolerant maize varieties in Zambia. Int J Clim Chang Strateg Manag. 2019;11(4):570–591. https://doi.org/10.1108/IJCCSM-03-2018-0024

12. Westengen OT, Haug R, Guthiga P, Macharia E. Governing seeds in East Africa in the face of climate change: Assessing political and social outcomes. Front Sustain Food Syst. 2019;3(1):53–67. https://doi.org/10.3389/fsufs.2019.00053

13. Fajardo Vizcayno J, Hugo W, Sanz Alvarez J. Appropriate seed varieties for small-scale farmers: Key practices for DRR implementers. Rome: FAO; 2014. Available from: https://agris.fao.org/agris-search/search.do?recordID=XF2015001488

14. Simtowe F, Marenya P, Amondo E, Worku M, Rahut DB, Erenstein O. Heterogeneous seed access and information exposure: Implications for the adoption of drought-tolerant maize varieties in Uganda. Agric Food Econ. 2019;7(1):1–23. https://doi.org/10.1186/s40100-019-0135-7

15. Dao A, Sanou J, Gracen V, Danquah EY. Indentifying farmers' preferences and constraints to maize production in two agro-ecological zones in Burkina Faso. Agric Food Secur. 2015;4(1):1–7. https://doi.org/10.1186/s40066-015-0035-3

16. Ariga J, Mabaya E, Waithaka M, Wanzala-Mlobela M. Can improved agricultural technologies spur a green revolution in Africa? A multicountry analysis of seed and fertilizer delivery systems. Agric Econ (United Kingdom). 2019;50(S1):63–74. https://doi.org/10.1111/agec.12533

17. Oluwatoyin BC. Factors influencing adoption of improved maize seed varieties among smallholder farmers in Kaduna State, Nigeria. J Agric Ext Rural Dev. 2021;13(2):107–114. https://doi.org/10.5897/JAERD2019.1032

18. Wossen T, Abdoulaye T, Alene A, Feleke S, Menkir A, Manyong V. Measuring the impacts of adaptation strategies to drought stress: The case of drought tolerant maize varieties. J Environ Manage. 2017;203:106–113. https://doi.org/10.1016/j.jenvman.2017.06.058

19. Kansiime MK, Mastenbroek A. Enhancing resilience of farmer seed system to climate-induced stresses: Insights from a case study in West Nile region, Uganda. J Rural Stud. 2016;47:220–230. https://doi.org/10.1016/j.jrurstud.2016.08.004

20. Maity A, Pramanik P. Climate change and seed quality: An alarming issue in crop husbandry. Curr Sci. 2013;105(10):1336–1338.

21. Gidi L, Hlongwane J, Nkoana M. Analysing factors limiting the adoption of improved maize varieties by small-scale maize farmers in Ga-Mamadila Village of Polokwane Municipality, Limpopo. In: Proceedings of the 56th Annual Conference of the Agricultural Economics Association of South Africa (AEASA); 2018 September 25–27; Cape Town, South Africa. Pretoria: AEASA; 2018. https://doi.org/10.22004/ag.econ.284760

22. Oluwayemisi IA, Olarinde LO, Fatunbi AO. Determinants of adoption of improved maize varieties in Kano-Katsina-Maradi, West Africa. African Crop Sci J. 2017;25(1):1–11. https://doi.org/10.4314/acsj.v25i1.1S

23. Setimela PS, Magorokosho C, Lunduka R, Gasura E, Makumbi D, Tarekegne AT, et al. On-farm yield gains with stress-tolerant maize in eastern and southern Africa. Agron J. 2017;109(2):406–417. https://doi.org/10.2134/agronj2015.0540

24. Elias SG. The importance of using high quality seeds in agriculture systems. Agric Res Technol Open Access J. 2018;15(4):1–2. https://doi.org/10.19080/ARTOAJ.2018.14.555961

25. Langyintuo A. Smallholder farmers' access to inputs and finance in Africa. In: Gomez y Paloma S, Riesgo L, Louhichi K, editors. The role of smallholder farms in food and nutrition security. Cham: Springer; 2020. p. 133–152. https://doi.org/10.1007/978-3-030-42148-9_7

26. Simtowe F, Amondo E, Marenya P, Rahut D, Sonder K, Erenstein O. Impacts of drought-tolerant maize varieties on productivity, risk, and resource use: Evidence from Uganda. Land Use Policy. 2019;88:104091–104101. https://doi.org/10.1016/j.landusepol.2019.104091

27. Kamau MW, Bagamba F, Riungu C, Mukundi J, Toel R. Early changes in farmers' adoption and use of an improved maize seed: An assessment of the impact of demos and field days. Afr Eval J. 2018;6(1):1–10. https://doi.org/10.4102/aej.v6i1.278

28. Cacho OJ, Moss J, Thornton PK, Herrero M, Henderson B, Bodirsky BL, et al. The value of climate-resilient seeds for smallholder adaptation in sub-Saharan Africa. Clim Change. 2020;162(3):1213–1229. https://doi.org/10.1007/s10584-020-02817-z

29. Sinyolo S. Technology adoption and household food security among rural households in South Africa: The role of improved maize varieties. Technol Soc. 2020;60:101214–101224. https://doi.org/10.1016/j.techsoc.2019.101214

30. Toenniessen G, Adesina A, DeVries J. Building an alliance for a green revolution in Africa. Ann N Y Acad Sci. 2008;1136(1):233–242. https://doi.org/10.1196/annals.1425.028

31. Danso-Abbeam G, Bosiako JA, Ehiakpor DS, Mabe FN. Adoption of improved maize variety among farm households in the northern region of Ghana. Cogent Econ Financ. 2017;5(1):1416896–1416910. https://doi.org/10.1080/23322039.2017.1416896

32. Cairns JE, Prasanna BM. Developing and deploying climate-resilient maize varieties in the developing world. Curr Opin Plant Biol. 2018;45:226–230. https://doi.org/10.1016/j.pbi.2018.05.004

33. James M. Low-cost drought and low nitrogen-tolerant maize hybrids for food security in South Africa [webpage on the Internet]. c201 [cited 2023 Jun 08]. Available from: https://www.grainsa.co.za/low-cost-drought-and-low-nitrogen-tolerant-maize-hybrids-for-food-security-in-south-africa

34. Myeni L, Moeletsi ME, Nyagumbo I, Modiselle S, Mokoena L, Kgakatsi IB. Improving the food and nutritional security of smallholder farmers in south africa: Evidence from the InnovAfrica project. Sustainability. 2021;13(17):9902–9913. https://doi.org/10.3390/su13179902

35. Adenle AA, Wedig K, Azadi H. Sustainable agriculture and food security in Africa: The role of innovative technologies and international organizations. Technol Soc. 2019;58:101143–101160. https://doi.org/10.1016/j.techsoc.2019.05.007

36. Atube F, Malinga GM, Nyeko M, Okello DM, Alarakol SP, Okello-Uma I. Determinants of smallholder farmers' adaptation strategies to the effects of climate change: Evidence from northern Uganda. Agric Food Secur. 2021;10(1):1–14. https://doi.org/10.1186/s40066-020-00279-1

37. Sigigaba M, Mdoda L, Mditshwa A. Adoption drivers of improved open-pollinated (OPVs) maize varieties by smallholder farmers in the Eastern Cape province of South Africa. Sustainability. 2021;13(24):13644–13718. https://doi.org/10.3390/su132413644

38. Mutanyagwa AP, Isinika A, Kaliba AR. The factors influencing farmers' choice of improved maize seed varieties in Tanzania. Int J Sci Res Manag. 2018;6(04):55–62. https://doi.org/10.18535/ijsrm/v6i4.ah01

39. Ngcinela S, Mushunje A, Taruvinga A, Ngarava S, Mutengwa CS. Determinants of genetically modified (GM) maize adoption and the intensity of adoption in OR Tambo District Municipality, Eastern Cape Province, South Africa. GM Crop Food. 2019;10(1):1–11. https://doi.org/10.1080/21645698.2019.1580127

40. Das B, Deventer FV, Wessels A, Mudenda G, Key J, Ristanovic D. Role and challenges of the private seed sector in developing and disseminating climate-smart crop varieties in eastern and southern Africa. In: The Climate-Smart Agriculture Papers. Cham: Springer; 2019. p. 67–78. https://doi.org/10.1007/978-3-319-92798-5_6

41. Senyolo MP, Long TB, Omta O. Enhancing the adoption of climate-smart technologies using public-private partnerships: Lessons from the WEMA case in South Africa. Int Food Agribus Manag Rev. 2021;24(5):755–776. https://doi.org/10.22434/IFAMR2019.0197

42. Alberts J, Rheeder J, Gelderblom W, Shephard G, Burger HM. Rural subsistence maize farming in South Africa: Risk assessment and intervention models for reduction of exposure to fumonisin mycotoxins. Toxins (Basel). 2019;11(6):334–354. https://doi.org/10.3390/toxins11060334

43. Mbiriri M, Mukwada G, Manatsa D. Influence of altitude on the spatiotemporal variations of meteorological droughts in mountain regions of the free state Province, South Africa (1960–2013). Adv Meteorol. 2018;2018:1–11. https://doi.org/10.1155/2018/5206151

44. Moeletsi ME, Walker S. Agroclimatological suitability mapping for dryland maize production in Lesotho. Theor Appl Climatol. 2013;114(1–2):227–236. https://doi.org/10.1007/s00704-012-0829-1

45. Botai CM, Botai JO, Dlamini LC, Zwane NS, Phaduli E. Characteristics of droughts in South Africa: A case study of Free state and North West provinces. Water (Switzerland). 2016;8(10):439–462. https://doi.org/10.3390/w8100439

46. Moswetsi G, Fanadzo M, Ncube B. Cropping systems and agronomic management practices in smallholder farms in South Africa: Constraints, challenges and opportunities. J Agron. 2017;16(2):51–64. https://doi.org/10.3923/ja.2017.51.64

47. InnovAfrica. Innovations in Technology, Institutional and Extension Approaches towards Sustainable Agriculture and Enhanced Food and Nutrition Security in Africa 2017–2018. EU H2020 Grant Support to InnovAfrica Project, Grant Agreement no 727201. Brussels: European Commission; 2016. Available from: https://cordis.europa.eu/project/id/727201

48. Schroeder C, Onyango TK, Nar RB, Jick NA, Parzies HK, Gemenet DC. Potentials of hybrid maize varieties for small-holder farmers in Kenya: A review based on SWOT analysis. Afr J Food Agric Nutr Dev. 2013;13(2):7562–7586. https://doi.org/10.18697/ajfand.57.11360

49. Jha S, Kaechele H, Lana M, Amjath-Babu TS, Sieber S. Exploring farmers' perceptions of agricultural technologies: A case study from Tanzania. Sustainability. 2020;12(3):1–21. https://doi.org/10.3390/su12030998

50. Kaliba AR, Mazvimavi K, Gregory TL, Mgonja FM, Mgonja M. Factors affecting adoption of improved sorghum varieties in Tanzania under information and capital constraints. Agric Food Econ. 2018;6(1):1–21. https://doi.org/10.1186/s40100-018-0114-4

51. Ademiluyi IO. Adoption of improved maize varieties among farmers in Bassa Local Government Area of Plateau State, Nigeria. Int J Innov Agric Biol Res. 2014;2(4):26–33.

52. Zahidul Islam KM, Sumelius J, Bäckman S. Do differences in technical efficiency explain the adoption rate of HYV rice? Evidence from Bangladesh. Agric Econ Rev. 2012;13(1):93–110. https://doi.org/10.22004/ag.econ.253500

53. Sibanda M, Mushunje A, Mutengwa CS. Factors influencing the demand for improved maize open pollinated varieties (OPVs) by smallholder farmers in the Eastern Cape Province, South Africa. J Cereal Oilseeds. 2016;7(2):14–26. https://doi.org/10.5897/JCO2015.0142

54. Ahmed H, Tetteh Anang B. Impact of improved variety adoption on farm income in Tolon District of Ghana. Agric Soc Econ J. 2019;19(2):105–115. https://doi.org/10.21776/ub.agrise.2019.019.2.5

55. Siyum N, Giziew A, Abebe A. Factors influencing adoption of improved bread wheat technologies in Ethiopia: Empirical evidence from Meket district. Heliyon. 2022;8(2):1–11. https://doi.org/10.1016/j.heliyon.2022.e08876

# Effect of manganese-rich solid waste on soil phosphorus availability applied as monopotassium and rock phosphate in two contrasting soils

**AUTHORS:**
Mohammed A. Elsheikh[1,2]
Khatab Abdalla[3]
Louis Titshall[2]
Pardon Muchaonyerwa[2]

**AFFILIATIONS:**
[1]Department of Soil and Environment Sciences, Faculty of Agriculture, University of Khartoum, Khartoum, Sudan
[2]School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa
[3]Agroecology Chair, Bayreuth Centre for Ecology and Environmental Research (BayCEER), University of Bayreuth, Bayreuth, Germany

**CORRESPONDENCE TO:**
Khatab Abdalla

**EMAIL:**
Khatab.Abdalla@Uni-Bayreuth.de

**DATES:**
**Received:** 27 Feb. 2023
**Revised:** 05 July 2023
**Accepted:** 07 July 2023
**Published:** 28 Sep. 2023

**HOW TO CITE:**
Elsheikh MA, Abdalla K, Titshall L, Muchaonyerwa P. Effect of manganese-rich solid waste on soil phosphorus availability applied as monopotassium and rock phosphate in two contrasting soils. S Afr J Sci. 2023;119(9/10), Art. #15689. https://doi.org/10.17159/sajs.2023/15689

**ARTICLE INCLUDES:**
☒ Peer review
☐ Supplementary material

**DATA AVAILABILITY:**
☐ Open data set
☐ All data included
☒ On request from author(s)
☐ Not available
☐ Not applicable

**EDITORS:**
Priscilla Baker
Amanda-Lee Manicum

**KEYWORDS:**
phosphate rock, $KH_2PO_4$, incubation experiment, mining waste, South Africa

**FUNDING:**
University of KwaZulu-Natal



Manganese (Mn) mining produces a large amount of Mn-rich solid waste contributing to soil and groundwater pollution. Adding Mn-rich waste to soils could reduce mining pollution by allowing specialised plants to take up this mineral for growth, that is, phytoremediation. However, Mn interacts with other soil major and trace minerals. The interaction with phosphorus (P), a key element for plant nutrition and growth, has received less attention. In this study, we aimed to evaluate the effects of P sources (i.e. rock phosphate and monopotassium [$KH_2PO_4$]) and application rates on the P and Mn relationships in clay and sandy soils mixed with Mn-rich waste. Soils differing in texture were incubated for 60 days at room temperature (±20 °C), and changes in available P, Mn and soil pH were determined at 0, 30 and 60 days. The addition of Mn-rich solid waste significantly decreased available soil P in both soils with the greatest reduction of 62% and 52% from the sandy soil subjected to $KH_2PO_4$ and rock phosphate, respectively. In the clayey soil, the reduction was higher for the rock phosphate source implying more P was released from the rock phosphate in Mn-rich soils. This explanation was supported by the significant positive correlation between P and Mn for both soils when P is added in the form of rock phosphate. Our results suggest that Mn-rich waste is better in clay soils subjected to rock phosphate addition. Further research is needed to control Mn solid waste pollution levels in soils using specific crops with known phytoremediation properties.

**Significance:**

South African mining and smelting processes produce a lot of Mn-rich waste as a by-product that harms the environment if not appropriately managed. The efficient use of Mn-rich solid waste in agricultural soils is poorly studied; hence, this study focused on the role of soil type and Mn-rich waste addition on phosphorus release and availability.

## Introduction

Mining and smelting activities pollute soil and groundwater resources with heavy metals that damage ecosystems and environments worldwide.[1,2] The mining-associated pollution is already taking its toll in many regions, such as Asia, South America and sub-Saharan Africa.[2,3] The problem is more acute in African countries such as South Africa, where mining of key minerals (e.g. gold, coal, diamonds, manganese (Mn), nickel and iron ore) contributes to the national economy.[4] The country has the largest known, natural deposits of Mn concentration of ~4 million tons, that is, about 75–80% of the world reservoir and most of it is (~99%) in the Kalahari basin, Northern Cape Province.[5-7] Mn mining is associated with high levels of Mn-rich solid waste which can be reused efficiently if appropriately managed. Therefore, the safe disposal and control of Mn mine solid waste is key for reducing negative mining impacts on soils, water, biodiversity and human well-being.

Mn is also an essential trace element for plant growth and soil microbial life that can be toxic when available in high amounts.[8,9] Mn can be used by plants and soil microorganisms in mineral-deficient soils, suggesting a possible biological removal of high levels of Mn in soils through land application by Mn-rich waste. However, Mn uptake by plants and soil microorganisms may be affected by interactions with other elements. For example, P is an important element for both agricultural and environmental sustainability, and it is deficient in most agricultural soils.[10,11] Also, Mn availability is considered to have a negative impact on P uptake and distribution in different plants.[12,13] The P shortage may be further aggravated by soil erosion and land degradation.[10,14] For example, Alewell et al.[10] estimated that up to 50% of global total P loss is caused by soil erosion. Therefore, understanding the interaction and behaviour of P in soils with high Mn concentration is a prerequisite for efficient P management in agricultural systems.

There are conflicting findings in studies concerning the interaction between P and Mn in soils and plants. For instance, Lindsay et al.[15] in a laboratory incubation study reported that soil Mn was lowered by P addition in acidic sandy loam soil (pH =4.6) and slightly calcareous (pH = 7.6) loamy soil from Nebraska, USA.[16] They also found a 14–21% increase of available P with the addition of Mn oxides in incubated soils collected from a rice field in the Texas Gulf Coast.[16] Moreover, P sorption may be affected by Mn oxides in rich-Mn soil as found in acid-sulfate soils of Thailand.[17] Contradictory findings over the interaction between Mn and P in the soils were also observed in the plant tissues. For example, elevated P supply was found to directly interfere with Mn uptake in barley roots which induced a negative impact on the barely shoots because of Mn deficiency.[13] Also, Marsh et al.[18] found that the Mn toxicity of potatoes increased by increasing phosphate levels, especially at higher temperatures (>25 °C). However, a reduction in Mn toxicity in some forage species, for example, perennial ryegrass and white clover, due to increased P supply was also reported.[19] However, Titshall[20] found Mn concentrations of ryegrass were lower than the control groups in soils mixed with Mn-rich waste after adding phosphate, which was attributed to the formation of Mn and P compounds, causing a reduction in Mn availability.

Contrasting findings about P and Mn interactions in plants and soils suggest the need for further work, especially in South Africa where Mn-rich waste is to be applied to land as a disposal method. More specifically, we seek to understand the interaction between Mn and P in soils in the absence of plants. Such an understanding of P behaviour from different sources with different levels of Mn-rich waste-amended soils will be essential for sustainable management of P, including optimisation of P fertiliser inputs for crop production on Mn waste-treated soils. Therefore, the objective of this study was to determine the effects of two phosphate sources and application rate on available P and extractable Mn in contrasting soils incubated with Mn-rich waste.

## Materials and methods

### *Electro-winning waste*

Electro-winning or electrolytic waste used in this study was collected from the top 20–30 cm from several locations of dedicated waste disposal sites (25°27'56.09"S; 30°57'4.70"E) of Mn mining and processing company near Mombela, Mpumalanga, South Africa.[20] The Mn-rich solid waste is produced as a by-product after Mn-rich ores are finely milled and dissolved in sulfuric acid.[20] Various conditioners (including ammonia and lime) were added during an electrolytic extraction process, resulting in a solid waste that contained high amounts of Mn residual and elevated levels of N, Ca and S. After electrolytic extraction of the Mn, the residue is passed through a belt-filter press to remove excess liquid and the solid waste is disposed of at a dedicated disposal site. The solid waste has a pH close to 7, and electrical conductivity of 1735 mS/m with 5702 mg/kg EDTA-extractable Mn (Table 1).

### *Soils and phosphate sources used in the study*

Two soils with different texture classes (Hutton and Cartref) were used for the current study. The A horizon of Hutton and Cartref soils were classified as Typic Haplustox and soil Typic Haplaquept, respectively.[21,22] The Hutton soil was collected from a maize farm near Howick, KwaZulu-Natal, South Africa (29°31′22.99″S; 30°13′11.83″E), whilst the Cartref was sourced from natural veld at Otto's Bluff near Pietermaritzburg, KwaZulu-Natal, South Africa (29°30′42.01″S; 30°22′51.99″E). The soils used for the incubation were collected from the upper soil layer (20 cm) using a spade from four randomly selected positions for each soil. For each soil, the four sampled positions were mixed together, thoroughly mixed and sieved through a 2-mm sieve to remove stone and plant debris. Two sources of phosphate were used, that is, sparingly soluble phosphate rock and readily available P as monopotassium phosphate ($KH_2PO_4$) fertiliser. The phosphate rock used in this study was obtained from the second richest reserve of phosphate rock, Langebaan town, South Africa. The chemical composition (%) of the phosphate rock was as follows: total P =10.0; citric acid soluble P = 3.3; Ca =20.1; Cu = 0.1; Mg = 1.30; Fe = 3.0; Mn = 9.0; Zn = 13; Mo = 10.0.

### *Laboratory incubation*

Solid waste was added at a rate of 40 g waste/kg soil (4 g waste in 100 g soil), which was found to be the upper application rate suitable for growing crops.[20] Approximately 100 g of soil was weighed in a 60-mL laboratory glass jars with screw caps. Phosphorus was added as either phosphate rock or $KH_2PO_4$ at four rates: control (0P), recommended dose (1P), double recommended dose (2P) and threefold recommended dose (3P). The P application rate was duplicated (2P) and triplicated (3P) to ensure a wide range of phosphorus availability in soils. A total of 48 jars were prepared. The recommended P fertilisation dose for the Hutton soil was 65 kg P/ha and Cartref soil was 80 kg P/ha as recommended by the Cedara College of Agriculture, Department of Agriculture, South Africa. These recommended rates translated to 0.0, 3.7, 7.4 and 11.1 mg of P per 100 g soil for Hutton soil and 0.0, 3.6, 7.2 and 10.8 mg of P per 100 g soil for Cartref soil, based on the estimation of 2 million kg soil per hectare furrow-slice. The mixtures were adjusted to field capacity moisture content and incubated at 20 °C for 60 days, with soil sampling collected at 0, 30 and 60 days of incubation. Previous work by Titshall[20] using this waste material indicated that the bulk of chemical changes occurred within the first 50–60 days of mixing soil with the waste, and thus 60 days was used as the maximum incubation period.

**Table 1:** Basic chemical profile of electro-winning waste, Hutton and Cartref soil ($n = 3$)

| Parameters | | Mn waste | Hutton | Cartref |
|---|---|---|---|---|
| pH | $H_2O$ | 6.89 | 5.34 | 5.17 |
| | KCl | 6.86 | | |
| Electrical conductivity (mS/m) | | 1735 | | |
| Organic carbon (g/kg) | | 5.6 | 34.4 | 1.8 |
| AMBIC P (mg/kg) | | 1.22 | 12.0 | 4.42 |
| Total N (g/kg) | | 9.3 | 2.1 | |
| Extractable cations (cmolc/kg) | Mg | 1.76 | 3.04 | 0.35 |
| | Na | 1.31 | 6.63 | 0.12 |
| | Ca | 71.0 | 6.63 | 1.28 |
| | K | 0.06 | 0.47 | 0.38 |
| Cation exchange capacity (cmolc/kg) | CEC | 9.50 | 12.7 | 6.10 |
| Exchangeable acidity (cmolc/kg) | | 4.47 | | |
| Exchangeable Al (cmolc/kg) | | 1.28 | | |
| Calcium carbonate (%) | $CaCO_3$ | 2.6 | | |
| Particle size distribution (%) | Clay | 29.4 | 54.0 | 19 |
| | Silt | 58.6 | 34.3 | 13 |
| | Sand | 12.0 | 11.7 | 68 |
| EDTA-extractable (mg/kg) | Fe | 1.02 | | |
| | Mn | 5708 | | |
| | Zn | 7.39 | | |
| | Co | 69.5 | | |
| | Cu | 17.9 | | |
| | Pb | 1.38 | | |
| | Ni | 5.57 | | |

### *Soil samples analysis*

Soil samples from the initial soil and during the incubation period (at 0, 30 and 60 days) were air-dried and stored for laboratory analysis. Soil pH was measured in both water and KCl (1 M) suspension (1 g soil: 2.5 mL of the solution) using a digital Metrohm Hersiau E396B pH meter. Plant available P was extracted with AMBIC solution (0.25 mol/L ammonium bicarbonate, pH 8.3) and measured colourimetrically using ultraviolet-visible (UV-vis) spectrophotometers.[23] Each sample was prepared for the analysis of diethylenetriaminepentaacetic acid (DTPA)-extractable Mn. Extractable Mn was determined using ammonium acetate, adjusted to pH 7, followed by quantification with atomic adsorption spectrophotometry (AAS, Varian 2600).[24]

### *Statistical analysis*

The experimental data were found to be normality distributed ($p > 0.05$) using the Shapiro–Wilk test. Factorial multivariate analysis was applied to all data sets to test the effect of soil type, Mn waste, phosphorus source and application rates on the available P and Mn. Follow-up tests using three-way variance analysis (ANOVA) analysis were done separately for incubation samples with and without Mn waste for each soil type. As sampling was repeated three times (0, 30 and 60 days) over the incubation period, a repeated two-way ANOVA was used to test for how soil type was affected by phosphate application rate and incubation periods. Relationships between available P against Mn and soil pH were assessed with linear regressions. Tukey's honestly significant difference post hoc multiple comparison tests were performed to compare means ($p < 0.05$ threshold) unless otherwise indicated. All the statistical analysis and graphs were conducted using Sigma Plot software (Version 14.5, Systat Software Inc., San Jose, CA, USA).

## Results

### *Descriptive statistics and analysis of variance*

Both soils have an acidic pH but have different organic carbon contents and soil textures (sandy loam and clayey texture for Hutton and Cartref soil, respectively) (Table 1). Hutton soil had higher extractable cations (i.e. Mg, Na, Ca and K) and cation exchange capacity than Cartref soil. Available P and extractable Mn in the two soils with different P as fertiliser sources showed that $KH_2PO_4$ had a greater overall mean in both soil types regardless of waste treatment (Table 2). On average, the $KH_2PO_4$ source increased available P by 5- and 1.8-fold in the Cartref and Hutton soils without Mn-rich solid waste, respectively. Extractable Mn was higher in soil amended with Mn-rich waste compared to soil without Mn-rich waste. Although the coefficient of variation (CV) of all parameters is consistently <1%, CV is always greater in available P from the $KH_2PO_4$ source than the rock phosphate source (Table 2). Multivariate factorial analyses showed that the effect of P application rate, incubation time, Mn-rich waste and P sources and their interaction on available P in both soils were highly significant ($p < 0.001$) (Table 3). However, incubation times and P dosage in the Cartref soil were only significant at $p < 0.05$ level. Extractable Mn was affected significantly by the soil type, presence or absence of Mn-rich waste, application rates, and the interaction of soils and application rates only (Table 3).

### *Available soil P from contrasting soil types with and without Mn-rich waste*

The available soil P increased significantly with the application rate of $KH_2PO_4$ and rock phosphate in Cartref and Hutton soils over the incubation time (Figure 1). The P concentration significantly rose ($p < 0.05$) with increasing $KH_2PO_4$ application rates (from 0P to 3P rate) in both soils (Figure 1a and 1b). The variation in available P between the P rates (0P, 1P 2P and 3P) was much higher in Cartref soil than in Hutton soil. For example, the largest difference in available was observed between P at the beginning of the incubation (0 days) between 0P with 4.42±0.38 mg/L and 3P with 94.82±2.78 mg/L in Cartef soil (Figure 1a). This difference was relatively lower in the Hutton soil (Figure 1b) despite the baseline available P comparable to Cartref soil. In terms of Rock phosphate source, variations in available P between the P rates were low (Figure 1c and 1d). Available P increased over time in the Cartref soil subjected to rock phosphate addition, with a higher P concentration observed at the 60 days of the incubation compared to the earlier sampling events (Figure 1c). In this soil, there was no significant difference between 2P and the 3P from the rock phosphate soil over the incubation time. On the other hand, significant variation in available P between the P rates was observed in the sampling at the 30 and 60 days, with clear variation between 0P on one hand and the other rates on the other hand at the end of the incubation (Figure 1d).

Concentrations of soil available P in soils mixed with Mn waste followed a similar pattern as the soils without Mn waste, with less available P from P doses applied as $KH_2PO_4$ (Figure 2). For example, the highest available P of 64.46±1.38 mg/L was recorded at the first sampling event (0 day) under the 3P application rate in Cartref soil, which was decreased by 45% compared to the average of a second (30 days) and last (60 days) sampling event (Figure 2a). A similar trend was also observed under Hutton soil subjected to P addition from $KH_2PO_4$, where the available P increased in the following order: 0P, 1P, 2P and 3P (Figure 2b).

The addition of rock phosphate at different rates had no consistent effect on available soil P in Cartref and Hutton soils in the presence of the Mn-rich wastes (Figure 2c and 2d). Although no significant difference between P rates was observed at the start of the incubation (immediately after the rock phosphate addition) in the Cartref soil, a significantly lower available P was found in 1P compared to other rates after 30 days

**Table 2:** Descriptive statistics of the available phosphorus (P) and extractable manganese (Mn) for Cartref and Hutton soil incubated with and without Mn-rich waste ($n = 36$)

| Soil type | Cartref | | Hutton | |
|---|---|---|---|---|
| P source | $KH_2PO_4$ | Rock phosphate | $KH_2PO_4$ | Rock phosphate |
| **Available P without Mn-rich waste** | | | | |
| Mean | 43.89 | 8.94 | 23.15 | 13.82 |
| SD | 30.78 | 4.06 | 9.46 | 2.55 |
| Min | 3.93 | 3.93 | 9.63 | 9.63 |
| Median | 44.91 | 8.49 | 22.29 | 12.84 |
| Max | 100.21 | 24.90 | 44.31 | 19.46 |
| SE | 5.13 | 0.68 | 1.58 | 0.42 |
| CV | 0.70 | 0.45 | 0.41 | 0.18 |
| **Available P with Mn-rich waste** | | | | |
| Mean | 27.03 | 5.90 | 18.83 | 10.65 |
| SD | 17.96 | 3.38 | 7.55 | 2.61 |
| Min | 3.73 | 2.16 | 7.86 | 3.73 |
| Median | 24.89 | 5.55 | 18.12 | 10.46 |
| Max | 66.05 | 24.02 | 36.85 | 18.63 |
| SE | 2.99 | 0.56 | 1.26 | 0.44 |
| CV | 0.66 | 0.57 | 0.40 | 0.25 |
| **Extractable Mn without Mn-rich waste** | | | | |
| Mean | 11.94 | 9.63 | 11.94 | 14.43 |
| SD | 7.34 | 5.86 | 6.75 | 6.49 |
| Min | 4.38 | 3.97 | 0.09 | 6.10 |
| Median | 8.52 | 7.53 | 9.51 | 13.00 |
| Max | 28.84 | 23.81 | 27.00 | 29.34 |
| SE | 1.22 | 0.98 | 1.13 | 1.08 |
| CV | 0.61 | 0.61 | 0.57 | 0.45 |
| **Extractable Mn with Mn-rich waste** | | | | |
| Mean | 748.18 | 732.73 | 457.77 | 517.97 |
| SD | 421.57 | 409.41 | 276.29 | 318.28 |
| Min | 134.40 | 122.05 | 64.55 | 75.57 |
| Median | 952.75 | 926.75 | 495.14 | 662.94 |
| Max | 1173.43 | 1186.20 | 819.04 | 879.70 |
| SE | 70.26 | 68.23 | 46.05 | 53.05 |
| CV | 0.56 | 0.56 | 0.60 | 0.61 |

*SD, standard deviation; Min, minimum; Max, maximum; SE, standard error; CV, coefficient of variation.*

(Figure 2c). Surprisingly, in the same soil, 0P induced similar available phosphorus to the 2P in the second sampling event, and no significant variation was observed at the end of the incubation. In terms of Hutton soil amended with Mn-rich waste and the P applied in the form of rock phosphate, significant variations were observed at 30 and 60 days,

**Table 3:** Four-way multivariate analysis (ANOVA) of the effect of soil types (soil), present and absence of Mn waste (waste), P sources (source) and P application rates (rate), and their interaction on soil available P and extractable Mn in Cartref and Hutton soils

| | | Available P | | Extractable Mn | |
|---|---|---|---|---|---|
| **Source of variation** | **DF** | **F** | ***p*** | **F** | ***p*** |
| Soil | 1 | 4142.92 | <0.001 | 804.35 | <0.001 |
| Waste | 1 | 41.04 | <0.001 | 14.55 | <0.001 |
| Rate | 1 | 82.48 | <0.001 | 337.01 | <0.001 |
| Source | 5 | 93.37 | <0.001 | 0.09 | ns |
| Soil × Waste | 3 | 596.05 | <0.001 | 0.12 | ns |
| Soil × Rate | 5 | 16.95 | <0.001 | 15.11 | <0.001 |
| Waste × Rate | 5 | 34.58 | <0.001 | 0.13 | ns |
| Soil × Source | 3 | 13.97 | <0.001 | 0.08 | ns |
| Waste × Source | 3 | 163.71 | <0.001 | 0.38 | ns |
| Rate × Source | 5 | 24.62 | <0.001 | 0.11 | ns |
| Soil × Waste x Rate | 11 | 80.61 | <0.001 | 0.03 | ns |
| Soil × Waste × Source | 7 | 6.14 | <0.001 | 0.11 | ns |
| Soil × Rate × Source | 11 | 17.65 | <0.001 | 0.29 | ns |
| Waste × Rate × Source | 6 | 29.46 | <0.001 | 0.04 | ns |
| Soil × Waste × Source × Rate | 17 | 5.58 | <0.004 | 0.03 | ns |

*P, phosphorus; Mn, manganese; DF, degrees of freedom.*

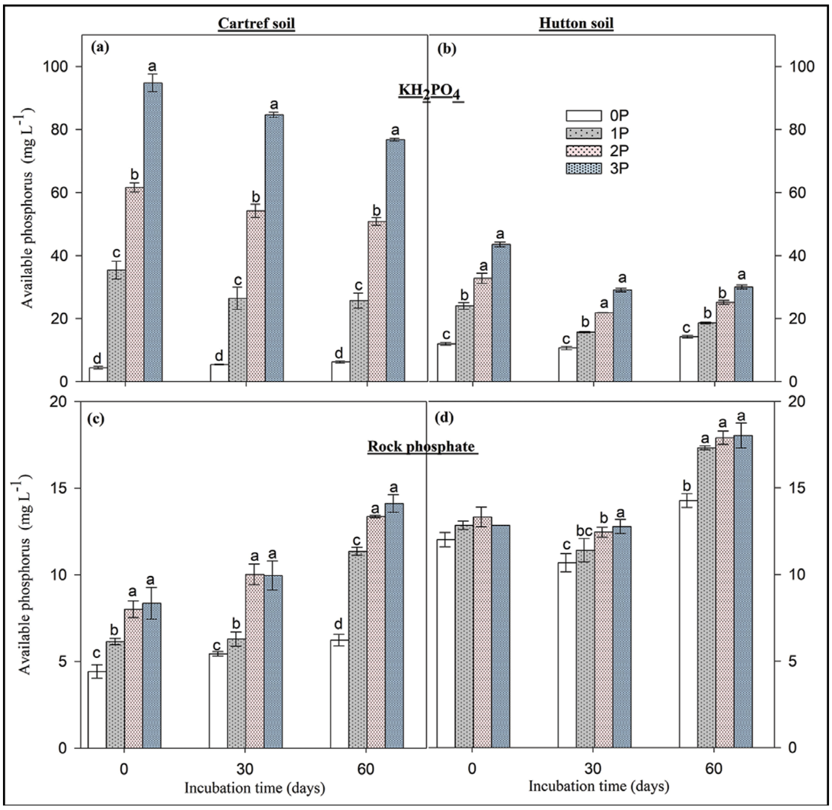



**Figure 1:** Mean ± standard error ($n = 3$) of soil available phosphorus (P) from soils (Cartref and Hutton) incubated with manganese-rich waste at an increasing dose of phosphorus (0P, 1P, 2P and 3P, corresponded to no addition, recommended P, double the recommended P and triple the recommended P, respectively) from $KH_2PO_4$ (a and b) and rock phosphate (c and d) sources. Within one P source at each time point, means followed by different letters are significantly different at $p \leq 0.05$.

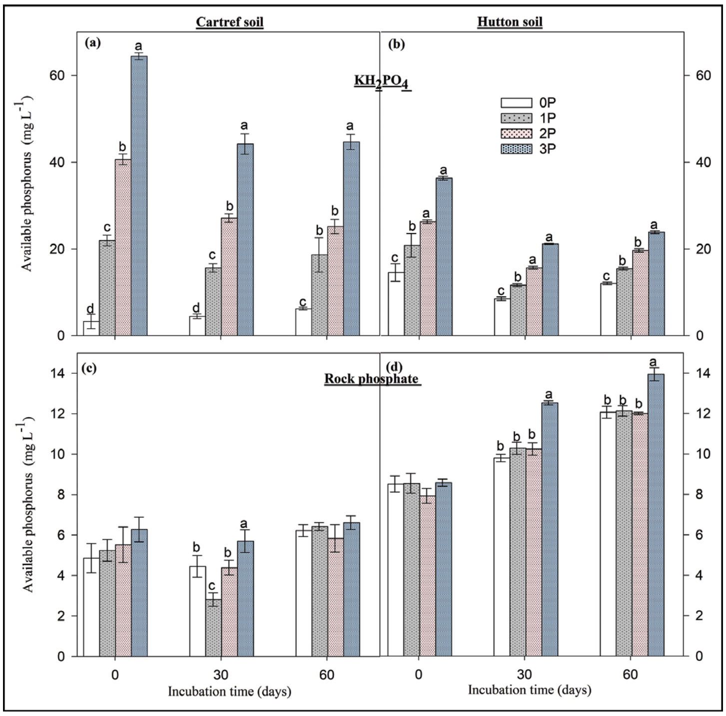



**Figure 2:** Mean ± standard error ($n = 3$) of soil available phosphorus (P) from soils (Cartref and Hutton) incubated with manganese-rich waste at an increasing dose of phosphorus (0P, 1P, 2P and 3P, corresponded to no addition, recommended P, double the recommended P and triple the recommended P, respectively) from $KH_2PO_4$ (a and b) and rock phosphate (c and d) sources. Within one P source at each time point, means followed by different letters are significantly different at $p \leq 0.05$.

where 3P induced the greatest available P in the soil (Figure 2d). Also, the available P increased with time in all the P application rates including 0P.

### *Extractable Mn in soils without and with Mn-rich wastes*

Extractable Mn in soil incubated without Mn-rich waste was the lowest at the start and the highest after 60 days of P dosing in both soil types (Figure 3). $KH_2PO_4$ P-form had significant variations in dose response at 30 and 60 days for Cartref, with 3P being highest after 30 days (Figure 3a). In contrast, Mn in Hutton soil was only significant at 30 days of the sampling event with the Mn being the highest (11.12±0.45 mg/L) at a 3P dose (Figure 3b). For the rock phosphate P source, a significant increase in Mn under Cartref soil was observed at the end of the incubation period (Figure 3c). Interestingly, 0P had the most Mn of 21.82±1.23 mg/L compared to the other P doses. In the Hutton soil, Mn was lowest in 0P dose and 1P being the highest for both 30 and 60 days. No significant differences were observed in Mn between 2P and 3P for 30 and 60 days. The addition of Mn-rich waste to both soil types increased Mn amounts with time for all cases (Figure 4). Although no significant variations were found in Mn between P doses from the $KH_2PO_4$ source in all the sampling events under Cartref soil (Figure 4b), this was not the case for Hutton soil (Figure 4b). In the Hutton soil amended with Mn-rich waste, Mn in the soil was highest in 3P at the start of the incubation and increased from 0P to the other P doses in the other sampling events; however, the variation was not significant at 60 days (Figure 4b). Under the rock phosphate source, Mn in Cartref soil followed a similar pattern to the $KH_2PO_4$ source (Figure 4c). In the Hutton soil amended with Mn waste and P addition from the rock phosphate source, Mn was significantly greater in 1P, 2P and 3P than 0P on the other at 30- and 60-day sampling events (Figure 4d).

### *Relationship between available P, extractable Mn and soil pH*

Figures 5 and 6 show the linear relationship between available P, extractable Mn and soil pH at 30 and 60-day sampling events from both soil types subjected to P additions from $KH_2PO_4$ and rock phosphate sources incubated with and without Mn-waste additions. No significant correlation was found in soils incubated without Mn waste (Figure 5) when the P source was $KH_2PO_4$ (Figure 5a and 5b). However, significant correlations were found when the P source added was rock phosphate in both soils, with positive trend relationships between P and Mn, and a negative relationship between P and pH (Figure 5c and 5d). The strongest positive correlation ($r^2 = 0.82$) was observed in the relationship between P and Mn in the Hutton soil subjected to P addition from a rock phosphate source (Figure 5d). Similar to the soil incubated without Mn waste, the soils incubated with Mn waste showed no significant correlations between the variables (P vs Mn and pH) when the P source was added as $KH_2PO_4$ (Figure 6a and 6b). However, the correlation between P and pH in the Cartref soil was not significant.

