South African Journal *of* Science

volume 115
number 9/10

Air pollution and health

Decoloniality of geography

Conservation of the Cape Floristic Region

Strength and conditioning in high school rugby

150 years of the periodic table

ASSAf
ACADEMY OF SCIENCE OF SOUTH AFRICA

Volume 115
Number 9/10
September/October 2019

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 & Technology.

Design and layout
SUN MeDIA Bloemfontein
T: 051 444 2552
E: publish@sunbloem.co.za

Correspondence and enquiries
sajs@assaf.org.za

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

Research Article

Research Letter

Cover caption
A Cape sugarbird *(Promerops cafer)* among the fynbos of the Cape Floristic Region. This species is one of several endemic to the Fynbos Biome. In an article on page 57, Allsopp and colleagues identify research questions for the conservation of the Cape Floristic Region.

Antecedents of sustainability

Hardly a day passes without media reference to 'sustainability' or 'sustainable development'. The three interlocked economic, social and environmental factors that sustain the earth, its population and its biosphere are familiar to many people working in both the natural and human sciences. The philosophy of sustainable development is usually expressed in phrases such as 'being able to meet the needs of the present without compromising those of the future' or 'maintaining an ecological balance by avoiding the depletion of natural resources'.

The 17 Sustainable Development Goals of the United Nations are clearly articulated, and they also form the national agenda for policy and action in many individual countries. This includes South Africa that has an official National Framework for Sustainable Development.[1]

It is often assumed that the idea of sustainable development emerged in the 1980s with the World Commission on Environment and Development chaired by Gro Harlem Brundtland, three times Prime Minister of Norway and Director-General of the World Health Organization (1998–2003). The Commission, and its report *Our Common Future*, rode the wave of environmentalism triggered by disappointment in the technological promise of the post-war world, the impact of writers like Rachel Carson[2], events such as the first Earth Day (1970), a growing appreciation of the gap between rich and poor and, particularly, as the report expressed it, the danger 'of creating a planet our ancestors would not recognise'[3]. The ideas in *Our Common Future* were reinforced with fanfare by the 1992 United Nations Conference on Environment and Development (Earth Summit) in Rio de Janeiro and by the 2002 World Summit on Sustainable Development in Johannesburg.

But, like many other ideas believed to be 'new', sustainability is not new, and many antecedents are worth revisiting. *The Invention of Sustainability: Nature and Destiny, c.1500–1870*, a recent book by Cambridge scholar Paul Warde, is indispensable for the topic[4], and reading it would productively be accompanied by *The Environment: A History of the Idea*[5].

The idea of sustainability has a very long history that stretches back more than 300 years; it emerged as a principle of forestry that required the attention of government even at that time. In England it was articulated by John Evelyn (1620–1706), founding member of the Royal Society of London. *Sylva, Or a Discourse of Forest-Trees and the Propagation of Timber in His Majesty's Dominions* was one of the first books published by the Society (in 1664).[6] Its purpose was to explain how to utilise, but not to deplete, the nation's trees so that the Royal Navy's ship-building programme might prosper, and with it the country's economy, military security and geographical expansion. Of particular urgency was the need to plant trees in the Forest of Deane, as Evelyn noted in his diary, 'with Oake now so much exhausted of the choicest ship-timber in the World'[7]. *Sylva* is one of the most influential texts on forestry ever published and is frequently referred to in many publications even today. So, too, is Evelyn's 1661 work on air pollution (London was notoriously polluted for centuries until the legislation of the 1950s) entitled *Fumifugium, Or the Inconvenience of the Aer and Smoak of London Dissipated, Together With Some Remedies Humbly Proposed*. This was the first publication to deal with urbanisation and pollutants, explaining the problem and suggesting a solution. Moreover, Evelyn's gardening manual of 1664, *Kalendarium Hortense, Or the Gard'ners Almanack, Directing what He is to do Monthly Throughout the Year*, has been the model for all books on gardening thenceforth.

In 2013, German Chancellor Angela Merkel officially celebrated the tricentenary of the German word for sustainability: *Nachhaltigkeit*, observing that it had been transformed into 'a principle of survival'[8]. The term appeared in *Sylvicultura oeconomica*, a book by Evelyn's contemporary, Hans Carl von Carlowitz (1645–1714), an aristocrat in the German kingdom of Saxony who was in charge of the region's silver mines. Vast amounts of charcoal were vital for mining and smelting ores. While the ore in the Saxony mines remained rich and abundant, in Von Carlowitz's time, output (and thus employment) were contracting owing to a timber shortage and escalating prices for what was available. Von Carlowitz wrote to the king, explaining how woodlands had disappeared, trees having been felled and replaced with cultivated crops and villages into which a growing population settled. He argued that national prosperity relied on good future planning, with land usage controlled and careful statistics maintained for ongoing monitoring of resources. Improvement in supply would not come about if the prevailing ethos was profligacy and Von Carlowitz urged that fitting policies be introduced for sustainable forestry.[9]

One might argue that despite centuries of discussion about the appropriate use of renewable resources and principles of sustainability, this vision will not materialise while the world remains dominated by policies predicated on constant improvement in the lives of 7.7 billion humans in a context of ever-expanding economic growth.

References

1. Department of Environmental Affairs and Tourism (DEAT). People, planet, prosperity: A national framework for sustainable development in South Africa. Pretoria: DEAT; 2008. Available from: https://www.gov.za/sites/default/files/gcis_document/201409/nationalframeworkforsustainabledevelopmenta0.pdf
2. Carson R. Silent spring. Boston, MA: Houghton Mifflin; 1962.
3. World Commission on Environment and Development. Our common future. Oxford: Oxford University Press; 1987. p. xi.
4. Warde P. The Invention of sustainability: Nature and destiny, c.1500–1870. Cambridge: Cambridge University Press; 2018.
5. Warde P, Robin L, Sörlin S. The environment: A history of the idea. Baltimore, MD: Johns Hopkins Press; 2018.
6. Evelyn J. Sylva, Or a discourse of forest trees and the propagation of timber in His Majesty's Dominions. 2 vols. London: John Martin for the Royal Society; 1664.
7. De Beer ES, editor. The diary of John Evelyn. London: Oxford University Press; 1959. p. 3447.
8. Mauch C. "But where the danger lies, also grows the saving power": Reflections on exploitation and sustainability. In: Mauch C, Robin L, editors. The edges of environmental history: Honouring Jane Carruthers. Munich: Rachel Carson Center; 2014. p. 129–135.
9. Von Carlowitz HC. Sylvicultura oeconomica, oder haußwirthliche Nachricht und Naturmäßige Anweisung zur wilden Baum-Zucht [Sylvicultura oeconomica or the instructions for wild tree cultivation]. Leipzig: Braun; 1713. German.

HOW TO CITE:
Carruthers J. Antecedents of sustainability. S Afr J Sci. 2019;115(9/10), Art. #6759, 1 page. https://doi.org/10.17159/sajs.2019/6759

AUTHORS:
Danielle Millar[1]
Caradee Y. Wright[1,2]

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

CORRESPONDENCE TO:
Caradee Wright

EMAIL:
cwright@mrc.ac.za

HOW TO CITE:
Millar D, Wright CY. Statement on Air Pollution and Health. S Afr J Sci. 2019;115(9/10), Art. #6823, 1 page. https://doi.org/10.17159/sajs.2019/6823

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
human health, well-being, climate change, environmental health

PUBLISHED:
26 September 2019

Statement on Air Pollution and Health

Clean air is vital to life. Suffering and death from polluted air are avoidable. Immediate, necessary action will prevent air pollution and its staggering toll on life and the fiscus.

In 2016, 91% of the world's population lived in places that did not meet the World Health Organization's (WHO) air quality guidelines.[1] One air pollutant of immense concern is atmospheric particulate matter (PM) which is the sum of the complex combination of solid and liquid particles of organic and inorganic substances suspended in the air. The key components of PM are present in different concentrations depending on the area of air which is tested. Particles with a diameter of 10 microns or less are known as PM_{10} and can affect human health. Particles with a diameter of 2.5 microns or less ($PM_{2.5}$) are more dangerous as they can penetrate the lung barrier and enter the bloodstream, disseminating to various organs. Chronic exposure to these particles increases the risk of developing multiple diseases and conditions.[2,3] There is evidence that air pollution affects human health at every stage of life, with the most vulnerable populations being the young, elderly and health-compromised. Evidence is mounting that associates air pollution with the premature deaths of at least five million people per year, as well as increasing susceptibility to and aggravating existing conditions.[1-3]

The main source of air pollution globally is the use and burning of biomass and fossil fuels for power, heat, transport and food production.[1-3] In South Africa in 2016 the death rate attributed to household air pollution was 34 per 100 000 population, calculated considering acute and chronic respiratory diseases linked to air pollution exposure, and cardiovascular diseases for which air pollution is a risk factor.[4] In South Africa, historically, low-cost residential areas were sited close to industrial zones. The continued influx of people to these urban and industrial areas has led to informal dwellings in and around the area boundaries. Providing basic services to these settlements is often delayed and/or frequently interrupted. Consequently, communities have limited resources: biomass and fossil fuels for cooking and for burning waste, often in illegal dumpsites. Household air pollution occurs from incomplete combustion of solid fuels, which generates smoke. Currently many areas in South Africa exceed National Ambient Air Quality Standards, and the geographical concentration of large population centres and industry have caused hotspots of air quality Priority Areas.[3]

In July 2019, the Academy of Science of South Africa (ASSAf) joined the science academies of Germany, Brazil and the USA, as well as the US National Academy of Medicine, at the United Nations headquarters in New York to issue an urgent call to citizens, governments and businesses to reduce global air pollution.[5] The delegation presented a science–policy statement[6] to senior UN representatives and high-level diplomats. ASSAf was represented at the event by Executive Officer, Professor Himla Soodyall, who was joined by Senior Specialist Scientist in the Environment and Health Research Unit of the South African Medical Research Council, Dr Caradee Wright.[5]

National academies are crucial as they are a forum in which scientists from every discipline can come together, share and reflect upon their findings, placing them in a unique position to address intricate issues such as the interplay between health and pollution. It is essential that the issue of pollution be moved up the policy agenda. Collaborations and continued strengthening of partnerships with other policy areas, such as climate change, sustainable development and food security, would expedite this process.

The Statement[6] appeals for emissions controls in all countries as well as proper monitoring of key pollutants, especially $PM_{2.5}$, and stresses that more funding is needed to invest in air pollution reduction measures to match the scale of the problem. Decisive action by stakeholders can result in cost-effective management of air pollution. Combatting air pollution will help fight climate change as the pivotal common source is the continued use of fossil fuels.

Many more stakeholders internationally will need to join the initiative to ensure and hasten its success. Policymakers and the public need to engage with researchers to improve the future health of people and the planet. September 2019 will see international action being taken. A full-text publication is planned for the September issue of *Annals of Global Health*, promoting global dissemination of the statement and ensuring it is indexed and accessible through PubMed and other databases. Furthermore, about 500 delegates from the Environmental Health fraternity are expected to attend a 2-day conference in 2019 – a collaboration between WHO and the Public Health Association of South Africa – for a national WHO-facilitated dialogue on air pollution and health, based on outcomes of the 1st WHO global conference on air pollution and health and the 3rd African Inter-Ministerial Conference on Health and Environment, both held in 2018, which also resolved to address air pollution as a regional priority.