## Discussion

Although soil available P is already considered a key limitation to plant and microbial metabolism and populations, its sources and the expected interaction between P and other soil minerals are another determinant influenced by local soil conditions and the abundance of the other soil minerals.[25-27] Mn-rich solid waste resulting from Mn mining could already tilt this delicate balance by altering available soil P.

### *Impact of P sources on P availability in soils*

The available soil P increased significantly with the increase of $KH_2PO_4$ application rates in both soils (Figure 1a and 1b), likely because of the high P released per unit weight of the applied $KH_2PO_4$. Except for the 0P dose (no P added), the available P in both soil types slightly decreased over time, probably due to the P adsorption on soil colloids over time. Time was an important factor for P adsorption in the soils, which increases rapidly with the reaction time at the beginning of the incubation and decreases over time without reaching a stable exact equilibrium.[28,29] Rock phosphate resulted in lower available P compared to $KH_2PO_4$ in both

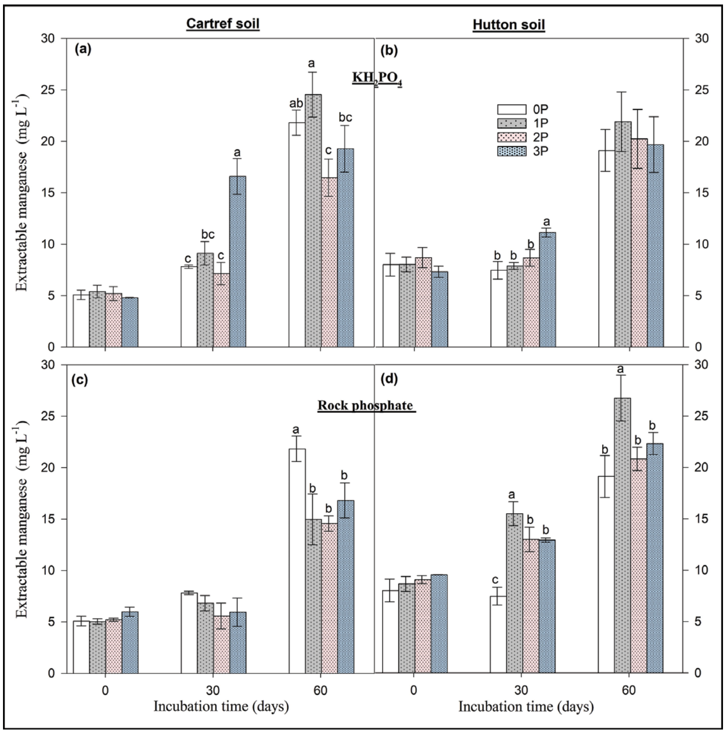



**Figure 3:** Mean ± standard error ($n = 3$) of soil extractable manganese (Mn) from soils (Cartref and Hutton) incubated without manganese-rich waste at an increasing dose of phosphorus (0P, 1P, 2P and 3P, corresponded to no addition, recommended P, double the recommended P and triple the recommended P, respectively) from $KH_2PO_4$ (a and b) and rock phosphate (c and d) sources. Within one P source at each time point, means followed by different letters are significantly different at $p \leq 0.05$.

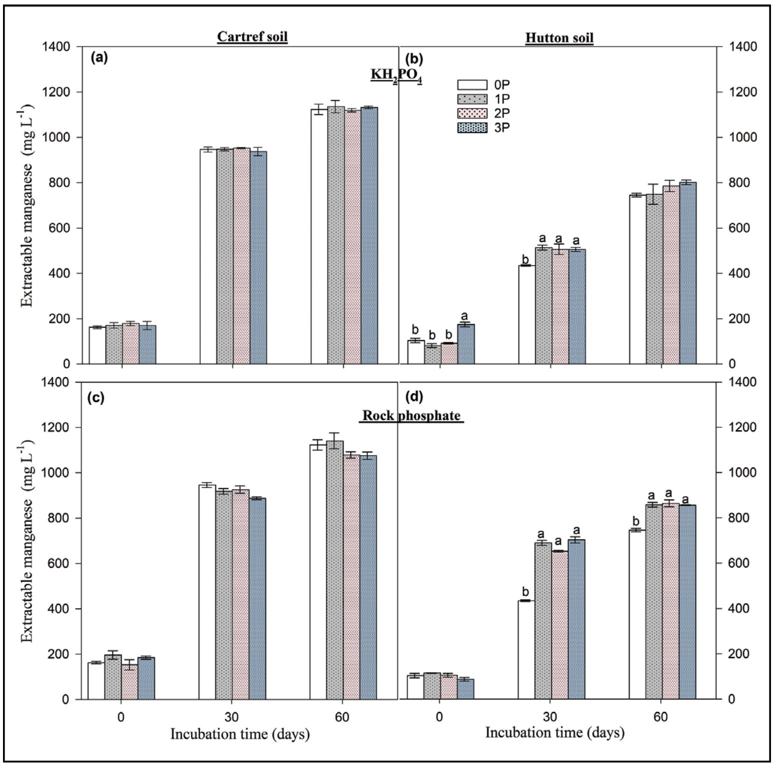



**Figure 4:** Mean ± standard error ($n = 3$) of soil extractable manganese (Mn) from soils (Cartref and Hutton) incubated with manganese-rich waste at an increasing dose of phosphorus (0P, 1P, 2P and 3P, corresponded to no addition, recommended P, double the recommended P and triple the recommended P, respectively) from $KH_2PO_4$ (a and b) and rock phosphate (c and d) sources. Within one P source at each time point, means followed by different letters are significantly different at $p \leq 0.05$.

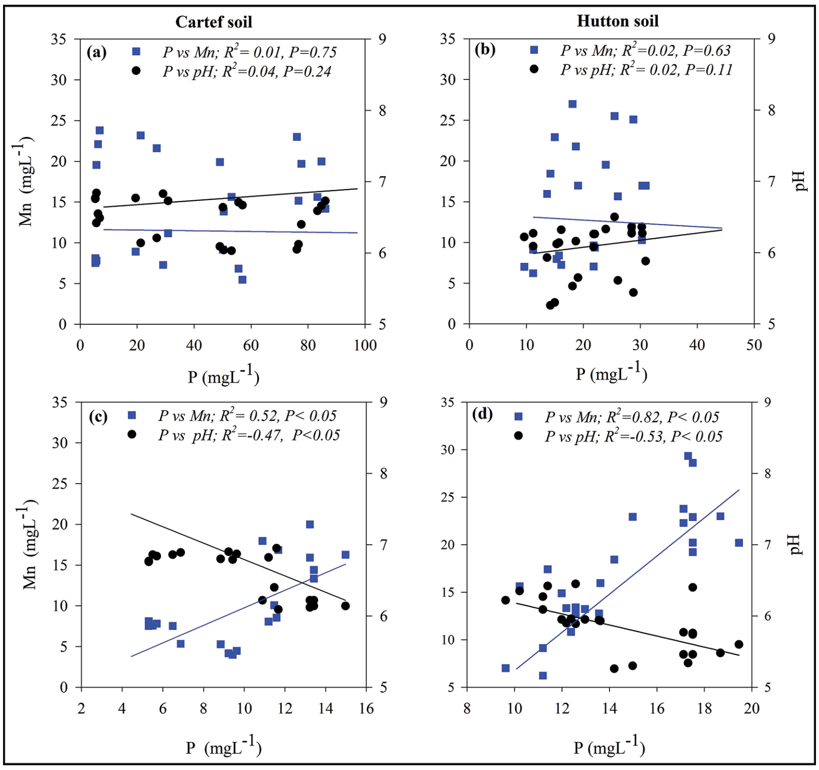



**Figure 5:** Relationship between soil available phosphorus (P) on one hand and soil manganese (Mn) and pH on the other hand in different soils (Cartref and Hutton) incubated without Mn waste with the phosphorus source of $KH_2PO_4$ (a and b) and rock phosphate (c and d) from 30- and 60-day incubation time (*n* = 24).

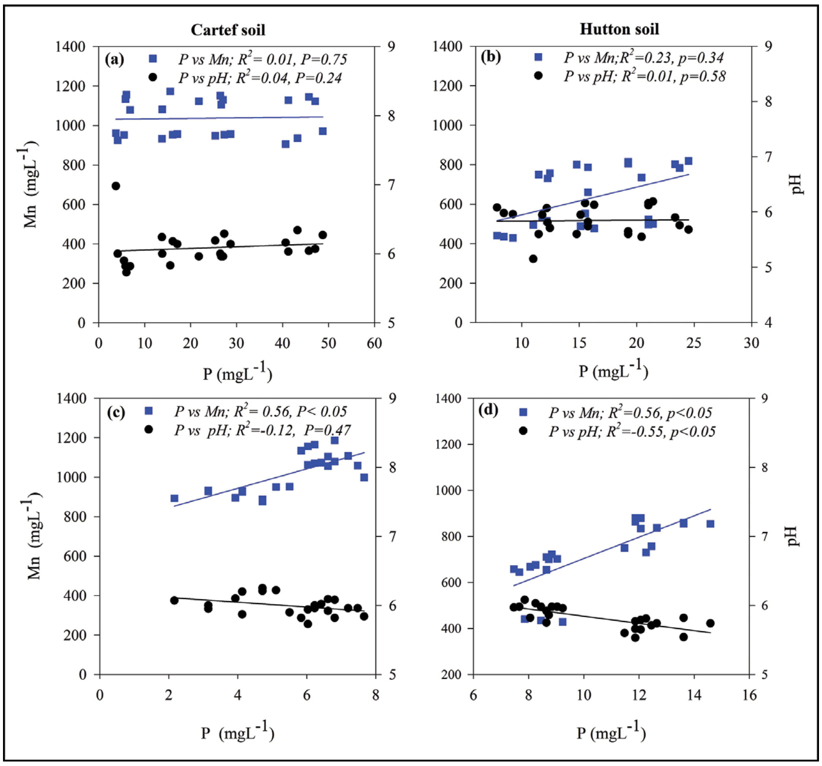



**Figure 6:** Relationship between soil available phosphorus (P) on one hand and soil manganese (Mn) and pH on the other hand in different soils (Cartef soil and Hutton soils) incubated with Mn waste with the phosphorus source of $KH_2PO_4$ (a and b) and rock phosphate (c and d) form 30- and 60-day incubation time (*n* = 24).

soils (Figures 1 and 2), suggesting slow P release from rock phosphate. Rock phosphate is characterised by low solubility[30,31], and therefore, the rock phosphate releases P slowly over time due to the continuation of the dissolution process, similar to the process that occurs in nature[32,33]. Increasing acidity, in general, will also improve P release from rock phosphate.[33,34] In the present study, the rock phosphate source showed an opposite pattern of available P where P release increased over the incubation time (Figure 1c and 1b) compared to the $KH_2PO_4$ source, indicating slow release of P from rock phosphate source.

The abundance of P increased by increasing the incubation period, not only in the initial P rate application, but also in the control soil without phosphate rock addition. Similar findings have been reported by many other studies.[35-37] In contrast, the available soil P from the $KH_2PO_4$ source is released immediately once the $KH_2PO_4$ is in contact with the soil and then decreases over time in both soils (Figure 1a and 1b). The interaction between incubation time and the rate of applied rock phosphate as well as other organic amendments has a positive impact on increasing the availability of P.[38,39] Furthermore, soil types that have different levels of carbon and clay content and incubation periods in soils have varying levels of microbes. Therefore, the augmentation of soils with Mn-rich waste may further limit microbial uptake of P over time. Despite the fact that the current study did not identify the possible role of soil microorganisms and their reaction to different P sources, their role is still important and cannot be ignored.

### *Impact of Mn-rich waste on P behaviour*

The decrease in available P in the soil treated with Mn-rich waste (Figure 2) can be explained by the sorption behaviour of P in a Mn-rich environment.[40] The P sorption can be influenced by various factors, such as soil pH, soil organic matter, ionic strength and cation types.[28,41,42] Accordingly, the reduction in the available P after the application of Mn-rich waste may be due to P absorption on the Mn oxide surface. Mustafa et al.[36] observed an increase in the sorption rate with the increase in P concentration; however, an opposite trend was observed with the soil pH. Another possible explanation for the reduction in available P in the soil amended with Mn-rich soils could be due to the dissolution and precipitation reactions.[15] It is well known that, in soils with high P content, the soluble P precipitates to form Fe, Al, Ca, Mg and Mn phosphate depending on the soil pH.[43,44]

In general, the reduction in P availability in the soils as a result of Mn-rich waste addition compared to soils without Mn waste agreed with other studies, for example, Lindsay et al.[15], Sample et al.[45] and Vassilev et al.[10] found that phosphate solubility was hindered due to the formation of Mn phosphate in slightly acidic soils. Moreover, other investigators found that the P adsorption was significantly affected by high soil Mn.[17,46] In fact, the Mn abundance in the soil does not only affect available soil P but also the P contents in plant tissues.[47,48]

### *Impact of Mn-rich waste on soil extractable Mn*

Application of Mn-rich waste causes higher Mn concentrations due to the nature of Mn-rich waste which contains a high amount of Mn (Table 1 and Figure 2). In Cartref soil (sandy soil), the Mn concentrations were higher than in Hutton soil which has a clay texture (Figure 4), expected due to lower reactivity, cation exchange capacity and organic matter than the more clay soil. These results clearly show that clay soil is much better than sandy soil regarding the amount of Mn released from the soil after applying Mn-rich waste, suggesting the high potential of clay soil for waste reclamation. In sandy texture, the abundance of Mn would result in Mn coating on the surface of sand grains as observed by previous studies.[49-51] Moreover, Ghasemi-Fasaei et al.[52] found an increase in Mn adsorption with increasing clay content, cation exchange capacity and organic matter in a highly calcareous soil in Southern Iran, which was partially explained by the low pH causing dissolution of oxide-bound Mn, leading to higher measured amounts of Mn.

### *Impact of soil texture on available P in soils amended with Mn-rich waste*

In the present study, the amount of available P in the sandy soil (Cartref) was found to be higher than in clayey soil (Hutton) regardless of Mn-rich waste or P sources (Figures 1 and 2), probably due to the higher specific surface area in clay soil compared to the sandy textural soil. Thus, the clay has more sorption capacity that reacts with P; consequently, more P adsorption and less P release in clayey soils. These results were in agreement with incubation experiments such as Kaloi et al.[37] and Rajput et al.[53] in India and Pakistan, respectively. Such findings suggest that sandy soil requires less amount of P fertiliser compared with clay soil.[37] The amounts of available P in soils subjected to $KH_2PO_4$ addition increased immediately at the beginning of the trial and then decreased by increasing incubation periods; this occurs because the absorption reaction slowly continued to increase with the increased contact time.[54,55] Therefore, clay content plays an important role in the retention of P in agricultural soils.[56,57]

Overall, the combined application of Mn-rich waste and rock phosphate to soils can be beneficial, not only for plant productivity but also for increasing the circular economy of mining by-product waste. However, caution is needed for determining the optimum dose of Mn-rich waste based on the soil texture to avoid groundwater contamination and Mn toxicity to plant and soil microbes.[48,58] However, Berríos et al.[59] found that Mn toxicity can be ameliorated by a high P supply using the Mn-resistant ryegrass genotype. Therefore, future research should consider an integrated approach using Mn-resistant plants, Mn-rich waste and sustainable P source in different soil textures to design an optimum dose that utilises the Mn-rich waste without compromising on the soil biodiversity and the environment. Another key limitation of the present study is not including possible microbial mechanisms that break down or use P for growth and activities as the microbial levels control.

## Conclusions

Sandy and clay soils subjected to different phosphorus doses from $KH_2PO_4$ and rock phosphate were incubated with and without manganese-rich solid waste for 60 days. Available phosphorus and extractable manganese were determined at three time intervals to evaluate the interaction effect of manganese abundance on soil phosphorus. Manganese-rich waste caused a reduction in the available soil phosphorus which was found to be greater from the readily available phosphorus source (i.e. $KH_2PO_4$) than the rock phosphate source. Therefore, the rock phosphate source is more beneficial in clayey soils amended with manganese-rich waste because of the slow release over time implying a better plant growth environment. Alternatively, the manganese-rich waste might be used to immobilise phosphorus in soils that contain excess available phosphorus. However, further research using crops known for their phytoremediation capacity to reduce the manganese abundance in the soil amended with manganese-rich waste is still required.

## Acknowledgements

We thank the technical staff of the Department of Soil Science, School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal for their support.

## Competing interests

We have no competing interests to declare.

## Authors' contributions

M.A.E: Conceptualisation, methodology, data collection, sample analysis, data analysis, validation, data curation and writing – the initial draft. K.A.: Data collection, sample analysis, data analysis, figure generation and writing – revisions. L.T.: Conceptualisation, methodology and writing – revision. P.M.: Conceptualisation, supervision, project leadership, project management and funding acquisition.

## References

1. Zhang X, Yang L, Li Y, Li H, Wang W, Ye B. Impacts of lead/zinc mining and smelting on the environment and human health in China. Environ Monit Assess. 2012;184(4):2261–2273. https://doi.org/10.1007/s10661-011-2115-6

2. Žibret G, Gosar M, Miler M, Alijagić J. Impacts of mining and smelting activities on environment and landscape degradation-Slovenian case studies. Land Degrad Dev. 2018;29(12):4457–4470. https://doi.org/10.1002/ldr.3198

3. Takam Tiamgne X, Kalaba FK, Nyirenda VR. Mining and socio-ecological systems: A systematic review of sub-Saharan Africa. Resour Policy. 2022;78:102947. https://doi.org/10.1016/j.resourpol.2022.102947

4. Ericsson M, Löf O. Mining's contribution to national economies between 1996 and 2016. Miner Econ. 2019;32(2):223–250. https://doi.org/10.1007/s13563-019-00191-6

5. Steenkamp JD, Basson J. The manganese ferroalloys industry in southern Africa. J South Afr Inst Min Metall. 2013;113(8):667–676. http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S2225-62532013000800014

6. Van Averbeke N. SAMI - South Africa's minerals industry 2004-2005. Pretoria: Department of Mineral Resources; 2005. p. 195. www.scielo.org.za/scielo.php?script=sci_nlinks&ref=2715819&pid=S2225-6253201300080001400029&lng=en

7. Beukes NJ, Swindell EPW, Wabo H. Manganese deposits of africa. Episodes J Int Geosci. 2016;39(2):285–317. https://doi.org/10.18814/epiiugs/2016/v39i2/95779

8. Li J, Jia Y, Dong R, Huang R, Liu P, Li X, et al. Advances in the mechanisms of plant tolerance to manganese toxicity. Int J Mol Sci. 2019;20(20):5096. https://doi.org/10.3390/ijms20205096

9. Millaleo R, Reyes- Diaz M, Ivanov AG, Mora ML, Alberdi M. Manganese as essential and toxic element for plants: Transport, accumulation and resistance mechanisms. J Soil Sci Plant Nutr. 2010;10(4):470–481. https://doi.org/10.4067/S0718-95162010000200008

10. Alewell C, Ringeval B, Ballabio C, Robinson DA, Panagos P, Borrelli P. Global phosphorus shortage will be aggravated by soil erosion. Nat Commun. 2020;11(1):4546. https://doi.org/10.1038/s41467-020-18326-7

11. Vassilev N, Vassileva M, Fenice M, Federici F. Immobilized cell technology applied in solubilization of insoluble inorganic (rock) phosphates and P plant acquisition. Bioresour Technol. 2001;79(3):263–271. https://doi.org/10.1016/S0960-8524(01)00017-7

12. Barben SA, Hopkins BG, Jolley VD, Webb BL, Nichols BA. Phosphorus and manganese interactions and their relationships with zinc in chelator-buffered solution grown russet burbank potato. J Plant Nutr. 2010;33(5):752–769. https://doi.org/10.1080/01904160903575964

13. Pedas P, Husted S, Skytte K, Schjoerring J. Elevated phosphorus impedes manganese acquisition by barley plants. Front Plant Sci. 2011;2. https://doi.org/10.3389/fpls.2011.00037

14. Kronvang B, Vagstad N, Behrendt H, Bøgestrand J, Larsen SE. Phosphorus losses at the catchment scale within Europe: An overview. Soil Use Manag. 2007;23(s1):104–116. https://doi.org/10.1111/j.1475-2743.2007.00113.x

15. Lindsay WL, Stephenson HF. Nature of the reactions of monocalcium phosphate monohydrate in soils: II. Dissolution and precipitation reactions involving iron, aluminum, manganese, and calcium. Soil Sci Soc of Am J. 1959;23(1):18–22. https://doi.org/10.2136/sssaj1959.03615995002300010013x

16. Shahandeh H, Hossner LR, Turner FT. Phosphorus relationships to manganese and iron in rice soils. Soil Sci. 2003;168(7):489–500. https://doi.org/10.1097/01.ss.0000080334.10341.6a

17. Krairapanond A, Jugsujinda A, Patrick WH. Phosphorus sorption characteristics in acid sulfate soils of Thailand: Effect of uncontrolled and controlled soil redox potential (Eh) and pH. Plant Soil. 1993;157(2):227–237. https://doi.org/10.1007/BF00011051

18. Marsh KB, Peterson LA, McCown BH. A microculture method for assessing nutrient uptake II. The effect of temperature on manganese uptake and toxicity in potato shoots. J Plant Nutr. 1989;12(2):219–232. https://doi.org/10.1080/01904168909363947

19. Rosas A, Rengel Z, Ribera A, de La Luz Mora M. Phosphorus nutrition alleviates manganese toxicity in *Lolium perenne* and *Trifolium repens*. J Plant Nutr Soil Sci. 2011;174(2):210–219. https://doi.org/10.1002/jpln.201000104

20. Titshall L. Revegetation and phytoremediation of tailings from a lead/zinc mine and land disposal of two manganese-rich wastes [PhD thesis],. Pietermaritzburg: University of KwaZulu-Natal; 2007: https://researchspace.ukzn.ac.za/xmlui/handle/10413/3512

21. Soil Classification Working Group. Soil classification: A taxonomic system for South Africa. Pretoria: Dept. of Agricultural Development; 1991.

22. IIUS-WRB. World reference base for soil resources 2014: International soil classification system for naming soils and creating legends for soil maps. Rome: FAO; 2014.

23. Non-Affiliated Soil Working Committee. Handbook of standard soil testing methods for advisory purposes. Soil Science Society of South Africa; 1990.

24. Liang J, Karamanos RE. DTPA-extractable Fe, Mn, Cu and Zn. In: Carter MR, editor. Soil sampling and methods of analysis. Ann Arbor, MI: Canadian Society of Soil Science, Lewis Publishers; 1993. p. 87–90.

25. Oliverio AM, Bissett A, McGuire K, Saltonstall K, Turner BL, Fierer N. The role of phosphorus limitation in shaping soil bacterial communities and their metabolic capabilities. mBio, ASM J. 2020;11(5), e01718-20. https://doi.org/10.1128/mBio.01718-20

26. Cui Y, Bing H, Moorhead DL, Delgado-Baquerizo M, Ye L, Yu J, et al. Ecoenzymatic stoichiometry reveals widespread soil phosphorus limitation to microbial metabolism across Chinese forests. Commun Earth Environ. 2022;3(1):1–8. https://doi.org/10.1038/s43247-022-00523-5

27. Bindraban PS, Dimkpa CO, Pandey R. Exploring phosphorus fertilizers and fertilization strategies for improved human and environmental health. Biol Ferti Soils. 2020;56(3):299–317. https://doi.org/10.1007/s00374-019-01430-2

28. Asomaning SK. Processes and factors affecting phosphorus sorption in soils. In: Kyzas G, Lazaridis N, editors. Sorption in 2020s. IntechOpen; 2020. https://doi.org/10.5772/intechopen.90719

29. Toreu BN, Thomas FG, Gillman GP. Phosphate sorption characteristics of soils of the north Queensland coastal region. Soil Res. 1988;26(3):465–477. https://doi.org/10.1071/SR9880465

30. Korzeniowska J, Stanisławska-Glubiak E, Hoffmann J, Górecka H, Jóźwiak W, Wiśniewska G. Improvement of the solubility of rock phosphate by co-composting it with organic components. Pol J Chem Techol. 2013;15(4):10–14. https://doi.org/10.2478/pjct-2013-0060

31. Nakamura S, Kanda T, Imai T, Sawadogo J, Nagumo F. Solubility and application effects of African low-grade phosphate rock calcinated with potassium carbonate. Soil Sci Plant Nutr. 2019;65(3):267–273. https://doi.org/10.1080/00380768.2019.1598236

32. FAO. Use of phosphate rocks for sustainable agriculture. In: Zapata F, Roy RN, editors. FAO Fertilizer and Plant Nutrition Bulletin. A joint publication of the FAO Land and Water Development Division and International Atomic Energy Agency (IAEA). Rome: FAO; 2004. https://www.fao.org/documents/card/en/c/a14279f8-af75-5253-9b57-65b7d5dc4069/

33. Roy T, Biswas DR, Datta SC, Sarkar A. Phosphorus release from rock phosphate as influenced by organic acid loaded nanoclay polymer composites in an alfisol. Proc Natl Acad Sci India Sect B Biol Sci. 2018;88(1):121–132. https://doi.org/10.1007/s40011-016-0739-6

34. Ivanova RP, Bojinova DY, Gruncharov IN, Damgaliev DL. The solubilization of rock phosphate by organic acids. Phosphorus Sulfur Silicon Relat Elem. 2006;181(11):2541–2554. https://doi.org/10.1080/10426500600758399

35. Dodor DE, Tokashiki Y, Oya K, Shimo M. Dissolution of phosphate rock fertilisers in some soils of Okinawa. Japan Soil Res. 1999;37(1):115–122. https://doi.org/10.1071/S98061

36. Hellal F, El-Sayed S, Zewainy R, Amer A. Importance of phosphate pock application for sustaining agricultural production in Egypt. Bull Natl Res Cent. 2019;43(1):11. https://doi.org/10.1186/s42269-019-0050-9

37. Kaloi GM, Bhughio N, Panhwar RN, Junejo S, Mari AH, Bhutto MA. Influence of incubation period on phosphate release in two soils of district Hyderabad. J Anim Plant Sci. 2011;21(4). http://www.thejaps.org.pk

38. Agbenin JO, Tiessen H. Phosphorus sorption at field capacity and soil ionic strength: Kinetics and transformation. Soil Sci Soc Am J. 1995;59(4):998–1005. https://doi.org/10.2136/sssaj1995.03615995005900040006x

39. Radwan SA, Shalaby MMH, Nada WA, Ashry SM, Seeda MA, Sisi NAE. Effect of added rock phosphate and compost on soil phosphorus fractions after different incubation periods. Menoufia J Soil Sci. 2020;5(6):149–157. https://doi.org/10.21608/mjss.2020.171482

40. Zhu Y, Du X, Gao C, Yu Z. Adsorption behavior of inorganic and organic phosphate by iron manganese plaques on reed roots in wetlands. Sustainability. 2018;10(12):4578. https://doi.org/10.3390/su10124578

41. Mustafa S, Zaman MI, Khan S. pH effect on phosphate sorption by crystalline $MnO_2$. J Colloid Interface Sci. 2006;301(2):370–375. https://doi.org/10.1016/j.jcis.2006.05.020

42. Ryden JC, Syers JK. Rationalization of ionic strength and cation effects on phosphate sorption by soils. J Soil Sci. 1975;26(4):395–406. https://doi.org/10.1111/j.1365-2389.1975.tb01963.x

43. Gichangi EM, Mnkeni PNS, Muchaonyerwa P. Phosphate sorption characteristics and external P requirements of selected South African soils. J Agric Rural Dev Trop Subtrop. 2008;109(2):139–149. https://www.jarts.info/index.php/jarts/article/view/65

44. Johan PD, Ahmed OH, Omar L, Hasbullah NA. Phosphorus transformation in soils following co-application of charcoal and wood ash. Agronomy. 2021;11(10):2010. https://doi.org/10.3390/agronomy11102010

45. Sample EC, Soper RJ, Racz GJ. Reactions of phosphate fertilizers in soils. In: The role of phosphorus in agriculture. John Wiley & Sons, Ltd; 1980. p. 263–310. https://doi.org/10.2134/1980.roleofphosphorus.c12

46. Jugsujinda A, Krairapanond A, Patrick WH. Influence of extractable iron, aluminium, and manganese on P-sorption in flooded acid sulfate soils. Biol Ferti Soils. 1995;20(2):118–124. https://doi.org/10.1007/BF00336590

47. Gunes A, Alpaslan M, Inal A. Critical nutrient concentrations and antagonistic and synergistic relationships among the nutrients of NFT-grown young tomato plants. J Plant Nutr. 1998;21(10):2035–2047. https://doi.org/10.1080/01904169809365542

48. Nogueira MA, Nehls U, Hampp R, Poralla K, Cardoso EJBN. Mycorrhiza and soil bacteria influence extractable iron and manganese in soil and uptake by soybean. Plant Soil. 2007;298(1):273–284. https://doi.org/10.1007/s11104-007-9379-1

49. Hu PY, Hsieh YH, Chen JC, Chang CY. Characteristics of manganese-coated sand using SEM and EDAX analysis. J Colloid Interface Sci. 2004;272(2):308–313. https://doi.org/10.1016/j.jcis.2003.12.058

50. Junta-Rosso JL, Hochella MF, Donald Rimstidt J. Linking microscopic and macroscopic data for heterogeneous reactions illustrated by the oxidation of manganese (II) at mineral surfaces. Geochim Cosmochim Acta. 1997;61(1):149–159. https://doi.org/10.1016/S0016-7037(96)00329-8

51. Yuan L, Sun L, Fortin D, Wang Y, Yin X. Microscale characterization and trace element distribution in bacteriogenic ferromanganese coatings on sand grains from an intertidal zone of the East China Sea. PLoS ONE. 2015;10(3), e0119080. https://doi.org/10.1371/journal.pone.0119080

52. Ghasemi-Fasaei R, Jarrah M, Mayel S. Dynamics of manganese adsorption onto highly calcareous soils of southern Iran. Commun Soil Sci Plant Anal. 2009;40(7–8):1171–1182. https://doi.org/10.1080/00103620902754333

53. Rajput A, Panhwar QA, Naher UA, Rajput S, Hossain E, Shamshuddin J. Influence of incubation period, temperature and different phosphate levels on phosphate adsorption in soil. Am J Agric Biol Sci. 2014;9(2):251–260. https://doi.org/10.3844/ajabssp.2014.251.260

54. Huang Q-h, Li D-m, Liu K-l, Yu X-c, Ye H-c, Hu H-w, et al. Effects of long-term organic amendments on soil organic carbon in a paddy rield: A case study on red soil. J Integr Agric. 2014;13(3):570–576. https://doi.org/10.1016/S2095-3119(13)60714-5

55. Indinti R, Coppola E, Buondonno A. Changes of soil phosphorus availability in Italian alfisols as estimated by short-term soil + phosphorus equilibration procedures using olsen, mehlich 3, and paper-strip methods. Commun Soil Sci Plant Anal. 1999;30(7-8):983–997. https://doi.org/10.1080/00103629909370262

56. Kanabo I-a. K, Gilkes RJ. The effect of soil texture on the dissolution of North Carolina phosphate rock. J Soil Sci. 1988;39(2):191–198. https://doi.org/10.1111/j.1365-2389.1988.tb01205.x

57. Suñer L, Galantini J. Texture influence on soil phosphorus content and distribution in semiarid pampean grasslands. Int J Plant Soil Sci. 2015;7:109–120. https://doi.org/10.9734/IJPSS/2015/16939

58. El-Jaoual T, Cox DA. Manganese toxicity in plants. J Plant Nutr. 1998; 21(2):353–286. https://doi.org/10.1080/01904169809365409

59. Berríos GA, Luengo Escobar A, Alberdi MR, Nunes-Nesi A, Reyes-Díaz MM. Manganese toxicity amelioration by phosphorus supply in contrasting Mn resistant genotypes of ryegrass. Plant Physiol Biochem. 2019;144:144–156. https://doi.org/10.1016/j.plaphy.2019.09.034

# Quantification of toxic metals in cropland soil using X-ray fluorescence

AUTHORS:
Mashinga J. Mvelase[1]
Paulus L. Masiteng[1]

AFFILIATION:
[1]Department of Physics, University of Johannesburg, Johannesburg, South Africa

CORRESPONDENCE TO:
Mashinga Mvelase

EMAIL:
emjay.mvelase@gmail.com

DATES:
**Received:** 17 Nov. 2022
**Revised:** 24 May 2023
**Accepted:** 25 May 2023
**Published:** 28 Sep. 2023

HOW TO CITE:
Mvelase MJ, Masiteng PL. Quantification of toxic metals in cropland soil using X-ray fluorescence. S Afr J Sci. 2023;119(9/10), Art. #15008. https://doi.org/10.17159/sajs.2023/15008

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☐ All data included
☒ On request from author(s)
☐ Not available
☐ Not applicable

EDITOR:
Michael Inggs

KEYWORDS:
XRF spectroscopy, heavy metal toxicity, eutrophication, soil pollution

FUNDING:
None



We aimed to assess toxic heavy metals in soil samples from cropland in Weenen (KwaZulu-Natal, South Africa) using X-ray fluorescence (XRF) spectroscopy. The metal contents in the soil samples were detected and quantified by wavelength dispersive XRF (WD-XRF) spectroscopy. On average, the values of all elements (mg/kg) were: Al (91.4 ± 6.9), Ba (0.488), Ca (16.8 ± 5), Fe (39.3 ± 0.8), K (15.7 ± 0.04), Mg (10.1 ± 0.3), Mn (0.6), Na (8.0 ± 1), P (1.3 ± 0.4), Si (458 ± 8) and Ti (5.6 ± 0.3). Toxic metals such as Hg, Cd, As, Pb, and Cr were not detected in the soil samples. The macronutrient P, which is capable of causing eutrophication in water bodies, was present at a low level in soil samples. The metal contents in both control and field samples were comparable, suggesting that the metals were mostly of lithogenic origin and not entirely influenced by anthropogenic activities. The metal levels we detected were within the limits reported to be safe by other studies.

**Significance:**

Although they are within the detection range, the toxic heavy metals mercury, cadmium, arsenic, lead and chromium, which enter the environment through the use of fertilisers on agricultural land, were not detected in this WDXRF spectrometry analysis. The remaining elements detected are lithogenic and non-anthropogenic, as the field and control samples had comparable concentrations. Nitrogen was not quantified, but phosphorus was present in a low concentration, so the field water run-off into the water bodies need not directly lead to water pollution in the area studied.

## Introduction

Soil is a product of weathering and decomposition of organic matter. The weathering process involves the interaction of the lithosphere, atmosphere, and hydrosphere, and is driven by solar energy.[1] Weathering and volcanic eruptions contribute to heavy metals in the environment.[2-5] Pollution contributes to health risks, as many diseases are airborne or waterborne, and spread from soil to humans through food.[6] Worldwide, air pollution causes about 3.1 million premature deaths.[7-9] These pollutants enter the soil through atmospheric deposition[10] and are subsequently absorbed by plants that are consumed by humans and other animals. Environmental impact assessments and pollution controls should be routinely carried out to protect new developments from pollution. The soil ecosystem can be damaged by high concentrations of toxic metals, so it is important to assess the ecological and environmental risks from heavy metals.[11]

Extensive use of fertilisers spreads heavy metals in the environment, as phosphate rock is heavily contaminated with heavy metals[10,12-14], especially with the radioactive decay series $^{238}U$, $^{232}Th$ and $^{40}K$. $^{238}U$, $^{232}Th$ and $^{40}K$ are ubiquitous in the earth's crust[10,15] and are used by organisms in small amounts; they are considered essential metals that can be naturally present in plants or can enter the soil through high concentrations. The distribution of these metals in the soil depends on the region and the type of parent rock from which the soil originates.[16] An assessment of heavy metal contamination of soil is very important, as this contamination can enter the food chain, causing toxicity in plants and animals. Soil contamination can spoil land for agricultural use[17,18], exacerbating poverty worldwide. Food grown on contaminated soil can be toxic for human and animal consumption.

Micronutrients such as iron (Fe), manganese (Mn), zinc (Zn) and copper (Cu)[19] are essential for plant growth, but their high content in plant tissues[20], coupled with the presence of mercury (Hg), cadmium (Cd), arsenic (As), lead (Pb) and chromium (Cr), have a toxic effect.[21-25] The heavy metals Cu, Zn, Pb, As and Cd[6,10,18-20,24] are termed toxic metals because gastrointestinal cancer and diseases of the skeletal, cardiovascular and nervous systems have been linked with their excessive intake[6,25]. Exposure to excessive Cu is linked with cell damage in humans.[26,27] Pb is not only toxic to the health of the soil ecosystem[11] but also poses a health risk to humans, causing brain damage and mental illness as well as kidney damage in children.[18,23,24] Cd and As enter agricultural soils through the main pathways, i.e. volcanic eruptions, soil erosion[10] and fertiliser application. Plants take up Cd through their roots, and this is a pathway through which non-smokers are exposed to this metal.[28] Phosphate deposits are divided into sedimentary and magmatic deposits. A total of 80% of phosphate rock comes from sedimentary deposits[29], from which phosphate fertilisers are produced. They contribute the most Cd[30] to the environment, and waste disposal also contributes significantly.

The anthropogenic increase of nitrogen in water bodies directly threatens human and aquatic ecosystems. Water bodies include streams, rivers, lakes, aquifers, and coastal waters, as well as groundwater. The accumulation of some nitrates from the soil in water bodies can lead to water pollution known as eutrophication.[22,31] Eutrophication is a global problem[32] in which water bodies receive excess nutrients from wastewater and agricultural sources based on, for example, nitrogen and phosphorus.[33] Eutrophication can have economic consequences as well as effects on tourism, human health, and water quality.[34] Eutrophication leads to hypoxia in water bodies, resulting in the death of plants and fish in the water and toxic algal blooms.[22,31] Eutrophication can occur in oligotrophic (low nutrient levels), mesotrophic (moderate nutrient levels), eutrophic (good nutrient levels), and hypertrophic (very good nutrient levels) water bodies. The South African Department of Water and Sanitation's Eutrophication National

Monitoring Programme (NEMP) has published data showing that about 28% of surface water samples are hypertrophic, 33% eutrophic, 37% mesotrophic and only 3% oligotrophic.[33]

Cyanobacteria (blue-green algae) thrive in eutrophic water[33], which can pose a health risk in underdeveloped areas because people still use water directly from the source for activities such as washing, drinking and food preparation[33,35] and cyanobacteria produce cyanotoxins[36]. Aquatic ecosystems can remove a significant amount of nitrogen enrichment, but this capacity has limits and depends on the characteristics of the ecosystem.[37] Fresh water is a limited resource that needs to be protected everywhere[33,38] and South Africa is no exception. In southern African regions, water is collected from springs in 20- to 25-litre plastic containers, which can lead to disease-causing bacteria being transported with the water or developing in the containers.[35]

Soil samples are analysed for the presence of toxic metals using X-ray fluorescence (XRF) spectroscopy. This technique has the advantage of being multi-elemental, non-destructive, fast, accurate and precise.[39,40] XRF can determine the chemical composition of all types of materials and determine the thickness and composition of layers and coatings.[41] These X-rays have characteristic energies related to the atomic number of an element, and each element, therefore, has a characteristic X-ray spectrum. Similar methods include atomic absorption spectroscopy (AAS), inductively coupled plasma atomic emission spectroscopy (ICP-AES), particle-induced X-ray emission (PIXE), inductively coupled plasma mass spectroscopy (ICP-MS) and inductively coupled plasma optical emission spectroscopy (ICP-OES).[1,42] XRF is capable of determining the broad content of elements in soil samples. The limitations are that XRF is not able to quantify light elements smaller than sodium but it is indeed fast and can process 1 million counts per second. Wavelength dispersive X-ray fluorescence (WD-XRF) spectrometry is preferred for its high-resolution applications (15–150eV)

ICP-MS is a highly sensitive method that gives quantitative results of elements and various isotopes in samples at milligram to nanogram levels per litreat a small per litre, but it is so expensive that many environmental laboratories cannot afford it, and daily operation is also costly.[40] ICP-OES quantification is based on the measurement of excited atoms and ions at the wavelength characteristics of the specific elements being measured, whereas ICP-MS measures the mass of an atom by mass spectrometry (MS). Both ICP-MS and ICP-OES are suitable for determining element concentrations in water, for testing water purity, and for use in materials science.[43] The problem again is affordability. The applications of all the techniques mentioned above overlap, therefore, any technique that is accessible to the researcher can be used for analysis.

PIXE, on the other hand, which measures elements from magnesium onwards, is generally better suited for determining the content of toxic metals and the portable unit costs about USD200. The cost of WD-XRF instruments is high, depending on which instrument you prefer. ICP-MS is a highly sensitive technique that provides quantitative results for various elements and isotopes in samples ranging from milligrams to nanograms per litre. On average, the initial cost of ICP-MS makes it the most expensive system to purchase. SPECTRAU is Spectrum's (The Central Analytical Facility of the Faculty of Science, University of Johannesburg) mission to provide state-of-the-art equipment and expertise to solve analytical challenges in academic research. As these methods work on the same bases, the availability and easy accessibility of WD-XRF at SPECTRAU made it the choice for this study.

In this study, we aimed to assess toxic heavy metals in soil samples from cropland using XRF spectroscopy. The lack of a similar study in this area was the reason for this work. Our results can serve as a basis for future anthropogenic environmental impact assessments on agricultural fields in Weenen and elsewhere in KwaZulu-Natal, South Africa.

## Materials and methods

### *Study area and soil sampling*

Batches of representative soil samples were collected randomly in Weenen, KwaZulu-Natal, South Africa, in January 2017. These batches consisted of a control sample spot (C17) and a field sample spot (A17). Figure 1 below shows a Google Earth view of the site at which samples were collected. Sample codes A1, A2, A3 and C18 (Control) represent spots or crop fields where additional sampling took place in August 2018. It is important to note that sample codes A17 and A1 represent the same sampling site, but the sampling took place in different years and seasons. Control samples were collected about 200–250 m from the crop field. Samples were collected at a depth of 15–20 cm and placed in 2-litre plastic containers. The study area was located on a farm about 17 km northeast of Weenen. The small town of Estcourt, with factories, a mine and businesses, is 57 km upstream.

### *Climate*

The climate in the study area is semi-arid with an average maximum temperature of about 27 °C and a wind direction from west to east. January, like December, is a moderately hot summer month in Weenen, with average temperatures ranging from a maximum of 27 °C to a minimum of 14.7 °C. The last month of winter, August, is pleasant in Weenen with an average temperature of 20.9 °C. There are about 128 days of rainfall and 409 mm of rainfall during the year.

Potatoes (*Solanum tuberosum*), harvested in November and December, and butternuts (*Cucurbita moschata*), harvested in April, are grown on the farm. Sometimes maize is also grown on the farm as part of crop rotation. The farm is not dependent on rain but it relies on an irrigation system that diverts water from the nearby river to the cultivated areas. Upstream, the river flows through the small town of Estcourt, which has some factories, a quarry mine and businesses. The farm is far from the mining and industrial areas. Most of the land is used for growing agricultural produce.

### *Sample preparation*

After collection, soil samples were oven dried at 105 °C at the SPECTRAU facility at the main campus of the University of Johannesburg, where the analysis of the samples took place. The dried samples were then crushed and pulverised at the University of Johannesburg's Mill Laboratory at the Doornfontein Campus to obtain a uniform powder. XRF was used to measure several elements that have a detection limit of about 0.05 mass per cent. The XRF technique scans for most elements from sodium (Na) tothose with higher atomic numbers. The scans are semi-quantitative. An amount of 1 g of the dried sample was then burnt or ignited in the air at 930 °C for 30 min and then the loss on ignition (LOI) was measured. LOI refers to the volatile content of the sample, such as organic compounds,


**Figure 1:** A Google Earth map showing the sites at which soil sampling took place.

$H_2O$ from hydrated phases and $CO_2$ from carbonates. The analysis was performed on the remaining 0.700 g after ignition removed 0.300 g of the 1.00 g soil sample.

### *Measurement technique and equations*

When an atom is irradiated with the appropriate amount of energy, an electron is ejected from the K shell. To fill the space, an electron falls out of the outermost shell into a quantum empty state. This descent is accompanied by the emission of an X-ray photon, which is determined by its destination and origin. That is, the K and L lines are formed when a hole is formed in the K and L layers. The electron jump leads to an X-ray absorption jump. The wavelength corresponding to this transition is called the absorption edge (band) of the respective element. An absorption edge is a feature of each chemical element that is used to uniquely identify the element based on its edge position.[43] When a vacancy in the K shell is filled with an electron in the L or M shell, $K_\alpha$ or $K_\beta$ rays are emitted.

In this study, the wavelength dispersive XRF technique is used to determine the elements in soil samples. X-rays are diffracted through single crystal or multilayer optics to select specific narrow wavelengths or energy ranges that often correspond to the characteristic X-rays of the element of interest. This allows for high specificity and sensitivity of the elements due to a high signal-to-background ratio.[41,44] The WD-XRF system physically separates the rays according to their wavelengths. The WD-XRF system is suitable for both major and trace elements.[44] Figure 2 shows the XRF system used for analysing samples.[45]

WD-XRF spectroscopy measures X-ray intensity as a function of wavelength. This is done by passing the radiation emitted by the sample through an analysing diffraction crystal mounted on a 2θ-goniometer. According to Bragg's law, the angle between the sample and the detector gives the wavelength of the radiation[46]:

$$2d \sin\theta = n\lambda \qquad \text{Equation 1}$$

where d is the d-spacing of the analysis crystal, θ is the half-angle between the detector and the sample, and n is the diffraction order. The elements in an unknown sample are identified by the energy they release when their atoms make a transition from a high-energy to a low-energy state. According to Bohr, the emission of radiation from the atom is due to the transition of the atom from a higher energy state to a lower energy state. In the simplest model of electronic transitions in hydrogen-like atoms, an electron gives off energy in moving between states with principal quantum numbers $n_i$ and $n_f$, and a photon with energy is emitted. The energy of the emitted photon is given by Equation 2:

$$\Delta E = E_i - E_f = \frac{me^4}{8\pi \varepsilon_o{}^2 h^3 c}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) Z^2 \qquad \text{Equation 2}$$

The integers $n_i$ and $n_f$ represent the angular momentum (units of ħ) of the initial and final states of the electron, and Z is an integer, called the atomic number, which represents the charge units of the nucleus. The coefficient of $Z^2$ is simplified to $R_\infty$, the Rydberg constant, and R∞ has the value 1.097 × 107 m$^{-1}$. According to Moseley, the electrons in the inner shells[47] repel those in the outer orbits. He showed that the frequencies of the characteristic X-rays emitted by chemical elements are proportional to the square of the atomic number of the element. According to Bohr's theory of atomic structure[48], the energy of an electron in its orbit n is given by:

$$E_n = -\frac{RhcZ^2}{n^2} \qquad \text{Equation 3}$$

The $K_\alpha$ X-ray emission is due to the transfer of an electron from the L-shell (n = 2) to the K-shell (n = 1), in which a vacancy has been created by the irradiation of the atom with X-rays before the transition. The energy of the K photon is therefore[47-50]:

$$E_{K\alpha} = -RhcZ^2\left(\frac{1}{2^2} - \frac{1}{1^2}\right) = \frac{3}{4}RhcZ^2 \qquad \text{Equation 4}$$

Moseley studied X-ray spectra at the same time as Bohr. He noticed that not all lines pass through the origin when the square root of the energy is plotted against Z, and to account for this fact he made E depend on $(Z-\sigma)^2$ instead of Z alone.[47,49] This leads Equation 4 to become Equation 5:

$$E_{K\alpha} = \frac{3}{4}hcR_\infty (Z - \sigma)^2 \qquad \text{Equation 5}$$

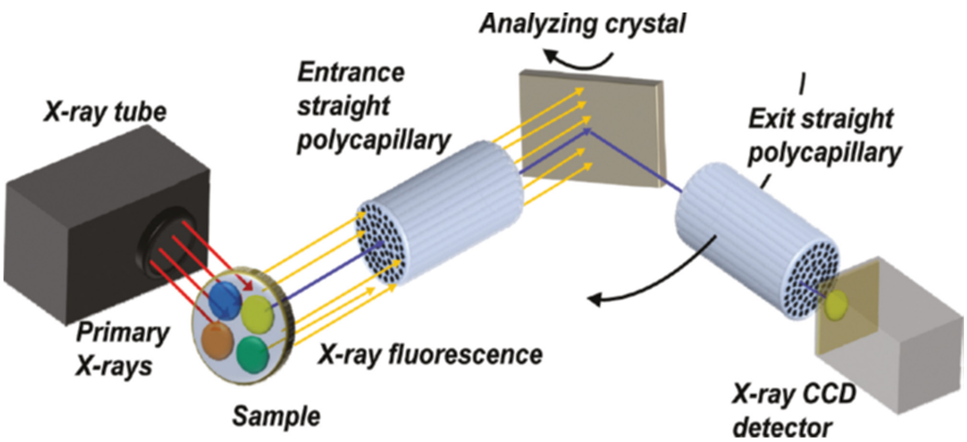
When E is small and Z is greater than 10, the typical photon energy is of the order of 1 keV, and this photon represents an X-ray. In Equation 5, $hcR_\infty$ amounts to 13.605 eV.[50] Dividing Equation 5 by h on both sides, we obtain an equation that establishes the relationship between the frequency $\nu$ and the atomic number Z. The wavelength of the $K_\alpha$ X-rays can be derived from Equation 5 and the result is given in Equation 6:

$$\lambda_{K_\alpha} = \frac{4}{3R}\frac{1}{(Z-1)^2} \qquad \text{Equation 6}$$

The following equation shows the relationship between energy, E and proton number, Z. Equation 7 represents the relationship between the square energy of a photon and the atomic number. Moseley's law says[49,50]:

$$\sqrt{E_{K\alpha}} = \sqrt{\frac{3E_o}{4}}(Z - \sigma) \qquad \text{Equation 7}$$

where $E_o = hcR_\infty$ = 13.6 eV. In measurements of Kα-X-rays for a range of elements, a curve of the square root of energy versus Z




*Source: Reproduced from Tsuji et al.[45] with permission. ©2011 American Chemical Society*

**Figure 2:** A typical XRF detection system for analysing samples.

should be a straight line, where a slope gives the ionisation energy of hydrogen and an intercept gives the screening factor $\sigma$ for 2→1 transition. With a screening factor known, the K$\alpha$ X-ray can be used to calculate the Z of an unknown sample. This is the principle that the XRF technique uses to distinguish different elements in the unknown sample.

The essential nutrients in columns 2 and 3 of Table 1 are required for various biochemical and physiological processes[51] in plants. The roots of plants often take up the macronutrients nitrogen (N), phosphorus (P), and potassium (K) in large quantities, so applying NPK fertilisers improves nutrient-poor soils and crop yields.[16] N is essential for stem growth and provides the rich green colour of plants. P is essential for root and flower growth, and the natural P derives from organic phosphates released when organic material in the soil decomposes. K derives from potash and contributes to the building of plant tissues, supports the production of chlorophyll and controls the opening and closing of stomata. After N, K is the element that plants take up in higher quantities.[52] Plants remove N from the soil in the form of nitrates. Some N degrades microbially through the denitrification process, in which N reduces to a gaseous form that diffuses out of the soil into the atmosphere.[22] K, P, calcium (Ca), magnesium (Mg), sulfur (S), Na and chlorine (Cl) occur in living organisms in trace and varying amounts[53] and are listed in Table 1 in their categories.[19,52]

Mineral plant nutrients are divided into two broad groups. Macronutrients are the ones needed in relatively large quantities by plants, while micronutrients are the nutrients needed in relatively small quantities by plants.

**Table 1:** Essential nutrients and non-nutrients required for plant growth

| From air and water | From soil and fertilisers | |
|---|---|---|
| **Non-nutrients** | **Macronutrients** | **Micronutrients** |
| Carbon (C) | Nitrogen (N) | Zinc (Zn) |
| Hydrogen (H) | Phosphorus (P) | Copper (Cu) |
| Oxygen (O) | Potassium (K) | Iron (Fe) |
| | Sulfur (S) | Manganese (Mn) |
| | Calcium (Ca) | Boron (B) |
| | Magnesium (Mg) | Molybdenum (Mo) |
| | Chlorine (Cl) | Cobalt (Co) |

## Results and discussion

The results of the 11 elements quantified in all soil samples are shown in Table 2. The system could not measure the elements below neon (Ne) as it measures from Na upwards. On average, the concentrations (mg/kg) of the elements found in the field samples are: Al (91.4 $\pm$ 7), barium (Ba; 0.488 $\pm$ 0.03), Ca (16.8 $\pm$ 5), Fe (39.3 $\pm$ 0.8), K (15.7 $\pm$ 0.04), Mg (10.1 $\pm$ 0.3), Mn (0.64 $\pm$ 0.04), Na (8.0 $\pm$ 1), P (1.3 $\pm$ 0.4), silicon (Si; 458 $\pm$ 8), and titanium (Ti; 5.6 $\pm$ 0.3).

In all soils, the elements are usually present in the form of compounds or other complex forms formed by previous oxidation or chemical reactions or formations.[16] Felsic rocks are usually supersaturated and contain free quartz ($SiO_2$). On average, felsic rocks contain high concentrations of Si, K, Ba, Pb and U.[54] These rocks contain minerals such as K-feldspar ($KAlSi_3O_8$) and mica ($KAl_2Si_3AlO_{11}$ $(OH)_2$) which contain more than one element in a complex compound.[55]

When these compounds decompose, their elements become available in the soil and they are then detected in large quantities because of their availability in the soil. It is important to note that the availability of a mineral in large quantities in the soil does not mean that it is readily available for plant uptake. A good example of this is iron oxides, which are present in the soil as $Fe^{3+}$, which is not readily available for plant uptake like $Fe^{2+}$; the solubility of $Fe^{3+}$ in soil depends on pH. A high pH impedes the availability of Fe to plants.

Comparability of potash ($K_2O$) in control and field samples shows that most of the $K_2O$ is predominantly lithogenic. The small difference, especially in field samples, could be because the roots of the crops absorbed $K_2O$ in soil water. On the other hand, rutile ($TiO_2$), which is very common in soil in KwaZulu-Natal, is harmful to the environment only at a very high content in the soil. There is a similarity between Ti control and field sample values. Based on this similarity of values, Ti is of natural origin. South Africa and Australia produced about 50% of the world's production in 2007.[56] Figure 3 shows the comparison of elements per sample in mg/kg.

In Figure 3, the samples include a high concentration of $SiO_2$, which is found in enormous quantities in the earth's crust but is not readily available for plant absorption. Figure 3 shows the elemental composition of the subsurface and nearby rocks from which the soils derive, i.e. Al, Fe and Si oxides, and their complex compounds. Studies have shown that the presence of Si in the soil increases the availability of nutrients (N, P and K), but the extent of this increase is not precisely known.[53] Therefore, it is not surprising that the P content in the samples is low and the N content is not present at all. In the geological sense, Si, Mn, and K are not trace elements, but major elements.

**Table 2:** Concentration of heavy metals (in mg/kg) in soil samples in this study

| Elements | Control soil samples | | Field soil samples | | | | |
|---|---|---|---|---|---|---|---|
| | **C17** | **C18** | **A17** | **A1** | **A2** | **A3** | **MV $\pm$ $\sigma$** |
| $Al_2O_3$ | 108 | 110 | 83.4 | 99.5 | 94.1 | 88.5 | 91.4 $\pm$ 7 |
| BaO | 0.440 | 0.450 | 0.470 | 0.480 | 0.480 | 0.52 | 0.488 $\pm$ 0.03 |
| CaO | 29.0 | 20.9 | 23.9 | 13.2 | 14.5 | 15.8 | 16.8 $\pm$ 5 |
| $Fe_2O_3$ | 45.2 | 55.8 | 38.8 | 39.0 | 40.5 | 39.0 | 39.3 $\pm$ 0.8 |
| $K_2O$ | 15.0 | 12.5 | 15.4 | 15.5 | 16.2 | 15.6 | 15.7 $\pm$ 0.04 |
| MgO | 14.3 | 13.2 | 10.1 | 9.70 | 10.3 | 10.0 | 10.1 $\pm$ 0.3 |
| MnO | 0.680 | 0.840 | 0.620 | 0.700 | 0.640 | 0.600 | 0.640 $\pm$ 0.04 |
| $Na_2O$ | 6.75 | 7.83 | 8.13 | 9.77 | 7.52 | 6.44 | 8.00 $\pm$ 1 |
| $P_2O_5$ | 1.33 | 0.570 | 1.53 | 1.28 | 1.72 | 0.850 | 1.30 $\pm$ 0.4 |
| $SiO_2$ | 342 | 421 | 468 | 456 | 458 | 450 | 458 $\pm$ 8 |
| $TiO_2$ | 5.33 | 7.03 | 5.43 | 6.01 | 5.73 | 5.41 | 5.60 $\pm$ 0.3 |

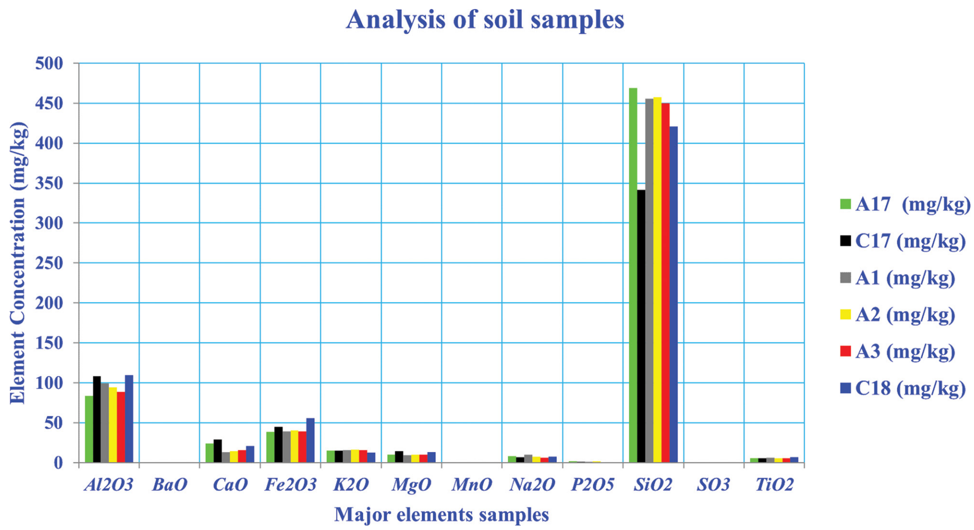



**Figure 3:** Trace elements in soil samples.

The oxides $Al_2O_3$, $Na_2O$, $K_2O$, $P_2O_5$ and $SiO_2$ are associated with feldspathic clays and mafic volcanic rocks[55], revealing the underlying rocks' mineral composition. Other metals present in large quantities are Ca, K, and Mg. For the significant concentrations, the concentrations of the metals in the soils ranged from low to high in the order Mg > K > Ca > Fe > Al >Si. The metal concentrations, in this study, compared with those reported in the literature are shown in Table 3.

As shown in Table 3, the Fe content found in this study is lower than that in a previous study,[17] which is not a negative finding as $Fe_2O_3$ cannot have toxic levels in soil[17] as it is not readily available for plant uptake in this insoluble form. The content of $SiO_2$, which is essential for certain plants, was higher than that reported previously.[54] The high content of Si also helps with the uptake of minerals like CaO and MnO in soil. The interaction of Si and essential nutrients is plant-specific. The application of Si releases the following elements for plant uptake: N, P, K, Ca, Mg, S, Zn, and Mn. Plants take up these nutrients well due to their availability resulting from the presence of Si in the soil.[53] Soils with neutral pH typically have a high K content. The K content is greater than that reported previously.[54]

Although the $Al_2O_3$ content in the samples is high, it is below the certified value of 116 mg/kg (16.62 wt.%). To counteract this, the availability of $Al_2O_3$ is reduced to a tolerable level by increasing the pH of the soil. $K_2O$ is higher than the certified value of 4.90 mg/kg (0.70 wt.%). $SiO_2$ is higher than the certified value of 318 mg/kg (45.42 wt.%). $Na_2O$ is lower than the certified value of 25.55 mg/kg (3.65 wt.%). $P_2O_5$ is very low compared to the certified value of 18.2 mg/kg (0.26 wt.%). The CaO content is also lower than the certified value of 76.51 mg/kg (10.93 wt.%). CaO is a macronutrient and is taken up by crops in large quantities; the field samples have a lower CaO content compared to the control samples, which could also indicate a low pH in the field samples. On the other hand, a high CaO content in the soil is often associated with a high pH value. The certified values, used for comparisons, were found in a study conducted at the University of Limpopo Experimental Farm.[55]

**Table 3:** Comparison of results from the literature and those from this study

| Elements | Concentration (mg/kg) | |
|---|---|---|
| | Present study | Reported |
| Mn | 0.64 ± 0.04 | (167; 377.61–499.68)[17,58] |
| K | 15.7 ± 0.04 | (2.53)[54] |
| Fe | 39.3 ± 0.8 | (8773)[17] |
| Al | 91.4 ± 7 | (12.3; 1.0)[40,52] |
| Si | 458 ± 8 | (31)[54] |

Both Mn and Fe contents were low compared to values reported in other studies.[17,54] The Mn content was similar in the field and control samples, but lower than the FAO/WHO recommended value of 437 mg/kg[57] and lower than the range of 377.61–499.68 mg/kg reported in a study conducted in Alice, in the Eastern Cape, South Africa[58]. The Al content is higher than that reported previously[40,52], but lower than the certified value of 116 mg/kg.[55] At a lower pH in the soil, Al becomes soluble and the amount of Al in the soil solution increases. Al is abundant in the earth's crust and tends to be toxic to some plants at concentrations of 2–3 mg/kg, especially when the soil pH is lower than 5.5.[59] The BaO content is similar in all samples but may be higher in the field samples than in the control samples due to the moderate pH and the application of fertilisers in the field soil samples. However, the BaO content is lower than the FAO/WHO recommended value of 100 mg/kg.

With the present technique, the toxic substances Hg, Cd, Pb, Cr, Sn and As were not detected in the samples. It is possible that their concentrations were below the limit of detection (LOD) and the limit of quantitation (LOQ). Of the toxic metals, the system only detected and quantified Fe and Mn.

## Future prospects and conclusion

The results of this study indicate that toxic metal contamination of the soils in the croplands investigated may not pose a significant risk. Al, Ba, Ca, Fe, K, Mn, Mg, Na, P, Si, and Ti were detected and quantified in the soil samples and they are largely related to the parent material of the soils in the area, so their uptake by plants is unavoidable as they are essential micronutrients for plants. However, some elements are not readily available for plants' uptake, regardless of how abundant they are in the soil, while some may also inhibit other elements' uptake.

Of the toxic elements Cd, Cu, Fe, Mn, Ni, Pb and Zn, only Fe and Mn were detected. The other elements were not detected, although they were within the detection range of the measuring system used. This study provides geological information on the origin of the soil on the croplands. The predominant elements such as $Al_2O_3$, $Fe_2O_3$, and $SiO_2$ are of lithogenic and non-anthropogenic origin, as the field and control samples have comparable contents for these elements. The advantage is that their availability to plants is pH-dependent. The high Si content usually favours the availability of other nutrients such as N, P, and K to plants, which can lead to lower levels of these macronutrients in the soil. The low $P_2O_5$ content in the soil in all samples could mean that the element was taken up by plants, run-off and erosion and leached into the subsoil. At this low concentration of $P_2O_5$, eutrophication is unlikely, especially in moving water streams. The study shows that neither the small towns upstream through industrial waste nor the fertilisers contribute to metal pollution in the soil and water.

The WD-XRF measurement technique could not detect most of the toxic metals. This work should serve as a baseline for future studies on toxic metals in this region. For future studies, using a different technique to compare and confirm the results might be useful. The use of a different technique could also help to verify whether the measurement system failed to detect and quantify toxic metals because they are not present in the soil or because their content is below the LOD or LOQ.

## Acknowledgements

We thank the SPECTRAU Lab, University of Johannesburg, for undertaking the analyses and Dr Christian Reinke for assistance with the analysis.

## Competing interests

We have no competing interests to declare.

## Authors' contributions

M.J.M: Conceptualisation, methodology, data collection, sample analysis, data analysis, validation, data curation, writing – the initial draft, writing – revisions. P.L.M.: Validation, data curation, student supervision, project leadership, funding acquisition.

## References

1. Kabata-Pendias A. Trace elements in soils and plants. 4th ed. Boca Raton, FL: CRC Taylor and Francis Group; 2011. p. 505. https://doi.org/10.1201/b10158
2. Bradl H, editor. Heavy metals in the environment: Origin, interaction and remediation. Amsterdam: Elsevier; 2005.
3. He ZL, Yang XE, Stoffella PJ. Trace elements in agroecosystems and impacts on the environment. J Trace Elem Med Biol. 2005;19(2–3):125–140. https://doi.org/10.1016/j.jtemb.2005.02.010
4. Nriagu JO. A global assessment of natural sources of atmospheric trace metals. Nature. 1989;338(6210):47–49. https://doi.org/10.1038/338047a0
5. Arruti A, Fernández-Olmo I, Irabien Á. Evaluation of the contribution of local sources to trace metals levels in urban $PM_{2.5}$ and $PM_{10}$ in the Cantabria region (Northern Spain). J Environ Monit. 2010;12(7):1451–1458. https://doi.org/10.1039/b926740a
6. Ding Z, Li Y, Sun Q, Zhang H. Trace elements in soils and selected agricultural plants in the Tongling mining area of China. Int J Environ Res Public Health. 2018;15(2):202. https://doi.org/10.3390/ijerph15020202
7. Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: An analysis of data from the Global Burden of Diseases Study 2015. Lancet. 2017;389(10082):1907–1918. https://doi.org/10.1016/S0140-6736(17)30505-6
8. Franchini M, Mannucci PM. Air pollution and cardiovascular disease. Thromb Res. 2012;129(3):230–234. https://doi.org/10.1016/j.thromres.2011.10.030
9. World Health Organization. Ambient air pollution: A global assessment of exposure and burden of disease. Geneva: World Health Organization; 2016. https://apps.who.int/iris/handle/10665/250141
10. Tchounwou PB, Yedjou CG, Patlolla AK, Sutton DJ. Heavy metal toxicity and the environment. In: Luch A, editor. Molecular, clinical and environmental toxicology. Basel: Springer; 2012. p.133–164. https://doi.org/10.1007/978-3-7643-8340-4_6
11. Xian Y, Wang M, Chen W. Quantitative assessment on soil enzyme activities of heavy metal contaminated soils with various soil properties. Chemosphere. 2015;139:604–608. https://doi.org/10.1016/j.chemosphere.2014.12.060
12. Fávaro DI. Natural radioactivity in phosphate rock, phosphogypsum and phosphate fertilizers in Brazil. J Radioanal Nucl Chem. 2005;264(2):445–458. https://doi.org/10.1007/s10967-005-0735-4
13. Hamamo H, Landsberger S, Harbottle G, Panno S. Studies of radioactivity and heavy metals in phosphate fertilizer. J Radioanal Nucl Chem. 1995;194(2):331–336. https://doi.org/10.1007/BF02038431
14. Hameed PS, Pillai GS, Mathiyarasu R. A study on the impact of phosphate fertilizers on the radioactivity profile of cultivated soils in Srirangam (Tamil Nadu, India). J Radiat Res Appl Sci. 2014;7(4):463–471. https://doi.org/10.1016/j.jrras.2014.08.011
15. Rout G, Samantaray S, Das P. Aluminium toxicity in plants: A review. Agronomie. 2001;21(1):3–21. https://doi.org/10.1051/agro:2001105
16. Hamidalddin SH. Determination of agriculture soil primordial radionuclide concentrations in Um Hablayn, north Jeddah west of Saudi Arabia. Int J Curr Microbiol Appl Sci. 2014;3(6):623–633.
17. Liphadzi MS, Kirkham MB. Partitioning and accumulation of heavy metals in sunflower grown at biosolids farm in EDTA-facilitated phytoremediation. Biorem Biodiv Bioavail. 2009;3:36–42.
18. Sabudak T, Seren G, Kaykioglu G, Dincer AR. Determination of trace elements in soil and sunflower (*Helianthus annuus* L.) plant parts. Fresenius Environ Bull. 2007;16(10):1274.
19. Hooda PS, editor. Trace elements in soils. Chichester: Wiley; 2010. https://doi.org/10.1002/9781444319477
20. Madejón P, Murillo JM, Marañón T, Cabrera F, Soriano MA. Trace element and nutrient accumulation in sunflower plants two years after the Aznalcóllar mine spill. Sci Total Environ. 2003;307(1-3):239–257. https://doi.org/10.1016/S0048-9697(02)00609-5
21. Alzahrani HR, Kumakli H, Ampiah E, Mehari T, Thornton AJ, Babyak CM, et al. Determination of macro, essential trace elements, toxic heavy metal concentrations, crude oil extracts and ash composition from Saudi Arabian fruits and vegetables having medicinal values. Arab J Chem. 2017;10(7):906–913. https://doi.org/10.1016/j.arabjc.2016.09.012
22. Hillel D. Introduction to environmental soil physics. London: Elsevier Academic Press; 2003.
23. Oluwole SO, Olubunmi MS, Kafeelah AY, Olusegun OF, Ayobami OO. Determination of heavy metal contaminants in leafy vegetables cultivated by the roadside. Int J Eng Res Dev. 2013;7(3):1–5.
24. Okeyode IC, Rufai AA. Determination of elemental composition of soil samples from some selected dumpsites in Abeokuta, Ogun State, Nigeria, using Atomic Absorption Spectrophotometer. Int J Basic Appl Sci. 2011;11:55–70.
25. Khan K, Lu Y, Khan H, Ishtiaq M, Khan S, Waqas M, et al. Heavy metals in agricultural soils and crops and their health risks in Swat District, northern Pakistan. Food Chem Toxicol. 2013;58:449–458. https://doi.org/10.1016/j.fct.2013.05.014
26. Faroon O, Ashizawa A, Wright S, Tucker P, Jenkins K, Ingerman L, et al. Toxicological profile for cadmium. Atlanta, GA: US Agency for Toxic Substances and Disease Registry; 2012. Available from: https://www.ncbi.nlm.nih.gov/books/NBK158838/
27. Tchounwou PB, Newsome C, Williams J, Glass K. Copper-induced cytotoxicity and transcriptional activation of stress genes in human liver carcinoma (HepG(2)) cells. Met Ions Biol Med. 2008;10:285–290.
28. Sharma H, Rawal N, Mathew BB. The characteristics, toxicity and effects of cadmium. Int J Nanotechnol Nanosci. 2015;3(10):1–9.
29. Stewart WM, Hammond LL, Van Kauwenbergh SJ. Phosphorus as a natural resource. In: Sims JT, Sharpley AN, Pierzynski GM, editors. Phosphorus: Agriculture and the environment. American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America; 2005. p. 1–22. https://doi.org/10.2134/agronmonogr46.c1
30. Louekari K, Mäkelä-Kurtto R, Pasanen J, Virtanen V, Sippola J, Malm J. Cadmium in fertilizers: Risk to human health and the environment webpage on the Internet]. c2000 [cited 2023 Aug]. Available from: https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160651/2000_4_Cadmium%20in%20fertilizers%20-%20Risks%20to%20human%20health%20and%20the%20environment.pdf
31. Oertli B, Auderset Joye D, Castella E, Juge R, Lehmann A, Lachavanne JB. PLOCH: A standardized method for sampling and assessing the biodiversity in ponds. Aquat Conserv Mar Freshw Ecosyst. 2005;15(6):665–679. https://doi.org/10.1002/aqc.744
32. Michalak AM, Anderson EJ, Beletsky D, Boland S, Bosch NS, Bridgeman TB, et al. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. Proc Natl Acad Sci USA. 2013;110(16):6448–6452. https://doi.org/10.1073/pnas.1216006110

33. Matthews MW, Bernard S. Eutrophication and cyanobacteria in South Africa's standing water bodies: A view from space. S Afr J Sci. 2015;111(5):1–8. https://doi.org/10.17159/sajs.2015/2014019

34. Wilson MA, Carpenter SR. Economic valuation of freshwater ecosystem services in the United States: 1971–1997. Ecol Appl. 1999;9(3):772–783. https://doi.org/10.1890/1051-0761(1999)009[0772:EVOFES]2.0.CO;2

35. Fosso-Kankeu E, Jagals P, Du Preez H. Exposure of rural households to toxic cyanobacteria in container-stored water. Water SA. 2008;34(5):631–636. https://doi.org/10.4314/wsa.v34i5.180660

36. Du Preez H, Swanepoel A, Van Baalen L, Oldewage A. Cyanobacterial Incident Management Frameworks (CIMFs) for application by drinking water suppliers. Water SA. 2007;33(5):643–652. https://doi.org/10.4314/wsa.v33i5.184085

37. Mulholland PJ, Helton AM, Poole GC, Hall RO, Hamilton SK, Peterson BJ, et al. Stream denitrification across biomes and its response to anthropogenic nitrate loading. Nature. 2008;452(7184):202–205. https://doi.org/10.1038/nature06686

38. Smith VH. Cultural eutrophication of inland, estuarine, and coastal waters. Successes, limitations, and frontiers in ecosystem science. 1998:7–49. https://doi.org/10.1007/978-1-4612-1724-4_2

39. Chauhan P, Chauhan RP, Gupta M. Estimation of naturally occurring radionuclides in fertilizers using gamma spectrometry and elemental analysis by XRF and XRD techniques. Microchem J. 2013;106:73–78. https://doi.org/10.1016/j.microc.2012.05.007

40. Isinkaralar K. The large-scale period of atmospheric trace metal deposition to urban landscape trees as a biomonitor. Biomass Convers Biorefin. 2022. https://doi.org/10.1007/s13399-022-02796-4

41. Brouwer P. Theory of XRF [document on the Internet]. c2006 [cited 2023 Aug 15]. Available from: https://home.iiserb.ac.in/~ramyasr/files/Manuals/XRF.pdf

42. Kramer MJ. Fertilizer effects on soil pH, soil nutrients, and nutrient uptake in swamp white and pin oak seedlings on an alkaline Missouri River bottomland. Columbia, MO: University of Missouri-Columbia; 2008.

43. Lider VV. X-ray fluorescence imaging. Physics-Uspekhi. 2018;61(10):980. https://doi.org/10.3367/UFNe.2017.07.038174

44. Chen ZW, Gibson WM, Huang H. High definition x-ray fluorescence: Principles and techniques. X-ray Optics Instrum. 2008;2008, Art. #318171. https://doi.org/10.1155/2008/318171

45. Tsuji K, Ohmori T, Yamaguchi M. Wavelength dispersive X-ray fluorescence imaging. Anal Chem. 2011;83(16):6389–6394. https://doi.org/10.1021/ac201395u

46. Humphreys CJ. The significance of Bragg's law in electron diffraction and microscopy, and Bragg's second law. Acta Crystallogr A. 2013;69(1):45–50. https://doi.org/10.1107/S0108767312047587

47. Conover CWS, Dudek J. An undergraduate experiment on x-ray spectra and Moseley's law using a scanning electron microscope. Am J Phys. 1996;64(3):335–338. https://doi.org/10.1119/1.18244 Tsuji

48. Khalid A, Khan AT, Anwar MS. X-ray fluorescence (XRF) spectrometry for materials analysis and "discovering" the atomic number. J Phys. 1996;64:335–338.

49. Berkowitz E. X-ray physics [document on the Internet]. c2006 [cited 2023 Aug 15]. Available from: http://www.phys.boun.edu.tr/~burcin/xray_physics.pdf

50. Soltis T, Folan LM, Eltareb W. One hundred years of Moseley's law: An undergraduate experiment with relativistic effects. Am J Phys. 2017; 85(5):352–358. https://doi.org/10.1119/1.4977793

51. World Health Organization. Trace elements in human nutrition and health. Geneva: World Health Organization; 1996.

52. Akpan IO, William ES. Assessment of elemental concentrations of roadside soils in relation to traffic density in Calabar, Nigeria. Int J Sci Technol Res. 2014;3:65–70.

53. Pavlovic J, Kostic L, Bosnic P, Kirkby EA, Nikolic M. Interactions of silicon with essential and beneficial elements in plants. Front Plant Sci. 2021;12:697592. https://doi.org/10.3389/fpls.2021.697592

54. Albright EI. Background concentrations of trace elements in soils and rocks of the Georgia piedmont [doctoral dissertation]. Athens, GA: University of Georgia; 2004.

55. Molepo KJ, Ekosse GE, Ngole VM. The role of potassium in the fertility status of soils: A case study of the University of Limpopo experimental Farm, South Africa. Res J Biotechnol. 2014;9(8):78–89.

56. Refiloe M. An overview of South Africa's titanium mineral concentrate industry. Report R71/2008. Pretoria: Department of Mineral Resources, Republic of South Africa; 2008. Available from: https://www.dmr.gov.za/LinkClick.aspx?fileticket=Yp939PhQ8IQ%3D&portalid=0

57. Påhlsson AM. Influence of aluminium on biomass, nutrients, soluble carbohydrates and phenols in beech (*Fagus sylvatica*). Physiol Plant. 1990;78(1):79–84. https://doi.org/10.1111/j.1399-3054.1990.tb08718.x

58. Bvenura C, Afolayan AJ. Heavy metal contamination of vegetables cultivated in home gardens in the Eastern Cape. S Afr J Sci. 2012;108(9/10), Art. #696. https://doi.org/10.4102/sajs.v108i9/10.696

59. FAO/WHO. Joint FAO/WHO Food Standards Programme Codex Committee on Contaminants in Foods. Fifth Session. Rome: FAO; 2011. p. 21–25.

Check for updates

# From waste cooking oil to oxygen-rich onion-like nanocarbons for the removal of hexavalent chromium from aqueous solutions

**AUTHORS:**
Themba D. Ntuli[1]
Ludwe L. Sikeyi[1]
Thomas H. Mongwe[1]
Orlette Mkhari[1]
Neil J. Coville[1]
Edward N. Nxumalo[2]
Manoko S. Maubane-Nkadimeng[1,3]

**AFFILIATIONS:**
[1]DSI-NRF Centre of Excellence in Strong Materials and the Molecular Sciences Institute, School of Chemistry, University of the Witwatersrand, Johannesburg, South Africa
[2]Institute for Nanotechnology and Water Sustainability, College of Science, Engineering and Technology, University of South Africa, Johannesburg, South Africa
[3]Microscopy and Microanalysis Unit, University of the Witwatersrand, Johannesburg, South Africa

**CORRESPONDENCE TO:**
Manoko Maubane-Nkadimeng

**EMAIL:**
manoko.maubane@wits.ac.za

**DATES:**
**Received:** 29 May 2022
**Revised:** 10 May 2023
**Accepted:** 30 May 2023
**Published:** 28 Sep. 2023

**HOW TO CITE:**
Ntuli TD, Sikeyi LL, Mongwe TH, Mkhari O, Coville NJ, Nxumalo EN, et al. From waste cooking oil to oxygen-rich onion-like nanocarbons for the removal of hexavalent chromium from aqueous solutions. S Afr J Sci. 2023;119(9/10), Art. #14006. https://doi.org/10.17159/sajs.2023/14006

**ARTICLE INCLUDES:**
☒ Peer review
☒ Supplementary material

**DATA AVAILABILITY:**
☐ Open data set
☒ All data included
☐ On request from author(s)
☐ Not available
☐ Not applicable

**EDITORS:**
Priscilla Baker
Amanda-Lee Manicum

**KEYWORDS:**
waste cooking oil, hexavalent chromium, onion-like nanocarbons, pyrolysis

**FUNDING:**
South African National Research Foundation (grant no. 138075), University of the Witwatersrand, DSI-NRF Centre of Excellence in Strong Materials



Vegetable cooking oil is used in domestic and commercial kitchens owing to its ability to modify and enhance the taste of the food through the frying process. However, as the oil is used through several frying cycles, it changes colour to dark brown and acquires an unpleasant smell. At this point, the waste oil is usually discarded, thereby finding its way into freshwater streams due to poor disposal and thus becoming an environmental pollutant. To provide an alternative, 'green' route to waste oil disposal, herein we report on the metal-free synthesis of onion-like nanocarbons (OLNCs) made from waste cooking oil via flame pyrolysis. The OLNCs were then applied in the removal of hexavalent chromium ions from aqueous solutions. The as-synthesised OLNCs were found to have similar properties (size, quasi-spherical shape etc.) to those synthesised from pure cooking oils. The Fourier-transform infrared spectroscopy data showed that the OLNCs contained C-O-type moieties which were attributed to the oxygenation process that took place during the cooking process. The OLNCs from waste oil were applied as an adsorbent for Cr(VI) and showed optimal removal conditions at pH = 2, t = 360 min, Co = 10 mg/L and $Q^0_{max}$ = 47.62 mg/g, superior to data obtained from OLNCs prepared from pristine cooking oil. The results showed that the OLNCs derived from the waste cooking oil were effective in the removal of hexavalent chromium. Overall, this study shows how to repurpose an environmental pollutant (waste cooking oil) as an effective adsorbent for pollutant (Cr(VI)) removal.

**Significance:**

- Waste cooking oil outperformed olive oil as a starting material for the production of OLNCs for the removal of toxic Cr(VI) from water.
- The superior performance of the OLNCs from waste cooking oil was attributed to the higher oxygen content found on their surface and acquired through the cooking process.
- Not only are the OLNCs produced from waste cooking oil effective in the removal of Cr(VI), but they can be used multiple times before replacement, which makes them sustainable.

## Introduction

Urbanisation has seen a rise in industries manufacturing chemicals and oil.[1-3] Some of these manufacturing industries produce or make use of compounds that include vegetable cooking oils which eventually enter freshwater streams via anthropogenic processes, resulting in aquatic environmental pollution.[4] Another toxic pollutant that can contaminate water systems is the Cr(VI) ion. It is known to be toxic to biota and humans alike due to its high mobility and biological accumulation.[5] Herein we provide a methodology to address these two issues.