The Statement is available at https://air-pollution.health/

References

1. World Health Organization. Ambient air pollution [webpage on the Internet]. c2019 [cited 2019 Jul 31]. Available from: https://www.who.int/airpollution/ambient/en/
2. Who.int. Ambient (outdoor) air quality and health [webpage on the Internet]. c2019 [cited 2019 Jul 31]. Available from: https://www.who.int/en/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health
3. South African Department of Environmental Affairs. Air quality [document on the Internet]. c2019 [cited 2019 Jul 31]. Available from: https://www.environment.gov.za/sites/default/files/reports/environmentoutlook_chapter10.pdf
4. Mortality attributed to household and ambient air pollution – Data by country [data set on the Internet]. Global Health Observatory data repository. Updated 2018 May 30 [cited 2019 Jul 31]. Available from: https://apps.who.int/gho/data/view.main.GSWCAH37v
5. South African Medical Research Council. Global Science Academies take a stand against air pollution [webpage on the Internet]. c2019 [cited 2019 Jul 31]. Available from: http://www.samrc.ac.za/media-release/global-science-academies-take-stand-against-air-pollution
6. Air pollution and health [webpage on the Internet]. No date [cited 31 Jul 2019]. Available from: https://air-pollution.health/

AUTHORS:
Valentine Saasa[1,2]
Nonkululeko Radebe[3]

AFFILIATIONS:
[1]Materials Science and Manufacturing, Council for Scientific and Industrial Research, Pretoria, South Africa
[2]Department of Biochemistry, University of Pretoria, Pretoria, South Africa
[3]Karlsruhe Institute of Technology, Karlsruhe, Germany

CORRESPONDENCE TO:
Valentine Saasa

EMAIL:
vsaasa@csir.co.za

HOW TO CITE:
Saasa V, Radebe N. 69th Lindau Nobel Laureate Meeting: Personal experiences of two young scientists. S Afr J Sci. 2019;115(9/10), Art. #a0315, 3 pages. https://doi.org/10.17159/sajs.2019/a0315

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
dark matter, diffraction limit, imposter syndrome, science for society

PUBLISHED:
26 September 2019

69th Lindau Nobel Laureate Meeting: Personal experiences of two young scientists

The annual Lindau Nobel Laureate Meeting is an extremely prestigious 1-week event at which bright young scientists from across the globe are able to meet Nobel laureates to discuss matters of science. These meetings were initiated in 1951 to liberate German scientists from their post-war isolation. The aim was to encourage and cement networks and reduce barriers between nations. Since that time, the meeting has taken place every year in the small Bavarian town of Lindau on Lake Constance, alternating among the disciplines of medicine and physiology, chemistry and physics. An interdisciplinary meeting revolving around all three disciplines is held every 5 years and a meeting on economics is held every 3 years.

This year's meeting attracted 580 young scientists – master's and doctoral students and postdoctoral researchers – from 89 countries and 39 Nobel laureates. The meeting was dedicated to physics, and revolved around topics such as cosmology, particle physics, laser physics, gravitational waves, quantum technologies, dark matter and the graphene flagship programme. Themes were addressed in the form of lectures, panel discussions, master classes and science breakfasts. The motto of the Lindau meeting is 'Educate, inspire and connect'.

In order to attend a Lindau Nobel Laureate Meeting as a young scientist, applications must be submitted to a host institution such as a national academy. In the case of South Africa, this is the Academy of Science of South Africa (ASSAf) which provides travel grants through support from the Department of Science and Innovation (DSI), formerly the Department of Science and Technology. The DSI and ASSAf cover the cost of travel and logistics in South Africa, while the Lindau Council co-funds accommodation and any medical expenses during the meeting. There is a variety of accommodation options, including hotels near the meeting venue and beyond the island, and even staying with a host family in Lindau.

This year, two of the young South African scientists who participated were Valentine Saasa and Nonkululeko Radebe. Valentine Saasa is a PhD candidate at the CSIR and is registered for her degree at the University of Pretoria. She works on the synthesis of nanostructured chemical sensors for non-invasive monitoring of diabetes mellitus. Nonkululeko Radebe is a PhD candidate at the Karlsruhe Institute of Technology in Germany. Her research involves combined rheo-spectroscopy techniques for hydration kinetic studies on cement paste. They respectively describe their experiences below.

Experience of the 69th Lindau Nobel Meeting: Valentine Saasa

The first time I heard of the Lindau Nobel Laureate Meeting was in 2016 at the University of Limpopo when I was hosting and facilitating a women in science communication event sponsored by ASSAf and the British Council. Stanley Maphosa, International Liaison Manager at ASSAf, was telling participants about this opportunity – dedicated to chemistry at that time. I did not fully understand what it was all about, until the 2018 meeting dedicated to physiology and medicine when I followed Edith Phalane's (a 2018 Lindau alumnus) twitter posts on her attendance. I was unsure about applying for the 2019 meeting dedicated to physics, but as my research involves physics and I am supervised by a physics chief researcher at the CSIR, I nonetheless applied.

Unbelievably, I was selected by ASSAf in the first stage and nominated by them to the Lindau Council. I could not believe it and I did not want to tell anyone yet, as I was not sure what was going to happen in the next stage. In January 2019, I received the email from the Lindau Council informing me that I had been selected to attend the meeting! I was excited, and yet nervous at the same time, wondering what I – a biochemist by training – might say to the greatest physicists in the world. But then I learned that there would be Nobel laureates who were biochemists by training but had won either chemistry or physics Nobel prizes. The pre-Lindau meeting with alumni in South Africa, organised by ASSAf, also served to calm my nerves and ignite excitement about the trip.

When we arrived, the meeting venue was breathtaking. It is situated on the island of Lindau. Even though there were many hotels, they could not accommodate all the young scientists and many had to stay outside the island, some with host families as I have mentioned. I was privileged to be accommodated just across from the meeting venue.

The second day at Lindau was remarkable for me, because South Africa, as host of the International Day, opened the ceremony with a Xhosa cultural dance. I felt at home, and it seemed that everyone from around the world was having fun. When your skin colour was noticed and you were asked whether you were a South African, there were many compliments on the opening ceremony. A second highlight of the opening day was when Nobel laureate Brian Schmidt (Nobel Prize in Physics 2011) delivered his keynote address entitled, 'Big questions for society, big questions for research'.

There were different talk formats during the week: lectures at which laureates presented their prize-winning work, open exchanges where questions were asked by young researchers and answered by laureates, and agora talks, rather more interactive and informal, during which the young scientists could ask laureates about their career paths and personal lives. I had the opportunity to talk to Harald zur Hausen, who presented on bovine products as the origin of infections linked to colon and breast cancers. He received the Nobel Prize in Physiology or Medicine in 2008 for his work on the role of papilloma viruses in cervical cancer.

We also had the opportunity to choose either to lunch or walk with a laureate. I had the honour of lunching with Hartmut Michel, together with nine other young scientists with a biochemistry or biophysics background. Michel won the Nobel Prize in Chemistry in 1988 for his work on the determination of the three-dimensional structure of a protein complex found in certain photosynthetic bacteria. We had a very relaxed lunch and conversation

ranged from the personal, the cultural and politics to academia. A take-home message from him was to focus on quality research rather than the impact factor and publish our work in open-access journals because they are the future of science publishing.

Another highlight was the breakfast session with Sir Konstantin Novoselov, who won the Nobel Prize in Physics in 2010 for discovering the material graphene. His work is closely related to mine, and I appreciated that he is very down to earth and approachable. Novoselov spoke on his experience in researching advanced materials like graphene, a single-layered carbon compound which seems to have an endless supply of applications. Graphene is just one of a plethora of new 'smart materials' which react to environmental changes such as pH, temperature or ultraviolet light. They form the basis of many modern sensors and are being used in fields from computing to medicine. Novoselov argues that the application of research is best achieved by commercial companies, not by universities and research institutions. Coming from both a research institute and a university, I couldn't agree more.

Young South African scientists at the International Day hosted by South Africa

How can science change the world for the better? A take home message for Valentine

Given the socio-economic status of our country, I was delighted to hear the Nobel laureates talk about science for society during a panel discussion. Our new White Paper, together with the National Development Plan, has identified science, technology and innovation (STI) as the primary drivers of economic growth, job creation and socio-economic reform. However, it is clear that South Africa is not yet fully benefitting from the potential of STI to address our socio-economic problems. This issue – is science solving societal issues – was addressed in detail by the panellists: Steven Chu, Brian Schmidt, Vinton Cerf, Tim Lauce and 2016 Lindau alumnus and South African Adriana Marais.

It is interesting to me that a scientist would ask how science can change the world for the better, because it indicates that, as scientists, we are uncertain if science is doing what it is supposed to. Schmidt argued that for science to serve all humanity, the issue of income distribution should be addressed. This, in turn, is dependent on science becoming more cooperative on a global scale.

'Scientific knowledge brings understanding, and that understanding can guide and inform how society can meet such great global challenges. So, let's use our scientific powers to be more active politically and in other ways', Chu (Nobel Prize in Physics 1997) appealed. The role of science is great, but in order to rise to the occasion, scientists need to think not only about how science could be done better but also about how its benefits might be felt by all humankind. The importance of curiosity-driven basic research translating research knowledge to technology in order to improve the way we do and communicate science is key to tackling our societal problems. However, according to Cerf, it is not the job of scientists to translate knowledge into technology, but rather the job of engineers and similar companies.

Attending the Lindau Nobel Laureate Meeting was a once-in-a-lifetime experience and I encourage other young researchers to apply. It exposes you to different kinds of research and the best research in the world – as well as a lasting network.

Experience of the 69th Lindau Nobel Laureate Meeting: Nonkululeko Radebe

I had been eager to attend a Lindau Nobel Laureate Meeting since speaking to Balindiwe Sishi, a postdoc at Stellenbosch University (my alma mater), who had attended the 2018 meeting on physiology and medicine. As a polymer chemist by training, I was curious about whether there was a meeting on chemistry. Although an Internet search revealed that the forthcoming meeting would be on physics, I was not discouraged because the boundaries between scientific disciplines are often not clear-cut. As it turns out, a part of my PhD project was based on principles of physics, namely the electromagnetic spectrum and nuclear magnetic resonance. So I applied as an open applicant, went through the two-phase process and was, to my delight, selected to participate in the 69th Meeting. Fortunately, after selection, two sponsors were found for me by the Nobel Laureate Meetings Committee: ASSAf and the Wilhelm and Else Heraeus Foundation; I am very grateful to both.

On my way to Lindau, it suddenly occurred to me that I had not prepared questions for the laureates I wanted to meet. Was I being paranoid or justifiably nervous? Was I under prepared? It quickly became apparent that there was no need to over-prepare. All participants had a specialised programme including talks and activities, depending on their choices during the application process. The majority of the programmes consisted of three consecutive lectures by three laureates or a panel discussion with laureates and senior students working separately on similar themes (i.e. gravitational waves, dark matter and lasers). One such panel discussion was on dark matter, titled 'The Dark Side of the Universe' with David Gross (Nobel Prize in Physics 2004), Adam Riess (Nobel Prize in Physics 2011), George Smoot (Nobel Prize in Physics 2006) and Brian Schmidt (Nobel Prize in Physics 2011). As a non-expert, I went with no expectations; I wanted to learn something on a topic in which I would otherwise have had no interest. Dark matter is anything that does not respond to electromagnetic radiation, which means it cannot be detected. More than anything, I wanted to know why it matters if it is undetectable. It matters because the composition of the universe is 25% dark matter, which is 20% more than ordinary matter. A second part of the topic was on the acceleration of the earth, which is a theory first considered in 1917 by Albert Einstein who later went on to dismiss it as improbable. Einstein allegedly referred to this as his 'greatest blunder'. I left this talk with a newfound interest in astrophysics.