Contamination of water with waste oil also leads to unsightly immiscible mixtures that are a threat to the environment and the ecosystem. This later issue arises from the lack of systems for the disposal of these oils.[6] Vegetable oil, which is mainly used in both industrial and household kitchens, is one such oil. This oil is often discarded in municipal drains and finds its way into groundwater or rivers. These oils increase the biological oxygen demand (BOD) and chemical oxygen demand (COD) of water, which could be detrimental to humans, micro-organisms and the quality of the water through de-oxygenation.[2,3]

One of the most used methods for trace metal removal is adsorption. Adsorption is typically achieved by using abundantly available carbonaceous materials which have functional groups consisting of oxygen, nitrogen, and sulfur moieties on their surface that attach to the trace metal ions.[7] One class of carbonaceous material that has been recently used for this purpose is based on nanomaterials of carbon. These materials include carbon nanofibres[8], carbon nanotubes[9], carbon dots[10], carbon spheres[11], graphene[12], carbon nano-onions (CNOs)[13], and onion-like nanocarbons (OLNCs)[14]. Their use is owing to their high surface-to-volume ratio, thermal stability, chemical robustness, and tunable surface functional groups.[15] CNOs and OLNCs can be viewed as fullerene types of carbon and have found application in numerous adsorption processes, in particular for the removal of Cr(VI) ions.[16] The high operational costs, high consumption of energy and the use of a metal catalyst to synthesise such carbons have to date limited their use in water application processes.[17]

Removal techniques for oils are different and include in-situ burning, containment booms, dispersants, and biodegradation, but all show an incomplete separation of oil from the water.[18] An alternative process is to convert these oils into useful carbonaceous products that could be used in various applications, thus putting value to waste. The use of oil for the synthesis of nano-carbons such as CNOs has been well documented.[19] For example, Shaku et al. used flame pyrolysis of grapeseed oil to synthesise CNOs that were applied in supercapacitors.[20] This method was used by Sikeyi et al. with olive oil as a starting material for CNOs that were applied in fuel cells.[21]

In the current study, we aimed to simultaneously address the aquatic environmental problem of trace metals (specifically Cr) and oil pollution (specifically cooking oil) in water streams. This was achieved by (1) using waste cooking oil as a precursor for the synthesis of OLNCs and (2) using the OLNCs for the removal of Cr(VI) ions from the aqueous solution. Although similar studies have been conducted, the current study adds to the literature of using these compounds in that converting waste cooking oil (an environmental pollutant) to a carbon nanomaterial (OLNCs) that can be used for the adsorption of Cr(VI) comes with several advantages over using pure oil and other methods to produce carbon nanomaterials. Firstly, waste cooking oil has low economic value and is an abundant source of carbon, making it an attractive option for producing carbon nanomaterials. Secondly, OLNCs have a high surface area to volume ratio compared to metal nanoparticles. In place of discarding the pollutant (waste cooking oil), we are repurposing it sustainably. This makes waste cooking oil attractive as a carbon- and oxygen-rich source for the synthesis of OLNCs via flame pyrolysis. For example, Khalisanni et al. reported that the chemical compounds found in waste cooking oils consist mainly of carboxylic acid derivatives.[22] As will be shown, the cooking process generates more oxygen groups than are found in pure oil and this leads to better Cr-OLNCs interactions. The OLNCs produced in this study were applied as adsorbents for the removal of Cr(VI) ions from aqueous solutions. The results were compared with those from our previous study in which we reported the use of olive oil as a starting material to make OLNCs for the removal of Cr(VI) in water.[23]

## Experimental

### *Materials and chemicals*

The waste cooking oil was donated to us by a local restaurant that uses a mixture of vegetable oils for the cooking of potato chips and fat cakes. All the chemicals used were obtained from Sigma-Aldrich (Johannesburg, South Africa), and were used as received unless otherwise stated. The pH adjustments were made by the dropwise addition of 0.1 M HCl and 0.1 M NaOH solutions. Stock solutions were prepared in the laboratory by dissolving appropriate amounts of $K_2Cr_2O_7$ in 1000 mL of distilled water.

### *Preparation of onion-like nanocarbons by flame pyrolysis*

The OLNCs adsorbent was synthesised using a method reported previously.[23] In the experiment, a custom-made glass container was filled with waste cooking oil and the wick stock was immersed in the oil and the tip of the wick was ignited. The flame was made to contact a brass collecting plate; at a 30 mm distance from the tip of the flame, where the black soot was collected, the flame can reach temperatures up to 950 °C.[24] The collected material was labelled onion-like nanocarbons (OLNCs) and was obtained with a yield of 8% (Equation 1).

$$\text{Yield} = \frac{\text{OLNCs produced (g)}}{\text{Oil used (g)}} \times 100\% \qquad \text{Equation 1}$$

where the OLNCs produced were 8 g and the oil used was 100 g.

### *Adsorbent characterisation*

The surface morphology of the adsorbent was determined by transmission electron microscopy (TEM; Jeol TEM-2100 F 200 kV) and scanning electron microscopy (SEM; ZEISS GeminiSEM 560 instrument). For thermal stability, a Perkin Elmer 6000 thermogravimetric analyser was employed. For the functional group's identification and surface composition, X-ray photoelectron spectroscopy (XPS; Thermo ESCAlab 250Xi) and Fourier-transform infrared (FTIR) spectroscopy (PerkinElmer Spectrum 100) were employed. A Micrometrics Tristar 3000 surface area and porosity analyser was used to determine the pore size, volume, and surface area.

### *Adsorption experiments*

All the Cr(IV) adsorption experiments were carried out in duplicate and in batch mode.[23] Various parameters that influenced the adsorption process were studied. These included the solution pH (2–8), contact time (5–720 min), initial concentration of Cr(VI) (10–50 mg/g), and the mass of the OLNCs (0.05–0.25 g). After the adsorption experiment, the adsorbent and adsorbate were separated via microfiltration, using a 0.45 μm filter. The Cr(VI) ions remained in solution and their concentration was determined by a Cary 100 ultraviolet-visible (UV–Vis) spectrometer. In the analysis, the Cr(VI) formed a complex with 1,5–diphenylcarbazide, detected at 540 nm, under acidic conditions.[25] For the determination of total Cr ions, we used an atomic absorption spectrometer (200 Series AA). The adsorption of both adsorbents was expressed by the amount of Cr(VI) ions adsorbed at equilibrium $q_e$ and the percentage of Cr(VI) ions removed (%R) as shown in Equations 2 and 3, respectively.

$$q_e = \frac{(C_o - C_e)V}{m} \qquad \text{Equation 2}$$

$$\%R = \frac{C_o - C_e}{C_o} \times 100 \qquad \text{Equation 3}$$

where $C_o$ (mg/L) and $C_e$ (mg/L) are the initial concentration of Cr(VI) and the equilibrium concentration of Cr(VI), respectively; V is the volume of Cr(VI) used; m is the mass of the OLNCs; and $q_e$ is the amount of Cr(VI) adsorbed at equilibrium. The equations for the adsorption isotherms, kinetic models, and thermodynamics are presented in the supplementary material as Supplementary equations 1–17.

## Results and discussion

### *Characterisation of the waste cooking oil and adsorbents*

The waste cooking oil was obtained from a local restaurant that used a mixture of vegetable oils for the cooking of potato chips and fat cakes. Compounds contained in the waste cooking oil were identified using gas chromatography–mass spectrometry (GC-MS) and the results were compared to different waste cooking oils found in the literature. The compounds for the waste cooking oil used for this study are presented in Supplementary figure 1.

Supplementary table 1 presents the major chemical compounds found in the waste cooking oil. These data are in agreement with data reported by Khalisanni et al. who noted that oleic acid and *n*-hexadecanoic acid were the two major compounds found in their study conducted at the Universiti Teknologi MARA, Malaysia,[22] and similar to data reported by Zayed et al.[26] for waste oils obtained from sunflower, cotton, and soybean oils. Any oil that is used at high temperatures will contain oxidised functional groups, which means that our study can be generalised to the synthesis of OLNCs from many similar waste oils.

During the cooking process, cooking oil undergoes three major processes – namely, hydrogenation, oxidation and polymerisation – which can result in the formation of volatile organic compounds (VOCs).[27] Some of these VOCs detected in the current study were nonanal, hexanoic acid, 2,4-decadienal, farnesene, propanoic acid, heptane, 2.4-heptadienal, and 2-decenal compounds, which are similar to those reported by Mannu et al. to be in their waste oil.[27] We also found oleic acid, hexadecenoic acid, palmitoleic acid, and 9-octadecenoic acid (Z)- in our waste oil – compounds also reported to be present in waste sunflower oil and waste cotton oil by Zayed et al.[26] This finding indicates that our data are similar to those observed in the literature.

Zhang et al. reported that VOCs such as furan, nonanal, hexanal, heptanal, octanal, and benzene are highly carcinogenic and that more VOCs are produced as the temperature of the oil is increased.[28] These carcinogenic compounds could be harmful to humans and biota should they find their way into water streams. Therefore, this study was motivated by seeking a method to convert waste cooking oil into a more valuable material. In our study, oil was used as a starting material for the metal-free synthesis of useful carbon nanomaterials via flame pyrolysis. The soot produced was collected in good yield (8%) as calculated by Equation 1. The OLNCs were then applied in the removal of Cr(VI) ions from the solution.

SEM micrographs of the adsorbent show that the materials are quasi-spherical and closely packed together (Figure 1a). The TEM micrographs show that the adsorbent consisted of interlinked spheres resembling a cut onion-like structure with multiple shells with an average external diameter of 41 nm. The OLNCs showed similar morphological characteristics and diameter with similar materials derived from pure oil.[17] The Brunauer–Emmett–Teller (BET) analysis results indicated that the synthesised material had a surface area of 69 $m^2$/g, a pore volume of 0.20 $cm^3$/g, and an average pore size of 11 nm. When compared to the BET results of the olive oil synthesised OLNCs, it was found that the olive oil materials had a larger surface area of 81.8 $m^2$/g. However, the olive oil material showed a smaller pore volume (0.08 $cm^3$/g) and a smaller average pore size (4 nm). The as-synthesised materials had surface areas that were in the range of those reported previously (60–85 $m^2$/g).[21,23] A closer look reveals the presence of cavities, another attribute in adsorption processes as the cavities could encapsulate adsorbate ions.

Thermal analysis showed a single decomposition step between 500 °C and 800 °C (Figure 2a), which indicates that the material was thermally stable and contained no impurities.[19] These temperatures are consistent with the decomposition temperature of graphitic carbon materials and were expected for the as-synthesised OLNCs as it can be seen From Figure 2c (Raman spectra) and Figure 3a (XPS spectra) that the OLNCs have a graphitic layer as part of their backbone structure.[29] The FTIR spectra showed that the adsorbent consisted mainly of carbon- and oxygen-containing functional groups and some O-H groups (Figure 2b). For example, the spectrum indicated O–H peaks at 3437 $cm^{-1}$, representing O-H stretching modes for alcohol moieties and a peak at 3153 $cm^{-1}$ for carboxylic acid moieties.[30] The O-H bending peaks at 1400 $cm^{-1}$ and 1100 $cm^{-1}$ are also consistent with the presence of carboxylic acid functional groups. The peak at 1635 $cm^{-1}$ signifies the C = C stretching mode for conjugated alkenes. The predicted surface moieties were similar to those reported previously on similar material synthesised via flame pyrolysis using a liquid paraffin precursor.[31] The Raman spectra of the OLNCs before and after the adsorption of Cr(VI) ions are depicted in Figure 2c. Two major peaks were identified between 1300 $cm^{-1}$ and 1650 $cm^{-1}$, due to the D band and G band. The presence of the D bands in the aforementioned range can be attributed to the amorphous nature of the carbon material with graphitic layers.[32,33] The $I_D/I_G$ ratios computed from the respective peak areas of the D and G bands were similar, indicating that the presence of Cr on the surface of the adsorbent did not disrupt the characteristics of the adsorbent. However, the peak position shifted, indicating possible Cr interaction with the surface (Table 1). The Raman spectra for the current materials were similar to those produced via flame pyrolysis of olive oil.[21]

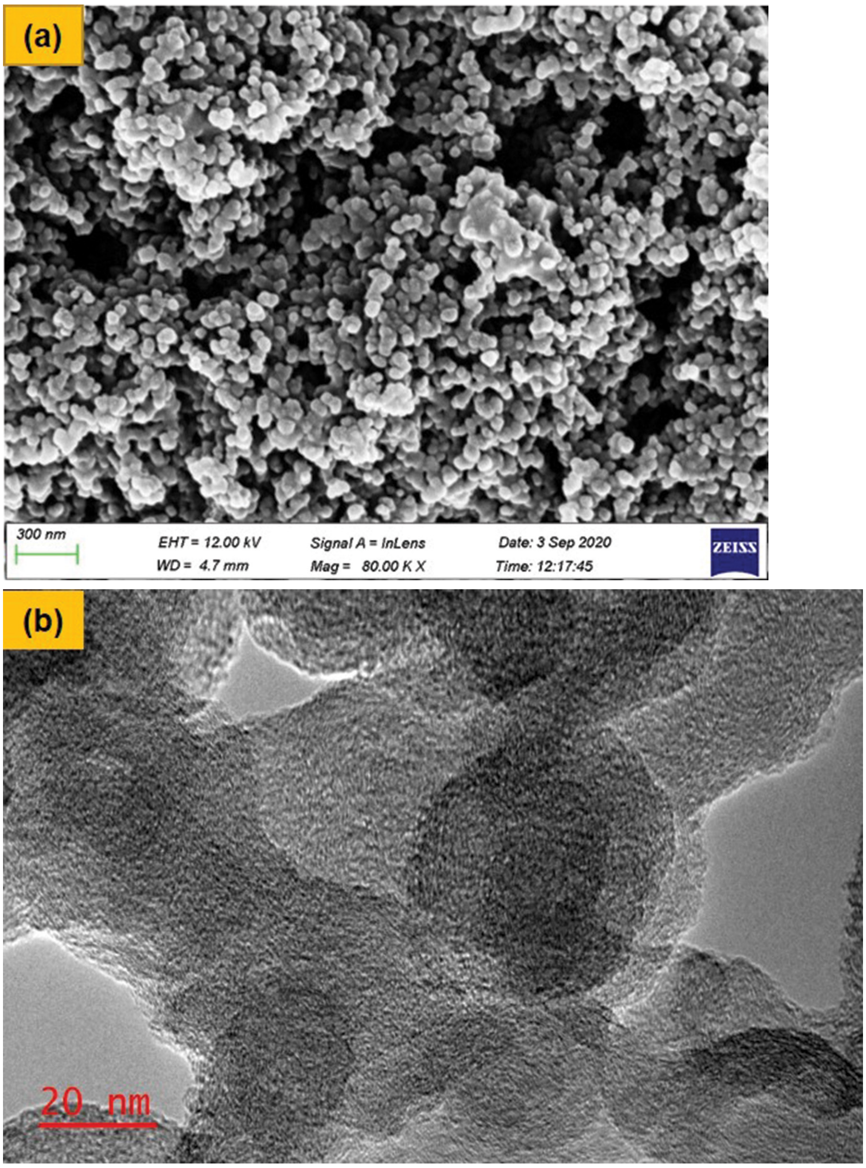



**Figure 1:** (a) Scanning electron micrographs and (b) transmission electron micrographs of the onion-like nanocarbons.

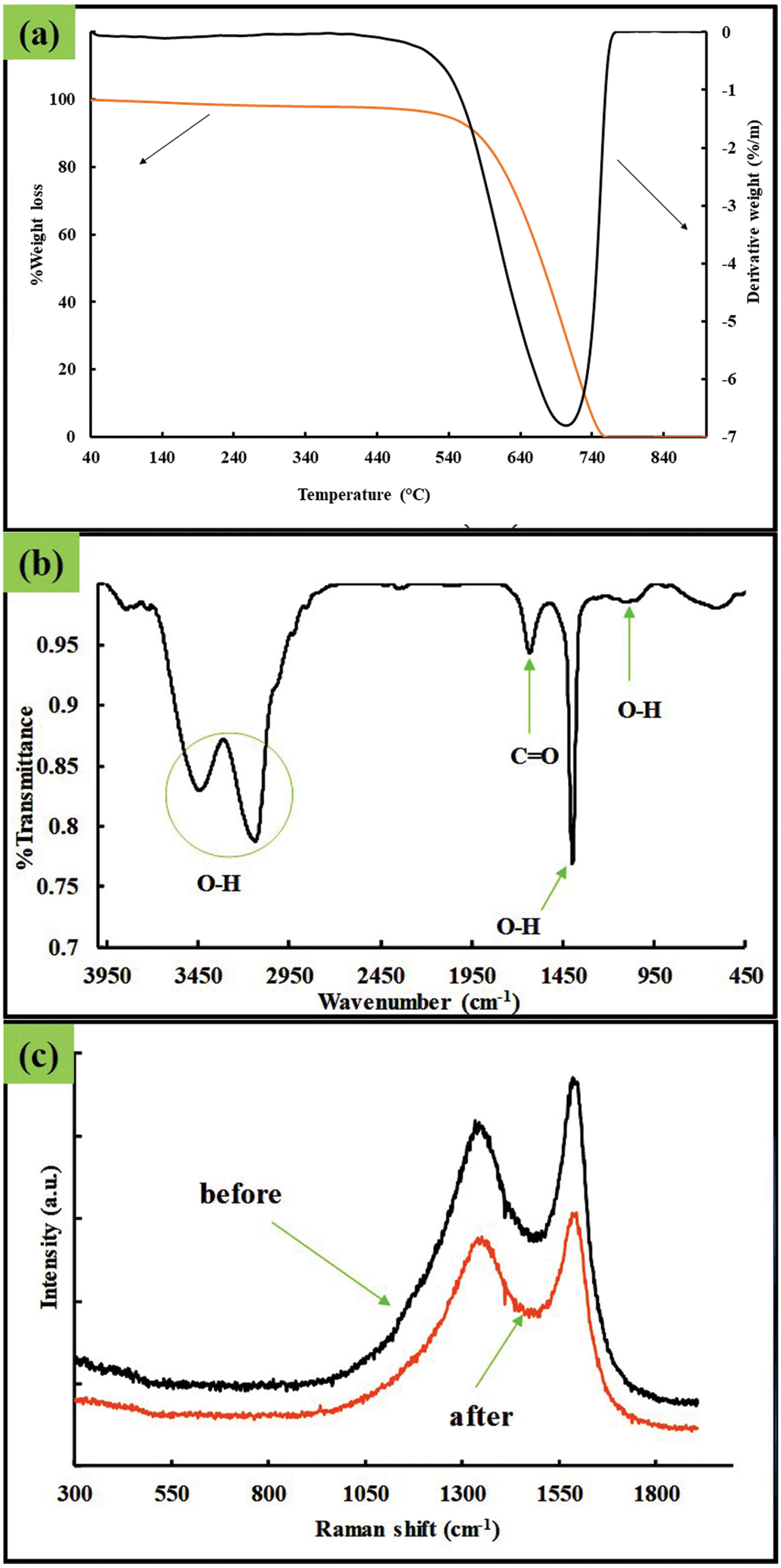



**Figure 2:** (a) Thermograms from thermogravimetric (TGA) and differential thermal analysis (DTA) of onion-like nanocarbons (OLNCs), (b) the Fourier transform infrared (FTIR) spectrum of OLNCs, and (c) the Raman spectra of OLNCs before and after adsorption of [Cr(VI)].

**Table 1:** Raman spectroscopy data of the onion-like nanocarbons (OLNCs) before and after removal of Cr(VI) ions

| Adsorbent | D-band | | G-band | | $I_D/I_G$ |
|---|---|---|---|---|---|
| | Raman shift ($cm^{-1}$) | Area | Raman shift ($cm^{-1}$) | Area | |
| OLNCs | 1352.6 | 409 981 | 1600.3 | 143 752 | 2.9 |
| OLNC-Cr | 1344.1 | 272 281 | 1593.7 | 97 375 | 2.8 |

deconvoluted spectra (Figure 3 and Table 2) showed threepeaks for C1s representing carbon bonded to oxygen at 288.7 eV, 288.0 eV, and 286.1 eV for O–C = O, C = O, and C–O moieties. The $sp^2$ and $sp^3$ hybridised carbon peaks were observed at 284.2 eV and 284.6 eV, respectively (Figure 3a). The material also showed two peaks for O1s representing carbon bonded to oxygen. These peaks were observed at 531.7 eV and 533.1 eV, as expected for carboxylic acid C = O and C–O moieties (Figure 3b). The XPS data for the OLCNs after the adsorption of Cr were also recorded. The data are shown in Supplementary figure 2. The total C (86.3%) and O (13.7%) surface concentrations did not change substantially (C = 87.3%; O = 12.6%) with the addition of a small amount of Cr to the OLCNs (0.1%) (Table 2). The oxidation state of the Cr could not be determined from the small amount of Cr detected. Of interest was the change in the types of surface C/O detected after adsorption of the Cr; a comparison of the data is shown in Supplementary table 2. The data suggest a change in the $Csp^2/Csp^3$ ratios, indicating coverage of the $Csp^3$ by the Cr. Interestingly, the increase in the amount of C attached to O increased on the Cr coverage, while the O content did not change significantly. This finding could be due to the presence of the Cr-O bonds on the surface of the OLCNs.

**Table 2:** Elemental identification and quantification of the onion-like nanocarbons (OLNCs) before and after adsorption of Cr

| Type of OLNCs | Element | Peak BE | Atomic % | Reference |
|---|---|---|---|---|
| OLNCs (waste oil) | C1s | 284.1 | 86.3 | This study |
| | O1s | 532.1 | 13.7 | |
| OLNCs (olive oil) | C1s | 284.1 | 88.2 | Ntuli et al.[23] |
| | O1s | 532.3 | 11.8 | |
| Cr-OLNCs (waste oil) | C1s | 248.2 | 87.3 | This study |
| | O1s | 532.3 | 12.6 | |
| | $Cr2p^3$ | 576.7 | 0.1 | |

It should be noted that the adsorbent had a higher oxygen content (13.7%) (see Table 2) than the OLNCs (11.8%) derived from olive oil.[23] This difference can be attributed to the oxidation of the starting material (waste cooking oil) during cooking. The surface moieties shown in the XPS data were similar to those reported by Ko et al. for carbon nano-onions (CNOs) synthesised after the annealing of nanodiamonds.[34]

Overall, the OLNCs synthesised via flame pyrolysis of waste cooking oil showed similar characteristics to those of OLNCs produced from pure oils such as olive oil, castor oil, paraffin oil, and ghee oil.[17,19,21] The results show that the hydrogenation, oxidation and polymerisation reactions that took place during the frying process did not negatively affect the quality of the OLNCs that were synthesised using the waste cooking oil. This finding offers the possibility of using waste oils in place of pure oils for the synthesis of OLNCs. More importantly, the as-synthesised OLNCs also showed similar characteristics to CNOs synthesised from more expensive methods such as the annealing of nanodiamonds and laser irradiation.[16,34] The synthesised OLNCs showed similar characteristics (quasi-spherical, multi-shells, size diameter >100 nm, disordered and having surface defects) to carbon nano-onions and carbon black.[35] The as-synthesised OLNCs did not show any fluorescent properties, unlike those reported by Tripathi et al. who synthesised OLNCs using flaxseed oil as a carbon precursor via flame pyrolysis, with the difference being their acid treatment of the OLNCs which was not done in our study.[36]

## *Cr(VI) ions removal studies*

### Effect of solution pH on Cr(VI) ions removal

The removal of Cr(VI) ions is interrelated with the solution pH of the Cr(VI) ions. The solution pH affects the speciation of Cr(VI) ions as well as the nature of the surface moieties on the adsorbent. Figure 4 presents the performance of the OLNCs in the removal of Cr(VI) ions in the pH range of 2–8. The removal efficiency decreased as the pH of the solution was increased from 2 to 8 (48% to 81%). This, as mentioned in the literature, can readily be attributed to an electrostatic attraction as a result of the nature of the adsorbate and adsorbent as Cr(VI) ions are present as anionic species (chromate and dichromate) in a basic and acidic environment.[37] In more acidic conditions, $Cr_2O_7^{2-}$ species are formed and undergo hydrolysis to form $HCrO_4^-$, while the surface moieties of the adsorbent predominantly undergo protonation. This results in an electrostatic attraction between the adsorbate and the adsorbent, producing good metal ion removal under acidic conditions.[38]

The electrostatic attraction was supported by the $pH_{PZC}$ of the adsorbent, which was found to be pH 7.2. The $pH_{PZC}$ was determined by measuring the initial pH and the final pH where the change in pH was plotted against the initial pH. The point at which the graph intersects the *x*-axis was taken as the $pH_{PZC}$. The adsorbent undergoes protonation at a pH below the $pH_{PCZ}$ and deprotonation above it. This arises because Cr ions have hard acid characteristics and may be bound to the surface of the adsorbent as a result of the hard base characteristics of the oxygen moieties.[7] The

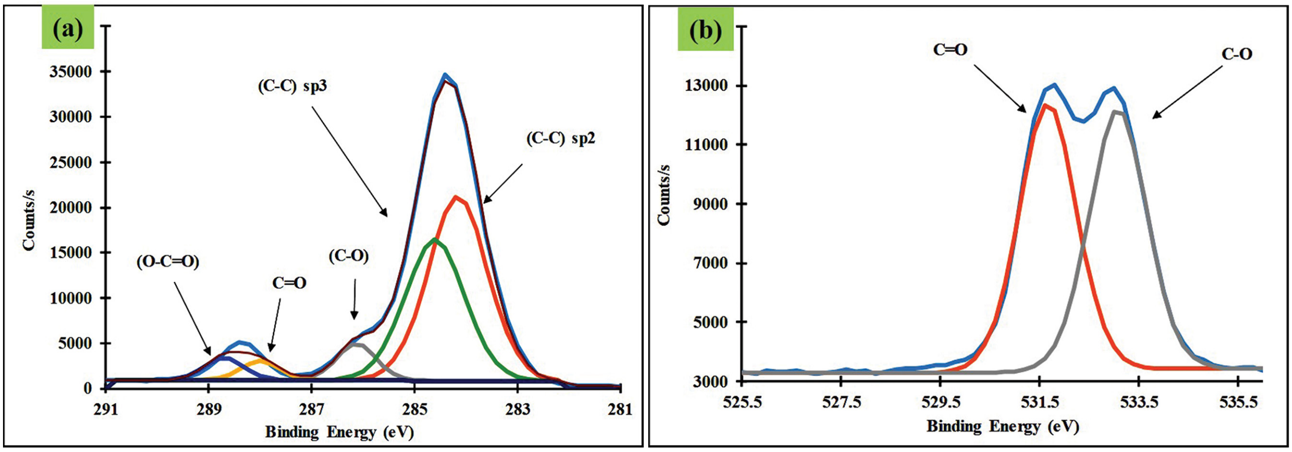



**Figure 3:** (a and b) The X-ray photoelectron spectra of onion-like nanocarbons adsorbent for C1s and O1s peaks.

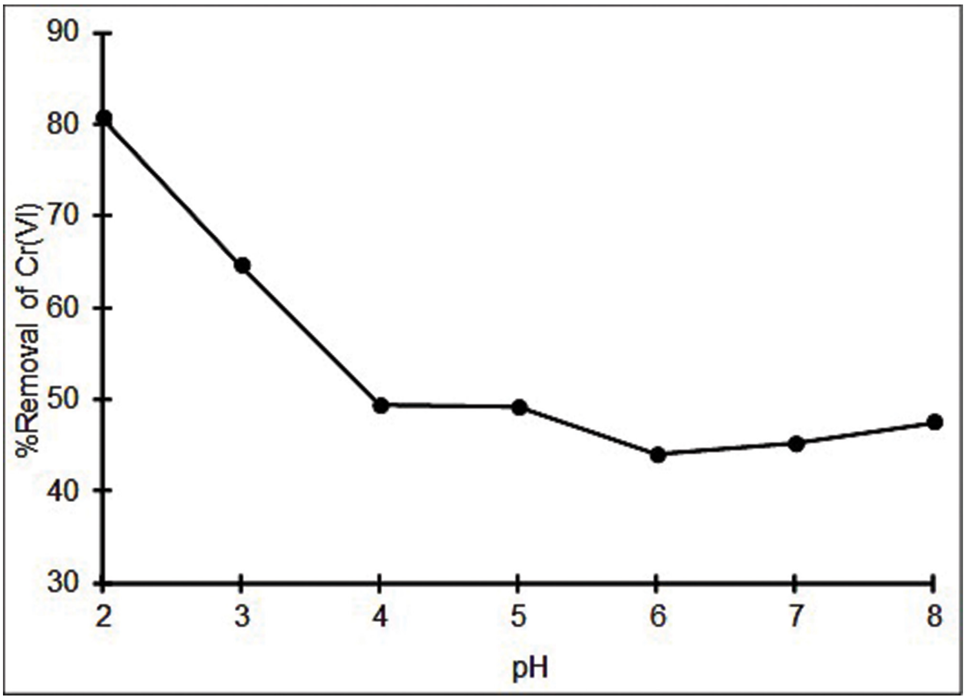



**Figure 4:** Performance of the onion-like nanocarbons as a function of solution pH (conditions: m = 0.05 g, t = 720 min and $C_0$ = 10 mg/L).

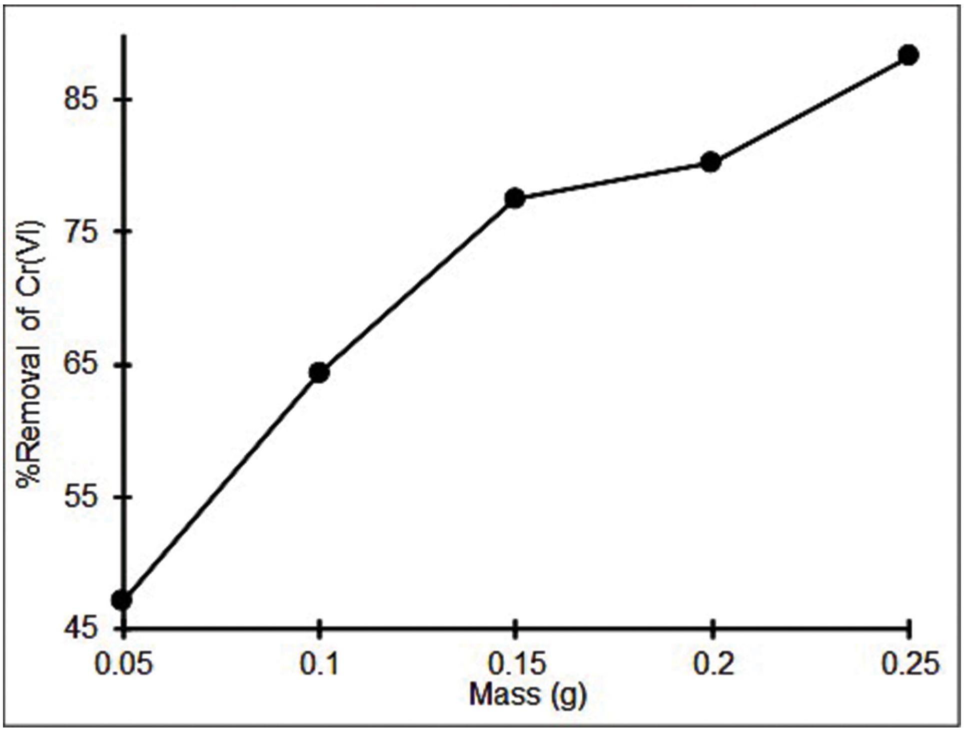



**Figure 5:** Performance of the onion-like nanocarbons as a function of adsorbent mass (conditions: pH = 2, t = 30 min and $C_0$ = 10 mg/L).

high removal of Cr(VI) ions under acidic media has also been extensively reported by other authors who have used carbon-rich adsorbents, and our data are in full agreement with the earlier studies.[39-42]

### Effect of adsorbent mass

The performance of the OLNCs as a function of adsorbent mass is depicted in Figure 5. From Figure 5, it is seen that an increase in the mass of the adsorbent leads to an increase in efficiency. This increase arises as more adsorption sites are available.[43] However, the efficiency profile did not reach equilibrium at the selected adsorbent mass range. This could be an indication that the surface of the adsorbent was not fully covered by the adsorbate ions and thus still had vacant active sites available for binding with the Cr ions.

### Effect of contact time

The duration for which the adsorbent and adsorbate are in contact plays a major role in the migration of the adsorbate ions to the surface of the adsorbent. Figure 6a shows the performance of

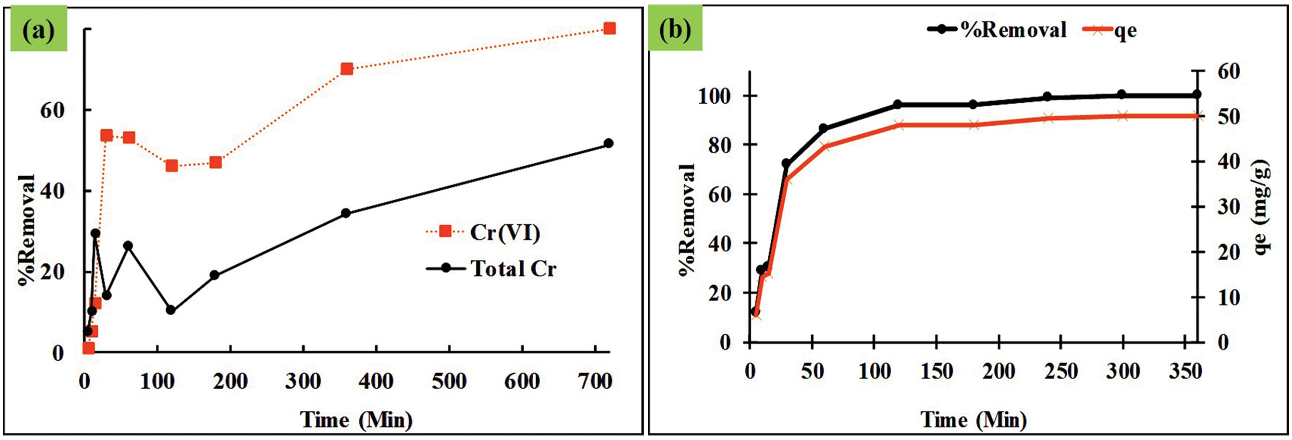



**Figure 6:** Performance of the onion-like nanocarbons as a function of time for (a) 0.05 g and (b) 0.1 g absorbent (conditions: pH = 2 and $C_0$ = 10 mg/L).

the adsorbent as a function of time (15 min to 720 min).[39] This adsorption capacity ($q_e$) changed from 0.96 mg/g (15 min) to 6.4 mg/g (720 min) with increasing time. This shows the removal of Cr(VI) ions from the total Cr and is thus able to demonstrate the reduction of Cr(VI) to Cr(III), as has been observed previously with carbon-rich materials.[44]

The removal of Cr(VI) ions was further tested as a function of time using two adsorbent masses (0.05 g and 0.1 g) (Figure 6b). As expected, as the adsorbent mass increased, the removal improved from 15.3 mg/g after 15 min to 48.0 mg/g after 180 min. This was attributed to the availability of more active sites as the adsorbent mass was increased because the number of active sites could be linked to the amount of the adsorbent.

### Kinetics

A summary of the adsorption kinetics for the removal of Cr(VI) ions using the as-synthesised OLNCs is presented in Table 3. The removal of Cr(VI) ions for the initial 30 min followed the pseudo-first-order (PFO) kinetic model. This could be attributed to a removal process taking place through interface diffusion with a rate constant ($k_1$) of 0.0345 1/min.[45] However, after 30 min of contact time, the adsorption kinetics followed a pseudo-second-order (PSO) route, indicating a chemisorption process with the rate constant ($k_2$) of 0.0024 g/mg x min.[46] Ho and McKay have suggested that the PFO may not be suitable for the whole range of contact times used and that only the first 20 to 30 min can be attributed to chemical bond formation between the adsorbate ions and the adsorbate surface.[47] The initial adsorption rate (*h*) for the reaction was found to be 2.074 mg/g x min and 0.190 mg/g x min from 30 min to 360 min (Table 3). This could be attributed to the surface coverage with more exposed sites filled first.

### Initial concentration

The performance of the OLNCs in the removal of Cr(VI) ions was tested as a function of the initial concentration of the Cr(VI) ions (Figure 7). It can be observed that as the initial concentration of the Cr(VI) ions increases from 10 mg/L to 100 mg/L, the removal percentage decreases. However, the adsorption capacity ($q_e$) increases from 7.92 mg/g to 50.40 mg/g, which is attributed to the possibility of enhanced interactions at the surface of the adsorbate as the initial concentration increased.[48] This phenomenon is in agreement with data that show that the initial concentration is directly proportional to the adsorption capacity (see Equation 1).

### Adsorption isotherms

The linear forms of the Langmuir and Freundlich adsorption isotherms were selected for modelling the Cr(VI) ions removal pathway using OLNCs as an adsorbent; the data are presented in Table 4. Both the dimensionless constants separation factor ($R_L$) and Freundlich intensity parameter (n) were between 0 and 1, indicating a favourable removal of Cr(VI) ions. The coefficient of determination ($R^2$) values for the Langmuir adsorption isotherm were nearer to unity compared to those of the Freundlich isotherm. This result corresponds with the chi-squared ($X^2$)

**Table 3:** A summary of the adsorption kinetics parameters

| Pseudo-first-order kinetics | | | Pseudo-second-order kinetics | | | |
|---|---|---|---|---|---|---|
| $k_1$ | $q_e$ | $R^2$ | $k_2$ | $q_e$ | *h* | $R^2$ |
| **Overall** | | | | | | |
| 0.012 | 22.60 | 0.7965 | 0.00069 | 54.65 | 2.074 | 0.9951 |
| **5 to 30 min** | | | | | | |
| 0.046 | 58.92 | 0.9545 | 344.379 | 1.24 | 526.316 | 0.0278 |
| **30 to 360 min** | | | | | | |
| 0.002 | 2.60 | 0.7050 | 0.00235 | 6.80 | 0.190 | 0.9881 |

values which were nearer to zero for the Langmuir isotherm and not the Freundlich isotherm. This suggests that the removal pathway was through a monolayer surface coverage of the adsorbent.[49] Moreover, the maximum adsorption capacity was 47.62 mg/g, which was close to the experimental adsorption capacity of 50.40 mg/g, in agreement with a Langmuir isotherm or monolayer surface coverage. The adsorption capacity of the adsorbent was significantly higher than that of similar material produced from olive oil (26.53 mg/g).[23] This could be due to the added adsorbent oxygen content due to oxidation through cooking. The results discussed thus far indicate that different oils (olive and waste cooking oil) that undergo flame pyrolysis yield carbon nanomaterials with similar characteristics. However, due to the difference in the oils, the surface moieties may be slightly different. Such differences may result in significant changes in the application of the materials.

## Removal mechanism of Cr(VI) ions

The XPS analysis of the OLNCs before and after the removal of Cr(VI) ions is presented in Supplementary figure 3. A new peak was observed at 577.9 eV for $Cr_{p2}$, representing Cr(III) ions bound to the surface of the adsorbent (Supplementary figure 3). This peak was absent before the material was used as an adsorbent for the removal of Cr(VI) ions. It can be noted that in the process of Cr(VI) ions removal, some ions were reduced to the less toxic Cr(III) ions. Figure 6a shows the percentage removal of Cr(VI) ions and of the total Cr ions. The thermodynamic parameters for the removal of Cr(VI) ions are depicted in Table 5. All the thermodynamic parameters returned negative values and were representative of a favourable and spontaneous ($-\Delta G$) removal of Cr. The removal of Cr(VI) ions at the adsorbate/adsorbent boundary layer involves an associative mechanism ($-\Delta S$).[50] The $-\Delta H$ value could indicate that the removal of Cr(VI) ions involved chemisorption, physisorption, or comprehensive removal (a mixture of chemisorption and physisorption). Therefore, the removal mechanism as described fits the proposed removal that involves an adsorption-coupled reduction mechanism between the Cr(VI) ions and the adsorbent.

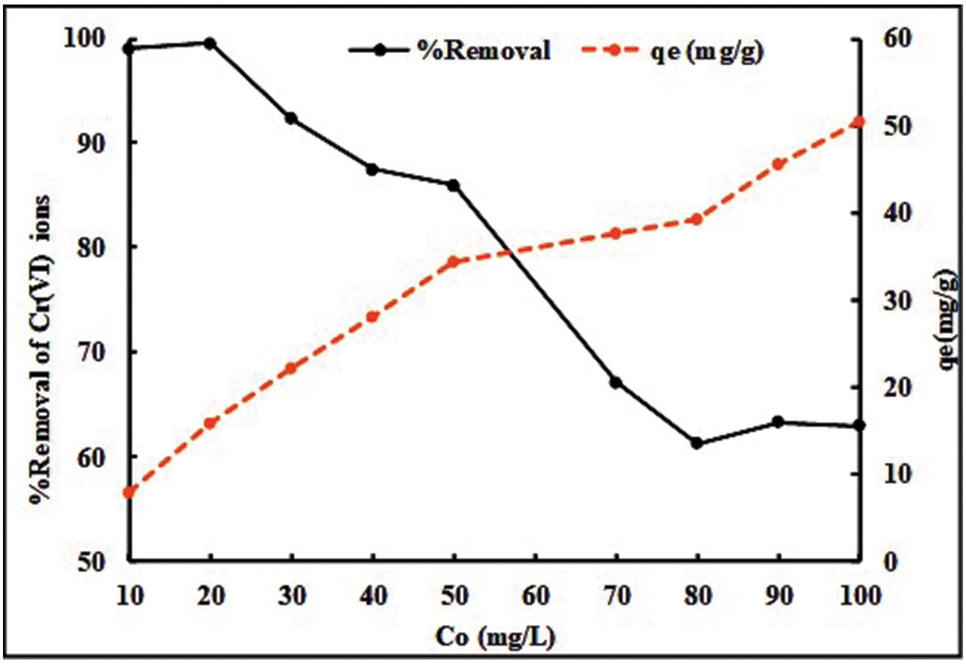



**Figure 7:** Performance of the onion-like nanocarbons as a function of the initial concentration of Cr(VI) ions (at pH = 2, t = 120 min, m = 0.025g and $C_o$ = 10–100 mg/L).

**Table 4:** Adsorption isotherm parameters

| Langmuir | | | | Freundlich | | |
|---|---|---|---|---|---|---|
| $Q_{max}$ (mg/g) | $K_L$ (L/mg) | $R^2$ | $X^2$ | n | $R^2$ | $X^2$ |
| 47.62 | 0.392 | 0.9453 | 0.162 | 0.384 | 0.8656 | 142478 |
| | | | | | | |
| $C_o$ (mg/L) | $R_L$ | | | | | |
| 10 | 0.2033 | | | | | |
| 20 | 0.1132 | | | | | |
| 30 | 0.0784 | | | | | |
| 40 | 0.0600 | | | | | |
| 50 | 0.0486 | | | | | |
| 70 | 0.0352 | | | | | |
| 80 | 0.0309 | | | | | |
| 90 | 0.0276 | | | | | |
| 100 | 0.0249 | | | | | |

### *Comparison of adsorption capacities for different adsorbents*

Direct comparison of different adsorption studies in the literature is challenging due to the different optimum adsorption conditions used in the respective studies. Therefore, the pH, contact time, adsorbent modification, and adsorption capacity were compared (Table 6).

It can be observed that the optimum pH for the removal of Cr(VI) ions was under acidic media (pH < 7). Most of the adsorbents in Table 6 underwent functionalisation to improve their adsorption capacity. In comparison, the as-synthesised adsorbent was used in its pristine state. Remarkably, the adsorption capacity of the OLNCs was higher than those of the pristine adsorbents as well as functionalised ones, except for a graphene nanocomposite (Table 6).

The graphene nanocomposite showed the highest adsorption capacity – an indication of the value of surface modification and the synergy between different materials to form a composite with high adsorption capacity.[51] However, the as-synthesised OLNCs from waste cooking oil showed a comparable adsorption capacity of 48 mg/g at a time of 360 min. This is an indication that the adsorbent has potential as an adsorbent for the removal of Cr(VI) ions. It should also be noted that the adsorbent had a higher adsorption capacity (48 mg/g) than that of pure oil (26.53 mg/g).[23] This was attributed to the higher oxygen content of the OLNCs from waste cooking oil, due to the oxidation process that takes place during cooking.

**Table 5:** Thermodynamic parameters calculated via the Langmuir constant $K_L$

| T (K) | van't Hoff equation | $K_c$ | ∆G° (kJ/mol) | ∆H° (kJ/mol) | ∆S° (J/mol) |
|---|---|---|---|---|---|
| 298.15 | y = 23092x – 61.419 | 6213254.8 | –38.8 | –192.0 | –510.6 |
| 303.15 | $R^2$ = 0.9333 | 4082853.0 | –38.4 | | |
| 308.15 | | 957665.7 | –35.3 | | |
| 313.15 | | 160788.5 | –31.2 | | |

**Table 6:** Comparison of different adsorbents

| Adsorbent | pH | Time (min) | Modification | $q_e$ (mg/g) | Reference |
|---|---|---|---|---|---|
| (3D) NiO/Ni | 4 | 120 | Composite | 25.94 | Wang et al.[52] |
| Activated carbon | 4 | 1440 | None | 3.47 | Mortazavian et al.[53] |
| | | 4320 | Polysulfide | 8.93 | |
| Graphene oxide | 2 | 1440 | Manganese ferrite | 34.02 | Shahrin et al.[54] |
| Multi-walled carbon nanotubes | 2 | 6000 | Oxidation | 31.95 | Gholipour and Hashemipour[55] |
| Carbon nano-onions | 2 | 720 | None | 26.53 | Ntuli et al.[23] |
| PANI/Ag/ GO CDs | 2 | 60 | Composite | 59.96 | Shokry et al.[51] |
| Onion-like nanocarbons | 2 | 360 | None | 48 | Current study |

## Conclusions

This study has shown that waste cooking oil be used as a starting material for the synthesis of high-quality OLNCs. The OLNCs obtained in this work had a higher oxygen content than pure oils due to the cooking process and this led to more sites for binding to metal ions.

The OLNCs were applied as an efficient adsorbent for the removal of Cr(VI) ions from an aqueous solution. With the maximum adsorption capacity of 47.62 mg/g at a pH of 2 and a contact time of 360 min, the OLNCs produced were better than those made from pure oils. The adsorption-coupled reduction mechanism was found to be the pathway for the removal of the Cr(VI) ions, as demonstrated by the presence of Cr(VI) and Cr(III) ions in the adsorption media. The removal was spontaneous and exothermic, and involved an associative mechanism. We have shown that waste material can be converted into a useful material for application in adsorption processes. This was achieved in this study by taking a toxic material (waste cooking oil) and converting it into a less harmful material that could be used in the removal of toxic Cr(VI) ions, thus putting value to waste. This then provides a means of solving the double aquatic environmental problem of trace metal removal and oil pollution.

## Acknowledgements

This work was supported by the National Research Foundation of South Africa (grant number: 138075), the University of the Witwatersrand, and the DSI-NRF Centre of Excellence in Strong Materials (CoESM). We thank Siyasanga Mpelane (University of Johannesburg, South Africa) for performing the high-magnification TEM analysis.

## Competing interests

We have no competing interests to declare.

## Authors' contributions

T.D.N: Conceptualisation, methodology, writing – original draft, investigation, formal analysis, writing – review and editing, validation. T.H.M: Data curation, formal analysis, investigation. L.L.S: Data curation, formal analysis, investigation. O.M.: Data curation, formal analysis, investigation. N.J.C.: Supervision, resources, writing – review and editing, validation. E.N.N.: Supervision, resources, writing – review and editing, validation. M.S.M-N: Conceptualisation, supervision, resources, project administration, writing – review and editing, validation.

## References

1. Pakade VE, Ntuli TD, Ofomaja AE. Biosorption of hexavalent chromium from aqueous solutions by macadamia nutshell powder. Appl Water Sci. 2017;7:3015–3030. https://doi.org/10.1007/s13201-016-0412-5
2. Sridhar S, Kale A, Khan AA. Reverse osmosis of edible vegetable oil industry effluent. J Memb Sci. 2002;205:83–90. https://doi.org/10.1016/S0376-7388(02)00065-0
3. Ahmad T, Belwal T, Li L, Ramola S, Aadil RM, Abdullah, et al. Utilization of wastewater from edible oil industry, turning waste into valuable products: A review. Trends Food Sci Technol. 2020;99:21–33. https://doi.org/10.1016/j.tifs.2020.02.017
4. Mofokeng M, Nthunya LN, Gutierrez L, Matabola P, Mishra S, Nxumalo EN. Perfluroooctyltriethoxy silane and carbon nanotubes-modified PVDF superoleophilic nanofibre membrane for oil-in-water adsorption and recovery. J Environ Chem Eng. 2020;8(6):104497. https://doi.org/10.1016/j.jece.2020.104497
5. World Health Organization (WHO). Guidelines for drinking-water quality [document on the Internet]. c2008 [cited 2022 May 29]. Available from: https://www.who.int/water_sanitation_health/dwq/2edvol3a.pdf
6. Zitte LF, AWaadu GDB. Used-oil generation and its disposal along East-West Road, Port Harcourt Nigeria. Int J Waste Resour. 2016;6(1):1–5. https://doi.org/10.4172/2252-5211.1000195
7. Lesaoana M, Mlaba RPV, Mtunzi FM, Klink MJ, Edijike P, Pakade VE. Influence of inorganic acid modification on Cr(VI) adsorption performance and the physicochemical properties of activated carbon. S Afr J Chem Eng. 2019;28:8–18. https://doi.org/10.1016/j.sajce.2019.01.001
8. Kakoria A, Sinha-Ray S, Sinha-Ray S. Industrially scalable chitosan/nylon-6 (CS/N) nanofiber-based reusable adsorbent for efficient removal of heavy metal from water. Polymer (Guildf). 2021;213:123333. https://doi.org/10.1016/j.polymer.2020.123333
9. Deng S, Liu X, Liao J, Lin H, Liu F. PEI modified multiwalled carbon nanotube as a novel additive in PAN nanofiber membrane for enhanced removal of heavy metal ions. Chem Eng J. 2019;375(May):122086. https://doi.org/10.1016/j.cej.2019.122086
10. Mkhari O, Ntuli TD, Coville NJ, Nxumalo EN, Maubane-Nkadimeng MS. A comparison of fluorescent N-doped carbon dots supported on the surface of hollow and solid carbon spheres, and solid silica spheres. Diam Relat Mater. 2021;118(February):108500. https://doi.org/10.1016/j.diamond.2021.108500
11. Mente P, Mashindi V, Phaahlamohlaka TN, Monyatsi TN, Forbes RP, Coville NJ. Oxidation of benzyl alcohol using cobalt oxide supported inside and outside hollow carbon spheres. ChemistryOpen. 2021;10(6):618–626. https://doi.org/10.1002/open.202000312
12. Majdoub M, Amedlous A, Anfar Z, Jada A, El Alem N. Engineering of amine-based binding chemistry on functionalized graphene oxide/alginate hybrids for simultaneous and efficient removal of trace heavy metals: Towards drinking water. J Colloid Interface Sci. 2021;589:511–524. https://doi.org/10.1016/j.jcis.2021.01.029
13. Mykhailiv O, Zubyk H, Plonska-Brzezinska ME. Carbon nano-onions: Unique carbon nanostructures with fascinating properties and their potential applications. Inorganica Chim Acta. 2017;468:49–66. https://doi.org/10.1016/j.ica.2017.07.021
14. Gunture, Dalal C, Kaushik J, Garg AK, Sonkar SK. Pollutant-soot-based nontoxic water-soluble onion-like nanocarbons for cell imaging and selective sensing of toxic Cr(VI). ACS Appl Bio Mater. 2020;3(6):3906–3913. https://doi.org/10.1021/acsabm.0c00456
15. Baby R, Saifullah B, Hussein ZM. Carbon nanomaterials for the treatment of heavy metal-contaminated water and environmental remediation. Nanoscale Res Lett. 2019;14(341):1–17. https://doi.org/10.1186/s11671-019-3167-8
16. Sakulthaew C, Chokejaroenrat C, Poapolathep A. Hexavalent chromium adsorption from aqueous solution using carbon nano-onions (CNOs). Chemosphere. 2017;184:1168–1174. https://doi.org/10.1016/j.chemosphere.2017.06.094

17. Makhongoana A, Matsoso BJ, Mongwe TH, Coville NJ, Wamwangi D, Maubane-Nkadimeng MS. The role of oxygen in a carbon source (castor oil versus paraffin oil) in the synthesis of carbon nano-onions. Nanotechnology. 2021;32(13):135603. https://doi.org/10.1088/1361-6528/abd0b1

18. Singh H, Bhardwaj N, Arya SK, Khatri M. Environmental impacts of oil spills and their remediation by magnetic nanomaterials. Environ Nanotechnol Monit Manag. 2020;14:100305. https://doi.org/10.1016/j.enmm.2020.100305

19. Mongwe TH, Matsoso BJ, Mutuma BK, Coville NJ, Maubane MS. Synthesis of chain-like carbon nano-onions by a flame assisted pyrolysis technique using different collecting plates. Diam Relat Mater. 2018;90(October):135–143. https://doi.org/10.1016/j.diamond.2018.10.002

20. Shaku B, Mofokeng TP, Mongwe TH, Coville NJ, Ozoemena KI, Maubane-Nkadimeng MS. Physicochemical properties of nitrogen doped carbon nano-onions grown by flame pyrolysis from grapeseed oil for use in supercapacitors. Electroanalysis. 2020;32(12):2946–2957. https://doi.org/10.1002/elan.202060383

21. Sikeyi LL, Ntuli TD, Mongwe TH, Maxakato NW, Carleschi E, Doyle BP, et al. Microwave assisted synthesis of nitrogen doped and oxygen functionalized carbon nano onions supported palladium nanoparticles as hybrid anodic electrocatalysts for direct alkaline ethanol fuel cells. Int J Hydrogen Energy. 2021;46(18):10862–10875. https://doi.org/10.1016/j.ijhydene.2020.12.154

22. Khalisanni K, Khalizani K, Rohani MS, Khalid PO. Analysis of waste cooking oil as raw material for biofuel production. Glob J Environ Res. 2008;2(2):81–83. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.414.3929&rep=rep1&type=pdf

23. Ntuli TD, Mongwe TH, Sikeyi LL, Mkhari O, Coville NJ, Nxumalo EN, et al. Removal of hexavalent chromium via an adsorption coupled reduction mechanism using olive oil derived carbon nano-onions. Environ Nanotechnol Monit Manag. 2021;16(April):100477. https://doi.org/10.1016/j.enmm.2021.100477

24. Makonese T, Pemberton-Pigott C, Robinson J, Kimemia D, Annegarn H. Performance evaluation and emission characterisation of three kerosene stoves using a Heterogeneous Stove Testing Protocol (HTP). Energy Sustain Dev. 2012;16(3):344–351. https://doi.org/10.1016/j.esd.2012.06.002

25. Greenburg A, Clescerl L, Eaton A. Standard methods for the examination of water and wastewater. Washington DC: American Public Health Association; American Water Works Association and Water Environment Federation; 1992.

26. Zayed MA, Abd El-Kareem MSM, Zaky NHS. Gas chromatography-mass spectrometry studies of waste vegetable mixed and pure used oils and its biodiesel products. J Pharm Appl Chem. 2017;3(2):109–116. https://doi.org/10.18576/jpac/030204

27. Mannu A, Vlahopoulou G, Urgeghe P, Ferro M, Del Caro A, Taras A, et al. Variation of the chemical composition of waste cooking oils upon bentonite filtration. Resources. 2019;8(2):1–15. https://doi.org/10.3390/resources8020108

28. Zhang DC, Liu JJ, Jia LZ, Wang P, Han X. Speciation of VOCs in the cooking fumes from five edible oils and their corresponding health risk assessments. Atmos Environ. 2019;211(January):6–17. https://doi.org/10.1016/j.atmosenv.2019.04.043

29. Fang L, Ohfuji H, Shinmei T, Irifune T. Experimental study on the stability of graphitic $C_3N_4$ under high pressure and high temperature. Diam Relat Mater. 2011;20:819–825. https://doi.org/10.1016/j.diamond.2011.03.034

30. Ntuli TD, Pakade VE. Hexavalent chromium removal by polyacrylic acid-grafted Macadamia nutshell powder through adsorption-reduction mechanism: Adsorption isotherms, kinetics and thermodynamics. Chem Eng Commun. 2019;207(3):1–16. https://doi.org/10.1080/00986445.2019.1581619

31. Venkatesan RA, Balachandran M. Novel carbon nano-onions from paraffinum liquidum for rapid and efficient removal of industrial dye from wastewater. Environ Sci Pollut Res. 2020;27(35):43845–43864. https://doi.org/10.1007/s11356-020-09981-w

32. Bartolome JP, Fragoso A. Preparation and characterization of carbon nano-onions by nanodiamond annealing and functionalization by radio-frequency $Ar/O_2$ plasma. Fullerenes Nanotubes Carbon Nanostructures. 2017;25(5):327–334. https://doi.org/10.1080/1536383X.2017.1303604

33. Seymour MB, Su C, Gao Y, Lu Y, Li Y. Characterization of carbon nano-onions for heavy metal ion remediation. J Nanoparticle Res. 2012;14(9):1–14. https://doi.org/10.1007/s11051-012-1087-y

34. Ko YJ, Choi K, Lee S, Cho JM, Choi HJ, Hong SW, et al. Chromate adsorption mechanism on nanodiamond-derived onion-like carbon. J Hazard Mater. 2016;320:368–375. https://doi.org/10.1016/j.jhazmat.2016.08.041

35. Zeiger M, Jäckel N, Mochalin VN, Presser V. Review: Carbon onions for electrochemical energy storage. J Mater Chem A. 2016;4(9):3172–3196. https://doi.org/10.1039/C5TA08295A

36. Tripathi KM, Tran TS, Kim YJ, Kim T. Green fluorescent onion-like carbon nanoparticles from flaxseed oil for visible light induced photocatalytic applications and label-free detection of Al(III) ions. ACS Sustain Chem Eng. 2017;5(5):3982–3992. https://doi.org/10.1021/acssuschemeng.6b03182

37. Namieśnik J, Rabajczyk A. Speciation analysis of chromium in environmental samples. Crit Rev Environ Sci Technol. 2012;42(4):327–377. https://doi.org/10.1080/10643389.2010.518517

38. Qhubu MC, Mgidlana LG, Madikizela LM, Pakade VE. Preparation, characterization and application of activated clay biochar composite for removal of Cr(VI) in water: Isotherms, kinetics and thermodynamics. Mater Chem Phys. 2021;260(February):124165. https://doi.org/10.1016/j.matchemphys.2020.124165

39. Thabede PM, Shooto ND, Xaba T, Naidoo EB. Magnetite functionalized Nigella sativa seeds for the uptake of chromium(VI) and lead(II) ions from synthetic wastewater. Adsorpt Sci Technol. 2021;2021(January):1–15. https://doi.org/10.1155/2021/6655227

40. Shooto ND. Removal of lead(II) and chromium(VI) ions from synthetic wastewater by the roots of *Harpagophytum procumbens* plant. J Environ Chem Eng. 2020;8(6):104541. https://doi.org/10.1016/j.jece.2020.104541

41. Nchoe OB, Ntuli TD, Klink MJ, Mtunzi FM, Pakade VE. A comparative study of acid, base and Fenton-like reagent treated biomass for Cr(VI) sequestration from aqueous solutions. Water Environ Fed. 2020;93:370–383. https://doi.org/10.1002/wer.1421

42. Pholosi A, Naidoo EB, Ofomaja AE. Batch and continuous flow studies of Cr(VI) adsorption from synthetic and real wastewater by magnetic pine cone composite. Chem Eng Res Des. 2020;153(Vi):806–818. https://doi.org/10.1016/j.cherd.2019.11.004

43. Gnanasundaram N, Loganathan M, Singh A. Optimization and performance parameters for adsorption of $Cr^{6+}$ by microwave assisted carbon from Sterculia foetida shells. IOP Conf Ser Mater Sci Eng. 2017;206(1):1–10. https://doi.org/10.1088/1757-899X/206/1/012065

44. Park D, Lim SR, Yun YS, Park JM. Reliable evidences that the removal mechanism of hexavalent chromium by natural biomaterials is adsorption-coupled reduction. Chemosphere. 2007;70(2):298–305. https://doi.org/10.1016/j.chemosphere.2007.06.007

45. Lagergren S. Zur Theorie der Sogenannten Adsorption GelosterStoffe [On the theory of the so-called adsorption of gelatinous substances]. Kungliga Svenska Vetenskapsakademiens. Handl Band. 1898;24(4):1–39. German.

46. Blanchard G, Maunaye M, Martin G. Removal of heavy metals from waters by means of natural zeolites. Water Res. 1984;18(12):1501–1507. https://doi.org/10.1016/0043-1354(84)90124-6

47. Ho YS, McKay G. The kinetics of sorption of basic dyes from aqueous solution by sphagnum moss peat. Can J Chem Eng. 1998;76(4):822–827. https://doi.org/10.1002/cjce.5450760419

48. Pakade VE, Maremeni LC, Ntuli TD, Tavengwa NT. Application of quaternized activated carbon derived from macadamia nutshells for the removal of hexavalent chromium from aqueous solutions. South Afr J Chem. 2016;69:180–188. https://doi.org/10.17159/0379-4350/2016/v69a22

49. Langmuir I. The adsorption of gases on plane surfaces of glass, mica and platinum. J Am Chem Soc. 1918;40(9):1361–1403. https://doi.org/10.1021/ja02242a004

50. Tran HN, You SJ, Hosseini-Bandegharaei A, Chao HP. Mistakes and inconsistencies regarding adsorption of contaminants from aqueous solutions: A critical review. Water Res. 2017;120:88–116. https://doi.org/10.1016/j.watres.2017.04.014

51. Shokry A, El Tahan A, Ibrahim H, Soliman M, Ebrahim S. The development of a ternary nanocomposite for the removal of Cr(VI) ions from aqueous solutions. RSC Adv. 2019;9(67):39187–39200. https://doi.org/10.1039/C9RA08298K

52. Wang Z, Shen Q, Xue J, Jia H, Xu B, Liu X, et al. Annealing temperature effect on 3D hierarchically porous NiO/Ni for removal of trace hexavalent chromium. Mater Chem Phys. 2020;240:1–10. https://doi.org/10.1016/j.matchemphys.2019.122140

53. Mortazavian S, Saber A, Hong J, Bae JH, Chun D, Wong N, et al. Synthesis, characterization, and kinetic study of activated carbon modified by polysulfide rubber coating for aqueous hexavalent chromium removal. J Ind Eng Chem. 2019;69:196–210. https://doi.org/10.1016/j.jiec.2018.09.028

54. Shahrin S, Lau WJ, Goh PS, Jaafar J, Ismaila AF. Adsorptive removal of Cr(VI) and Cu(II) ions from water solution using graphene oxide-manganese ferrite (GMF) nanomaterials. Int J Eng Trans B Appl. 2018;31(8):1341–1346. https://doi.org/10.5829/ije.2018.31.08b.24

55. Gholipour M, Hashemipour H. Evaluation of multi-walled carbon nanotubes performance in adsorption and desorption of hexavalent chromium. Chem Ind Chem Eng Q. 2012;18(4-I):509–523. https://doi.org/10.2298/CICEQ111104025G

# One-pot hydrothermal green synthetic approach of fluorescent carbon dots as optical probes for 2-nitrophenol

**AUTHORS:**
Pogisego Dinake[1]
Gothatamang N. Phokedi[1]
Mbhatshi M. Keetile[1]
Mmamiki A. Botlhomilwe[1]
Mogomotsi Tlhako[1]
Bokang Present[2]
Janes Mokgadi[1,3]
Rosemary Kelebemang[1,4]

**AFFILIATIONS:**
[1]Department of Chemical and Forensic Sciences, Botswana International University of Science and Technology, Palapye, Botswana
[2]Department of Physics and Astronomy, Botswana International University of Science and Technology, Palapye, Botswana
[3]Chemical, Biological Nuclear and Radiological Weapons Management Authority, Ministry of Defence, Justice and Security, Gaborone, Botswana
[4]National Environmental Laboratory, Department of Waste Management and Pollution Control, Gaborone, Botswana

**CORRESPONDENCE TO:**
Pogisego Dinake

**EMAIL:**
dinakep@biust.ac.bw

**DATES:**
**Received:** 11 May 2022
**Revised:** 07 Feb. 2023
**Accepted:** 09 Feb. 2023
**Published:** 28 Sep. 2023

**HOW TO CITE:**
Dinake P, Phokedi GN, Keetile MM, Botlhomilwe MA, Tlhako M, Present B, et al. One-pot hydrothermal green synthetic approach of fluorescent carbon dots as optical probes for 2-nitrophenol. S Afr J Sci. 2023;119(9/10), Art. #13921. https://doi.org/10.17159/sajs.2023/13921

**ARTICLE INCLUDES:**
☒ Peer review
☐ Supplementary material

**DATA AVAILABILITY:**
☐ Open data set
☐ All data included
☐ On request from author(s)
☐ Not available
☒ Not applicable

**EDITORS:**
Priscilla Baker
Amanda-Lee Manicum

**KEYWORDS:**
2-nitrophenol, fluorescence, carbon dots, green synthesis, biomass, *Citrullus vulgaris*

**FUNDING:**
Botswana International University of Science and Technology (BIUST/ds/R&I/26/2016)



The pursuit of a cost-effective and green synthetic approach to chemical sensors and their application in the sensing of toxic and harmful substances is a never-ending exercise for scientists and researchers. Preparation of fluorescent carbon dots (C-dots) from biomass using water as a solvent and a hydrothermal autoclave to provide the required synthesis temperature offers a cheap and environmentally friendly synthetic approach. Herein, we report a faster, less costly and ecofriendly hydrothermal synthetic approach of carbon dots from *Citrullus vulgaris* peels as a precursor. The as-prepared carbon dots exhibited hydroxyl, carbonyl and amide functional groups on the surface and an amorphous structure with a particle size distribution of 1.7–3.0 nm. Moreover, the carbon dots displayed intense blue emission fluorescence at 470 nm after excitation at 400 nm. The as-prepared carbon dots demonstrated effective application without further modification towards the selective and sensitive optical recognition of 2-nitrophenol used in the manufacture of explosives. A limit of detection of $2.28\times10^{-7}$ M was achieved, and no fluorescence quenching was observed in the presence of other nitroaromatic and benzene derivatives indicating excellent selectivity towards 2-nitrophenol. Finally, further studies are required to investigate the potential for the as-prepared carbon dots to monitor nitroaromatic pollutants in real environmental systems.

**Significance:**

- Terrorism is an ever-increasing problem, and law enforcement agencies are continuously searching for and detecting explosives hidden in travel luggage, mail packages, vehicles and aircrafts using sophisticated equipment which are not available in developing countries such as Botswana.
- This work unveils a facile and environmentally friendly approach towards the detection of 2-nitrophenol used in the manufacture of explosives by employing highly luminescent C-dots obtained from locally available agricultural waste.
- The utilisation of agricultural waste can help advance a sustainable waste management programme and promote a circular economy.

## Introduction

Terrorism is an ever-increasing problem due to an array of chemicals that can be employed as precursors in the manufacture of explosives.[1] Human social security is under constant threat due to a surge in well-organised and sophisticated terrorist bomb attacks. Law enforcement and security officers are continuously searching for and detecting hidden explosives in travellers, travel luggage, mail packages, vehicles and aircrafts. Therefore, analytical field and laboratory detection and unearthing of explosives are of utmost importance in bomb neutralisation and protection of lives and property.[2] Nitroaromatic compounds are common major components of explosives.[3] Their substantial use has resulted in the release of elevated levels of these organic compounds into the environment leading to pollution of soil and water.[4] Additionally, the nitroaromatic compounds are poorly biodegradable, highly stable and result in adverse health effects in humans due to their toxicity.[5] Nitroaromatic compounds such as 2-nitrophenol (2-NP) are not only used in the manufacture of explosives but have also found extensive use in the production of pesticides, herbicides, insecticides, dyes, rubber chemicals, textile, petroleum solvents, pharmaceuticals and as intermediates in the synthesis of other chemicals.[6] The hydrolysis of 2-NP containing explosive residues results in the release of 2-NP into the environment causing contamination of the environment and public health risk owing to its acute toxicity and mutagenic capabilities.[7] The adverse health effects associated with exposure to 2-NP include kidney and liver damage, cancer and blood disorder.[8] As a result, 2-NP has been included in the United States Environmental Protection Agency (USEPA) list of priority pollutants owing to its carcinogenic and bioaccumulation effects.[9] On account of its highly toxic effects, this nitroaromatic compound has a maximum contaminant limit (MCL) of 4.8 μg/L as set by the USEPA.[10] Therefore, it is of paramount importance that highly sensitive and selective methods are available for the low-level detection of this toxic organic chemical.

Presently, various conventional methods and techniques have been applied in the analysis of 2-NP such as chromatography, electrochemistry and spectroscopic techniques.[11-13] Regrettably, these techniques generally require sophisticated and costly equipment accompanied by complex sample preparation and pre-treatment steps that are time consuming.[6] Some of these methods lack low analyte discriminating ability towards nitroaromatic isomers such as 4-nitrophenol, 2-nitrophenol and 3-nitrophenol which restricts their broad utilisation.

This has necessitated scientists and researchers to continuously search for new methods and techniques that are not only cost effective but also sensitive and selective towards 2-NP. Amidst these several techniques, photoluminescence has been extensively explored owing to its simplicity and high sensitivity. Another important attribute of luminescent-based recognition and sensing techniques is the capability to detect explosives from afar.[2] A great number of studies have been carried out on changes in the photoluminescence properties of certain substances upon exposure to 2-NP. For example, in a study by Chaudhary et al. 2019, fluorescent europium oxide ($Eu_2O_3$) nanoparticles that were surface modified with (3-aminopropyl)triethoxysilane (APTES) successfully detected 2-NP in aqueous media achieving detection limits of $4.6\times10^{-6}$ M.[14] The optical recognition of 2-NP was based on fluorescence quenching of the $Eu_2O_3$ nanoparticles by the 2-NP. The mechanism of this interaction involves donor–acceptor interaction between the electron-deficient 2-NP and the electron-rich primary amines on the surfaces of the $Eu_2O_3$ nanoparticles leading to the formation of Meisenheimer complexes (MHCs).[14] In a similar study, luminescent silicon nanoparticles based on *N*-[3-(trimethoxysilyl) propyl]-ethylenediamine (DAMO) and functionalised with dopamine moiety were used for selective detection of 2-NP in the presence of other nitroaromatic compounds such as 3-nitrophenol, 4-nitrophenol, 2,4-dinitrophenol, 2,4,6-trinitrophenol, nitrotoluene, 3,5-dinitrobenzoic acid, 2,4-dinitrotoluene, *p*-nitrobenzoic acid, 1,3-dinitrobenzene, 2,4,6-trinitrotoluene, nitrobenzene and metal cations such as $Cu^{2+}$, $Zn^{2+}$, $Mn^{2+}$, $Al^{3+}$, $K^{+}$, $Ca^{2+}$, $Mg^{2+}$ and $Ba^{2+}$.[6] A detection limit of $2.9\times10^{-8}$ M was achieved for sensing of 2-NP using the luminescent silicon nanoparticles. Photoluminescent quantum dots doped containing heavy metals have also been applied as optical probes for nitroaromatic compounds such as 2-NP.[15,16] Liu and co-workers have investigated the use of fluorescent hybrid copper (I) iodine cluster-based sensor ([Cu4I4(RTBT)4], where ETBT is 2-ethylbenzo[d]thiazole) for optical recognition of 2-NP in aqueous media achieving high selectivity and low detection limits of $2.3\times10^{-6}$ M.[17] Similarly, Uddin et al.[15] utilised modified glassy carbon electrode bearing $ZnO/RuO_2$ nanoparticles for the ultrasensitive and excellent selective detection of 2-NP with a detection limit of $5.22\times10^{-13}$ M. Even though the use of luminescent quantum dots has been found to be efficient and effective as sensitive optical probes for 2-NP, they have rather shown some unappealing attributes such as high toxicity due to the use of toxic metals and solvents, high cost of starting materials and laborious preparation processes.[16] Therefore, there is a continuous need to develop rapid, sensitive and selective methods that are simple and cost effective for the detection of 2-NP.

The discovery of carbon dots (C-dots) in 2004 has brought about a new perspective in chemical sensors and the attention of scientists and researchers has shifted towards exploiting the photoluminescence properties of these carbon nanomaterials.[18] C-dots possess fascinating photoluminescent characteristics that are controlled by their size, shape, defects and surface functionalisation.[19] The massive interest in carbon dots is believed to stem from the ease of their preparation from abundant precursors, functionalisation, low cost, low toxicity, biodegradability, tunable optical properties, biocompatibility, high chemical and photostability, excellent water solubility and immense potential for considerable applications.[20] This broad spectrum of attractive properties of carbon dots has resulted in their versatile application potential such as chemical sensing, drug delivery, catalysis, bioimaging and forensic applications.[21-25] To reap the full benefits of carbon dots, scientists and researchers are continuously searching for cheaper methods with high specificity and excellent yields for their synthesis. Two synthetic pathways have been exploited in the preparation of carbon dots, being the top-down and bottom-up approaches.[26] The top-down strategy employs the preparation of C-dots from the breakdown of large carbon source precursors such as multi-walled carbon nanotubes (MWCNTs), carbon fibres and graphite under aggressive physical or chemical reaction conditions.[27] In contrast, bottom-up synthetic strategies involve flexible synthesis procedures that can be modified to give an array of C-dots with distinct features.[28] They use countless molecular organic precursors which encompass smaller substances such as carbohydrates and more diverse carbon-containing materials such as biomass.[29] Compared to top-down synthesis strategies, bottom-up approaches are favoured because of low cost and potential for mass production.[30] Previous studies have predominately used commercially available carbon-containing materials as precursors in the preparation of C-dots as optical sensors of 2-NP.[20] Commercially available carbon sources such as ethylene diamine tetraacetic acid (EDTA) and glutathione have been utilised as carbon sources in the synthesis of C-dots used as optical sensors for 2-NP.[20,31]

Few studies have reported taking advantage of the abundant biomass as carbon sources for the preparation of C-dots and application in the sensing and detection of 2-NP. Recently, there was a preparation of C-dots using bamboo leaves through a thermal pyrolysis synthetic procedure in a muffle furnace.[32] The biomass-obtained C-dots displayed excellent photoluminescent characteristics such as a high fluorescence quantum yield of up to 5.18%. The as-prepared C-dots also exhibited high sensitivity and selectivity towards detection of 2-NP achieving a quenching constant of 102 L/mol. In another recent study, C-dots derived from celery leaves showed high selectivity and sensitivity towards sensing of 2-NP and a detection limit of $3.9\times10^{-10}$ M was determined.[31] Therefore, the use of biomass as precursors for the synthesis of C-dots gives the advantage of low cost and ecofriendliness, and the procedures are easily scalable.

The present work continues the quest to look for alternative sources of precursors in the synthesis of C-dots, especially locally available biomass. This work unveils a facile and environmentally friendly approach with potential scalability for the fabrication of highly luminescent C-dots using locally available agricultural waste as the carbon precursor. Peels from the fruit of *Citrullus vulgaris* were used in the preparation of the C-dots following the green hydrothermal synthetic approach. The utilisation of agricultural waste can help advance a sustainable waste management programme. The as-prepared C-dots exhibited excellent water solubility, low toxicity and good chemical and photostability. Furthermore, the as-prepared C-dots showed high sensitivity and selectivity towards sensing of 2-NP without further modification and functionalisation. This makes the biomass-derived C-dots a good candidate for application in the detection of 2-NP in real samples such as water and soil.

## Materials and methods

### *Reagents and materials*

The peels from the fruit of *C. vulgaris* were collected from hawkers selling the melons in the local market of Palapye, Botswana, after the harvest season. Double-deionised water was produced in the Chemistry laboratory using the Milli-Q system and was used throughout this study without any further purification. A 500-mL hydrothermal synthesis autoclave reactor with a Teflon chamber was purchased from Techinstro (Nagpur, India). A dialysis membrane with a flat width of 43 mm, average diameter of 27 mm and molecular weight cut-off (MWCO) of 14 000, potassium phosphate monobasic (≥99.0%), potassium dibasic (≥98.0%), anthracene (≥97.0%), 2-NP (≥99.0%) and dinitrophenylhydrazine (≥97.0%) were all purchased from Sigma Aldrich (St Louis, USA). All chemicals used were of analytical grade and used as received without any further purification.

### *Preparation of carbon dots*

The C-dots were synthesised via a simple, green and one-pot hydrothermal synthesis procedure. The fruit of *C. vulgaris* peels was washed with plenty of deionised water, and this process was repeated several times to obtain clean material. The peels were then cut into small pieces and dried in an oven at 60 °C for 36 h. The dried material was then ground into powder using a pestle and mortar. About 2 g of the powder was dissolved into 40 mL of deionised water using a magnetic stirrer for about 10 min. After the complete dissolution of the powder, the solution was then placed into an autoclave and placed in the oven. The time and temperature for the hydrothermal synthesis procedure were varied until an optimum time and temperature of 1 h and 100 °C were achieved. After cooling the solution to room temperature, the resultant solution was then homogenised under sonication (SonicClean, Labotec) for 20 min and placed into a centrifuge (Thermo Scientific, Heraeus, Megafuge 40 Centrifuge) at 8000 rpm for another 20 min. The solution was then filtered using a 0.22-μM syringe filter. The filtered solution was dialysed (retained molecular weight: 14 000 Da) against deionised water for 24 h. The final C-dots light brown solution was stored at 4 °C for further characterisation and application. The synthesis process of the as-prepared C-dots is depicted in Figure 1.

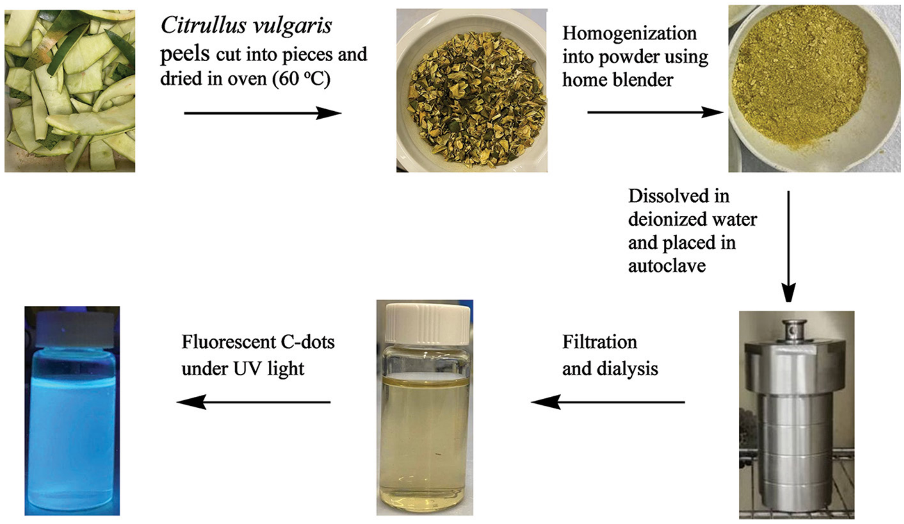



**Figure 1:** Schematic representation of the preparation of C-dots via the hydrothermal method.

### *Optimisation of preparation time and temperature*

The optimum conditions for the preparation of the C-dots such as time and temperature were determined experimentally. The *C. vulgaris* peels aqueous solutions were subjected to a temperature range of 25–250 °C. The temperature at which the as-prepared C-dots produced the most intense fluorescence was taken as the optimum temperature. Similarly, a time in the range of 0–4 h was selected for the hydrothermal preparation of the carbon dots at a fixed temperature of 100 °C. The optimum C-dots preparation time was taken as the time that resulted in the most intense fluorescence.

### *Equipment for characterisation of the as-prepared C-dots*

The size and morphology of the as-prepared C-dots were determined using an atomic force microscopy (AFM), (Bruker AFM dimension edge in Scan Asyst Peak Force), in the tapping mode. The AFM sample was prepared by placing a drop of the purified C-dots solution on a clean and dry AFM mica slide. The solution was dried at room temperature in a desiccator for 2 h. Fluorescence studies were carried out using a Perkin–Elmer LS 55 spectrofluorometer. All fluorescence measurements were taken at room temperature. The existence of different functional groups on the surface of the as-prepared C-dots was determined using a Thermo Scientific Nicolet iS10 FTIR spectrometer operated at 32 (0.50) $cm^{-1}$ resolution. The FTIR spectra of the pure C-dots were obtained after filtration and dialysis of the aqueous solution. The absorption characteristics of the as-prepared C-dots were established using a Thermo Scientific Evolution 201 UV–Vis spectrophotometer. All the measurements were carried out at room temperature.

### *Photoluminescent quantum yield of the as-prepared C-dots*

The photoluminescent quantum yield (FQY) of the as-prepared C-dots was evaluated using a method reported in the literature.[20] The FQY of the as-prepared C-dots was estimated using Equation 1 using anthracene dissolved in ethanol as a reference compound.

$$\phi_{\text{C-dot}} = \phi_{\text{std}} \times \frac{I_{\text{C-dot}} A_{\text{std}} \eta^{2}_{\text{C-dot}}}{A_{\text{C-dot}} I_{\text{std}} \eta^{2}_{\text{std}}}$$ Equation 1

where $\phi_{\text{C-dot}}$ and $\phi_{\text{std}}$ are the respective photoluminescent quantum yields of the as-prepared C-dots and anthracene standard compound. $I_{\text{C-dot}}$ and $A_{\text{C-dot}}$ represent the integrated fluorescence intensity and absorbance of the as-prepared C-dots whilst $\eta_{\text{C-dot}}$ is the refractive index of water used to dissolve the C-dots. On the other hand, $I_{\text{std}}$, $A_{std}$ are the integrated fluorescence intensity and absorbance of the anthracene standard material whilst $\eta_{\text{std}}$ is the refractive index of ethanol in which anthracene is dissolved. The FQY of anthracene in ethanol was taken as 0.27 after excitation at 350 nm as reported in the literature.[33] Anthracene was a preferred reference compound because its emission falls within the range of the emission spectrum of the as-prepared C-dots (360–480 nm). Plots of $I_{\text{C-dot}}$ versus $A_{\text{C-dot}}$ and $I_{\text{std}}$ versus $A_{std}$ were made and their slopes were used in the estimation of FQY for the as-prepared C-dots in Equation 1. The absorbance of the as-prepared C-dots and that of the anthracene in the 1 cm cuvette were kept under 0.1 so that suppression of re-absorption effects inside the solutions on the obtained emissions can be achieved.

### *Optimisation of sensing capability of as-prepared C-dots*

Optimisation of conditions for the optical sensing capabilities of the as-prepared C-dots was also carried out. The effect of pH on the photoluminescence properties of the as-prepared C-dots was established. The as-prepared C-dots were dissolved in phosphate buffer in the pH range of 5–13 and the pH at which the fluorescence intensity of the as-prepared C-dots was the highest was taken as the optimised pH for sensing applications.