Young South African scientists Sinenhlanhla Sikhosana and Nonkululeko Radebe chat with Nobel laureate Brian Schmidt

Another engaging part of the day's activities were the agora talks. The format was a lecture followed by a 20-min question-and-answer session. This gave a great opportunity to engage with the laureates specifically on the work that won them their Nobel prize. Stefan Hell, from the Max Plank Institute for Biophysical Chemistry in Germany, presented one of the most exciting talks I attended. He received the Nobel Prize in Chemistry in 2014 together with Eric Betzig and William Moerner.

The title of his talk was 'Reaching Molecular Size Resolution in Lens-Based Microscopy: The Diffraction Limit Blown Away' which neatly summarised for what his joint prize was awarded. The talk started with him introducing an image of molecules observed under a confocal microscope during the 20th century. At that time, the resolution was limited to 200 nm. Using principles in physics, they overcame the diffraction barrier to obtain a spatial resolution of 20 nm – 10 times more than that previously possible. I was intrigued by his answer to the question of whether he knew that his research would be awarded a Nobel prize: yes. He explained that he knew that this work was groundbreaking and was in no way surprised when he received the call from Stockholm, where the Nobel Foundation is situated.

Apart from the many scientific talks and presentations, there was time to have more relaxed conversations with the laureates. I had the honour of talking to Vinton Cerf, who was a Turing Award winner in 2005 for his work on Internet protocols. Although a Turing award is not a Nobel Prize, it is considered one of the highest honours for a computer scientist. I caught him in the middle of a conversation with three other students taking about how multinational Internet service providers like Google can help in ensuring the integrity of research that is published on the Internet. I suspect this was triggered by the fact that we live in an era of 'fake news' and anyone can publish content as 'scientific' whether true or not. Another concern raised was on how impact factors of journals are calculated and that they do not reflect the impact of a specific paper. That means it is possible to have an 'okay' paper in a high impact factor journal and after a few decades no one would be the wiser that your paper was in fact not very relevant. We discussed how to measure longevity of research using something more than citations because we all know that it is possible to cite a paper you have not read. Some suggestions were tracking whether the scientific principles used in the cited paper were used in the work that cited it, beyond the section on literature. This was an extremely important, relevant and evolving conversation, to which I will refer for many years in the future.

The biggest obstacle of the week for me was to get over the feeling of being an imposter. I discovered this feeling was not unique to me after conversations with other students, especially those from South Africa. I felt that we needed a workshop devoted to how to own your space and believe in your science as valid and on par with the rest of the world. In retrospect, it was quite bold of me to apply to attend the Meeting because it suggests that I considered myself sufficiently capable to compete with students and postdocs around the world who are pure physicists. I suspect that I was naïve in thinking that the pool of applicants was small and I was surely not competing with the 'cream of the crop'. It turns out, however, that thousands of hopefuls applied and fewer than 600 participants were selected. I am honoured and humbled to have been given the opportunity to be among some of the most brilliant minds of our time.

AUTHOR:
Lawrence Malinga[1]

AFFILIATION:
[1]Secretariat: Southern and Eastern African Cotton Forum, Agricultural Research Council, Rustenburg, South Africa.

CORRESPONDENCE TO:
Lawrence Malinga

EMAIL:
lawrencem@arc.agric.za

HOW TO CITE:
Malinga L. Southern and Eastern African Cotton Forum: Platform for the advancement of cotton production in Africa. S Afr J Sci. 2019;115(9/10), Art. #6358, 3 pages. https://doi.org/10.17159/sajs.2019/6358

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
SEACF, ICAC, cotton, farmer training, yield, germplasm database

PUBLISHED:
26 September 2019

Southern and Eastern African Cotton Forum: Platform for the advancement of cotton production in Africa

Background

The Southern and Eastern African Cotton Forum (SEACF) was established in 1996 under the auspices of the International Cotton Advisory Committee (ICAC) that had supported three African regional meetings prior to the formation of SEACF: in Sudan (1982), Tanzania (1984) and Togo (1989). In 1994, researchers from the southern and eastern African countries set up the African Cotton Research Network, which later merged with SEACF. Initially, the main purpose of the forum was for each country to present national reports on cotton activities, but in 2010 the format was altered to a scientific symposium at which the regional network of researchers would present research results to their peers. The objective of SEACF is to promote cotton production through collaborative research and technology transfer and bi-annual research symposia are held in different member countries. Since its inception, SEACF has held 14 meetings. The chair of SEACF rotates with the country that hosts these symposia; member countries are Ethiopia, Kenya, Mozambique, Namibia, South Africa, Sudan, Tanzania, Uganda, Zambia and Zimbabwe.

Since 2006, the SEACF Secretariat has been based at the Agricultural Research Council – Industrial Crops in Rustenburg, South Africa. The Secretariat organises meetings and workshops and maintains a database of relevant scientists, currently numbering almost 100. Cotton industry players like Monsanto, Bayer and Cotton SA also attend meetings and make presentations, while after each meeting, visits to farmers' fields, research institutions or cotton ginneries are arranged. Membership is free and SEACF is self-funding; it receives some financial support from the ICAC. The SEACF Secretariat and the local organising committee of the hosting country raise additional funds for special events at meetings.

In 2001, SEACF initiated a collaborative project among South Africa, Zimbabwe, Tanzania, Uganda, Ethiopia and Sudan to determine the importance of cotton diseases in the region that was successfully completed the following year. This project was funded through ICAC and the fast-track programme of the Common Fund for Commodities. Regional germplasm exchange programmes have been established to allow cotton breeders to exchange cotton breeding lines from their respective countries.

The SEACF network has an electronic mailing list, hosted by ICAC, for communications among researchers.

Report on the 14th Meeting in Zimbabwe 2018

The most recent SEACF meeting was that in Harare, Zimbabwe, the theme of which was 'Global best practices for cotton yield enhancement in Africa'[1]. The meeting was hosted by the Zimbabwe Cotton Research Institute under the leadership of Dr Dumisani Kutywayo; 77 researchers from seven countries (Bangladesh, China, India, Kenya, Mozambique, South Africa and Zimbabwe) attended. Topics prioritised are discussed below.

Enhancing the competitiveness of southern and eastern African cotton

Of the total world cotton production, 10% is grown in sub-Saharan Africa and 18 million people in the region rely on cotton production either directly or indirectly. Input costs are supported by ginners and contract production and are later recovered through sales. The benefits are that farmers receive inputs and advisory support; ginners are ensured production and quality; and the export earnings of the country are improved through quality cotton to boost the textile industry and the cotton lint price. Mozambique cotton production is, however, unique in the region as it is entirely based on a concession set-up ('One Zone, One Gin Concept') that allows the cotton sector to operate in a monopsony system, where ginning companies are granted rights as exclusive buyers of cottonseed in their respective areas of the concession. Tanzania is currently investigating this system at a district level. There are benefits to the concession system, the most important of which are the exclusive right to purchase all cotton grown within the concession area; farmer support; extension training; and the supply of seed and crop chemicals.

In Zimbabwe, however, the Agricultural Marketing Authority regulates the production and marketing of cotton and administers matters like registration of growers, licensing of contractors, buyers and ginners; monitoring of seed cotton grading; and classification of lint. The Agricultural Marketing Authority also monitors the marketing and export of cotton and fosters the growth of the sector by providing a level playing field for all players in the cotton industry in Zimbabwe.

Increasing cotton productivity and public/private sector interventions

An important session focused on the ecological, technological and social environment of cotton production. In Africa, cotton is cultivated exclusively by smallholder farmers, and in Zimbabwe, it is grown by more than 200 000 smallholder farmers on an average plot of about 1 ha. The norm is that cotton is produced based on agreements signed by farmers and contractors who buy the cotton. Contractors supply the farmers with inputs as a loan, from which deductions are made when the product is delivered. In 2015, the Zimbabwean government approved a 3-year plan to support the smallholder farmers with free production inputs with the aim of reviving cotton production in the country. While the plan has resuscitated production, challenges remain: increasing costs of production; low cotton yields and poor-quality seed cotton production; input distribution based on generalised recommendations; side selling of free inputs from government; and the limited involvement of the private sector. Various recommendations were made to address these challenges.

Technology transfer

The experiences in India were highlighted in a paper on front line demonstrations. Field demonstrations, conducted by the National Agriculture Research System, the Indian Council of Agricultural Research (ICAR) and the State Agricultural Universities, provided effective learning. ICAR introduced the 'Lab to Land' programme that tested viable technologies developed by the researchers' on-farm field to convince farmers to adopt them. The programme further enabled scientists to obtain direct feedback from cotton farmers and create effective linkage among scientists, extension personnel and farmers. To date, 1157 demonstrations have been conducted and the analysis of yield parameter over 20 years has revealed an average yield increase of 18% compared to the farmers' previous practices, as well as a reasonable reduction in the cost of cultivation. Considering the profile of African cotton growers and the Indian experiences, conducting front line demonstrations is suggested for improving the socio-economic status of African cotton growers. Replicating the success of front line demonstrations would pave the way for profitable and sustainable cotton farming in the future.

An interesting study was done in Zimbabwe on value-added products that can be obtained from cotton stalk by-products. An average of 3 tonnes of the cotton stalk is generated per hectare of land; cotton stalk is an appropriate raw material for manufacturing bio-composite products. Due to the high deforestation rate, cotton stalks can be used as an alternative for the paper and pulp industry, as well as for providing crop residues to cultivate oyster mushrooms. This initiative is a valuable addition to the cotton-farming process and more research into natural resins with cotton stalk fibres will be done.

Despite water scarcity and soil degradation, most parts of Africa continue to use traditional and inefficient tools for cotton production. Over 9 years, the ICAR Central Institute for Cotton Research in India conducted a research study on conservation agriculture as best management practices for sustainable cotton production. The research revealed that the three key principles in conservation agriculture were minimum soil disturbance, increased soil cover and crop rotation. Reduced tillage systems with crop residue recycled resulted in increased boll retention, high seed cotton yields and profitability. The question was posed that if conservation agriculture is so beneficial, why is it not followed? After raising some of the challenges, it was clear that the solution lay in collaborative research, efficient technology transfer, appropriate mechanisation to improve farm labour efficiency, and the identification of crops that are compatible with cotton for intercropping.

Plant breeding

Due to low cotton yields of 500 kg/ha in Mozambique, the cotton research programme in that country has been developing and introducing new genotypes to discover suitable varieties for local production. As many as 18 cotton genotypes, including 16 imported, were tested in different localities around the country. From these genotypes, however, only three presented acceptable adaptability and potential stability, perhaps indicating that the seed cotton yield was affected more by the environmental complex than by the genotypes.

In Zimbabwe, 11 genotypes (8 experimental lines and 3 commercial varieties) were evaluated in order to identify superior genotypes for the Lowveld region conditions. Three genotypes had good yield and stability, and it was recommended that they be further tested for distinctiveness and uniformity based on field performance and fibre qualities before their release.

Eight Mahyco cotton hybrids were registered for commercial production in Zimbabwe and Malawi in 2017 and in Zambia the following year. Two of the tested hybrids yielded 5500 kg/ha dryland production, had an early maturity index of above 60%, high boll retention, and tolerance to jassids (leafhoppers). Seven other African countries are currently in the process of registering the hybrids, which have the potential to transform cotton production through improved farmer viability, increasing the area under cotton and cotton output.

The Institute of Cotton Research in China has done research on genome-wide quantitative trait locus (QTL) mapping for resistance to Verticillium wilt, fibre quality and yield traits in cotton chromosome segment substitution lines. In total, 251 QTLs have been detected, among them, 98 are of the fibre-quality traits, 93 of the yield-related traits, and 60 are Verticillium wilt resistant. Of these, 86 QTLs were consistent and three chromosomes contained more QTLs. Another study was presented from China which focused on cloning and expression of drought- and salt-tolerant genes on cotton. The study suggested that, on average, saline stress resulted in 70–80% loss of productivity in cotton, and that the Chinese cotton germplasm had 8873 cotton accessions. Few of the accessions that were tested were resistant to salinity as well as drought. The genetic diversity of cotton germplasm was also analysed among the salinity-tolerance relevant accessions, which showed that most of the germplasm had a closer genetic relationship.

Cotton agronomy

South African national cotton cultivar trials are conducted in different localities under irrigation and dryland conditions. These annual trials are aimed at evaluating cultivar performance that will be recommended to farmers in their respective areas. Characteristics that are taken into consideration include yield, fibre percentage, length, strength and micronaire. Even though a cultivar is identified as very stable, it does not necessarily mean that it will always give the highest yields, although it will perform better in unfavourable climatic conditions.

Cotton production is mainly dependent on the conventional tillage system in Zimbabwe which exposes the soil to degradation. Conservation agriculture is premised on the principles of reduced or no-soil disturbance. A study was conducted to determine the effects of conservation tillage technologies on seed cotton yield under Zimbabwean rainfed conditions. The lowest seed cotton yield was 511 kg/ha while the highest yield was 3000 kg/ha. The study recommended that the project continue and focus on low rainfall areas. It was also recommended that a cost–benefit analysis and crop rotation should be taken into account.

Crop protection

In Kenya, 80% of the population live in rural areas and depend on agriculture, and the cotton industry has been identified as one of the sub-sectors for alleviating poverty. However, cotton is characterised by low production per unit area. This low productivity is attributed to poor-quality seeds, poor land preparation, declining soil fertility, inadequate pest control, low adoption of technologies and inadequate technical support. A study was conducted in six districts to examine the transfer of integrated crop and pest management strategies based on an approach that meets the needs and circumstances of target farmers. The outcomes included increases in cotton yield, an improvement in cotton incomes, demonstration of the use of good production practices, an improvement in post-harvest handling, better communication among farmers, and collective action by farmers.

Verticillium wilt is one of the most important diseases of cotton worldwide, and affects yield and fibre quality. It is caused by *Verticillium dahliae*, a soil-borne fungus. There is no effective chemical control, so the use of tolerant cultivars is of great importance in controlling the disease. The objective of the research that was conducted in Zimbabwe was to determine the tolerance levels of new cotton genotypes to Verticillium wilt. Verticillium wilt screening indicated that the varieties used had different tolerance levels and most genotypes which were tolerant to the disease produced high yields. It was recommended that further research was required to determine the mode of tolerance.

Best practices for yield enhancement in Africa

Yield enhancement in Africa is an extremely important topic. It was agreed that to a greater extent cotton research should address the cotton value chain and the development of technologies that are affordable, viable, sustainable and easy to use. Moreover, researchers need to ensure that those technologies reach the end-user. In addition, the technology, knowledge and information generated by research must reach the grower in a practical manner. To enhance production yield, effective weed and pest management is vitally important, and timely planting in rainfed areas is

critical to ensure that the crop takes full advantage of rainfall. As smallholder farmers form a larger percentage in cotton farming, there is a need for specially designed mechanisation that caters for smallholdings.

There are 24 countries that have adopted biotech crops, of which the USA constitutes 40% of the global share. Cotton is planted in 15 of those countries – covering 81% of global planting. The major traits in biotech crops are those for insect and herbicide tolerance. By 2015, the adoption rate of GM cotton was almost 80% worldwide. The highest year-on-year increment in the biotech cotton area was obtained in South Africa (a remarkable 315%), followed by the USA (24%) and Brazil (19%).[2] Although technology has increased the yield and reduced the production costs, challenges are the minor pests that have become major pests and the development of insect resistance. Some solutions towards challenges in increasing yield include reduction of plant density per hectare, lower crop duration, short critical window and efficient management of insect pests, nutrients, water and light. The training workshop was led by Dr Keshav Kranthi, head of the Technical Information Section at the ICAC, and he also conducted participatory practical sessions on Bt-detection using immuno-chromatographic strips and Bt-quantification through enzyme-linked immunosorbent assays.[3]

Recommendations

The following are some of the key recommendations from the 2018 meeting in Zimbabwe:

- Improve research and development in the region
- Introduce farmer training programmes to improve production
- Manage high production costs due to the increase in input cost and labour wages
- Mitigate the low viability of cotton production due to international prices and its volatility
- Consider the role of climate change on low cotton yields through drought and the recurrence of diseases and pests
- Improve collaboration among the member countries and information exchange
- Develop a regional germplasm database and exchange programme

Uganda has been proposed as the host country for the 15th Meeting of the SEACF in 2020.

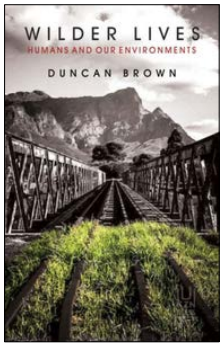
Acknowledgements

SEACF meetings are undertaken under the auspices of the ICAC, and the Secretariat expresses heartfelt thanks to Mr Kai Hughes (Executive Director, ICAC) and Dr Keshav Kranthi (Head Technical Section, ICAC) for all their efforts and funding to ensure the success of the Harare meeting. The Secretariat is further appreciative to Dr Graham Thompson who served as the SEACF coordinator for more than 10 years. Many thanks go to the Zimbabwean Local Organising Committee under the leadership of Dr Dumisani Kutywayo (Director, Crops Research Division – Department of Research and Specialist Services) and the organiser Mr Washington Mubvekeri. The Secretariat is also grateful to all the presenters at the 14th Meeting in Zimbabwe, to the Zimbabwean government for permission to hold the event in Harare, to the South African Agricultural Research Council for being the Secretariat home, and the sponsors who enabled the meeting to run smoothly and funded some of the activities.

References

1. Business writer. Zimbabwe to host African Cotton Forum. Business Times (Zimbabwe). 2018 June 22;Agriculture/Economy. Available from: https://businesstimes.co.zw/zimbabwe-to-host-african-cotton-forum/
2. International Service for the Acquisition of Agri-biotech Applications (ISAAA). Global status of commercialized biotech/GM crops in 2017: Biotech crop adoption surges as economic benefits accumulate in 22 years. ISAAA brief no. 53 [document on the Internet]. c2017 [cited 2019 May 29]. Available from: https://www/isaaa.org/resources/publications/briefs/53/.../isaaa-brief-53-29017.pdf
3. Kranthi KR, Malinga L, Mubvekeri W. Perspectives on cotton research and ideas for Africa – Proceedings & recommendations of the XIV Meeting of the Southern & Eastern Africa Cotton Forum (SEACF). The ICAC Recorder. 2018;36(3):4–13.

BOOK TITLE:
Wilder lives: Humans and our environments



AUTHOR:
Duncan Brown

ISBN:
9781869144210 (softcover)

PUBLISHER:
UKZN Press, Pietermaritzburg; ZAR285

PUBLISHED:
2019

REVIEWER:
Adam Cruise

AFFILIATION:
Department of Philosophy, Stellenbosch University, Stellenbosch, South Africa

EMAIL:
cruiseadam1970@gmail.com

HOW TO CITE:
Cruise A. Let nature decide. S Afr J Sci. 2019;115(9/10), Art. #6442, 1 page. https://doi.org/10.17159/sajs.2019/6442

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

PUBLISHED:
26 September 2019

Let nature decide

Wilder Lives uses ideas of 'wildness' and 'rewilding' to rethink the human relationship with our environment. As a species we have single-handedly destroyed our planet's ecosystems in the short space of a few hundred years. Now we urgently need to reconsider and redefine our behaviour. This excellent and timely book explores ways to do so.

Brown's approach is wide-ranging, inquiring and thought-provoking. Brown, Professor of English at the University of the Western Cape, expertly takes us through such topics as wildness and conservation, wild cities, rewilding language, wildness and food, wild animals, wild margins, and wildness in the ethics of human–animal relations. He uses dozens of his own experiences throughout the book in order to illustrate an alternative approach to how we might view the natural environment.

At the heart of the book is the notion that humans, try as we may to separate ourselves from wild nature, remain integrated and inseparable from it.

Ecological rewilding

Brown introduces the notion of 'rewilding' as developed by environmental writer George Monbiot, author of *Feral: Rewilding the Land, the Sea, and Human Life* (University of Chicago Press, 2017). From an ecological perspective, rewilding is not about keeping ecosystems in a state of arrested development, like preserving a jar of pickles; or forcibly restoring environments that have already been degraded into a human-crafted area. Instead, rewilding is an approach that permits ecological processes to resume by themselves, like allowing trees to naturally repopulate deforested land; or mangroves, salt marshes and seagrass beds to grow back after being destroyed.

Rewilding is a concept that fundamentally recognises that the ecosystem restores itself. It is not simply adding a collection of animals and plants to an area, as is often the case with private reserves and game ranches in South Africa, but rather of allowing their shifting relationships with each other and their surrounding environment to re-develop. In other words, an ecosystem is not a finished static product but a growing, continuous, shifting process that must be left to evolve and adapt to changes. In essence, rewilding is about resisting the urge to control nature and allowing it to find its own way. At times, it may involve some assistance, like reintroducing absent plants and animals (and in a few cases culling exotic species which cannot be contained by native wildlife), pulling down fences, blocking drainage ditches, but otherwise stepping back. At sea, it means excluding commercial fishing and other forms of exploitation.

The ecosystems that result, or rewildings, are best described not as a wilderness but as self-willed – governed not by human management but by their own processes. Rewilding has no end points, no view about what a 'right' ecosystem or 'right' assemblage of species looks like. It does not strive to produce a heath, a meadow, a rainforest, a kelp garden, or a coral reef. It lets nature decide.

There are hundreds of such cases where similar aspects of rewilding have occurred. A particularly interesting, if not macabre, example Brown uses is the 4700 km^2 Chernobyl disaster zone between the Ukraine and Belarus. Here rare and endangered animals have thrived since the area was evacuated by humans in 1986.