### *Procedure for optical sensing of 2-nitrophenol using the as-prepared C-dots*

The sensing of 2-nitrophenol was carried out by adding 2.0 mL of varying concentrations of the 2-nitrophenol after dissolution in phosphate buffer (pH 10.0) to 1.0 mL of the as-prepared C-dot contained in a 5.0-mL centrifuge tube. The total volume of the solutions was maintained at 3.0 mL. The solution mixtures were then agitated for 1 min in a multi-pulse vortexer machine and transferred into a quartz cuvette. The fluorescence spectra of the solution mixture were then recorded under the experimentally established excitation wavelength of 400 nm. Method validation was carried out by determining the limit of detection (LOD) and limit of quantification (LOQ) of the as-prepared C-dots towards 2-nitrophenol. Selectivity studies were carried out by introducing potential interferences into the C-dot solution in place of 2-nitrophenol and obtaining the fluorescence spectra.

## Results and discussion

### *Structure and morphological characteristics of the as-prepared C-dots*

The particle size, spatial distribution, morphology and crystal lattice of the as-prepared C-dots were explored using an atomic force microscope (AFM) at room temperature. As shown in Figure 2(a) and (b), the C-dots prepared from biomass indicated an amorphous structure with no sign of crystallographic order. The particles are small and irregularly shaped and mainly distributed between 1.7 and 3.0 nm with an average particle size of 2.5 nm (Figure 2(c)). The 2-D and 3-D AFM images (Figure 2(a)

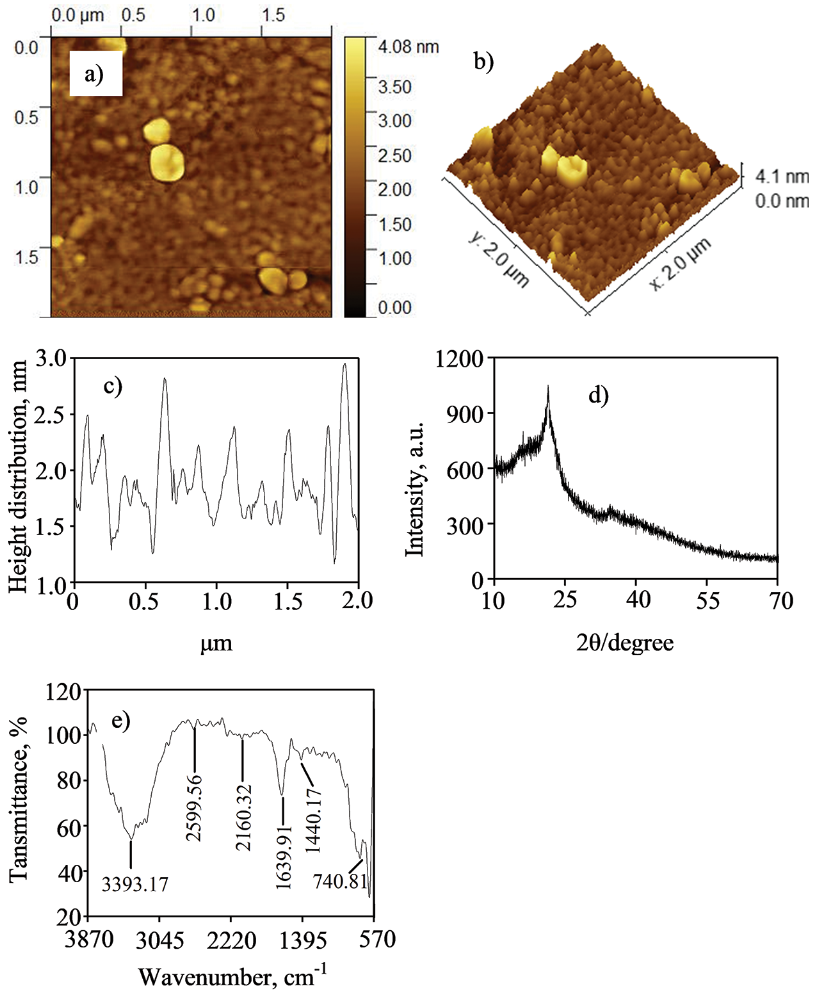



**Figure 2:** Characterisation of the as-synthesised C-dots using AFM and FTIR: (a) two-dimensional image of the C-dots; (b) three-dimensional image of the C-dots; (c) height distribution of the as-prepared C-dots; (d) XRD pattern of the as-prepared C-dots and (e) FTIR spectrum of the as-prepared C-dots.

and (b)) show that the particles are mono-dispersed with moderate aggregation which indicates good water solubility. The diameter of the C-dots indicates that the nanoparticles comprise several layers, demonstrating graphite-like structure of the C-dots.[34] The structural pattern and morphology of the as-prepared C-dots were also studied using the XRD technique. XRD pattern of the as-prepared C-dots exhibits an amorphous carbonaceous core (Figure 2(d)).[35] The XRD pattern revealed a distinct peak at $2\theta = 24^{o}$ which suggests intergranular accumulation of carbon atoms in the carbon amorphous phase.[36] This confirms the non-crystalline amorphous nature of the C-dots obtained using AFM (Figure 2(a) and (b)). In addition, the peak at $2\theta = 24°$ confirms the graphitic character of the amorphous carbon core.[36] The different functional groups present on the surface of the as-prepared C-dots were determined using FTIR. As shown in Figure 2(e), the broad peak centred at 3393 $cm^{-1}$ suggests the presence of the hydroxyl (–OH) and amine (–NH) groups, and the peaks at 1639 $cm^{-1}$ and 1440 $cm^{-1}$ are assigned to the carbonyl (–C=O) and C–N stretching vibrations, respectively. The weak absorption band at 2160 $cm^{-1}$ represents the stretching vibrations of the –C–H and –C=CH.[37] In addition, the absorption band at 740 $cm^{-1}$ is assigned to $NH_2$ wagging band and the bending vibration band of the aromatic C–H bond.[38] The presence of the hydroxyl and amine functional groups on the surface of the C-dots is responsible for their solubility in water.

## *Optimisation of hydrothermal synthesis conditions of C-dots*

The optimum conditions at which the C-dots revealed the most intense photoluminescence were determined experimentally. The hydrothermal preparation conditions were based on synthesis time and temperature (Figure 3(a) and (b)). In both cases, the photoluminescence of the as-prepared C-dots was determined at an excitation wavelength of 400 nm. Initially, the fluorescence intensity of the freshly ground *C. vulgaris* peels was determined and revealed very low fluorescence (Figure 3(a); fresh precursor). The hydrothermal synthesis temperature that produced C-dots with the most intense photoluminescence was 100 °C after exposing the *C. vulgaris* peels to different temperatures in the range 25–250 °C. It should be noticed that the fluorescence intensity of the as-prepared C-dots decreased when the temperature was raised beyond 100 °C. A further increase of temperature up to 250 °C resulted into complete decomposition of melon peels into ashes instead of the carbon-rich residue. Alternatively, the optimum time required for the hydrothermal synthesis of the C-dots was determined to be 2 h (Figure 3(b)). This is the synthesis time that resulted in the C-dots bearing the most intense photoluminescence. Likewise, increasing synthesis time beyond 2 h caused deterioration of the C-dots photoluminescence (Figure 3(b)). This was due to the complete conversion of the carbon source material into ashes and rather than the carbon-rich nanomaterial.[39] Another important point to note is that there is a red shift of the emission wavelength when the synthesis temperature is increased from room temperature to 250 °C (Figure 3(a)). A similar observation was made when the synthesis time was increased from 1 h to 4 h (Figure 3(b)). A plausible explanation for the red shift in the emission wavelength is the broad size distribution of the as-prepared C-dots.[40] C-dots that possess small sizes have been reported to exhibit blue shift emissions whilst larger ones show red shift emissions.[40] There was no noticeable change in the emission wavelength maxima for C-dots solutions obtained at higher temperatures of 200 and 250 °C from that of

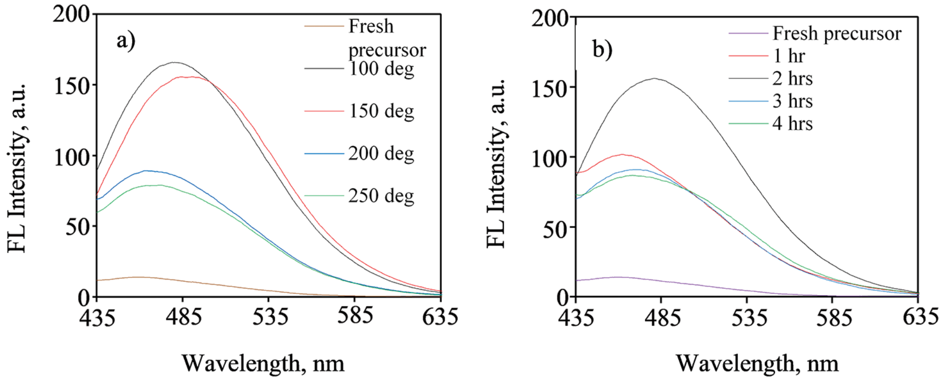



**Figure 3:** Optimisation of synthesis conditions: (a) fluorescence spectra of the as-prepared C-dots at different synthesis temperature and (b) fluorescence spectra of the as-prepared C-dots at different synthesis times ($\lambda_{ex}$ = 400 nm).

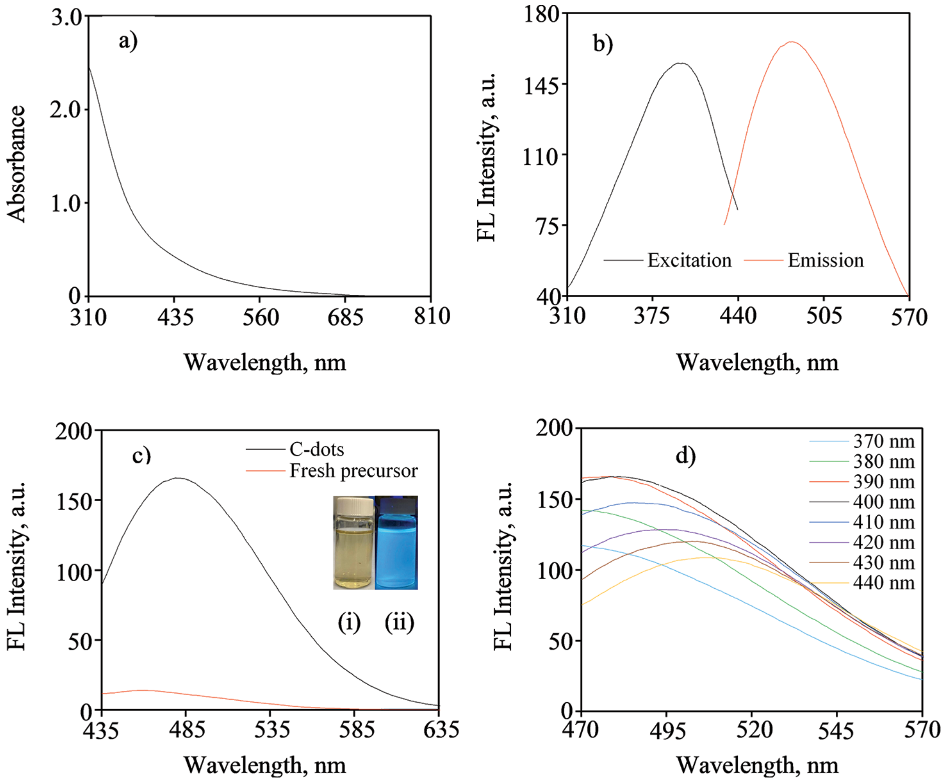



**Figure 4:** Optical characteristics of C-dots: (a) UV–Vis spectrum of the as-prepared C-dots; (b) excitation-emission spectra of the as-prepared C-dots ($\lambda_{ex}$ = 400 nm and $\lambda_{em}$ = 470 nm); (c) emission spectra of the fresh precursor (red curve) and C-dots (black curve) ($\lambda_{ex}$ = 400 nm); inset (i) C-dots in deionised water under daylight and (ii) C-dots in deionised water under UV light at 365 nm; and (d) emission spectra of the as-prepared C-dots at different excitation wavelengths ($\lambda_{ex}$ = 370–440 nm).

the fresh precursor. This could be due to the complete decomposition of the precursor instead of the carbon-rich nanomaterial resulting in similar optical characteristics of the decomposed material to that of the starting material.[40] A similar shift to a longer emission wavelength was observed when heating time of the carbon source material was increased from 1 to 2 h with a more pronounced shift at 2 h. (Figure 3(b)). A small change in the emission wavelength maxima was observed when heating time was increased to 3 h and 4 h, an indication of decomposition of the carbon source material and loss of optical characteristics.

### *Photoluminescence and absorption properties of the as-prepared C-dots*

The absorption spectrum of the C-dots solution derived from *C. vulgaris* peels did not reveal any distinctive sharp peak but rather a broad absorption band in the range of 310–685 nm (Figure 4(a)). The absence of a distinct absorption peak corroborates the XRD pattern of the possibility of a broad particle size distribution of the as-prepared C-dots.[41] The excitation wavelength ($\lambda_{ex}$) responsible for the photoluminescence of the as-prepared C-dots was determined experimentally by running the absorption spectrum of the C-dots using spectrofluorometer (Figure 4(b)). The excitation wavelength that resulted in the most intense fluorescence of the as-prepared C-dots was found to be 400 nm. The emission spectrum revealed a broad peak with an emission wavelength ($\lambda_{em}$) maxima of 470 nm and a full width at half maxima (FWHM) of about 100 nm which supports the XRD results of a possible broad particle size distribution of the as-prepared C-dots.[42] However, the excitation-emission spectra showed a relatively broad Stokes shift of around 70 nm which suggests the photoluminescence of the as-prepared C-dots may be originating from their small emission energy gap and are free from the self-quenching effect (Figure 4(b)).[43] As shown in Figure 4(c), the as-prepared C-dots solution exhibited a

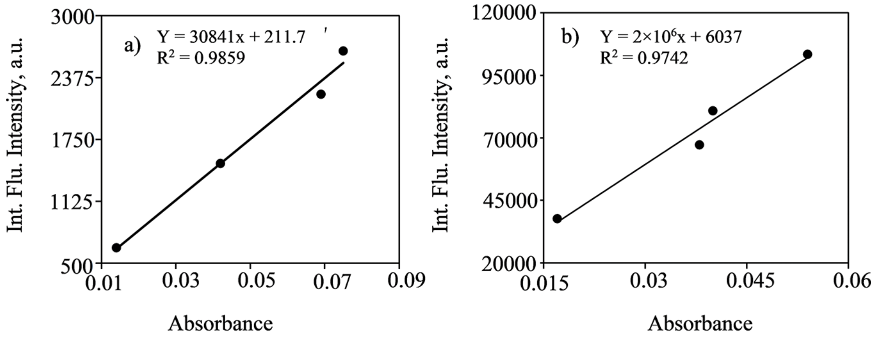



**Figure 5:** Determination of the fluorescence quantum yield of the as-prepared C-dots. (a) Linear plot of integrated fluorescence intensity versus absorbance for as-prepared C-dots ($\lambda_{ex}$ = 400 nm). (b) Linear plot of integrated fluorescence intensity versus absorbance for the anthracene standard in ethanol ($\lambda_{ex}$ = 350 nm).

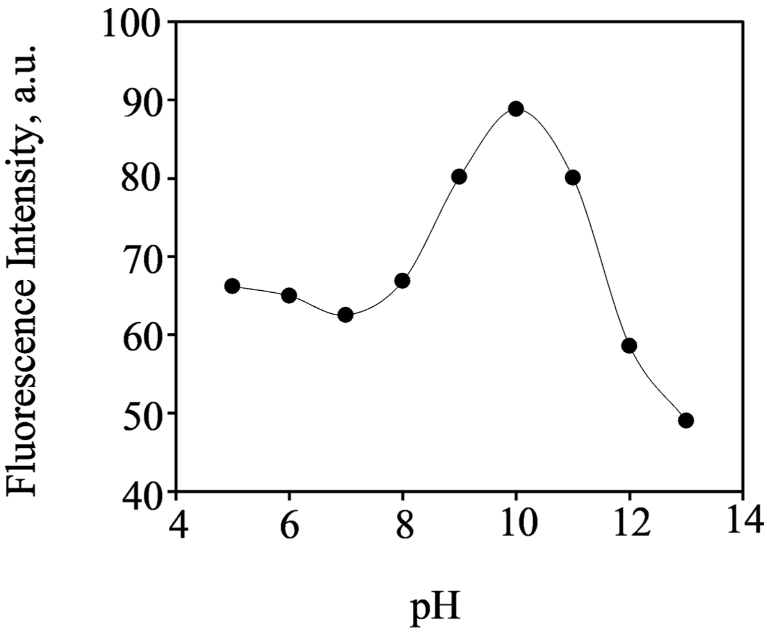



**Figure 6:** Determination of the effect of pH on the fluorescence intensity of the as-prepared C-dots: $C_{phosphate\ buffer}$ = 0.1 M; $\lambda_{ex}$ = 400 nm.

bright blue colour under UV irradiation at 365 nm. The inset of Figure 4(c) demonstrates the colour of the as-prepared C-dots solution under daylight (pale brown) and UV lamp at 365 nm (bright blue). It is clear from Figure 4(c) that the as-prepared C-dots exhibit intense fluorescence as compared to the starting material. This implies that hydrothermal heating at optimum conditions of temperature and time results in an efficient and effective carbonisation process through a well-organised structure disintegration and depolymerisation of the precursor carbon polymer sequences into sub-microscopic aggregates.[40] The photoluminescence properties of the as-prepared C-dots were further investigated by obtaining the emission spectra by varying the excitation wavelength from 370 nm to 440 nm (Figure 4(d)). A red shift in the emission spectra of the as-prepared C-dots was observed as the excitation wavelength was increased progressively from 370 to 400 nm with a concomitant increase in fluorescence intensity. The fluorescence intensities then decreased after excitation at 410–420 nm whilst maintaining a red shift of the emission wavelength. The as-prepared C-dots exhibited the most intense fluorescence after excitation at 400 nm which is in agreement with the excitation-emission spectra in Figure 4(b). A similar pattern of the excitation-dependent fluorescence intensities exhibited by the as-prepared C-dots was observed in a similar study.[44] The shift in the emission maxima as the excitation wavelength is progressively increased reveals the potential luminescence tuning of the as-prepared C-dots. The excitation-dependent fluorescence of the as-prepared C-dots that is accompanied by changes in their emission maxima corroborates the XRD and absorption studies of a possible broad particle size distribution and diverse functional groups on the surface states.[44] In a similar study, the emission maxima occurred at a fixed wavelength when the excitation wavelength was progressively increased from 350 to 450 nm, suggesting that the C-dots synthesised were of high purity and uniform size distribution.[38] The optical properties of C-dots have been proposed to emanate from the size of the C-dots, structural defects, aromatic conjugate assembly and the presence of $sp^2$ sites.[36] The fluorescence quantum yield (FQY) of the as-prepared C-dots in aqueous solution (refractive index~1.33) was established experimentally against a reference standard of anthracene exhibiting FQY of 0.27 in ethanol (refractive index ~1.36) (Figure 5). The FQY of the as-prepared C-dots was calculated to be 0.40. The low FQY of the as-prepared C-dots can be attributed to the broad particle size distribution.[45] The FQY of the as-prepared C-dots can be improved through surface passivation and optimisation of other reaction conditions.[46]

### *Effect of pH on optical properties of C-dots*

The effect of pH on the optical characteristics of the as-prepared C-dots was determined experimentally in a pH range of 5–13 using 0.1 M phosphate buffer solutions (Figure 6). The as-prepared C-dots exhibited the most intense fluorescence at pH 10. The fluorescence intensity decreased slightly as the pH was increased from 5 to 7 and

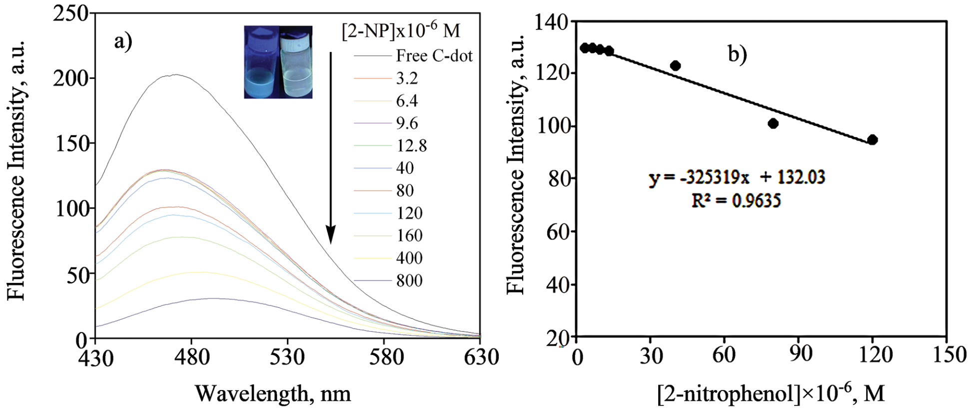



**Figure 7:** (a) Fluorescence spectra of the as-prepared C-dots in different concentrations of 2-NP in phosphate buffer (pH 10.0). (b) Plot of linear relationship between *I* and concentration of 2-NP (*I* = fluorescence of C-dot in the presence of 2-NP). $\lambda_{ex}$ = 400 nm, slit size = 2.5 nm.

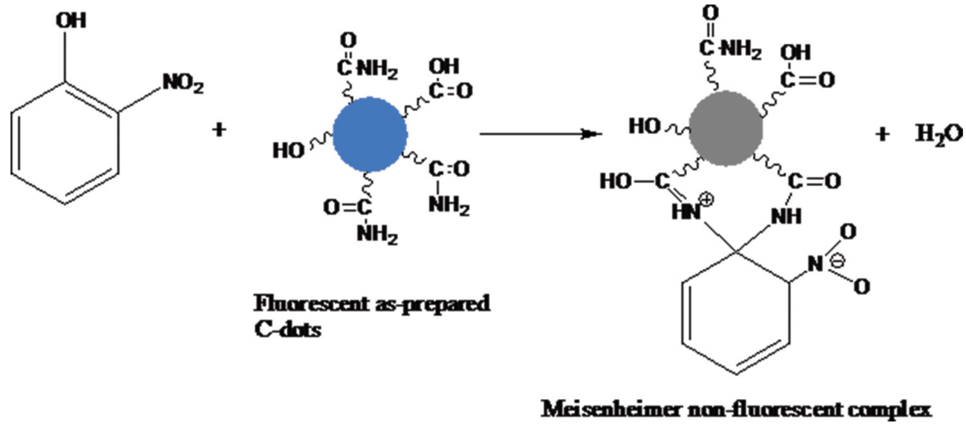



**Figure 8:** Proposed possible fluorescence quenching mechanism of C-dots in the presence of 2-NP.[50]

then increased substantially from pH 8 reaching a maximum at pH 10. A drastic decrease in the fluorescence intensity was observed between pH 11 and 13. There are several explanations for these observations: firstly, there is the possibility of protonation of functional groups on the surface of C-dots such as carboxyl and amine moieties leading to aggregation which may result in fluorescence quenching of the as-prepared C-dots in highly acidic media.[47] In more basic conditions (pH 11–13), the presence of hydroxyl anion ($OH^-$) leads to the abstraction of the protons from the carboxyl and amine groups forming carboxylate and amide ionic groups which result in the quenching of the fluorescence of the as-prepared C-dots.[48] Subsequently, the optical sensing studies of the as-prepared C-dots were carried out at an optimal pH of 10.

## *Application for optical detection of 2-nitrophenol*

The intense luminescence of the as-prepared C-dots was exploited towards sensing of the 2-NP explosive additive in phosphate buffer at pH 10. Successive concentrations of 2-NP (0–8×10$^{-4}$ M) were added to the C-dots solution (Figure 7(a)). The fluorescence intensity of the C-dots decreased substantially upon the addition of increasing concentrations of 2-NP. A fluorescence quenching efficiency of about 85% was determined after the addition of 8×10$^{-4}$ M of 2-NP to the C-dots solution. In addition, a red shift in the emission wavelength was observed with the concomitant addition of increasing concentrations of 2-NP at a fixed excitation wavelength of 400 nm. The insert of Figure 7(a) shows the visual observation of the quenching efficiency of 2-NP under UV light (365 nm). To obtain a quantitative understanding of the sensitivity of the as-prepared C-dots towards 2-NP, the LOD and LOQ for the explosive additive were determined. A Stern-Volmer plot was obtained, and its slope was used to estimate the LOD and LOQ at $\lambda_{ex}$ of 400 nm (Figure 7(b)). The LOD and LOQ were, respectively, calculated using the formulae $3\sigma/s$ and $10\sigma/s$ where $\sigma$ is the standard deviation of the blank ($n$ = 6), obtained from measurement of fluorescence intensities of the as-prepared C-dots in the absence of any analyte and $s$ is the magnitude of the slope of fluorescence intensity versus concentration of 2-NP plot. The as-prepared C-dots were able to detect 2-NP with a linear range of 3.2×10$^{-6}$ to 1.2×10$^{-4}$ M. A low LOD of 2.28×10$^{-7}$ M and LOQ of 7.60×10$^{-7}$ M were calculated for the sensitive optical recognition of the as-prepared C-dots towards 2-NP. The LOD of 2-NP obtained in this study is similar to the results obtained elsewhere for studies of the sensitivity of C-dots towards other nitrophenols.[20,31,49] An LOD of 7.7×10$^{-8}$ M towards detection of 2-NP was achieved in a previous study where ethylenediaminetetraacetic acid (EDTA.2Na) and urea were used to prepare C-dots under hydrothermal conditions.[20] In a similar study, C-dots derived from celery leaves have been reported to detect 2-NP to an LOD of 3.9×10$^{-8}$ M.[31] This shows that the results obtained in the present study are comparable to those reported in the literature. A possible mechanism for the fluorescence quenching of C-dots by the nitrophenol compounds has been proposed before based on energy transfer action.[50] An overlap between the maximum absorption band of 2-nitrophenolate ion that is centred at 369 nm and the emission band of the as-prepared C-dots results in a resonant energy transfer process (association quenching).[50] This mechanism is referred to as competitive absorption or the inner filter effect (IFE).[20] The contact quenching is believed to occur via the formation of a spirocyclic zwitterionic Meisenheimer complex.[50] In such a complex, the negative charge is thought to be delocalised over the nitro group and the cyclohexadienone ring whilst the positive charge is spread over the iminium group (Figure 8).

Interferences from other nitroaromatic and benzene derivatives were evaluated. The fluorescence quenching efficiency was highest for 2-NP compared to all other benzene derivatives studied (Figure 9(a) and (b)). A fluorescence quenching efficiency of 85% for 2-NP was determined

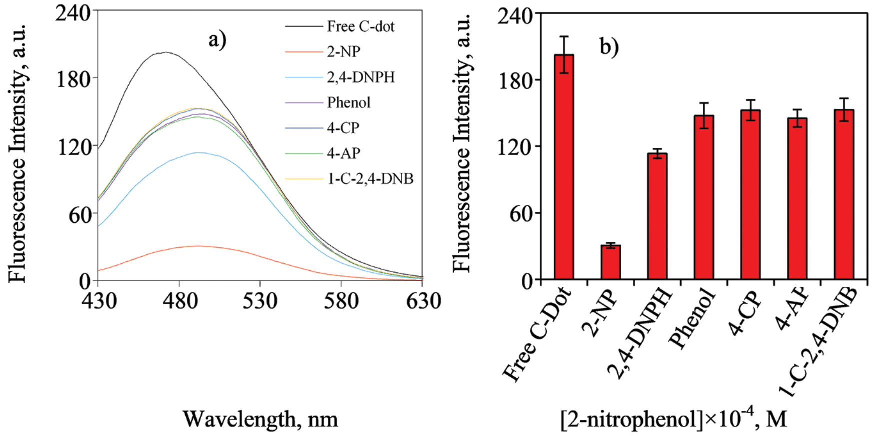



**Figure 9:** Selectivity studies of the as-prepared C-dots towards 2-NP. (a) Fluorescence spectra of the free C-dots and in the presence of 2-NP and various possible interferences. (b) A plot of fluorescence intensity of free C-dots after the addition of 2-NP and different possible interferences. 2-NP = 2-nitrophenol; 2,4-DNPH = 2,4-dinitrophenylhydrazine; 4-CP = 4-chlorophenol; 4-AP = 4-aminophenol and 1-C-2,4-DNB = 1-chloro-2,4-dinitrobenzene. $\lambda_{ex}$ = 400 nm, [Interferent] = $8\times10^{-4}$ M and [C-dot] = 0.05 g/ml.

whilst the other five benzene derivatives achieved only 24–44% fluorescence quenching efficiency. The low fluorescence quenching efficiencies of the other nitrobenzene and benzene derivatives were observed even after the addition of a large concentration of the possible interferant ($8\times10^{-4}$ M). Contrastingly, 2,4-dinitrophenylhydrazine bearing two nitro groups showed a moderate fluorescence quenching efficiency of about 44%. This could be attributed to the presence of two nitro groups that are electron withdrawing which possibly enhanced the competitive absorption of the nitroaromatic compound against that of the as-prepared C-dots. The as-prepared C-dots displayed excellent selectivity towards 2-NP under controlled conditions of pH 10, and this indicates that the C-dots prepared from biowaste can be used effectively for monitoring 2-NP in aqueous media. This study has shown that *C. vulgaris* peels are not only limited to the synthesis of C-dots and their application towards detection of metal ion and cellular bioimaging[41,51] or detection of toxic chemicals such as ethyl carbamate in alcoholic beverages[52] or photocatalytic degradation of methyl orange[53] but that they can also be used to prepare C-dots for the detection of toxic environmental pollutants such as 2-nitrophenol in aqueous media. In a previous study, non-environmentally friendly chemicals such as organic solvent, dimethylformamide (DMF), have been used for the synthesis of C-dots and application towards detection of 2-nitrophenol.[20] This work rather provided an environmentally friendly alternative for the synthesis of C-dots that are responsive towards the detection of 2-nitrophenol.

## Conclusions

In summary, we have synthesised hydroxyl and amide surface-bearing C-dots from *C. vulgaris* biomass as demonstrated by the characterisation studies. The as-prepared C-dots exhibited broad particle size distribution. After optimisation of the synthesis and application conditions, the as-prepared C-dots displayed excellent optical recognition of 2-NP in aqueous media. The effective fluorescence quenching of the C-dots by 2-NP is postulated to occur through energy transfer mechanism. The C-dots prepared from biomass have proven to have potential probes for 2-NP as revealed by their analytical benefits such as quick detection, excellent sensitivity, considerable selectivity and ease of use equipment. The photoluminescence features of the C-dots can be enhanced through surface passivation and functionalisation. This work provides a synthetic approach that is simple, economical, scalable and ecofriendly. Finally, the synthetic strategy of fluorescent C-dots obtained from biomass demonstrates a green approach for converting biowaste into valuable and beneficial optical chemical sensors for environmental pollutants monitoring. The novelty of this research work is the provision of an alternative, cost-effective and ecofriendly method for the potential detection of 2-nitrophenol.

## Acknowledgements

We thank the Botswana International University of Science and Technology (BIUST) for the laboratory facilities used for sample analysis.

## Competing interests

We have no competing interests to declare.

## Authors' contributions

P.D.: Conceptualisation; methodology; student supervision; manuscript writing; data analysis. G.N.P.: Conceptualisation; student supervision; methodology; manuscript writing and revisions. M. M. K.: Data collection; data analysis; writing – initial draft. M. A. B.: Data analysis; manuscript reviewing. M. T.: Materials characterisation; manuscript reviewing. B. P. Materials characterisation; data analysis. J. M.: Conceptualisation; student supervision; methodology; validation; manuscript revisions. R. K.: Materials characterisation, manuscript reviewing.

## References

1. Singh S. Sensors – An effective approach for the detection of explosives. J Hazard Mater. 2007;144:15–28. https://doi.org/10.1016/j.jhazmat.2007.02.018

2. Yinon J. Field detection and monitoring of explosives. Trends Anal Chem. 2002;21(2002):292–301. https://doi.org/10.1016/S0165-9936(02)00408-9

3. Kumar D, Elias AJ. The explosive chemistry of nitrogen, a fascinating journey from 9th century to the present. Resonance. 2019;24:1253–1271. https://doi.org/10.1007/s12045-019-0893-2

4. Das P, Mandal SK. Understanding the effect of an amino group on the selective and ultrafast detection of TNP in water using fluorescent organic probes. J Mater Chem C. 2018;6:3288–3297. https://doi.org/10.1039/C7TC05852G

5. Siddique AB, Pramanick AK, Chatterjee S, Ray M. Amorphous carbon dots and their remarkable ability to detect 2,4,6-trinitrophenol. Sci Rep. 2018;8:9770. https://doi.org/10.1038/s41598-018-28021-9

6. Han Y, Chen Y, Feng J, Na M, Liu J, Ma Y, et al. Investigation of nitrogen content effect in reducing agent to prepare wavelength controllable fluorescent silicon nanoparticles and its application in detection of 2-nitrophenol. Talanta. 2019;194:822–829. https://doi.org/10.1016/j.talanta.2018.11.008

7. Alam MK, Rahman MM, Ahmed MAA, Torati SR, Asiri AM, Kim D, et al. Ultra-sensitive 2-nitrophenol detection based on reduced graphene oxide/ZnO nanocomposites. J Electroanal Chem. 2017;788:66–73. https://doi.org/10.1016/j.jelechem.2017.02.004

8. Adeosun WA, Asiri AM, Marwani HM. Sensitive determination of 2-nitrophenol using electrochemically deposited polymethyl red film for healthcare and environmental safety. Synth Met. 2020;261:116321. https://doi.org/10.1016/j.synthmet.2020.116321

9. Keith L, Telliard W. ES&T Special report: priority pollutants: I-a perspective view. Environ Sci Technol. 1979;13:416–423. https://doi.org/10.1021/es60152a601

10. Ammar S, Oturan N, Oturan MA. Electrochemical oxidation of 2-nitrophenol in aqueous medium by electro-fenton technology. J Environ Eng Manage. 2007;17:89–96.

11. Scanlon JJ, Falquer PA, Robinson GW, O'Brien GE, Sturrock PE. High-performance liquid chromatography of nitrophenols with a swept-potential electrochemical detector. Anal Chim Acta. 1984;158:169–177. https://doi.org/10.1016/S0003-2670(00)84825-4

12. Liu J, Chen Y, Guo Y, Yang F, Cheng F. Electrochemical sensor for o-nitrophenol based on β-cyclodextrin functionalized graphene nanosheets. J Nanomater. 2013;2013:632809. https://doi.org/10.1155/2013/632809

13. Niazi A, Yazdanipour A. Spectrophotometric simultaneous determination of nitrophenol isomers by orthogonal signal correction and partial least squares. J Hazard Mater. 2007;146:421–427. https://doi.org/10.1016/j.jhazmat.2007.03.063

14. Chaudhary S, Kumar S, Mehta SK, Ahmad Umar A, Khan MA. Fabrication of water soluble and luminescent $Eu_2O_3$ nanoparticles for specific quantification of aromatic nitrophenols in aqueous media. Chem Phys Lett. 2019;736:136799. https://doi.org/10.1016/j.cplett.2019.136799

15. Uddin MT, Alam MM, Asiri AM, Rahman MM, Toupance T, Islam MA. Electrochemical detection of 2-nitrophenol using a heterostructure $ZnO/RuO_2$ nanoparticle modified glassy carbon electrode. RSC Adv. 2019;10:122–132. https://doi.org/10.1039/C9RA08669B

16. Deng P, Wang W, Liu X, Wang L, Yan Y. A hydrophobic polymer stabilized $CsPbBr_3$ sensor for environmental pollutant detection. New J Chem. 2021;45:930–938. https://doi.org/10.1039/D0NJ04498A

17. Liu GN, Xu RD, Zhao RY, Sun Y, Bo QB, Duan ZY, et al. Hybrid copper iodide cluster-based pellet sensor for highly selective optical detection of o-nitrophenol and tetracycline hydrochloride in aqueous solution. ACS Sustainable Chem Eng. 2019;7:18863–18873. https://doi.org/10.1021/acssuschemeng.9b03963

18. Xu X, Ray R, Gu Y, Ploehn HJ, Gearheart L, Raker K, et al. Electrophoretic analysis and purification of fluorescent single-walled carbon nanotube fragments. J Am Chem Soc. 2004;126:12736–12737. https://doi.org/10.1021/ja040082h

19. Dhenadhayalan N, Lin KC, Saleh TA. Recent advances in functionalized carbon dots toward the design of efficient materials for sensing and catalysis applications. Small. 2020;16:1905767. https://doi.org/10.1002/smll.201905767

20. Qu Y, Ren G, Yu L, Zhu B, Fang Chai F, Chen L. The carbon dots as colorimetric and fluorescent dual-readout probe for 2-nitrophenol and 4-nitrophenol detection. J Lumin. 2019;207:589–596. https://doi.org/10.1016/j.jlumin.2018.12.017

21. Wang Y, Hu A. Carbon quantum dots: Synthesis, properties and applications. J Mater Chem C. 2014;2:6921–6939. https://doi.org/10.1039/C4TC00988F

22. Wang Q, Huang X, Long Y, Wang X, Zhang H, Zhu R, et al. Hollow luminescent carbon dots for drug delivery. Carbon. 2013;59:192–199. https://doi.org/10.1016/j.carbon.2013.03.009

23. Han Y, Huang H, Zhang H, Liu Y, Han X, Liu R, et al. Carbon quantum dots with photo enhanced hydrogen-bond catalytic activity in aldol condensations. ACS Catal. 2014;4:781–787. https://doi.org/10.1021/cs401118x

24. Du Y, Guo SJ. Chemically doped fluorescent carbon and graphene quantum dots for bioimaging, sensor, catalytic and photoelectronic applications. Nanoscale. 2016;8:2532–2543. https://doi.org/10.1039/C5NR07579C

25. Verhagen A, Kelarakis A. Carbon dots for forensic applications: A critical review. Nanomaterials. 2020;10:1535. https://doi.org/10.3390/nano10081535

26. Roy P, Chen PC, Periasamy AP, Chen YN, Chang HT. Photoluminescent carbon nanodots: Synthesis, physicochemical properties and analytical applications. Mater Today. 2015;18:447–458. https://doi.org/10.1016/j.mattod.2015.04.005

27. Sun YP, Zhou B, Lin Y, Wang W, Fernando KAS, Pathak P, et al. Quantum-sized carbon dots for bright and colorful photoluminescence. J Am Chem Soc. 2006;128:7756–7757. https://doi.org/10.1021/ja062677d

28. Crista DMA, Esteves da Silva JCG, Pinto da Silva L. Evaluation of different bottom-up routes for the fabrication of carbon dots. Nanomaterials. 2020;10:1316. https://doi.org/10.3390/nano10071316

29. Sharma V, Tiwari P, Mobin SM. Sustainable carbon-dots: Recent advances in green carbon dots for sensing and bioimaging. J Mater Chem B. 2017;5:8904–8924. https://doi.org/10.1039/C7TB02484C

30. Choi Y, Choi Y, Kwon O-H, Kim B-S. Carbon dots: Bottom-up syntheses, properties and light-harvesting applications. Chem Asian J. 2018;13:586–598. https://doi.org/10.1002/asia.201701736

31. Qu Y, Yu L, Zhu B, Chai F, Su Z. Green synthesis of carbon dots by celery leaves for use as fluorescent paper sensors for the detection of nitrophenols. New J Chem. 2020;44:1500–1507. https://doi.org/10.1039/C9NJ05285B

32. Yang X, Wang D, Luo N, Feng M, Peng X, Liao X. Green synthesis of fluorescent N,S-carbon dots from bamboo leaf and the interaction with nitrophenol compounds. Spectrochim Acta A Mol Biomol Spectrosc. 2020;239:118462. https://doi.org/10.1016/j.saa.2020.118462

33. Melhuish WH. Quantum efficiencies of fluorescence of organic substances: Effect of solvent and concentration of the fluorescent solute. J Phys Chem. 1961;65(2):229–235. https://doi.org/10.1021/j100820a009

34. Bhatt M, Bhatt S, Vyas G, Raval IH, Haldar S, Paul P. Water-dispersible fluorescent carbon dots as bioimaging agents and probes for Hg2+ and Cu2+ ions. ACS Appl Nano Mater. 2020;3:7096–7104. https://doi.org/10.1021/acsanm.0c01426

35. Ansi VA, Renuka NK. Table sugar derived carbon dot - a naked eye sensor for toxic Pb2+ ions. Sens Actuators B. 2018;264:67–75. https://doi.org/10.1016/j.snb.2018.02.167

36. Prasannan A, Imae T. One-pot synthesis of fluorescent carbon dots from orange waste peels. Ind Eng Chem Res. 2013;52:15673–15678. https://doi.org/10.1021/ie402421s

37. Wang BB, Jin JC, Xu ZQ, Jiang ZW, Li X, Jiang FL, et al. Single-step synthesis of highly photoluminescent carbon dots for rapid detection of Hg with excellent sensitivity. J Colloid Interface Sci. 2019;551:101–110. https://doi.org/10.1016/j.jcis.2019.04.088

38. Dang DK, Sundaram C, Ngo Y-LT, Choi WM, Chung JS, Kim EJ, et al. Pyromellitic acid-derived highly fluorescent N-doped carbon dots for the sensitive and selective determination of 4-nitrophenol. Dyes Pigm. 2019;165:327–334. https://doi.org/10.1016/j.dyepig.2019.02.029