Human rewilding

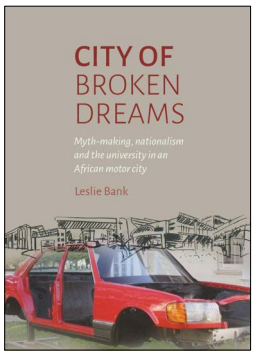
Rewilding does not exclude humans. Neither is it about conflict between humans and an ecosystem. Brown is at pains throughout the book to argue that modern humans can still participate in a meaningful way with the natural environment while enjoying the benefits of technology and civilisation. Humans need to relearn how to connect with the natural environment – the source of vitality – by engaging in activities that bring them into bodily contact with it. This might include a walk along a trail, a swim in a lake or ocean, a ski down a snowy slope, a climb up a mountain, a safari, trout fishing in the mountains or a camping trip.

Activities like these bring peace of mind, a psychological escape from the humdrum of the built environment with its stress, pollution and noise. How many of us come back from such experiences in nature feeling revitalised? It is no coincidence that almost all major cities have set aside bits of nature for human well-being. One only has to think of Central Park in New York, Hyde Park in London and, best of all, Table Mountain National Park in Cape Town. The latter is a perfect example of how human civilisation remains interconnected to the wild environment. Almost half of the city remains natural, thanks to the wisdom of the early town planners to not build above a certain altitude. Today, thousands of people access this natural environment, either hiking up and along the hundreds of trails, or, for the more sedate, by road or cable way. All do so to benefit from the spectacular views and to breathe in the fresh fynbos-scented air.

Yet, preserving the natural environment goes so much further than psychological, spiritual and physical well-being. The biosphere is essential for the survival of all life. It is imperative we set enough space aside for biodiversity, and us, to thrive. Failing to do so will not only doom natural life, but all life including our own. Given that much land has to remain under the plough or grazed by livestock to feed the 7-billion plus humans, the land available for rewilding is quite modest, but Brown believes it is sufficient to constitute an important ecological and existential resource.

Ultimately, Brown does not try to offer a definitive account or exhaustive set of arguments. Instead, this book is suggestive, and cumulatively sketches the possibility of 'wilder lives'. 'Rewilding – including of one's own life,' writes Brown, 'is a lifelong commitment.'

BOOK TITLE:
City of broken dreams: Myth-making, nationalism and the university in an African motor city



AUTHOR:
Leslie J. Bank

ISBN:
9780796924544 (softcover)

PUBLISHER:
HSRC Press, Cape Town; ZAR302

PUBLISHED:
2019

REVIEWER:
Alan Mabin

AFFILIATION:
School of Architecture and Planning, University of the Witwatersrand, Johannesburg, South Africa

EMAIL:
alan@alanmabin.org

HOW TO CITE:
Mabin A. The university and South Africa's 'Motor City'. S Afr J Sci. 2019;115(9/10), Art. #6473, 1 page. https://doi.org/10.17159/sajs.2019/6473

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

PUBLISHED:
26 September 2019

The university and South Africa's 'Motor City'

Recent research and writing on economy, society and politics in South Africa's 'secondary' cities are scarce relative to the literature on Johannesburg and Cape Town, and to a slightly lesser extent Durban (Ethekwini). The two other sizeable harbour cities, Port Elizabeth (Nelson Mandela Bay) and East London (Buffalo City), have been the subjects of a few PhD theses and occasionally books, but remain little explored. They therefore offer rich fields for scholarship.

East London – *iMonti* and its hinterland, including areas variously named 'Border', 'Ciskei', 'Transkei' – gained the attention of an earlier generation of scholars. Leslie Bank, who lived and worked for some decades in the region (he was Director of the Fort Hare Institute for Social and Economic Research) has, in previous books – *Home Spaces, Street Styles*[1] and *iMonti Modern*[2], taken on some of the older directions. However, in *City of Broken Dreams*, his canvas expands to include the long-term histories of East London through colonialism, segregation, apartheid and democracy. Informed by archival, participant, media, and other sources, Bank sweeps through shifting ideologies, industrial development and decline, the patchy record of recent government, and deep social changes. His account portrays how both prosperity and broken dreams coexist today, offering fresh ideas on how to move beyond the present contrast between the global reach of an apparently successful car exporting economy and the deteriorating conditions that face the majority of citizens.

The text is organised in four parts, preceded by a powerful introduction that sets out the main challenges of the city and the questions of the book, and concludes with ideas on 'remapping the city'. Among the major themes, a key thread concerns the history of different regimes of power in East London, through segregation and the earlier growth machine, apartheid and its capture of many kinds of change in the Eastern Cape, and then democracy with its disappointments but also its triumphs. The book deals with the rise and decline of manufacturing, linking those developments with the shifting regimes. Moreover, and importantly, Bank argues a case for a new role for the university in making something better of this place of 'dreams'. The book concludes with a deep account of university power politics as well as student revolt, right up to the present, presenting readers with an argument as to how the concentration and enhancement of the university presence might lead to greater hope.

Bank also accounts for the significance of the car in East London life: from the appearance of the first vehicles, through local initiative in establishing car assembly plants many decades ago, to the enormous but difficult significance of a major German company continuing to produce for a global market. In this theme there are parallels between the experience of East London and Detroit. East London has been and remains a particularly motor oriented city – home of the South African Grand Prix for a generation, and the present centrality of cars to the weekend celebration that takes place along the beach and that could hardly have been imagined 30 years ago.

There are several new ways in which the city is occupied and used by its present population. Bank refers to 'occupy urbanism' in this context, touching on the double rootedness between the 'rural' and the city, with particular attention to the huge weekend gatherings known as *Ebuhlanti* on the beach front. There are certainly struggles on the part of newer, black African middle- and upper-class individuals to control existing and persistent elements of the city, including modes of production and distribution – yet an apparent lack of interest or recognition of possibilities towards creating something really new.

Some readers may find it surprising that relatively little attention is given to significant moments in the make-up of the city and its region. For example, there are numerous former 'homeland' civil servants and politicians who populate the bureaucracies of East London; the famous strikes at Wilson Rowntree and at Mercedes Benz have their place, but not the 1983 bus boycott. And Mdantsane, the major township of the region, central to the bus boycott and strikes and home to a majority of East London's people, is scarcely mentioned. This omission suggests that there are opportunities for further work.

However, the ultimate purpose of the book is to argue for a transformed role for universities in city development, a theme that Bank has previously explored,[3] and in particular for finding new pathways for building upon the combined presence of the University of Fort Hare, Walter Sisulu University and Unisa in East London. Through his review of student circumstances and protest, Bank points towards suggestions on how these major elements of urban society might more effectively contribute to redeveloping *iMonti* as a prosperous city that could offer a better life to a larger population than it does at present. In sum, Bank calls for a 'greater appreciation of the globalisation of higher education and its connection to city building and urban opportunity [which] might offer hope and insight for the future development of this and other struggling "rust belt" cities in South Africa'.

City of Broken Dreams is highly readable, based on very substantial knowledge and research, and original in its portrayal of a South African urban situation, but it may also resonate with many other struggling cities in the world. It is an important contribution and should be widely read.

References

1. Bank L. Home spaces, street styles: Contesting power and identity in a South African city. London: Pluto Press; 2011.
2. Bank L, Qebeyi MK. Imonti modern: Picturing the life and times of a South African location. Cape Town: HSRC Press; 2017.
3. Bank L, Cloete N, Van Schalkwyk F, editors. Anchored in place: Rethinking universities and development in South Africa. Cape Town: African Minds; 2018.

AUTHOR:
Jules Skotnes-Brown[1]

AFFILIATION:
[1]Department of History and Philosophy of Science, Christ's College, University of Cambridge, Cambridge, United Kingdom

CORRESPONDENCE TO:
Jules Skotnes-Brown

EMAIL:
jasb2@cam.ac.uk

HOW TO CITE:
Skotes-Brown J. Recording the past, predicting the future: The periodic table 150 years on. S Afr J Sci. 2019;115(9/10), Art. #6835, 2 pages. https://doi.org/10.17159/sajs.2019/6835

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
chemistry, elements, atomic weight, Mendeleev

PUBLISHED:
26 September 2019

Recording the past, predicting the future: The periodic table 150 years on

The International Year of the Periodic Table of Chemical Elements in 2019 marks the 150th anniversary of the publication of Dmitri Mendeleev's periodic table. The periodic table – a classificatory tool, educational diagram and representation of elements – has since become one of the most ubiquitous and successful productions of European science, and is to chemistry as Darwin's theory of natural selection is to biology.[1] The periodic table is not only found in every chemistry laboratory and high school science classroom, but has also inspired a plethora of literary and cinematic interpretations. Acclaimed works of literature such as holocaust survivor, writer and chemist Primo Levy's *The Periodic Table,* as well as Oliver Sacks's much acclaimed *Uncle Tungsten* were inspired by it. Elements bromine and barium form the opening credits to the TV series *Breaking Bad*. More recently, the periodic table has even garnered its own song in a viral YouTube video.[2]

Aside from being an educational, organisational and predictive tool for aspiring and professional chemists, the periodic table is also a highly mediated record of the history of chemistry: the elements are named to commemorate its long genesis. Reflecting its indebtedness to ancient Greek alchemists, helium takes its name from Helios, god of the sun. Promethium is named after Prometheus, who stole heavenly fire and gave it to human beings – a substance without which the discipline of chemistry would never have existed. Nihonium was named in honour of Japan, the nationality of the team of scientists who first observed it. Americium, francium, darmstadtium and californium among others, likewise, take their names from places in which they were discovered, or that their discoverers wished to commemorate. Some elements have been named after individuals who made major contributions to chemistry, such as Glenn Seaborg, Albert Einstein, Nils Bohr, Ernest Rutherford, Alfred Nobel and, of course, Mendeleev.

Behind this neat version of the periodic table's history, lies centuries of discoveries, errors, controversies, successes and failures.[3] Within its own historical period, each iteration of the periodic table offers a record of scientific consensus, which has continuously changed as new findings have been made, errors been corrected and elements have been renamed or removed. To name one example, in the late 18th century, chemist Antoine Lavosier regarded light and heat to be elements.[1] The major player in this history of the modern periodic table is, of course, Dmitri Mendeleev, who organised the known elements by atomic weight and chemical affinities and is often credited as the 'discoverer' of the periodic system.

The word 'discovery', however, masques a long history of laboratory work, publication and debates, and myriad attempts at representing the elements. Mendeleev's periodic system was a complex organisational scheme that synthesised centuries of work in chemistry into a highly successful organisational, classificatory and predictive framework. It was one in a long lineage of attempts to 'make sense of the way in which particular elements enter into chemical bonding'[1]. The earliest attempts to classify the elements can be traced as far back as ancient Greece, in which the term 'element' denoted 'a "tendency" or "potentiality" that gave rise to the observable properties of the element', a definition which drew a distinction between its abstract and observable form.[1] In the 18th century, this abstract definition gave way to an empirical understanding, advocated by Robert Boyle and Antoine Lavoisier, who argued that an element should be defined as 'a material substance that has yet to be broken down into any more fundamental components by chemical means'[1].

Attempts to place the elements into tables are more recent. Since the 18th century, as Benjamin Cohen has argued, chemists have utilised tables as organisational tools for 'collecting the known and practical guides for directing work toward the unknown'[4]. In this period, they were indispensable tools in the laboratory, which mostly acted as guides that assisted chemists by making predictions based on known chemical affinities.[4] Perhaps the most famous of these early systems, is Antoine Lavosier's 1789 list of elements, and atomic theorist John Dalton's 1808 table of relative atomic weights.[4]

Learning to read these earlier tables often meant learning a new visual language.[5] Before 1813, many systems for organising the elements placed arcane alchemical symbols in relation to one another, but these were ultimately discarded due to the problems they posed for publication and communicability.[1,4] Right up to (and after) the publication of Mendeleev's system, the table itself as a diagram was not the only form of representation. Gmelin's 'remarkable system', published 26 years before Mendeleev's, for example, was also ordered according to atomic weight, but took a V shape.[1]

The 1860s: Development of the modern periodic system

The critical decade for the development of the modern periodic table was the 1860s. In this decade, following the 'rationalization of atomic weights' at a seminal international congress of chemists in Karlsruhe, there were at least six 'discoveries' of the periodic system.[1] At this conference, more 'precise definitions of the concepts of atom, molecule, equivalent, atomicity, alkalinity' were discussed.[6] The six discoveries published in the years following Karlsruhe, were not revolutionary, but dependent upon the scaffolding provided by earlier chemists.[1] While some took the form of tables, others opted for entirely different visual strategies. The first was published by French geologist Alexandre-Émile Béguyer de Chancourtois in 1862. His model was three dimensional: a cylinder with elements arranged in descending vertical groups.[7] Unfortunately for him, printing a three-dimensional model on a sheet of paper proved challenging, which made it relatively incommunicable to large audiences.[1] Next, came William Odling and Julius Lothar Meyer, who both published their own tables in 1864. Perhaps most interestingly, in 1866, John Newlands utilised the model of a musical scale to organise the elements in a 'law of octaves'. Like the rhyming sound produced by each eighth note on a sequence, so too did every eighth element seem to correspond.[8] After Newlands, came Gustave Hinrichs's system, which depicted relationships between elements within a spiral.[1,9]

Mendeleev's system: Publication and reception

Mendeleev's first periodic table was published in 1869. As is well known, Mendeleev organised elements according to atomic weight, and proposed that they exhibited 'an apparent *periodicity* of properties' when arranged in such a way.[1] An important and unique feature of his table were the gaps it left, which could be filled with future discoveries. As a recent article in the *Economist* put it, if 19th-century chemists could be considered as stamp collectors, Mendeleev provided 'an album in which to stick their discoveries'.[8]

As historians of science have long argued, scientific discoveries are never automatically accepted as 'breakthroughs' and subsequently diffused amongst the scientific community: their authority as 'truth' or 'facts' has to be constructed in a crucible of social, political, economic and epistemological forces.[10,11] Any theory also has to 'prove its value as a guide to ongoing research, and only after it does will it be incorporated into textbooks as part of the established knowledge of a science'[12]. Mendeleev's periodic table was no different. It took time for his system to gather attention, even in his native Russia. In Autumn 1869, for example, the president of the Russian Chemical Society, N.N. Zinin advised him 'to do "[real] work," meaning do something experimental, preferably on organic chemistry, which was the mainstream research discipline at that time'[12].

Despite this critique, over the next 20 years, Mendeleev's was to become the dominant periodic system. Key to its success was his treatment of time, and capacity to speculate.[6] Like his contemporaries, he drew upon past discoveries in order to organise his table, but unlike many others, he also looked towards the future, leaving spaces in which new elements could be added.[12] In November and December of 1870, he published two papers on 'the Natural System of Elements', which successfully predicted the 'properties of some undiscovered elements in detail'. This attracted significant attention within the local and international communities of chemists.[1] Mendeleev was fortunate to have access to the support of and inspiration provided by the newly formed Russian Chemical Society, which encouraged his continued research on a periodic system, and stimulated him to write textbook-length syntheses of the field, including *The Principles of Chemistry*.[13] It also provided resources to propagate his ideas outside of Russia. The presence of German speakers within this society such as Viktor von Richter, a correspondent of the German Chemical Society, facilitated the circulation and translation of his research within Germany in the 1870s.[13]

1875 proved a watershed year for Mendeleev's table, with the successful prediction of the chemical properties of the newly discovered element gallium. News of his system spread fast, and by the late 1880s, the majority of British and US chemistry textbooks were discussing his periodic law.[12] In 1891, when his periodic system was in its fifth edition in Russia, it was translated into English for the first time.[1] Finally, the discovery of noble gases in the 1890s to 1900s, which initially appeared to contradict his system, ultimately cemented its success: he was able to fit these into the final group in his table.[1,14]

Mendeleev's success was not necessarily a result of him discovering a yet unknown feature of nature. On one level, he created a useful predictive, classificatory and representational system, which provided a record of the past, and a depiction of the present, which also looked towards the future. On another level, he was fortunately positioned within a network of Russian and German chemists, who facilitated the circulation and translation of his work. Ultimately, his system rapidly became the standard depiction of the periodic system, which has constantly been added to and adapted ever since. Since 1869, a cascade of at least 700 representations of the periodic system have been produced, and the flow shows no sign of stopping as new elements are discovered.[15] One recent article published in *Nature* even suggested that it should be turned upside down! Inverting the periodic table, the authors argue, would not only give it 'legs', but offer psychological and pedagogical advantages. In high school classrooms, the bottom area of the table is least often referred to by teachers, but the easiest for desk-bound teenagers to see. Turning it upside down would also apparently be 'consistent with psychological evidence that people associate greater magnitudes (numbers) with higher vertical positions'[16]. With or without 'legs', rotated or not, the periodic table has grounded itself as one of the most influential representations in the history of science, and it seems its stability shows no signs of faltering.

References

1. Scerri E. The periodic table: Its story and significance. Oxford: Oxford University Press; 2007.
2. AsapSCIENCE. The periodic table song [video on the Internet]. c2015 [cited 2019 Aug 05]. Available from: https://www.youtube.com/watch?v=VgVQKCcfwnU
3. Fontani M, Costa M, Orna M. The lost elements: The periodic table's shadow side. Oxford: Oxford University Press; 2014.
4. Cohen B. The element of the table: Visual discourse and the preperiodic representation of chemical classification. Configurations. 2004;12(1):41–75. https://doi.org/10.1353/con.2005.0001
5. Rudwick MJS. The emergence of a visual language for geological science 1760–1840. Hist Sci. 1976;14(3):149–195. https://doi.org/10.1177/007327537601400301
6. Niaz M, Rodríguez MA, Brito A. An appraisal of Mendeleev's contribution to the development of the periodic table. Stud Hist Philos Sci A. 2004;35(2):271–282. https://doi.org/10.1016/j.shpsa.2003.12.014
7. Scerri ER. The evolution of the periodic system. Sci Am. 1998;279(3):78–83.
8. The periodic table is 150 years old this week. The Economist. 2019 February 28. Available from: https://www.economist.com/science-and-technology/2019/02/28/the-periodic-table-is-150-years-old-this-week
9. Stewart PJ. A century on from Dmitrii Mendeleev: Tables and spirals, noble gases and Nobel prizes. Found Chem. 2007;9(3):235–245. https://doi.org/10.1007/s10698-007-9038-x
10. Schaffer S, Shapin S. Leviathan and the air pump. Princeton, NJ: Princeton University Press; 1985.
11. Shapin S. A social history of truth: Civility and science in seventeenth-century England. Chicago, IL: University of Chicago Press; 1994.
12. Brush S. The reception of Mendeleev's periodic law in America and Britain. Isis. 1996;87(4):595–628. https://doi.org/10.1086/357649
13. Kaji M. Mendeleev's discovery of the periodic law: The origin and the reception. Found Chem. 2003;5(3):189. https://doi.org/10.1023/a:1025673206850
14. Periodic table [webpage on the Internet]. No date [cited 2019 Aug 05]. Available from: http://www.rsc.org/periodic-table/history/about
15. Mazurs E. Graphic representations of the periodic system during one hundred years. Tuscaloosa, AL: University of Alabama Press; 1974.
16. Poliakoff M, Makin ADJ, Tang SLY, Poliakoff E. Turning the periodic table upside down. Nat Chem. 2019;11(5):391–393. https://doi.org/10.1038/s41557-019-0253-6

AUTHORS:
Marius Claassen[1]
Gonçalo Zagalo-Pereira[2]
Ana Sofia Soares-Cordeiro[2]
Nikki Funke[3]
Karen Nortje[3]

AFFILIATIONS:
[1]Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa
[2]Foundation for Science and Technology, Lisbon, Portugal
[3]Natural Resources and the Environment, Council for Scientific and Industrial Research, Pretoria, South Africa

CORRESPONDENCE TO:
Marius Claassen

EMAIL:
mclaasse86@gmail.com

HOW TO CITE:
Claassen M, Zagalo-Pereira G, Soares-Cordeiro AS, Funke N, Nortje K. Research and innovation cooperation in the South Atlantic Ocean. S Afr J Sci. 2019;115(9/10), Art. #6114, 2 pages. https://doi.org/10.17159/sajs.2019/6114

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
SDG 17, sustainable economic growth, strategic partnerships, marine resources

PUBLISHED:
26 September 2019

Research and innovation cooperation in the South Atlantic Ocean

The ocean provides a diversity of services, which range from food, minerals, energy and transportation to biodiversity and cultural services such as tourism and heritage. Achieving a balance between harnessing these services to support social and economic development and protecting the resource base remains a challenge. For example, the Food and Agriculture Organization of the United Nations reports that the percentage of stocks fished at biologically unsustainable levels increased from 10% in 1974 to 33% in 2015.[1] According to the Ocean Health Index, natural products, coastal protection and carbon storage at a global level all saw steady declines between 2012 and 2016.[2]

The South Atlantic Ocean is described by the International Hydrographic Organization as stretching from the equator in the north to the Antarctic continent in the south, and is broadly bound by South America to the west and Africa to the east.[3] According to the Ocean Health Index, South Africa, bordering the South Atlantic in the east, is ranked 45th (out of 221) on the overall score, whereas Brazil and Trindade on the other side of the South Atlantic, are ranked 126th.[2] South Africa's highest ranking (21st) is on the 'Carbon storage' metric and its lowest ranking (144th) is on the 'Clean water' metric, whereas Brazil and Trindade's highest rankings (26th) are on the 'Sense of place' metric and their lowest rankings (162nd) are on the 'Livelihoods' metric.[2] With the UN General Assembly Resolution 71/312 calling for, among others, action to 'Develop comprehensive strategies to raise awareness of the natural and cultural significance of the ocean…'[4], the above metrics provide guidance on areas of strength on which to build and areas in which to improve.

World Wildlife Fund South Africa reports that 14% of South African linefish species are overexploited and that the stocks of 52% of linefish species have collapsed. This proportion is significant, as small-scale fisheries in South Africa support the livelihood of more than 28 000 households directly, whereas indirect employment is provided by the commercial fisheries to more than 100 000 employees.[5] In Brazil, catches peaked at just more than 950 000 tonnes in 1985, whereas sharp declines on the abundance of the main target stocks caused catches to drop to around 600 000 tonnes/annum between 1990 and 1999, before recovering to just more than 700 000 tonnes in 2016.[6,7]

Sustainable management of ocean resources is a global issue that requires cooperation. The strongest recognition of this need is the inclusion of a Sustainable Development Goal on the need to strengthen the means of implementation and revitalise partnerships (SDG 17) in the UN 2030 Agenda for Sustainable Development.[8] Research and innovation are key areas where partnerships can play a crucial role to sustainably manage resources to support the aspirations of the present generation and maintain the development potential for future generations.