39. Das P, Ganguly S, Maity PP, Srivastava HK, Bose M, Dhara S, et al. Converting waste *Allium sativum* peel to nitrogen and sulphur co-doped photoluminescence carbon dots for solar conversion, cell labeling, and photobleaching diligences: A path from discarded waste to value-added products. J Photochem Photobiol B Biol. 2019;197:111545. https://doi.org/10.1016/j.jphotobiol.2019.111545

40. Tan XW, Romainor ANB, Chin SF, Ng SM. Carbon dots production via pyrolysis of sago waste as potential probe for metal ions sensing. J Anal Appl Pyrolysis. 2014;105:157–165. https://doi.org/10.1016/j.jaap.2013.11.001

41. Zhou J, Sheng Z, Han H, Zou M, Li C. Facile synthesis of fluorescent carbon dots using watermelon peel as a carbon source. Mater Lett. 2012;66:222–224. https://doi.org/10.1016/j.matlet.2011.08.081

42. Zhi-hao Gao Z-H, Lin Z-Z, Chen X-M, Lai Z-Z, Huang Z-Y. Carbon dots-based fluorescent probe for trace Hg2+ detection in water sample. Sens Actuators B. 2016;222:965–971. https://doi.org/10.1016/j.snb.2015.09.032

43. Ye Q, Yan F, Shi D, Zheng T, Wang Y, Zhou X, et al. N,B-doped carbon dots as a sensitive fluorescence probe for $Hg^{2+}$ ions and 2,4,6-trinitrophenol detection for bioimaging. J Photochem Photobiol B Biol. 2016;162:1–13. https://doi.org/10.1016/j.jphotobiol.2016.06.021

44. Xu H, Yang X, Li G, Zhao C, Liao X. Green synthesis of fluorescent carbon dots for selective detection of tartrazine in food samples. J Agric Food Chem. 2015;63:6707–6714. https://doi.org/10.1021/acs.jafc.5b02319

45. Dinake P, Phokedi GN, Janes Mokgadi J, Ntshekisang A, Botlhomilwe MA, Kelebemang R, et al. A facile microwave-assisted green synthetic approach of solid-state fluorescent carbon-dot nanopowders derived from biowaste for potential latent-fingerprint enhancement. Int J Nanosci. 2021;20:2150051. https://doi.org/10.1142/S0219581X21500514

46. Zhuo K, Sun D, Xu P, Wang C, Cao Y, Chen Y, et al. Green synthesis of sulfur- and nitrogen-co-doped carbon dots using ionic liquid as a precursor and their application in $Hg^{2+}$ detection. J Lumin. 2017;187:227–234. https://doi.org/10.1016/j.jlumin.2017.03.022

47. Yang Z, Xu M, Liu Y, He F, Gao F, Su Y, et al. Nitrogen-doped, carbon-rich, highly photoluminescent carbon dots from ammonium citrate. Nanoscale. 2014;6:1890–1895. https://doi.org/10.1039/C3NR05380F

48. Sun X, He J, Meng Y, Zhang L, Zhang S, Ma X, et al. Microwave-assisted ultrafast and facile synthesis of fluorescent carbon nanoparticles from a single precursor: Preparation, characterization and their application for the highly selective detection of explosive picric acid. J Mater Chem. 2016;4:4161–4171. https://doi.org/10.1039/C5TA10027E

49. Tammina SK, Yang Y. Highly sensitive and selective detection of 4-nitrophenol, and on-off-on fluorescence sensor for Cr (VI) and ascorbic acid detection by glucosamine derived n-doped carbon dots. J Photochem Photobiol A: Chem. 2020;387:112134. https://doi.org/10.1016/j.jphotochem.2019.112134

50. Ahmed GHG, Rosana Badía Laino RB, Calzon JAG, Garcia MED. Highly fluorescent carbon dots as nanoprobes for sensitive and selective determination of 4-nitrophenol in surface waters. Microchim Acta. 2015;182:51–59. https://doi.org/10.1007/s00604-014-1302-x

51. Rodwihok C, Tam TV, Choi WM, Suwannakaew M, Woo SW, Wongratanaphisan D, et al. Preparation and characterization of photoluminescent graphene quantum dots from watermelon rind waste for the detection of ferric ions and cellular bio-imaging applications. Nanomaterials. 2022;12:702. https://doi.org/10.3390/nano12040702

52. Han L, Zhu P, Liu H, Sun B. Molecularly imprinted bulk and solgel optosensing based on biomass carbon dots derived from watermelon peel for detection of ethyl carbamate in alcoholic beverages. Microchim Acta. 2022;189:286. https://doi.org/10.1007/s00604-022-05388-1

53. Ramzilah UR, Aziz A. Removal of methyl orange (MO) using carbon quantum dots (CQDs) derived from watermelon rinds. Int J Eng Technol Sci. 2019;6:91–99. https://doi.org/10.15282/ijets.v6i1.2226

Check for updates

**AUTHORS:**
Glynn K. Pindihama[1]
Mugera W. Gitari[1,2]
Rabelani Mudzielwana[1]
Ntakadzeni E. Madala[3]

**AFFILIATIONS:**
[1]Environmental Remediation and Nano Sciences Research Group, Department of Geography and Environmental Sciences, Faculty of Science, Engineering and Agriculture, University of Venda, Thohoyandou, South Africa
[2]Department of Chemical Sciences and Technology, School of Chemistry and Material Sciences, Technical University of Kenya, Nairobi, Kenya
[3]Department of Biochemistry and Microbiology, Faculty of Science, Engineering and Agriculture, University of Venda, Thohoyandou, South Africa

**CORRESPONDENCE TO:**
Glynn Pindihama

**EMAIL:**
gpindihama@gmail.com

**DATES:**
**Received:** 15 Sep. 2022
**Revised:** 16 Mar. 2023
**Accepted:** 30 Apr. 2023
**Published:** 28 Sep. 2023

**HOW TO CITE:**
Pindihama GK, Gitari MW, Mudzielwana R, Madala NE. Development of a chitosan-multi-walled carbon nanotubes composite for application in solid-phase adsorption toxin tracking of microcystins. S Afr J Sci. 2023;119(9/10), Art. #14786. https://doi.org/10.17159/sajs.2023/14786

**ARTICLE INCLUDES:**
☒ Peer review
☐ Supplementary material

**DATA AVAILABILITY:**
☒ Open data set
☐ All data included
☐ On request from author(s)
☐ Not available
☐ Not applicable

**EDITORS:**
Priscilla Baker
Amanda-Lee Manicum

**KEYWORDS:**
chitosan, multi-walled carbon nanotubes, solid-phase adsorption toxin tracking, microcystins

**FUNDING:**
Water Research Commission (K5/2972)



# Development of a chitosan-multi-walled carbon nanotubes composite for application in solid-phase adsorption toxin tracking of microcystins

Contamination of water and food with cyanotoxins poses human health risks, and hence the need for sensitive early warning tools to monitor these in water. A composite of glutaraldehyde-crosslinked chitosan and multi-walled carbon nanotubes (ChMWCNTs) was synthesised and tested for potential use as a solid-phase adsorption toxin tracking (SPATT) adsorbent for monitoring microcystins (MCs) in fresh water. The composite was characterised by Fourier transform infrared spectroscopy, Brunauer–Emmett–Teller theory and scanning electron microscopy. Batch adsorption experiments to assess the effect of contact time, adsorbent dosage and initial microcystin-LR (MC-LR) concentration were conducted. The composite was found to be efficient in adsorbing MC-LR, showing 97% removal and a maximum adsorption capacity of 4.639 μg/g under optimised conditions of 5 μg/L of MC-LR, adsorbent dose of 0.03 g/5 mL and 30 min contact time. The adsorption kinetics were better explained by a pseudo-second-order model, inferring chemisorption adsorption. The isotherm data better fitted the Langmuir isotherm model, thus inferring monolayer surface adsorption. For desorption, 100% methanol was the most effective, with an efficiency of 84.71%. The composite effectively adsorbed and desorbed three congeners of MCs (–LR, –RR and –YR) when tested in raw dam water, regardless of its lower maximum adsorption capacity compared to those of other adsorbents used for similar purposes.

**Significance:**

- Monitoring of microcystins is problematic in large reservoirs and rivers.
- Chitosan can be crosslinked and modified to enhance its adsorption properties.
- Composites of chitosan and carbon nanotubes efficiently adsorb and desorb microcystins.
- This study is possibly the first to apply a chitosan-based sorbent in solid-phase toxin tracking (SPATT) to be used as an early warning tool in passive monitoring of microcystins in water resources.

## Introduction

The global increase in incidents of toxin-producing cyanobacterial blooms has gained international attention in recent years. These increases have been attributed to a wide variety of environmental factors including nutrient pollution, increased temperature and salinity, many of which will likely be exacerbated by climate change.[1,2] Toxins produced by cyanobacteria (cyanotoxins) pose a significant risk for humans, livestock and wildlife as they can cause impairment and mortality.[3] Amongst the many types of cyanotoxins that have been documented, microcystins (MCs) are the most frequently occurring in the freshwater environment, and hence, have been widely investigated. Due to the chronic toxicity of MCs, the World Health Organization (WHO) set a provisional threshold of 1 μg/L for MC-LR concentration in drinking water[4] and a tolerable daily intake (TDI) of 0.04 μg MC-LR/kg body weight (BW) in food.[1] MC levels above this WHO threshold have been reported in South Africa, with levels in the range of 2 μg/L to 23.7 μg/L being reported in various water bodies in the country.[5]

Several countries have routine programmes to monitor these toxins and test possible contamination of food crops and fish. However, the drawback of the sampling and monitoring of these toxins in rivers and large water bodies is that their levels can vary rapidly.[6] To overcome the drawbacks of grab sampling, MacKenzie et al.[7] came up with the solid-phase adsorption toxin tracking (SPATT) technology, for possible use for the detection and early warning of the presence of cyanotoxins. SPATT involves suspending small bags containing adsorbent which accumulate toxins in the water body. The toxins can then be extracted and measured, providing information on extracellular toxins over an extended period.

Different sorbents have been used for SPATT, including the polymeric adsorbents Diaion® HP-20 to SEPABEADS-type resins for the accumulation of cyanobacterial toxins of different polarities.[8] Although passive sampling has been successfully used several times to monitor cyanotoxins using different bulk polymeric sorbents,[8] most of these sorbents are synthetic and relatively costly to buy. Studies on the characterisation and mechanism of SPATT resins for the adsorption of lipophilic and hydrophilic cyanotoxins are also limited.[1]

Recent studies, for example those of Gomez-Maldonado et al.[9] and Tran et al.[10], have demonstrated that chitosan can be modified and be applied for the adsorption of MC-LR in water purification, but no studies have evaluated its use as a sorbent in the passive sampling of MCs in SPATT. Insertion of multi-walled carbon nanotubes (MWCNTs) into the chitosan hydrogel to form a chitosan-multi-walled carbon nanotube (ChMWCNT) composite was hypothesised to have the synergistic effect of improving the physical properties of chitosan and improve its adsorptive characteristics. Thus the aim of this work was to insert multi-walled carbon nanotubes into the chitosan

12 RESPONSIBLE CONSUMPTION AND PRODUCTION

hydrogel structure and evaluate the subsequent effects on the formed composite material and its capabilities for MC adsorption and desorption for possible use in SPATT samplers.

## Materials and methods

### *Chemicals and reagents*

Analytical-reagent grade glutaraldehyde (GLA), acetic acid and powdered high molecular weight chitosan (Ch) were purchased from Rochelle Chemicals (Johannesburg, South Africa). Hydroxyl functionalised multi-walled carbon nanotubes (MWCNTs, purity: >98%, average diameter: 10–20 nm) were purchased from SabiNano (Pty) Ltd (Johannesburg, South Africa). Synthetic aromatic adsorbent Diaion® HP-20 resin was purchased from Rochelle Chemicals (Johannesburg, South Africa). De-ionised water from a Milli-Q water purification system (Merck-Millipore Ltd., Germany) of 18.2 MΩ/cm quality was used to prepare all the solutions. Microcystin-LR (MC-LR) standards, 0.5 mg films (Eurofins Scientific, USA) were purchased from ToxSolutions, Kits and Services (South Africa).

### *Preparation of the ChMWCNT composite*

Chitosan was crosslinked with GLA using ratios optimised by Gonçalves et al.[11] and chitosan-to-MWCNTs ratios applied by Alves et al.[12] In brief, chitosan (1 g) was dissolved in 50 mL of 1% v/v acetic acid. After the complete dissolution of chitosan, carbon nanotubes (CNT) (10 % wt) were added to the solution. Then, GLA (2% (v/v)) was used as a crosslinking agent and slowly added to the MWCNT-chitosan solution under mechanical stirring (50 revolutions per minute [rpm]) until it formed a gel. The formed hydrogel was freeze dried for 48 h at −48 °C under a constant vacuum of 44 μmHg (Telstar Lyoquest Freeze Dryer, Terrassa, Spain). The freeze-dried material was then ground to powder using a mortar and pestle, followed by sieving using a 250 μm sieve to get particles above 250 μm only (suitable to be used in SPATT bags).

### *Characterisation*

The Fourier transform infrared (FTIR) spectroscopy spectra of the chitosan (CH), glutaraldehyde-crosslinked chitosan (ChGLA) and the chitosan/MWCNT composite (ChMWCNT) were recorded using a Bruker Alpha-P FTIR spectrometer equipped with a diamond ATR window (Bruker Optik GmbH, Ettlingen, Germany). All spectra were recorded in the spectral range of 4000–400/cm with a resolution of 4/cm. Surface area measurements were conducted on a Micromeritics TriStar II 3020 Surface Area and Porosity Analyser (Norcross, Georgia, USA). The surface morphology of the samples was characterised using scanning electron microscopy (SEM; FEI Nova Nano SEM 230, USA).

### *Adsorption and desorption studies*

#### Optimisation of equilibration time

A 5 μg/L MC-LR solution was used to investigate the effect of contact time. The toxin solutions (5 mL) were placed in 15 mL glass centrifuge tubes with 0.01 g of the ChMWCNT composite and the mixture was shaken at 145 rpm (Stuart Reciprocating Shaker, SSL2, UK) for 5, 10, 15, 30 and 60 min at room temperature and immediately after, filtered through 0.22 μm pore membrane filters and the filtrate analysed for MCs using microcystins/nodularins (ADDA) enzyme-linked immunosorbent assay (ELISA) kits, EUROFINS (Kit Lot No: 19I1120: PN 520011, Eurofins Scientific, USA) and a SPECTROstar Nano Plate Reader (BMG LABTECH, 601-1106, Germany) for quantification.

#### Optimisation of dosage

A 5 μg/L of MC-LR (5 mL) and composite dosages 0.005 g; 0.01 g; 0.02 g; 0.03 g; 0.05 g and 0 g (control) were investigated by placing the composite in 15 mL glass centrifuge tubes and shaken at 145 rpm at room temperature for 30 min (as determined from the previous experiment on optimisation of equilibration time). Samples were then filtered through 0.22 μm pore membrane filters immediately after agitation and the residual toxin levels in the supernatant determined as described in the section of optimisation of equilibration time.

#### Adsorption and equilibrium studies

Adsorption experiments were conducted to investigate the adsorption isotherms at equilibrium and the kinetics. The effect of contact time and the kinetics studies were investigated by placing 0.03 g of ChMWCNT (the optimised unit dose) in 15 mL glass centrifuge tubes with 5 mL of 5 μg/L MC-LR solution. This was then followed by agitation for 5, 10, 15, 30 and 60 min; then it was filtered through 0.22-μm pore membrane filters, and the residual toxin levels in the supernatant were determined as described in the section of optimisation of equilibration time.

For the equilibrium study, 0.03 g of adsorbent was introduced into seven 15-mL centrifuge tubes with 5 mL of different MC-LR concentrations (within the range commonly found in South African water bodies[5]) (i.e. 2.5, 5, 7.5, 10, 15, 20, and 25 μg/L). The tubes were agitated for 30 min at room temperature. Afterwards, samples were filtered through 0.22 μm pore membrane filters immediately after agitation and the residual toxin levels in the supernatant were determined. Two isotherm models of Freundlich and Langmuir were employed to analyse the equilibrium data.

#### Adsorption data modelling

The adsorption capacity ($q_t$) and equilibrium adsorption capacity ($q_e$, $\mu$g/g) were calculated using Equation 1 and Equation 2, respectively:[13]

$$q_t = V\ \frac{C_o - C_t}{m} \qquad \text{Equation 1}$$

$$q_e = V\ \frac{C_o - C_e}{m} \qquad \text{Equation 2}$$

where $V$ is the solution volume (L); $m$ is the weight of the adsorbent (g); $C_o$, $C_t$ and $C_e$ (μg/L) are the concentrations of the adsorbate at the initial, certain and equilibrium times.

To better understand the factors and mechanisms controlling the adsorption of MC-LR onto ChMWCNT, the experimental data obtained at various contact times were fitted into kinetics models namely, pseudo first order (PFO) and pseudo second order (PSO) and the Weber–Morris intra-particle diffusion models. The mathematical expressions (nonlinear fits) of the PFO and PSO kinetic models are shown in Equations 3 and 4, respectively.[14]

$$q_t = q_e(1 - e^{-k_1 t}) \qquad \text{Equation 3}$$

$$q_t = \frac{(k_2 q_e^2 t)}{(1 + k_2 q_e t)} \qquad \text{Equation 4}$$

where $q_t$ (mg/g) and $q_e$ (mg/g) are the amounts of MC-LR adsorbed at time $t$ (h) and equilibrium, respectively, and $k_1$ (/h) and $k_2$ (g/mg.h) are the PFO and PSO adsorption rate constants, respectively.

Equation 5 shows the Weber–Morris intra-particle diffusion model:

$$q_t = k_{id} t^{1/2} + C \qquad \text{Equation 5}$$

where $k_{id}$ is the intra-particle diffusion coefficient (mg/g. min$^{1/2}$), and $C$ (g/g) a constant associated with diffusion resistance

Adsorption isotherm models, namely, Langmuir and Freundlich adsorption isotherm models, were used to give more insight into the relationship between the adsorbate concentration and the solid phases in the aqueous solution at equilibrium.[15] The nonlinear mathematical expressions of the Langmuir and Freundlich isotherm models are shown in Equation 6 and Equation 7, respectively.[14]

$$q_e = \frac{Q_{max} K_L C_e}{1 + K_L C_e}$$ Equation 6

$$q_e = K_F C_e^{\frac{1}{n}}$$ Equation 7

where $Q_{max}$ (mg/g) and $K_L$ (L/mg) are the maximum adsorption capacity of the adsorbent and the Langmuir constant related to the adsorption energy of the adsorbent, respectively. $K_F$ (mg/g (mg/L)$^{-1/n}$ and $n$ (–) are the Freundlich constant and adsorption strength, respectively.

For the Langmuir isotherm, the adsorption process efficiency was determined using a dimensionless constant separation factor $R_L$, calculated using the calculated $K_L$ from Equation 6 as follows (Equation 8):

$$R_L = \frac{1}{1 + K_L * C_e}$$ Equation 8

If the $R_L$ values are found to be between 0 and 1, then the process of adsorption is deemed favourable at the temperatures studied.[15]

### Desorption optimisation

Supernatants were discarded after adsorption equilibration was achieved (based on optimum equilibration time and dosage in sections on optimisation of equilibration time; optimisation of dosage and adsorption and equilibrium studies). The residual adsorbent was extracted with varying percentages of methanol (0%, 20%, 50%, 80% and 100%) through sonication (SCIENTEC Ultrasonic Cleaner, Model 705, South Africa) for 5 min to release toxins before centrifugation at 570 rcf for 10 min. The whole cycle of sonication and centrifuging was repeated three times and the supernatant pooled to give approximately a 15-mL extract of each sample. The pooled samples were then evaporated to dryness under a gentle stream of nitrogen ($N_2$) gas, with the residue suspended in 1 mL of Milli-Q water followed by analysis using the ELISA method.

### Adsorption and desorption in SPATT bag format

#### SPATT construction and activation

SPATT bags were constructed from nylon mesh with approximately 95–100 μm pore size. The nylon mesh was sewn on three sides using a sewing machine and the bag was filled with 0.2 g of the ChMWCNT composite, then the fourth side was sewn to form a finished SPATT bag of approximately 60 × 60 mm dimension. The SPATT bags were activated by soaking in 100% methanol (MeOH) for 48 h, then rinsed thoroughly in Milli-Q water for removal of any MeOH residues and kept in Milli-Q water at 4–6 °C before use.

#### Adsorption experiment with field water

Raw dam water collected from Roodeplaat Dam was used for the SPATT adsorption experiment with field water. Before the experiment, the raw dam water was analysed for pH, total dissolved solids (TDS), electrical conductivity (EC) and turbidity. TDS, pH and EC were determined using a Jenway pH/Cond meter model (430), and turbidity was determined using a TB200 portable turbidity meter model (#TB200-10). The raw dam water (1 L) was placed in 1 L amber bottles, and the SPATT bags packed with 0.2 g of ChMWCNT were introduced to each bottle and the solutions were shaken at 145 rpm for 30 min at room temperature. The SPATT bags were removed after 30 min, rinsed in Milli-Q water and kept at 4–6 °C before extraction.

#### SPATT extraction

The SPATT bags were oven dried at 40 °C, then cut open and the material was placed in 15 mL centrifuge tubes. The toxins were extracted using 100% MeOH through sonication (5 min) to release toxins and centrifuged at 1750 rpm as described under 'Desorption optimisation'. The pooled supernatant was collected and dried at 50 °C under a stream of nitrogen. The dried material was then reconstituted with 1 mL of Milli-Q water and analysed using the ELISA method (ELISA test kits), EUROFINS (Kit Lot No: 19I1120: PN 520011) and liquid chromatography–mass spectrometry (LC–MS).

#### Determination of MC levels in the residues and extracted material

EUROFINS microcystin/nodularin ELISA plate kit were used for the analysis of total MCs in the samples following the manufacturers' instructions. The method is based on a direct competitive ELISA for quantitative detection of MCs and nodularins on the polyclonal antibodies. To determine the levels of some of the congeners of MCs (LR, RR and YR) adsorbed by the composites, a chromatographic method using a Shimadzu triple quadrupole LC–MS/MS system (model LCMS-8045, Shimadzu Corporation, Japan) was used.

#### Chromatographic conditions

Levels of microcystins (LR, RR and YR) in the samples were determined on a triple quadrupole LC–MS/MS system (model LCMS-8045, Shimadzu Corporation, Japan) with a Shim-pack Velox SP-C18, 2.7 μm, with dimensions 2.1 × 100 mm (Shimadzu, Japan). The injection volume was set at 10 μL and the mobile phases used were 0.1% formic acid (FA) in water (A) and 0.1% FA in acetonitrile (B). A flow rate of 0.5 mL/min and a 5-min binary gradient were used with an elution profile of 2% B (0.4 min), linear gradient to 70% B (3.1 min), 100% B (0.5 min), and, finally, 2% B (1 min). Detection limits were 0.5 μg/L for all the three MC congeners used.

The final concentration of the toxins in the samples was determined by using Equation 9:

$$\text{Conc in sample}\,(\mu\text{g/L}) = \left(\frac{C_o \times \text{Vol of extract used (L)}}{\text{Volume of sample used (L)}}\right)$$ Equation 9

where $C_o$ = the concentration of the sample determined from the calibration curve (μg/L)

To determine the microcystin concentrations adsorbed to the SPATT composite/resin, Equation 10 was used:

$$\mu\text{g/g (resin)} = \frac{(\mu\text{g/L in extract} \times 0.001\,\text{L extract vol})}{0.2\,\text{g (composite)}}$$ Equation 10

where the extract volume is 0.001 L and is 0.2 g of ChMWCNT.

To compare the levels of MCs adsorbed and desorbed by different adsorbents and solutions, analysis of variance (ANOVA) and/or the Kruskal–Wallis tests were used at 95% confidence interval (CI) using GraphPad InStat 3 (GraphPad Software, California, USA).

## Results and discussion

### *Physicochemical characterisation*

#### Fourier transform infrared analysis

The FTIR spectra for raw chitosan, ChGLA and ChMWCNT composite are shown in Figure 1. The notable peaks of chitosan can be attributed to: the broad and strong band ranging from 3200 to 3600/cm corresponds to the presence of –OH and $–NH_2$ groups, 2927/cm ($CH_3$ symmetric stretch), 1634/cm (C = O stretching vibration), 1154/cm (C–O–C bending vibration), and 1066/cm (C–OH stretching vibration) as also observed by Li et al.[16] Major differences in the spectra of chitosan compared to ChGLA and ChMWCNT were the presence of C–O–C band at 1154/cm in chitosan but not in the two crosslinked products and the presence of the amide II at 1554/cm in the crosslinked material (ChGLA and ChMWCNT) but not in the spectra of chitosan (Figure 1). Of importance to note were the broader peaks of the ChMWCNT compared to the other two materials, which probably indicated the coating of the chitosan by the multi-walled carbon nanotubes as was also observed by Liu et al.[17]

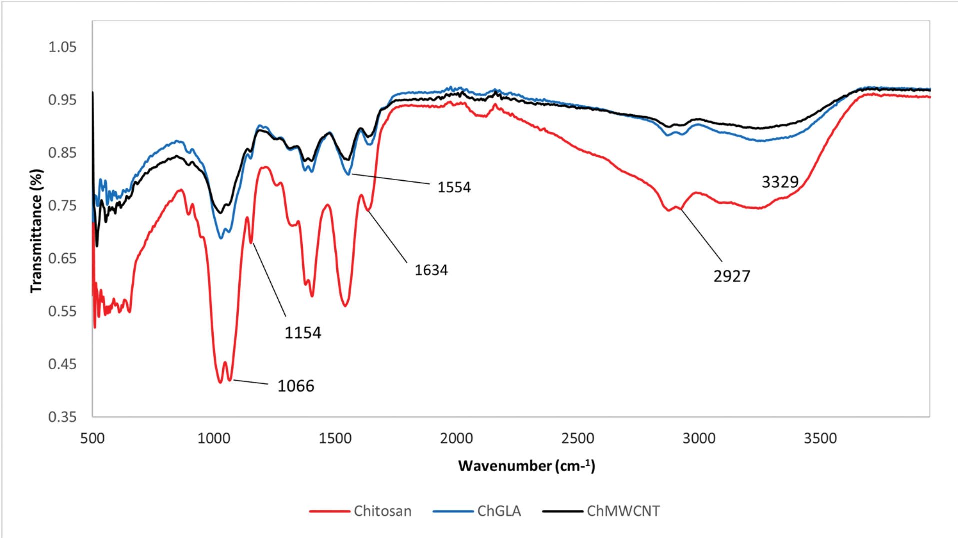



**Figure 1:** FTIR spectra for chitosan, ChGLA and ChMWCNT.

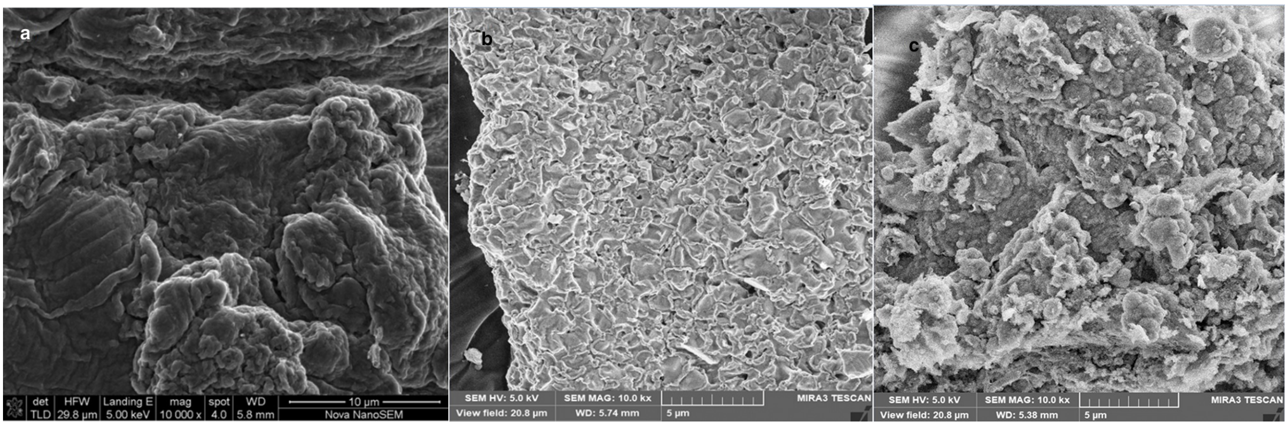



**Figure 2:** SEM images for (a) raw chitosan, (b) ChMWCNT before adsorption and (c) ChMWCNT after adsorption x10 000 magnification.

## Morphological analysis

Figure 2a shows the surface morphology of raw chitosan at x10 000 magnification. The SEM images show a slightly rough and dense surface with large clusters of particles for raw chitosan. After the crosslinking and introduction of multi-walled CNTs (Figure 2b), the surface structure appears more porous and flaky. The introduction of multi-walled CNTs seems to have enhanced pore formation. A clear distinction of the ChMWCNT morphology before and after adsorbing MCs can be seen in Figure 2b vs 2c with the structure becoming deformed and irregular upon adsorption. Figure 2c shows the formation of granules in the used material, which could be attributed to the swelling of the material during adsorption with the pores appearing saturated after use.

## Surface area, pore volume and pore size

Specific surface area and pore volume are very important aspects for any material to be used for the adsorption of MCs, as MCs are large molecules and cannot easily enter into the micropores of materials with low micropore sizes.[1] Table 1 shows the surface area and pore size of the raw chitosan, synthesised ChGLA, ChMWCNT and the Diaion® HP-20 resin, which were evaluated using the Brunauer–Emmett–Teller (BET) method.

The results indicate that the Diaion® HP-20 resin had surface area, pore volume and pore size comparable to what was reported by Li et al.[18] In terms of the surface area, pore volume and pore size of the synthesised materials, raw chitosan had the lowest, followed by ChGLA, ChMWCNT. The Diaion® HP-20 resin displayed far more superior surface area and pore volume compared to any of the synthesised material, but the ChMWCNT had much greater pore sizes compared to Diaion® HP-20.

Introduction of multi-walled CNTs onto the ChGLA, improved the surface area, pore volume and pore size of the chitosan, and hence ChMWCNT seemed more suitable for the adsorption of MCs compared to ChGLA. Regardless of the lower surface area of the ChMWCNT compared to Diaion® HP-20, its higher pore sizes make it ideal for the adsorption of MCs. Li et al.[18] reiterate the importance of materials' pore size instead of surface area in determining the equilibration rates and abilities to adsorb the toxins.

## Adsorption and desorption characteristics of the composite for microcystins

Results of batch adsorption experiments showing the effects of the effect of contact time, initial adsorbate (MC-LR) concentration and adsorbent (ChMWCNT) dosage are presented in Figure 3a, b and c, respectively.

### Effect of contact time

The effect of contact time was evaluated to establish the equilibrium time for optimum adsorption of MC-LR onto ChMWCNT and establish the adsorption kinetics. The results in Figure 3a show that MC-LR was rapidly adsorbed in the first 15 min of the reaction and then slowed down as the reaction time increased to 30 min where optimum uptake of MC-LR was recorded. Thereafter, removal efficiency remained almost constant as the time increased to 60 min indicating that equilibrium had been reached.

The rapid adsorption of MC-LR in the initial stages (5–15 min) of contact could be attributed to binding of adsorbate onto the readily available active sites on the outer surface of the ChMWCNT. In the early stages, there are many

**Table 1:** Surface area, pore volume and pore size of HP20, ChGLA*, ChMWCNT and Chitosan

| | HP20 | ChGLA | ChMWCNT | Chitosan |
|---|---|---|---|---|
| **Surface area (m²/g)** | | | | |
| Single point surface area | 618.0379 | 1.7616 | 8.0174 | 0.264 |
| BET surface area | 653.6041 | 1.8978 | 8.377 | 0.3241 |
| BJH adsorption cumulative surface area of pores | 481.58 | 0.506 | 8.424 | 0.158 |
| BJH desorption cumulative surface area of pores | 475.1997 | 0.49 | 10.5357 | 0.9885 |
| **Pore volume (cm³/g)** | | | | |
| Single point adsorption total pore volume of pores | 0.74539 | – | 0.02248 | – |
| BJH adsorption cumulative volume of pores | 0.999586 | 0.00022 | 0.029747 | 0.001056 |
| BJH desorption cumulative volume of pores | 0.983416 | – | 0.029914 | 0.001138 |
| Pore size (Å) | | | | |
| Adsorption average pore width (4V/A by BET) | 45.6172 | – | 107.3432 | – |
| BJH adsorption average pore diameter (4V/A) | 83.026 | 17.397 | 141.254 | 267.817 |
| BJH desorption average pore diameter (4V/A) | 82.779 | – | 113.57 | 46.043 |

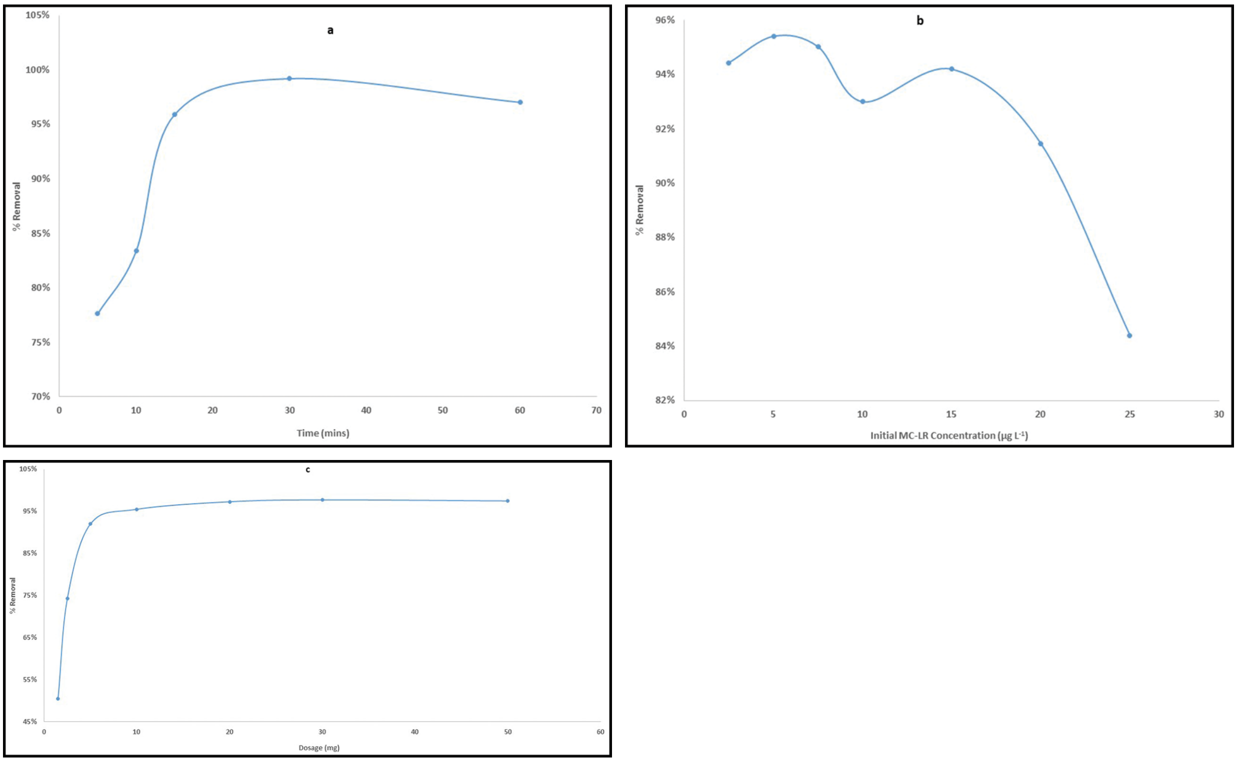



**Figure 3:** (a) Effect of contact time on the adsorption of MC-LR onto ChMWCNT. (b) Adsorption of MC-LR onto ChMWCNT as a function of initial concentration (5 µg/L initial concentration, 10 mg/5 mL adsorbent dosage. (c) Adsorption of MC-LR onto ChMWCNT as a function of adsorbent (pH 6.5 ± 0.5 and shaking speed of 145 rpm).

sorption sites available for occupation by the adsorbate, thus higher initial rates; then as the reaction continues, the sorption sites and concentration of the adsorbate decreases and the rate of adsorption also lowers.[14]

Based on that, 30 min was adopted as the reaction time for subsequent experiments. Our findings were similar to those of Zhao et al.[1] who found the equilibration times of MCs onto the Diaion® HP-20 resin to be 30 min, but that of SP700 to be 15 min. This implies that the synthesised composite adsorbs MCs at comparable rate with other adsorbents being used for the same purpose.

To further explain the adsorption mechanisms and the adsorption rate controlling factors, the experimental data were fit into reaction kinetics models namely PFO, PSO and intra-particular diffusion.[19] The nonlinear plot for PFO and PSO are presented in Figure 4a whilst the constant parameters are presented in Table 2. The PSO models showed a higher correlation coefficient ($R^2 = 0.875$); this means that the experimental data are better explained by the PSO model, thus showing that the adsorption was occurring through chemisorption. In a related study locally, Mashile et al.[20] also found their adsorption data to be fitting the fitted the

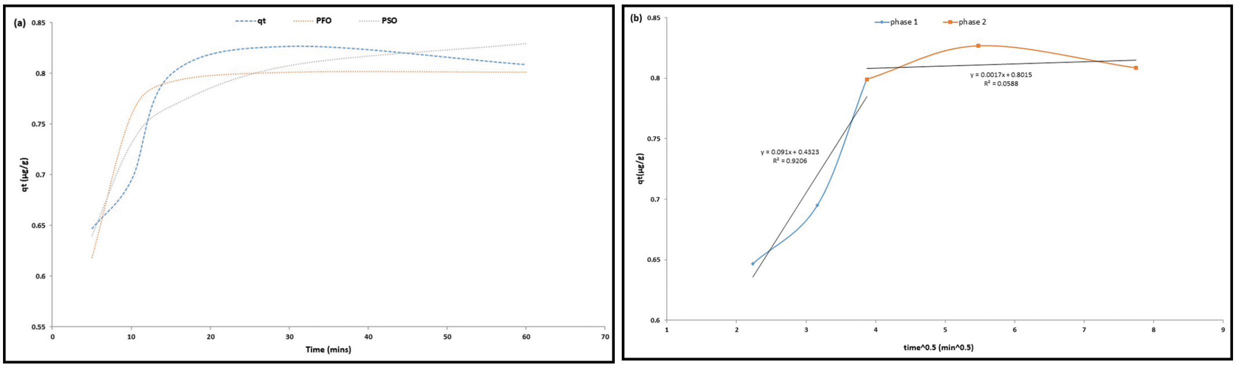



**Figure 4:** (a) Results of kinetics experiment on ChMWCNT and nonlinear regression analysis of PFO and PSO models. (b) Intra-particle diffusion of MC-LR onto ChMWCNT (initial MC-LR concentration: 5 µg/L, adsorbent dose: 30 mg/5 mL, time: 5, 10, 15, 30 and 60 min).

**Table 2:** Parameters of pseudo-first-order and second-order reaction kinetics

| | Concentration (µg/L) | Pseudo-first-order model | | | Pseudo-second-order model | | |
|---|---|---|---|---|---|---|---|
| | | $q_e$ (µg/g) | $K$ (L/mg) | $R^2$ | $q_e$ (µg/g) | $K$ (L/mg) | $R^2$ |
| **ChMWCNT** | 5 | 0.801 | 0.295 | 0.773 | 0.852 | 0.710 | 0.875 |

pseudo-second-order model when applying tyre-based activated carbon for the removal of MC-LR via adsorption. The adsorption mechanism of MC-LR onto the composite can be explained in terms of the functional groups in chitosan. Chitosan has numerous amino ligands and hydroxyl groups on its surface,[21] which are all considered to be active sites for the sorption of MCs.

The experimental data were also fitted into intra-particle diffusion plot to understand the rate controlling factors limiting the adsorption of MC-LR onto ChMWCNT (Figure 4b). Based on this model, if a graph/plot displays multi-linearity, this is an indication that there are more than one adsorption processes taking place and the rate-limiting step is not the intra-particle diffusion.[22] In this study, a multi-linear plot was obtained (Figure 4b) thus showing a two-step diffusion process. Figure 4b shows a sharp first step (in the first 15 min) showing external mass transfer through instantaneous sorption. The second and last step (the equilibrium stage) shows intra-particle diffusion and a plateau demonstrating slow diffusion as the levels of the adsorbate decrease. These findings thus show that the adsorption of MC-LR onto ChMWCNT is not solely controlled by intra-particle diffusion. Similar findings were also reported by Mashile et al.[20] who also found that the adsorption of MC-LR onto tyre-based activated carbon was not controlled by intra-particle diffusion as the sole rate-determining step.

### Effect of adsorbent dosage

Dosage of adsorbent is known to have an effect on removal efficiency of adsorbates because increasing the adsorbent dosage increases the available active sites for adsorption. Findings in Figure 3c show that the adsorption of MC-LR onto ChMWCNT was increasing with increasing dose of the adsorbent. This is because increasing the dosage of the adsorbent increased the active sites for the MC-LR to be adsorbed thus an increase in removal efficiency. Maximum adsorption was reported to be 97% at a dose of 30 mg per 5 mL of the solution, and this was adopted as the optimum dose and used in the subsequent experiments.