Regional and national efforts to promote African Atlantic cooperation

The 'Convention for Cooperation in the Protection, Management and Development of the Marine and Coastal Environment of the Atlantic Coast of the West, Central and Southern Africa Region' known as the Abidjan Convention, was entered into force on 5 August 1984 and has been ratified by Benin, Cameroon, Congo, Côte d'Ivoire, Democratic Republic of Congo, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Mauritania, Nigeria, Senegal, Sierra Leone, South Africa and Togo. The Convention addresses the need for cooperation to ensure sustainable and environmentally sound development through a coordinated and comprehensive approach, and the need for a carefully planned research, monitoring and assessment programme.[9]

The Benguela Current Large Marine Ecosystem (BCLME) spans the exclusive economic zones of Angola, Namibia and South Africa and beyond. The Benguela Current Commission was established in March 2013 by the Republics of Angola, Namibia and South Africa to coordinate the long-term conservation, protection, rehabilitation, enhancement and sustainable use of the BCLME, in order to provide economic, environmental and social benefits. The concepts of cross-sectoral marine spatial planning, environmental stress reduction and climate change have emerged strongly over recent years.[10]

In South Africa, Operation Phakisa was launched in 2014 as a results-driven approach, involving clear plans and targets, ongoing monitoring of progress and communication of results to the public. The Marine Transport and Manufacturing work stream builds on South Africa's strategic location, infrastructure and skills base to accelerate growth of the sector.[11] By 2016, the initiative had unlocked investments of about ZAR17 billion (USD1.2 billion) in the oceans economy and created over 4500 jobs in the various sectors.[12]

The priorities and actions in the 'Research Innovation and Knowledge Management Road Map for the South African Maritime Sector' take a broader perspective, with the objectives including the establishment of a maritime culture, an enabling governance framework, structured financing of initiatives, and a system of coordination, collaborating and knowledge sharing. Objectives related to ocean resources include sustainable utilisation and protection of natural resources and promoting safety, security and military protection.[13]

Trans-Atlantic cooperation

Brazil's Minister of Foreign Affairs (Antonio Patriota) highlighted the commitment of African and South American countries to a South Atlantic identity at the Seventh Ministerial Meeting of the Zone of Peace and Cooperation of the South Atlantic in 2013 and added that the South Atlantic is a bridge between brother continents.[14] Convergent aspirations from across the Southern Atlantic paved the way for the 'South-South Framework for Scientific and Technical Cooperation in the South and Tropical Atlantic and Southern Oceans'.[15] This high-level scientific research agenda emphasises the need for research in three themes: climate variability and change; ecosystems variability

and controlling processes; and living and non-living resources and biodiversity. The framework sets out to promote scientific cooperation and capacity building among South Atlantic countries, through the exchange of expertise and knowledge, in order to bring about environmental and socio-economic benefits for these countries. The principles of the framework are entrenched in the 'Belém Statement on Atlantic Research and Innovation Cooperation', which was signed by South African and Brazilian Ministers (Naledi Pandor and Gilberto Kassab) and the EU Commissioner (Carlos Moedas) in Lisbon on 13 July 2017.[16] The agreement promotes sustainable cooperation on marine science, research and innovation by linking research activities with those in the North Atlantic, exploring synergies with other initiatives, and optimising the use and sharing of research infrastructure, data and platforms. It further sets out to develop a common understanding and deepening of scientific knowledge of marine ecosystems and the interrelations between oceans and climate change, oceans and food, and oceans and energy systems. The envisaged outcomes are: better monitoring and forecasting capacities; improved safety at sea, human health and well-being; sustainable use of marine resources; new and emerging technologies to service societal needs and new value chains; and ocean-engaged citizens through enhanced ocean literacy activities.

Implementation of the Belém Statement is facilitated through the newly established Cooperation and Support Action, entitled 'All AtlaNtic Co-operation for Ocean Research and innovation' (AANChOR; www.AllAtlanticOcean.org), under the EU Horizon 2020 programme. The South African Department of Science and Innovation, National Research Foundation and Council for Scientific and Industrial Research are among the 17 project partners from Portugal, Belgium, Spain, France, Germany, South Africa, Brazil, Argentina and Cape Verde. AANChOR's mission is to bring together all relevant research and innovation actors around the Atlantic Ocean to build a long-lasting Atlantic Research and Innovation Ocean Community to address grand challenges and opportunities. AANChOR works closely with policymakers to inform them on long-lasting, high-potential, new actions to further enhance the sustainable economic growth and well-being of the Atlantic society. These long-lasting and impactful initiatives will be discussed and identified by the Atlantic Research and Innovation Ocean Community created within the scope of the project and will be based on the numerous existing initiatives on research and innovation in the Atlantic Ocean, including the Atlantic Ocean Research Alliance (AORA project; www.atlanticresource.org). The most exciting ones will be appraised by the policymakers and some will be supported through the Support Action. For their continued growth, AANChOR will also help in finding the best tools and funding opportunities. The initiative will run for 4 years (until September 2022), during which engagements with partners and stakeholders will provide key inputs to the process. Among the engagement opportunities will be 10 multi-stakeholder platform workshops that will take place across the different regions.

By promoting the implementation of the Belém Statement, it is envisaged that the AANChOR project will upscale research and innovation cooperation within the Atlantic basin, from Antarctica to the Arctic.

The AANChOR Coordination and Support Action received funding from the EU's Horizon 2020 research and innovation programme under grant agreement no. 818395.

References

1. Food and Agriculture Organization of the United Nations (FAO). The state of world fisheries and aquaculture 2018 – Meeting the sustainable development goals. Rome: FAO; 2018.
2. Ocean Health Index Global Scores: South Africa [webpage on the Internet]. No date [updated 2016; cited 2018 Oct 26]. Available from: http://www.oceanhealthindex.org/region-scores/scores/south-africa
3. International Hydrographic Organization (IHO). Limits of oceans and seas. 3rd ed. Special publication no. 28. Monte Carlo: IHO; 1953. Available from: https://epic.awi.de/id/eprint/29772/1/IHO1953a.pdf
4. UN General Assembly. Our ocean, our future: Call for action: Resolution adopted by the General Assembly. Resolution 71/312 of 6 July 2017. New York: UN; 2017. Available from: https://digitallibrary.un.org/record/1291421
5. World Wildlife Fund South Africa (WWF-SA). Oceans facts and futures: Valuing South Africa's ocean economy. Cape Town: WWF-SA, 2016.
6. World Bank. Capture fisheries production [webpage on the Internet]. No date [cited 2018 Oct 26]. Available from: https://data.worldbank.org/indicator/ER.FSH.CAPT.MT?locations=BR&view=chart
7. Arana PM, Pezzuto PR, Ávila-da-Silva AO, Queirolo D, Perez JAA, Arfelli CA. Pathways for sustainable industrial fisheries in southeastern and southern Brazil. Latin Am J Aquat Res. 2016;44(5):875–881. https://doi.org/10.3856/vol44-issue5-fulltext-1
8. UN General Assembly. The 2030 Agenda for Sustainable Development. UN General Assembly Resolution 70/1 of 25 September 2015. New York: UN; 2015. Available from: https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_RES_70_1_E.pdf
9. Abidjan Convention [homepage on the Internet]. No date [cited 2018 Nov 07]. Available from: https://abidjanconvention.org/
10. The Benguela Current Commission. Strategic Action Programme 2015 – 2019 [webpage on the Internet]. No date [updated 2014 Nov 03; cited 2018 Nov 07]. Available from: http://www.benguelacc.org/index.php/en/publications
11. South African Department of Environmental Affairs. Operation Phakisa – Oceans economy [webpage on the Internet]. No date [cited 2018 Oct 26]. Available from: https://www.environment.gov.za/projectsprogrammes/operationphakisa/oceanseconomy
12. Zuma J. Address by His Excellency President Jacob Zuma on progress made in respect of the implementation of the Operation Phakisa: Oceans Economy initiatives in Port Elizabeth, Eastern Cape, 8 April [webpage on the Internet]. c2016 [cited 2019 Jan 30]. Available from: http://www.dpe.gov.za/resourcecentre/archive/Pages/Address-by-His-Excellency-President-Jacob-Zuma-on-progress-made-in-respect-of-the-implementation-of-the-Operation-Phakisa-O.aspx
13. Funke N, Claassen M, Nortje K, Meissner R. A research, innovation and knowledge management road map for the South African maritime sector: Charting a course to maritime excellence by 2030. Report no. CSIR/NRE/WR/ER/2016/0044/A. Pretoria: Council for Scientific and Industrial Research; 2016.
14. Mattos BRB, Matos FEL, Kenkel KM. Brazilian policy and the creation of a regional security complex in the South Atlantic: Pax Brasiliana? Contexto Int. 2017;39(2):263–280. https://doi.org/10.1590/s0102-8529.2017390200004.
15. South African Department of Science and Technology, Brazilian Ministry of Science, Technology, Innovation and Communications. South-South Framework for Scientific and Technical Cooperation in the South and Tropical Atlantic and Southern Oceans [document on the Internet]. c2017 [cited 2019 Jan 30]. Available from: https://www.atlanticresource.org/aora/sites/default/files/GalleryFiles/AtlanticFacts/South--South-Framework-for-Scientific-and-Technical-Cooperation-in-the-S....pdf
16. Moedas C, Pandor N, Kassab G. Belém Statement on Atlantic Research and Innovation Cooperation: Conference in Lisbon on 13–14 July 2017 [document on the Internet]. c2017 [cited 2019 Jan 30]. Available from: https://ec.europa.eu/research/iscp/pdf/belem_statement_2017_en.pdf

AUTHORS:
Dianne Long[1]
Mwazvita S. Dalu[2]
Reuben L. Lembani[2]
Ashley Gunter[2]

AFFILIATIONS:
[1]School of Education, Social and Economic Science Division, University of the Witwatersrand, Johannesburg, South Africa
[2]Department of Geography, College of Agriculture and Environmental Science, University of South Africa, Pretoria, South Africa

CORRESPONDENCE TO:
Ashley Gunter

EMAIL:
gunteaw@unisa.ac.za

HOW TO CITE:
Long D, Dalu MS, Lembani RL, Gunter A. Shifting sands: The decoloniality of geography and its curriculum in South Africa. S Afr J Sci. 2019;115(9/10), Art. #5738, 3 pages. https://doi.org/10.17159/sajs.2019/5738

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
African collaboration, post-colonial theory, indigenous knowledge

PUBLISHED:
26 September 2019

Shifting sands: The decoloniality of geography and its curriculum in South Africa

Since the era of European exploration, the world has been conquered and nations subdued by imperial powers.[1] The evidence of colonisation is explicitly seen in the territorial demarcations on the global map, borders that at times seem to follow little more than natural boundaries chosen to divide territories between competing powers. The borders that demarcate African countries and separate its people are a clear illustration of this. It is within these borders that sovereign states now function and across these borders that international interactions occur. Geography, as an academic discipline which is concerned with both the physical environment and human interactions that occur within and between the borders of nation states, cannot ignore the politics of space. To date, the understanding and production of geographical knowledge continue to be informed by the European colonial modalities of power.[1,2] Even though decolonial scholars acknowledge that the modern episteme is saturated with coloniality[1], the geography curricula in African higher education does not show or encourage a different way of imagining or describing indigenous places[3]. One example of this is how indigenous knowledge, and African and indigenous scholarships, continue to be erased or subsumed.[4] Much of the current curricula content continue to exclusively credit only international explorers, e.g. Admiral Antonio with Table Mountain, yet most explorers used local field guides who introduced them to most spaces on the continent. Another example is how the curricula are still referring only to colonial names of places despite many indigenous names of places being known in local languages. Curricula also have a long way to go to sufficiently incorporate local ecological knowledge and practices that have wider implications and practical application, especially in topics of relevance to the higher education geography curricula. Evidence already exists that supports that African mythology can be credited for conservation[5,6], yet this is not adequately reflected in African geography curricula. It can, however, be justifiably argued that this omission may be a mere consequence of the mediation of knowledge by different communication technologies, which have uneven power dynamics, especially at the global scale.[5] It is therefore imperative that geography as a discipline, and geographers as scholars, take it upon themselves to understand the centrality of this discipline within the call to de-colonise higher education and make an ardent effort towards building de-colonial curricula.[7] Such acknowledgement, however, can only support the proposition of Bloor[8], which is to interrogate popular discourses and established bodies of knowledge through careful historical geographical scholarship to the African geography higher education curricula. One way of doing so, could be by considering the application of post-colonial theory in tackling the decolonisation of the geography curricula.