### Effect of adsorbate concentration

The adsorbates' initial concentration has a direct effect on the rate of adsorption, and this makes it a parameter of paramount importance to investigate to understand the adsorption process. The effect of initial concentration on the adsorption of MC-LR by ChMWCNT was evaluated using seven initial solution concentrations of MC-LR (2.5, 5, 7.5, 10, 15, 20 and 25 µg/L) and a dosage of 30 mg per 5 mL solution. Results in Figure 3b show that the removal efficiency (as % removal) of MC-LR was declining as the concentration of the adsorbate was increased. The trend observed is due to adsorbent's active site being exhausted by the adsorbate with an increase in the initial MC-LR concentrations. Highest MC-LR removal was reported when a 5 µg/L solution was used. This was then selected as the optimum concentration in the subsequent experiments.

The adsorption data obtained when investigating the effect of the adsorbate initial concentration on the process were fitted into nonlinear equations of Langmuir and Freundlich isotherm models. Analysis of the equilibrium isotherms gives data on the adsorption capacity, which is crucial when investigating adsorption systems.[14] The Langmuir model assumes that the adsorption process happens on a homogenous surface with a single layer adsorption taking place at all the available adsorption sites. The model also assumes no interaction between the adsorbent and the adsorbate.[22] On the other hand, the Freundlich model assumes that adsorption occurs in a multi-layer on a heterogeneous adsorbent. The isotherm model assumes adsorption sites to be having unequal energy which differs exponentially, and this results in numerous adsorbate layers forming onto the surface of the adsorbent.[19]

The nonlinear plots for Langmuir and Freundlich isotherms are shown in Figure 5 and their respective constant parameters are shown in Table 3. Based on the correlation coefficient values ($R^2$) (Table 3), the data better fitted the Langmuir isotherm model, thus inferring a monolayer uniform adsorption onto the surface of the ChMWCNT composite and no interaction of the adsorbate molecules.

The calculated $R_L$ (dimensionless equilibrium parameter) values for the Langmuir isotherm and the Freundlich adsorption intensity ($1/n$) in Table 3 were all between 0 and 1, thus showing that the adsorption of MC-LR onto the ChMWCNT composite is favourable. However, when compared to the adsorption capacities of other adsorbents applied in SPATT for microcystins as reported by Zhao et al.[1] (Table 4), the adsorption capacities of Diaion® HP-20 and SEPABEADS SP700 were found to be much higher than the developed composite. Even though superior adsorption capacities were observed by Zhao et al.[1] for the resins Diaion® HP-20 and SEPABEADS SP700, the huge differences with the developed composite could also be due to the huge differences in the experimental conditions, with Zhao et al.[1] having used microcystin concentrations (MC-LR 1 980–1.488 6 × $10^5$ µg/L) which were at least a thousand times more and adsorbent dosages (0.1 g) which were at least three times more than applied in this study.

### Desorption study on ChMWCNT composite

The desorption of MC-LR from the ChMWCNT composite was tested in five different aqueous MeOH solutions (0%, 20%, 50%, 80% and 100%). Recoveries of MC-LR from the composite using these solutions are shown in Figure 6. Pure water (0%), 20% and 50% MeOH solutions proved to be very ineffective in desorbing MC-LR. Higher mean recoveries of 71.99 ± 4.47% and 84.71 ± 6.47% were observed for

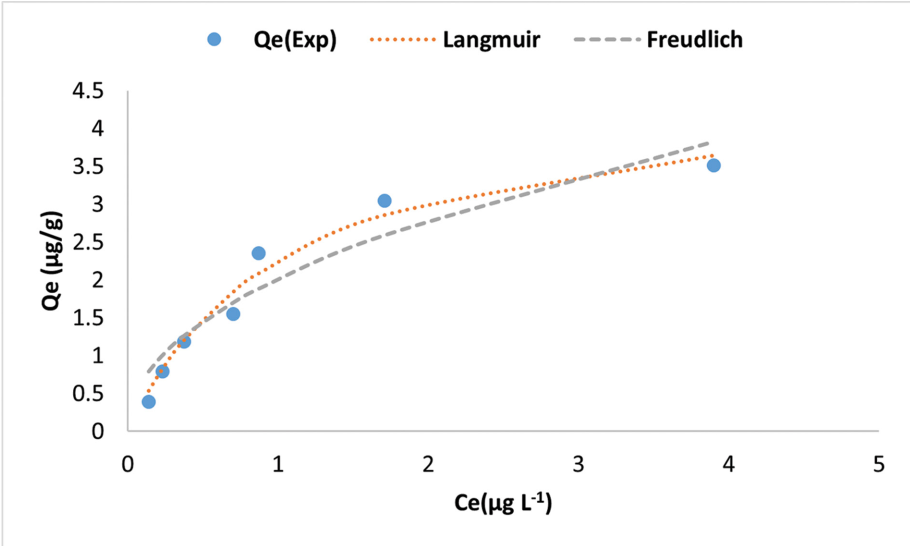



**Figure 5:** Adsorption equilibrium isotherms for MC-LR on ChMWCNT.

**Table 3:** Parameters of Langmuir's and Freundlich's isotherm models at 25 °C

| | Langmuir's isotherm model | | | Freundlich's isotherm model | | |
|---|---|---|---|---|---|---|
| | $Q_{max}$ (μg/g) | $K_L$ (L/mg) | $R^2$ | $K_F$ (L/mg) | $1/n$ | $R^2$ |
| **ChMWCNT** | 4.639 | 0.938 | 0.972 | 2.012 | 0.473 | 0.908 |

**Table 4:** List of adsorbents used in SPATT for microcystins

| Adsorbent | Type of microcystin | Adsorption capacity (μg/g) | Reference |
|---|---|---|---|
| HP20 | MC-LR | 335 | Zhao et al.[1] |
| SP700 | MC-LR | 828 | Zhao et al.[1] |
| ChMWCNT | MC-LR | 4.639 | This study |

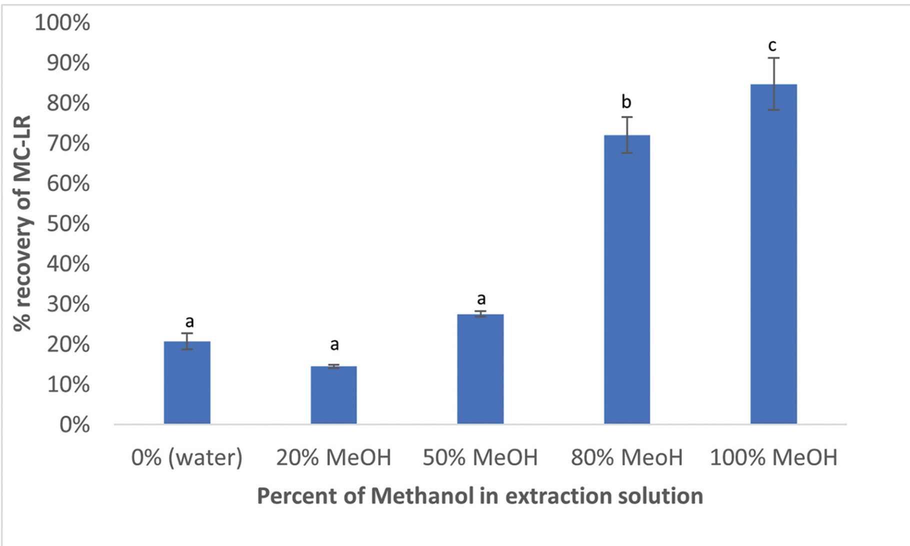



**Figure 6:** Desorption efficiency of MC-LR from ChMWCNT using various percentages of methanol as the extraction solution. Data labelled with different small letters (a, b, c) differed significantly at $p < 0.05$ in each bar ($n = 3$).

80% and 100% MeOH, respectively. Similar findings were reported by Zhao et al.[1] who reported the highest recovery of 78.1 ± 4.1% for MC-LR from the resin SEPABEADS SP700 using 100% MeOH. However, Zhao et al.[1] found better recoveries of 91.5 ± 4.6% for MC-LR with 75% MeOH compared to 100% methanol (± 78% recovery).

Such findings imply that, compared to Diaion® HP-20, the synthesised composite has a stronger affinity for MC-LR. However, the fact that some MCs were desorbed when 100% water was used suggests a possibility of MCs leaching from the composite in solution when this material is used for SPATT.

**Application in raw dam water and SPATT bag format**

To determine the potential of the composite to adsorb and desorb different congeners of MCs (MC-LR, RR and YR), we exposed SPATT samplers loaded with 0.2 g of the composite to 1 L of raw dam water with pH of 7.29 ± 0.71, EC of 878.67 ± 42.44 μS/cm and MCs with the following mean concentrations: MC-LR 25.14 ± 2.34; MC-YR 10.21 ± 0.41 and MC-RR 7.92 ± 0.10 μg/L. Similar congeners were (MC-LR, –YR and –RR), were reported by Pedro et al.[23] when using passive sampling devices (PSDs) followed by LC–MS when monitoring the toxins in Southern Mozambique. Levels of MCs in the range 2.1–159.4 ng/g of PSDs reported by (23) were consistent to the levels reported here (Figure 7).

The findings in Figure 7a and 7b show that except for MC-YR, where the recovery seemed disproportional to the levels in solution, the levels of MC-LR and MC-RR recovered were representative/proportional to the levels in the dam water. This seems to imply less affinity or recoveries of MC-YR compared to the other two congeners monitored. Such findings are consistent with Cook and Newcombe[24] who found the adsorption of different MC congeners in the order MC-RR > MC-YR > MC-LR > MC-LA, when investigating the performance of charcoal and wood powdered activated carbons (PACs) using the surface diffusion model. Similar adsorption performances MC-RR > MC-YR > MC-LR > MC-LA were also reported by Ho et al.[25] for adsorption behaviours of PACs. It is also highly possible that the water had other congeners of MCs not monitored here as reported by Mbukwa et al.[26] for the same catchment, which could have outcompeted MC-YR for adsorption sites onto the composite.

## Conclusions

In this study, a glutaraldehyde-crosslinked chitosan-multi-walled carbon nanotube (ChMWCNT) composite was synthesised and evaluated for its adsorption and desorption of MC-LR in batch experiments and then later applied in raw dam water to assess its applicability in the passive sampling of MCs in SPATT samplers. The optimum conditions for MC-LR adsorption were 30 min contact time, 5 μg/L dosage and 30 mg per 5 mL dosage at ambient room temperatures and natural pH water. The composite was found to be efficient in adsorbing MC-LR, with up to 97% removal of the toxin under optimised conditions. The kinetics data for the adsorption were better explained by the pseudo-second-order model, meaning that the adsorption

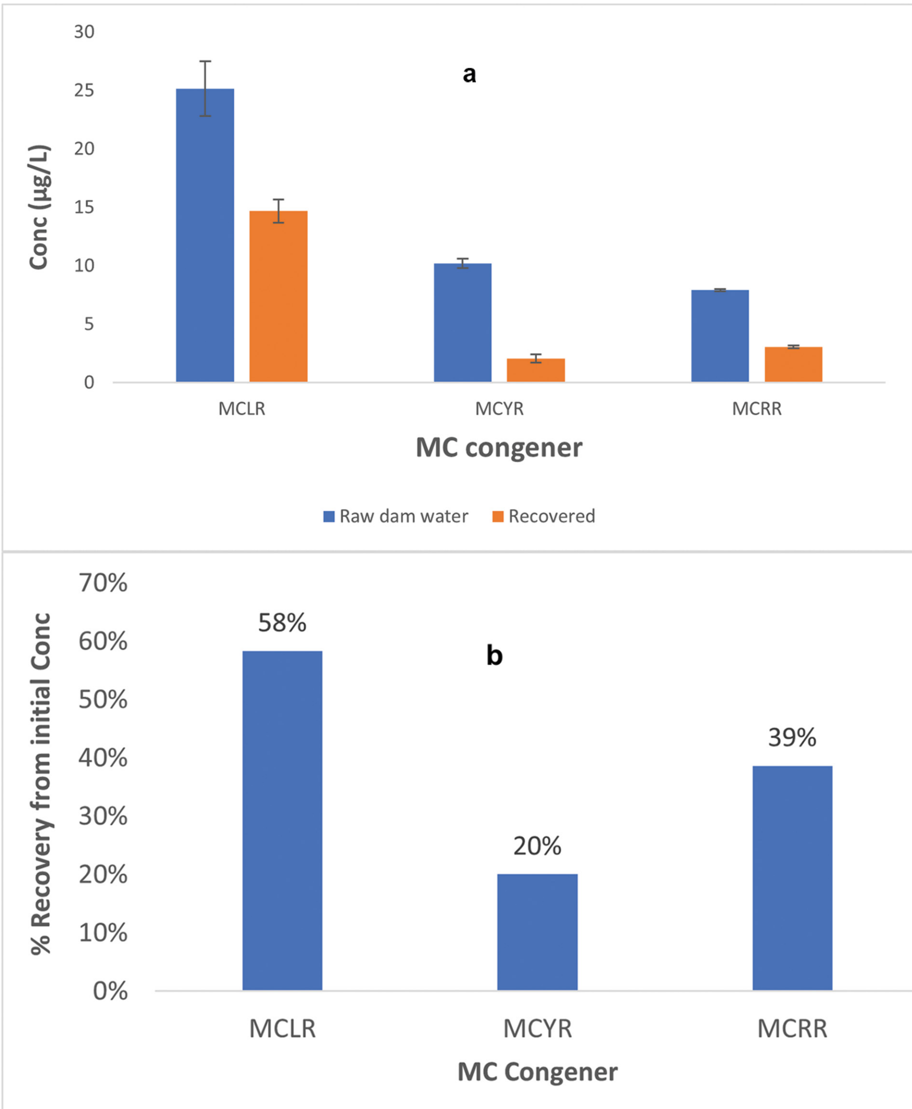



**Figure 7:** (a) The different levels of MCs (LR, YR, RR) adsorbed onto ChMWCNT. (b) Percentage Levels of MC (LR, YR, RR) adsorbed from the raw dam water.

was occurring through chemisorption. The experimental data were fitting better into the Langmuir isotherm compared to the Freundlich isotherm, thus inferring a monolayer surface adsorption of MC-LR onto ChMWCNT. Adsorption capacity of 4.639 μg/g was reported, and $R_L$ (dimensionless equilibrium parameter) values for the Langmuir isotherm were between 0 and 1, inferring that adsorption of MC-LR onto the ChMWCNT composite is favourable. In terms of desorption, 100% methanol was found to be most effective with a highest mean desorption efficiency of 84.71 ± 6.47% reported. When applied for the adsorption in raw dam water, the composite was saturated within 2 days of exposure and adsorbed and desorbed the three congeners of MCs (–LR, –RR and –YR) relatively well. Based on the findings, we recommend further studies looking into the effect of other parameters such as pH, effect of co-existing pollutants and ions and field SPATT monitoring of MCs using the developed composite to assess its performance on different cyanotoxins. The results show that the chitosan-based composite, which can be derived from the deacetylation of chitin from seafood waste, which is abundant in South Africa, can be a cheaper ingredient for a sorbent to be used in the SPATT for microcystins in water bodies used for drinking water and agricultural purposes.

## Acknowledgements

We acknowledge the Department of Nutrition (Faculty of Health Sciences) at the University of Venda for the use of the Shimadzu High-Performance Liquid Chromatography Triple Quadrupole Mass Spectrometer (LCMS-8045). Funding for this study was granted by the South African Water Research Commission (WRC) Project No: K5/2972.

## Competing interests

We have no competing interests to declare.

## Authors' contributions

G.P.K.: Conceptualisation, methodology, data curation, writing – the initial draft and revisions. M.W.G.: Conceptualisation, validation, student supervision, funding acquisition, writing – revisions. N.E.M.: Data curation, writing – revisions.

## References

1. Zhao H, Qiu J, Fan H, Li A. Mechanism and application of solid phase adsorption toxin tracking for monitoring microcystins. J Chromatogr A. 2013;1300:159–164. http://dx.doi.org/10.1016/j.chroma.2013.02.048
2. Díez-Quijada L, Guzmán-Guillén R, Prieto Ortega AI, Llana-Ruíz-Cabello M, Campos A, Vasconcelos V, et al. New method for simultaneous determination of microcystins and cylindrospermopsin in vegetable matrices by SPE-UPLC-MS/MS. Toxins (Basel). 2018;10(10):1–16. https://doi.org/10.3390/toxins10100406
3. Howard MDA, Nagoda C, Kudela RM, Hayashi K, Tatters A, Caron DA, et al. Microcystin prevalence throughout lentic waterbodies in coastal southern California. Toxins (Basel). 2017;9(7):1–21. https://doi.org/10.3390/toxins9070231
4. Buratti FM, Manganelli M, Vichi S, Stefanelli M, Scardala S, Testai E, et al. Cyanotoxins: Producing organisms, occurrence, toxicity, mechanism of action and human health toxicological risk evaluation. Arch Toxicol. 2017;91(3):1049–1130. https://doi.org/10.1007/s00204-016-1913-6
5. Pedro O, Correia D. Occurrence of microcystins in freshwater bodies in Southern Mozambique. J Res Environ Sci Toxicol. 2012;1(14):58–65. https://www.interesjournals.org/articles/occurrence-of-microcystins-in-freshwater-bodies-in-southern-mozambique.pdf
6. Wood SA, Holland PT, MacKenzie L. Development of solid phase adsorption toxin tracking (SPATT) for monitoring anatoxin-a and homoanatoxin-a in river water. Chemosphere. 2011;82(6):888–894. http://dx.doi.org/10.1016/j.chemosphere.2010.10.055
7. MacKenzie L, Beuzenberg V, Holland P, McNabb P, Selwood A. Solid phase adsorption toxin tracking (SPATT): A new monitoring tool that simulates the biotoxin contamination of filter feeding bivalves. Toxicon. 2004;44(8):901–918. https://doi.org/10.1016/j.toxicon.2004.08.020
8. Zendong Z, Herrenknecht C, Abadie E, Brissard C, Tixier C, Mondeguer F, et al. Extended evaluation of polymeric and lipophilic sorbents for passive sampling of marine toxins. Toxicon. 2014;91:57–68. http://dx.doi.org/10.1016/j.toxicon.2014.03.010
9. Gomez-Maldonado D, Filpponen I, Vega Erramuspe IB, Johansson L, Mori MF, Jayachandra Babu R, et al. Development of a β-cyclodextrin-chitosan polymer as active coating for cellulosic surfaces and capturing of microcystin-LR. Surf Interfaces. 2022;33:102192. https://doi.org/10.1016/j.surfin.2022.102192
10. Tran QN, Jin X, Doan NQH. Enhanced removal of extracellular microcystin-LR using chitosan coagulation-ultrafiltration: Performance and mechanisms. J Environ Chem Eng. 2022;10(3):107902. https://doi.org/10.1016/j.jece.2022.107902
11. Gonçalves JO, Santos JP, Rios EC, Crispim MM, Dotto GL, Pinto LAA. Development of chitosan based hybrid hydrogels for dyes removal from aqueous binary system. J Mol Liq. 2017;225:265–270. http://dx.doi.org/10.1016/j.molliq.2016.11.067
12. Alves DCS, Gonçalves JO, Coseglio BB, Burgo TAL, Dotto GL, Pinto LAA, et al. Adsorption of phenol onto chitosan hydrogel scaffold modified with carbon nanotubes. J Environ Chem Eng. 2019;7(6). https://doi.org/10.1016/j.jece.2019.103460
13. Lim DT, Tuyen TN, Nhiem DN, Duc DH, Chuc PN, Bac NQ, et al. Fluoride and arsenite removal by adsorption on $La_2O_3$-$CeO_2$/laterite. J Nanomater. 2021;2021. https://doi.org/10.1155/2021/9991050
14. Choi M, Lee C. Enhanced uoride adsorption on aluminum- impregnated kenaf biochar: Adsorption characteristics and mechanism [preprint]. Res Square. 2022; version 1. https://doi.org/10.21203/rs.3.rs-1504419/v1
15. Braik S, Amor TB, Michelin L, Rigolet S, Bonne M, Lebeau B, et al. Natural water defluoridation by adsorption on Laponite clay. Water Sci Technol. 2022;85(6):1701–1719. https://doi.org/10.2166/wst.2022.091
16. Li B, Shan CL, Zhou Q, Fang Y, Wang YL, Xu F, et al. Synthesis, characterization, and antibacterial activity of cross-linked chitosan-glutaraldehyde. Mar Drugs. 2013;11(5):1534–1552. https://doi.org/10.3390/md11051534
17. Liu AH, Sun KN, Li AM. Preparation of chitosan/carbon nanotubes composites. Adv Compos Lett. 2007;16(4):143–148. https://journals.sagepub.com/doi/pdf/10.1177/096369350701600403
18. Li A, Ma F, Song X, Yu R. Dynamic adsorption of diarrhetic shellfish poisoning (DSP) toxins in passive sampling relates to pore size distribution of aromatic adsorbent. J Chromatogr A. 2011;1218(11):1437–1442. http://dx.doi.org/10.1016/j.chroma.2011.01.043
19. Mudzielwana R, Gitari WM, Ndungu P. Evaluation of the adsorptive properties of locally available alumino-silicate clay in As(III) and As(V) remediation from groundwater. Phys Chem Earth. 2019;112(2018):28–35. https://doi.org/10.1016/j.pce.2018.11.008
20. Mashile PP, Mpupa A, Nomngongo PN. Toxicon adsorptive removal of microcystin-LR from surface and wastewater using tyre-based powdered activated carbon : Kinetics and isotherms. Toxicon. 2018;145:25–31. https://doi.org/10.1016/j.toxicon.2018.02.044
21. Sanford S, Singh KS, Chaini S, LeClair G. Study of natural adsorbent chitosan and derivatives for the removal of caffeine from water. Water Qual Res J Canada. 2012;47(1):80–90. https://doi.org/10.2166/wqrjc.2012.021
22. Zaidi R, Khan SU, Farooqi IH, Azam A. Investigation of kinetics and adsorption isotherm for fluoride removal from aqueous solutions using mesoporous cerium-aluminum binary oxide nanomaterials. RSC Adv. 2021;11(46):28744–28760. https://doi.org/10.1039/d1ra00598g
23. Pedro O, Correia D, Lie E, Skåre JU, Leão J, Sandvik M, et al. Polymerase chain reaction (PCR) detection of the predominant microcystin-producing genotype of cyanobacteria in Mozambican lakes. Afr J Biotechnol. 2011;10(83):19299–19308. https://doi.org/10.5897/AJB11.1521
24. Cook D, Newcombe G. Removal of microcystin variants with powdered activated carbon. Water Supply. 2002;2(5-6):201–207. https://doi.org/10.2166/ws.2002.0170
25. Ho L, Lambling P, Bustamante H, Duker P, Newcombe G, Supe E. Application of powdered activated carbon for the adsorption of cylindrospermopsin and microcystin toxins from drinking water supplies. Water Res [Internet]. 2011;45(9):2954–2964. http://dx.doi.org/10.1016/j.watres.2011.03.014
26. Mbukwa EA, Msagati TAM, Mamba BB. Quantitative variations of intracellular microcystin-LR, -RR and -YR in samples collected from four locations in Hartbeespoort Dam in North West Province (South Africa) during the 2010/2011 summer season. Int J Environ Res Public Health. 2012;9(10):3484–3505. https://doi.org/10.3390/ijerph9103484

Check for updates

# Concentrations of lead in ceramic tableware in South Africa

AUTHORS:
Angela Mathee[1,2]
Louise Renton[1]
Renée Street[1,2]

AFFILIATIONS:
[1]Environment and Health Research Unit, South African Medical Research Council, Johannesburg, South Africa
[2]Department of Environmental Health, Faculty of Health Sciences, University of Johannesburg, Johannesburg, South Africa

CORRESPONDENCE TO:
Angela Mathee

EMAIL:
angie.mathee@mrc.ac.za

DATES:
**Received:** 28 Apr. 2022
**Revised:** 07 Nov. 2022
**Accepted:** 12 July 2023
**Published:** 28 Sep. 2023

HOW TO CITE:
Mathee A, Renton L, Street R. Concentrations of lead in ceramic tableware in South Africa. S Afr J Sci. 2023;119(9/10), Art. #13853. https://doi.org/10.17159/sajs.2023/13853

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☐ All data included
☒ On request from author(s)
☐ Not available
☐ Not applicable

EDITOR:
Michael Inggs

KEYWORDS:
lead, ceramic ware, food, South Africa

FUNDING:
South African Medical Research Council

Ceramic ware is used around the world, usually daily. In the past, lead was used in the glazes and decorative paints applied to ceramic ware, mainly to increase durability, impart a smooth, glasslike finish to glazes and intensify decorative pigments. However, this use of lead at times contributed to lead exposure and poisoning. While measures have been put in place to limit the use of lead in ceramic ware in well-resourced countries, there is relatively little information on the situation in poorly resourced settings. In the current preliminary South African study, we assessed the lead content and leaching rates from newly purchased ceramic ware. The majority of the 44 ceramic ware items had lead levels $\geq$ 90 ppm. Elevated lead concentrations were found in the leachate from only one item. The findings indicate a need for further research on the potential for lead exposure from ceramic wares, and support calls for increased attention to the many potential sources of lead exposure in poorly resourced settings.

**Significance:**

- The study reveals the potential for lead contamination of certain types of ceramic ware available in South Africa.
- Daily use of lead-contaminated ceramic ware may increase the risk of lead exposure, especially among the poorest.
- The study findings are indicative of a need for further research to fully characterise the extent of lead in ceramic ware.

## Introduction

Around the world, ceramic ware is often used daily. Lead has been used for decades in the glazes and decorative paints applied to ceramic ware, mainly to increase durability, impart a smooth, glasslike finish to glazes and intensify decorative pigments. Because of the associated health implications, there has been long-standing concern about the potential for lead to leach from ceramic ware and contaminate foodstuffs.[1] Some ceramic wares, especially artisanal products, have been found to leach significant quantities of lead into the foods they contain[2-4]; for example, when the glaze has been incorrectly fired, it becomes degraded over time, or when chips or cracks form as a consequence of daily wear and tear. Lead leaching rates have also been shown to be affected by the temperature and acidity of foodstuffs (for example foods containing vinegar or fruit juice), as well as the duration of cooking or food storage, amongst other factors.[5-7] At high concentration levels, lead leached from artisanal ceramic ware has been associated with cases of serious or fatal lead poisoning, as well as lead poisoning outbreaks.[8-11] Even low lead levels in blood have been linked to a range of detrimental effects, especially in children, including reductions in intelligence scores, changes in behaviour and increased aggression and violence.[12] In adults, lead poisoning has been associated with kidney damage, hypertension and cardiac disease, amongst other effects.[13,14]

There is a paucity of published research on lead exposure from ceramic ware in South Africa, of both artisanal and commercial origin. However, an instance is known of the US Food and Drug Administration (FDA) issuing an import alert in relation to certain types of commercially produced ceramic ware from South Africa.[15] To obtain information in this regard, a small-scale, preliminary study was undertaken to determine the lead content of commercially available ceramic ware available in Johannesburg, South Africa.

## Methods

Randomly selected ceramic ware was purchased from six large retail chains, with a minimum of four samples per retail chain. A total of 44 items, including side plates and shallow bowls with a capacity of less than 1.1 L, comprised the sample. The items were subjected to X-ray fluorescence (XRF) analysis using a portable, hand-held XRF device, set for elemental ceramic analysis. In addition, a 4% acetic solution was left standing in the ceramic ware items for 24 h at room temperature. The lead content of the leachate was determined by inductively coupled plasma-atomic emission spectroscopy (ICP-AES) according to the protocols of the FDA.[16] A reference level of 2 μg/mL was used as a cut-off point for unacceptable concentrations of ceramic ware leaching lead.

## Results and discussion

Lead levels in the ceramic ware ranged from undetectable (0.0 parts per million; ppm) to a maximum of 64 668.4 ppm, with mean and median concentrations of 4 248.2 ppm and 107.5 ppm, respectively (Table 1). More than 59% of measurements exceeded 90 ppm, which is the level that countries around the world, including South Africa, have agreed to as legally binding with respect to lead in paint, glazes and other coatings.[17] The majority of items included in the sample had been manufactured in China (64%), with smaller proportions originating from Thailand (7%), Portugal (5%), Indonesia (2%) and South Africa (2%). A further 20% of items did not have the country of origin specified. At least a proportion of the products from all countries, apart from Thailand, had lead levels exceeding 90 ppm. Variation in ceramic ware lead levels may include glaze type or firing conditions.

**Table 1:** Lead content and leachate concentrations in commercially available ceramic wares in South Africa

| | *n* | Lead content analysis using XRF | | | | Lead content of leachate (ICP-AES) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Range | Mean (SD) | Median | Above 90 ppm | Range | Mean (SD) | Median | Above FDA reference level |
| **Supplier** | | | | | | | | | |
| 1 | 13 | 0 – 64 668 | 8768 (19 489) | 101 | 54% | 0.0 – 38.8 | 3.1 (10.7) | 0.0 | 7.7% |
| 2 | 7 | 24 – 52 038 | 7525 (19 629) | 108 | 57% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| 3 | 6 | 297 – 4802 | 1526 (1905) | 393 | 100% | 0.0 – 0.2 | 0.0 (0.1) | 0.0 | 0.0% |
| 4 | 6 | 332 – 1577 | 565 (497) | 368 | 100% | 0.0 – 0.2 | 0.0 (0.1) | 0.0 | 0.0% |
| 5 | 4 | 0 – 107 | 53 (45) | 52 | 25% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| 6 | 8 | 0 – 7300 | 940 (2570) | 28 | 25% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| **Country of origin** | | | | | | | | | |
| Not specified | 9 | 67 – 52 038 | 6142 (17 217) | 332 | 89% | 0.0 – 0.2 | 0.0 (0.1) | 0.0 | 0.0% |
| South Africa | 1 | 64 668 – 64 668 | 64 668 (0) | 64 668 | 100% | 38.0 – 38.0 | 38.8 (0.0) | 38.8 | 100.0% |
| Thailand | 3 | 0 – 59 | 31 (30) | 33 | 0% | 0.0 – 0.1 | 0.0 (0.1) | 0.0 | 0.0% |
| Portugal | 2 | 0 – 35 811 | 19 635 (22 877) | 19 635 | 100% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| Indonesia | 1 | 653 – 653 | 653 (0) | 653 | 100% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| China | 28 | 0 – 8949 | 963 (2276) | 82 | 50% | 0.0 – 1.8 | 0.1 (0.3) | 0.0 | 0.0% |
| **Background colour** | | | | | | | | | |
| White | 11 | 0 – 8949 | 2278 (3289) | 360 | 73% | 0.0 – 1.8 | 0.2 (0.5) | 0.0 | 0.0% |
| Cream/beige/yellow | 11 | 23 – 653 | 168 (202) | 101 | 55% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| Orange/red | 4 | 28 – 392 | 195 (194) | 180 | 50% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| Shades of blue | 4 | 0 – 35 811 | 9046 (17 844) | 186 | 50% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| Shades of green | 5 | 40 – 64 668 | 24 096 (31 619) | 3459 | 80% | 0.0 – 38.8 | 7.8 (17.4) | 0.0 | 20.0% |
| Brown/black | 7 | 33 – 1577 | 352 (558) | 67 | 43% | 0.0 – 0.2 | 0.0 (0.1) | 0.0 | 0.0% |
| Multiple colours | 2 | 0 – 108 | 54 (76) | 54 | 50% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| **Pattern colour** | | | | | | | | | |
| White | 7 | 0 – 2920 | 570 (1048) | 133 | 71% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| Cream/beige/yellow | 3 | 23 – 24 | 23 (0.7) | 23 | 0% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| Orange/red | 5 | 28 – 7300 | 1622 (3178) | 332 | 60% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| Shades of blue | 8 | 0 – 35 811 | 5274 (13 468) | 52 | 43% | 0.0 – 0.0 | 0.0 (0.0) | 0.0 | 0.0% |
| Shades of green | 6 | 40 – 64 668 | 20 148 (29 889) | 1933 | 83% | 0.0 – 38.8 | 6.6 (15.8) | 0.0 | 16.7% |
| Brown/black | 9 | 33 – 4804 | 820 (1572) | 108 | 56% | 0.0 – 0.2 | 0.1 (0.1) | 0.0 | 0.0% |
| Multiple colours | 6 | 0 – 8949 | 1593 (3605) | 107 | 83% | 0.0 – 1.8 | 0.3 (0.7) | 0.0 | 0.0% |
| **Total sample** | 44 | 0 – 64 668 | 4248 (13 259) | 108 | 59% | 0.0 – 38.8 | 0.9 (5.8) | 0.0 | 2.3% |

The product with the highest lead measurement was manufactured in South Africa; South Africa was also the country of origin of the only item from which lead leached at a concentration exceeding the reference level of 2 μg/mL. While it is reassuring that lead did not leach from the majority of surveyed products at the time of purchase, from a public health perspective, concern remains over what might occur over time, with erosion of the glazes and pigments through general wear and tear, and with chipping and cracking of ceramic ware surfaces. While we focused on new ceramic ware in this study, the use of old, antique dinner ware has also been found to be of concern.[18] Microwave ovens have also been associated with increased leaching of lead from ceramic ware[19]. as has the use of acidic ingredients, longer contact times between ceramic ware and foodstuffs, and higher temperatures[20].

Lead exposure, as well as the associated detrimental health and social outcomes, are well established to be elevated in settings of poverty.[21] With regard to ceramic ware, poor people are unlikely to be able to afford frequent replacement of cracked or chipped ceramic ware, and in this regard too, may be at elevated risk of exposure and the concomitant health effects. Reductions in IQ points associated with lead exposure in South Africa has been calculated to cost the national economy around USD17.7 billion annually[22]; in this light, uncompromising action to reduce lead exposure, especially in the most vulnerable communities in South Africa, is warranted.

When lead paint was first identified as a public health concern in South Africa[23-25], the government response was to draft regulations to limit the use of lead in paint intended for household use in the country. Initially promulgated in 2009, those regulations are on the verge of being strengthened by reducing the maximum permissible concentration of lead in paint from 600 ppm to 90 ppm, and by making the regulations applicable to all paints and coatings, including those used in the commercial and industrial sectors. These developments ought to provide a higher level of public health protection against lead exposure from ceramic ware, but, importantly, need to be accompanied by a programme of monitoring and surveillance of the lead content of paint and other coatings manufactured, imported, sold and used in South Africa. This should include both the formal and informal sectors, with concomitant punitive action, where warranted. Given that levels of awareness of lead hazards in South Africa are very low[26], a public education campaign is also key.

## Acknowledgements

We gratefully acknowledge the support received by Modiegi Mogotsi-Maakwe in preparation and analysis of the samples. This research was funded by the South African Medical Research Council (Baseline Grant 2016/17).

## Competing interests

We have no competing interests to declare.

## Authors' contributions

A.M.: Conceptualisation, methodology, formal analysis, writing – original draft preparation, writing – review and editing, supervision. L.R.: Methodology, formal analysis, investigation, writing – original draft preparation, writing – review and editing, supervision, project administration. R.S.: Formal analysis, writing – review and editing.

## References

1. Wallace DM, Kalman DA, Bird TD. Hazardous lead release from glazed dinnerware: A cautionary note. Sci Total Environ. 1985;44:289–292. https://doi.org/10.1016/0048-9697(85)90101-9

2. Spielholtz GI, Kaplan FS. The problem of lead in Mexican pottery. Talanta. 1980;27:997–1000. https://doi.org/10.1016/0039-9140(80)80134-2

3. Aderemi TA, Adenuga AA, Oyekunle JAO, Ogunfowokan AO. High level leaching of heavy metals from colorful ceramic foodwares: A potential risk to human. Environ Sci Pollut Res Int. 2017;24:17116–17126. https://doi.org/10.1007/s11356-017-9385-7

4. Mandal PR, Das S. Leachable lead and cadmium in microwave-heated ceramic cups: Possible health hazard to human. Environ Sci Pollut Res Int. 2018;25:28954–28960. https://doi.org/10.1007/s11356-018-2944-8

5. Sheets RW. Extraction of lead, cadmium and zinc from overglaze decorations on ceramic dinnerware by acidic and basic food substances. Sci Total Environ. 1997;197:167–175. https://doi.org/10.1016/S0048-9697(97)05431-4

6. Mejía EG, Craigmill A. Transfer of lead from lead-glazed ceramics to food. Arch Environ Contam Toxicol. 1996;31:581–584. https://doi.org/10.1007/BF00212443

7. Lynch R, Elledge B, Peters C. An assessment of lead leachability from lead-glazed ceramic cooking vessels. J Environ Health. 2008;70(9):36–41. http://www.jstor.org/stable/26328206

8. Fralick M, Thompson A, Mourad O. Lead toxicity from glazed ceramic cookware. Can Med Assoc J. 2016;188:E521–E524. https://doi.org/10.1503/cmaj.160182

9. Klein M, Namer R, Harpur E, Corbin R. Earthenware containers as a source of fatal lead poisoning: Case study and public-health considerations. N Engl J Med. 1970;283:669–672. https://doi.org/10.1056/NEJM197009242831302

10. Manor E, Freundlich E. Lead poisoning from ceramics in the Arab population. Arch Toxicol Suppl. 1983;6:275–278. https://link.springer.com/chapter/10.1007/978-3-642-69083-9_51

11. Matte TD, Proops D, Palazuelos E, Graef J, Hernandez Avila M. Acute high-dose lead exposure from beverage contaminated by traditional Mexican pottery. Lancet. 1994;344:1064–1065. https://doi.org/10.1016/S0140-6736(94)91715-9

12. Reyes JW. Lead exposure and behavior: Effects on antisocial and risky behavior among children and adolescents. Econ Inq. 2015;53:1580–1605. http://dx.doi.org/10.1111/ecin.12202

13. Betts KS. CDC updates guidelines for children's lead exposure. Environ Health Perspect. 2012;120:Art. #a268. https://doi.org/10.1289/ehp.120-a268

14. Bellinger D, Malin A, Wright RO. The neurodevelopmental toxicity of lead: History, epidemiology, and public health implications. In: Aschner M, Costa LG, editors. Advances in Neurotoxicology. Cambridge, MA: Academic Press; 2018. p. 1–26. https://doi.org/10.1016/bs.ant.2018.03.009

15. US Food and Drug Administration (FDA). Import Alert 52-08 [webpage on the Internet]. c2023 [cited 2023 Aug 17]. Available from: https://www.accessdata.fda.gov/CMS_IA/importalert_122.html

16. US Food and Drug Administration (FDA). Elemental analysis manual for food and related products [webpage on the Internet]. c2010 [cited 2021 Aug 23]. Available from: https://www.fda.gov/food/laboratory-methods-food/elemental-analysis-manual-eam-food-and-related-products

17. United Nations Environment Programme (UNEP), World Health Organization (WHO). Global Alliance to Eliminate Lead Paint/Lead Paint Alliance advisory council meeting [webpage on the Internet]. c2018 [cited 2021 Aug 25]. Available from: https://www.unep.org/explore-topics/chemicals-waste/what-we-do/emerging-issues/global-alliance-eliminate-lead-paint

18. Sheets RW, Turpen SL. Lead hazards from old ceramic dinnerware. Stud Environ Sci. 1997;66:327–333. https://doi.org/10.1016/S0166-1116(97)80053-2

19. Sheets RW, Turpen SL, Hill P. Effect of microwave heating on leaching of lead from old ceramic dinnerware. Sci Total Environ. 1996;182:187–191. https://doi.org/10.1016/0048-9697(95)05053-1

20. Bolle F, Fekete V, Demont M, Boutakhrit K, Petit D, Brian W, et al. Lead migration from ceramicware in contact with foodstuff: Effect of glaze, temperature, pH and food simulant. J Food Sci Eng. 2012;2:301–313. https://doi.org/10.17265/2159-5828/2012.06.001

21. Leech TG, Adams EA, Weathers TD, Staten LK, Filippelli GM. Inequitable chronic lead exposure. Fam Community Health. 2016;39:151–159. https://doi.org/10.1097/FCH.0000000000000106

22. Attina TM, Trasande L. Economic costs of childhood lead exposure in low-and middle-income countries. Environ Health Perspect. 2013;121:1097–1102. https://doi.org/10.1289/ehp.1206424

23. Mathee A, Röllin H, Levin J, Naik I. Lead in paint: Three decades later and still a hazard for African children? Environ Health Perspect. 2007;115:321–322. https://doi.org/10.1289/ehp.9575

24. Mathee A, Singh E, Mogotsi M, Timothy G, Maduka B, Olivier J, et al. Lead-based paint on playground equipment in public children's parks in Johannesburg, Tshwane and Ekurhuleni. S Afr Med J. 2009;99(11):819–821. Available from: http://www.scielo.org.za/pdf/samj/v99n11/v99n11a26.pdf

25. Montgomery M, Mathee A. A preliminary study of residential paint lead concentrations in Johannesburg. Environ Res. 2005;98:279–283. https://doi.org/10.1016/j.envres.2004.10.006

26. Haman T, Mathee A, Swart A. Low levels of awareness of lead hazards among pregnant women in a high risk Johannesburg neighbourhood. Int J Environ Res Public Health. 2015;12(12):15022–15027. https://doi.org/10.3390/ijerph121214968