Post-colonial theory: Adding value and reshaping the discipline of geography

The term post-colonial is often misconceived to mean that colonialism has come to an end and we are now in an era that has proceeded colonial rule.[9] Post-colonial theory has very little to do with temporality but rather with the state of mind of people who have to question and respond to the 'theoretical underpinnings of disciplinary knowledge that has privileged some knowledge, societies and cultures, and have been used indiscriminately outside of the context and conditions and voices of people'[5]. Therefore, post-colonial theory implies resistance to imperialism, a critique of the colonial, and a rediscovery of indigenous histories and heritage.[5,10,11] The idea of critiquing and answering back to the dominant narrative that the imperial is somehow superior and represents the normative for all people groups around the world has been questioned within higher education.[5,9,11] De-linking decolonial thinking from the global hegemony of the West and her canon has been key in post-colonial critiques on higher education[11], and forms the basis of liberation from the colonial oppression of the West[3]. Answering back and questioning the supremacy of imperial knowledge has been utilised to great effect in the education system of New Zealand, a former British colony.[5]

Scholars have begun to make inroads into how academics can incorporate aspects of post-colonial theory into curriculum design that may refine understandings of Africa's particularities in disciplines such as geography.[12] In the sciences, Carter[13] has formulated three aspects that may be useful in bringing post-colonial theory into the curriculum as a means of extricating the post-colonial societies from the Western hegemony of knowledge. The first one is termed 'representations' and calls for scholars to try and understand how the Western hegemony of knowledge and influence is maintained in the academy, and through understanding we can begin to question this process of legitimising the colonial voice in our curriculum. Second, incorporating post-colonial theory in the academy involves the idea of recuperation. Here what is necessary is the building of a local subjective understanding of the history of the nation. By doing this, it is argued that local histories, experiences and languages can be used to write back to the dominant colonial narrative, and in doing so legitimise the local voice, and transcend the binary thinking that currently dominates the curricula of so many nations.[4,10] Translating difference is the third step in Carter's[13] methodology for incorporating post-colonial theory in the sciences. Translating difference refers to the act of pushing back against the Western assimilation of all societies into a Western norm which allows the West to speak for all others while really only speaking for itself. Incorporating local and indigenous knowledge into the canon of teaching is a means by which to push back against this Western norm that has been created for all other societies.[7]

It is argued here that geography as a discipline, so entrenched in its very foundations in the Western imperial canon, should start to engage with post-colonial theory and the efforts that are being made to at least try to decolonise the curriculum. By questioning the supremacy of the British canon and the means by which we perpetuate this supremacy even in our language of instruction we begin to address the call to decolonise.[4] Whether this culminates in a radical change with regard to all teaching materials, pedagogy and research is yet to be seen. Through engaging with post-colonial theory in higher education teaching, we can begin to build our own ideas and thoughts and legitimise our positions as African scholars in the academy with a voice of our own built upon our worldview and not that of the West.

Post-colonial theory, therefore, could be the means by which geographers begin to question the history of the discipline, the history of the country and the politics that shape it. It would evolve into a discipline that legitimises the African voice in scholarship rather than parroting the colonial voice.[1] Post-colonial geography could become the discipline in which students who are infuriated with the colonial dominance of the current curriculum can begin to voice their opinions and begin to write back to the imperial narrative. It is these students – students who have yet to be trained in the tradition of the current curriculum, but have encountered it and seen that it is not representative of us as Africans – who have the most potential to decolonise the curriculum. As a means of fostering this transition, current educators should foster the dissent that has been brought to the fore through protests. Instead of studying ourselves and our world from an academic tradition that is from the outside, we need to look from within to critique the outside. This necessitates bringing the resistance slogans into the classroom and starting to theorise, to build the African identity in the academy and legitimise it as different to the West and yet just as valuable.

Geography and its centrality in the decolonising debate

In both the physical and the human environment, geography engages with the politics of space and power. Climate change is a key example of a discourse that can be studied entirely as a scientific endeavour, but once solutions are proposed the geography of space and power come into play[14]: who governs which territory, who has power over local and multinational polluters, who signs agreements and how we should work together as global citizens towards a more sustainable future. When human geography is examined, interactions of humans with the environment are examined; issues of class, race, and power are all examined within a given geography that has oftentimes been demarcated by colonial powers. Thus, the need to decolonise geography raises important questions. Why decolonise? And for whose benefit is this imperative?

Within the South Africa context, these interactions are not only examined within predetermined geographies, but are also examined using Western theories and ideologies. The result of this is that we are studying Africa from within using a lens that is without. In contrast, critiques made from this space emphasise the decolonisation of geography curricula in higher education as a commitment to centre and empower marginalised groups.[7,12] As the question of decolonising higher education becomes more and more pertinent in our society[7], we need to consider whether or not we as geographers are mere stewards of the colonial mantle of knowledge, passing it on to the students we influence through our writings, teaching and supervision. When we examine the history of the academy and the discipline of geography and her academics in South Africa, it is not surprising that this is the case.

As geography academics in South Africa are either educated locally or abroad, there is limited distinction between the theoretical and knowledge learnt in the two instances. The universities and their subsequent departments and curricula in South Africa have a long colonial history.[15] A most obvious example is Rhodes University, named after a British imperialist, which was established to teach white colonisers using imperial curricula. The same is true for the establishment of the Universities of Cape Town, Witwatersrand, Stellenbosch and more. The aforementioned universities are also located in affluent spatial pockets of the country, in contrast to the historically disadvantaged universities that are located in impoverished spaces, even today. The colonial history of the establishment of these institutions was further cemented in the apartheid era with race rather than academic ability being the key determining factor for student placement. These universities, which now educate students of all races, are still teaching predominantly from the dominant Western canon, perhaps as a legacy of the colonial curriculum taught to those who now carry the mantle of lecturer. Aside from this, the importing of teaching staff and qualifications from overseas has only meant that we have further imported our curriculum from foreign nations.

Not only has the establishment of the academy and the politics of the country influenced the degree of coloniality in the discipline of geography in South Africa, but it is the history of the discipline as a whole that also needs to be questioned. Geography became recognised as a valid discipline in the 19th century in Western Metropolitan countries and through colonialism was exported as such.[2] Furthermore, it was in these nations that the very divide was made between the sub-disciplines of human and physical geography.[2] This divide persists in South African geography even though the divide between the human and physical, and the way in which the two interact, is less clear-cut in the local culture. Furthermore, our very own rather prestigious Society of South African Geographers was modelled on the colonial Royal Geographical Society.[2] The Royal Geographical Society being the follow on from the Association for Promoting the Discovery of the Interior Parts of Africa, established to explore and map out the African continent according to a British cartographer.[15] Furthermore, it is a mark of pride for the *South African Geographical Journal*, established through this society, to enjoy an international standing.[2] This international standing is determined by criteria of excellence not set by ourselves.

The sum of this is that we have a geography canon that does not reflect the local but rather the Western. We are studying from within the country as if we have a deficit of our own ideas and thoughts. We should be theorising from within, our studies should reflect the world as we understand it to be as Africans rather than through the lens of the West. If we as geographers can do this, we will make a far greater contribution to the academy and global discourse as a whole. We need to better understand how Africans can respond to issues of climate change, and need to further build how we theorise geography from African roots and gain a better understanding of our society.[16] Responding to student calls to abolish colonial teaching cannot be ignored. Geography as a discipline, and geography academics as curriculum planners, are not ignoring this call, but equally, more needs to be done. When we consider that we study both the physical and human environment and how both of these spheres have been influenced by the advent of colonialism, it becomes an imperative that we decolonise our curriculum and teach the history and the impacts of colonialism not from the narrative of the coloniser but rather from the colonised.

A way forward for geography curricula

The addition of post-colonial theory into the geography curriculum is essential as a starting point for decolonising the curriculum. It is this voice that writes back to the West that we must begin to foster in the discipline of African geography. However, the question of who should be allowed to speak on issues of decoloniality remains pertinent in this debate. If we are to be critical of our positionality, let us question the legitimacy of white geographers to speak on issues of decoloniality. Many white South Africans now have no affiliation with the colonial powers that saw them becoming African decedents, and yet white South Africans are in many ways affiliated with the colonial history of the country. This means that the position of white South African geographers and academics as a whole may be questioned, leaving them as outsiders to the debates on decolonisation. There is, however, some concern as to how this exclusion will foster decoloniality when a large portion of the staffing base in South African universities is still white. The top five universities in South Africa, as ranked by the Times Higher Education World University ranking, all have a typical staffing ratio of white South Africans of European descendent to black South Africans of over 70%, with all senior positions within the departments being held almost exclusively by white South Africans.[17-19]

Furthermore, as geographers, we need to destabilise the conception that the most highly ranked universities and their academics are able to voice the African experience and theorise from Africa. These universities are far too entrenched in their colonial pasts. Rather, South African universities that still conform to colonial structures and curricula, who also house most of the funding, should begin to partner with other universities who have students and academics who are truly able to be the African voice that can speak back to the colonial curriculum.[12,20] Through partnerships, the discipline of geography can be transformed to reflect Africa and its indigenous people as a whole. With partnerships being key to knowledge collaboration and transmission, South African geographers must expand existing partnerships and forge new ones with other African universities across the continent. These Africa partnerships can foster the transmission of African knowledge building and sharing, which should be one of the mandates for geography in South Africa, as

noted by Sithole[17], and echoed once again by Dalu[12] in considering the decolonisation agenda in South Africa specifically.

Conclusion

South African geography – with its very colonial roots – has the potential to be a discipline that brings to the fore the call for decoloniality. However, this will not happen without the academics who very much control and disseminate the curricula they design; the theories with which they choose to engage; the partnerships they make and the ranking they espouse to. In questioning our positionality we need to question how we as geographers feel about these various aspects and how we wish to engage moving forward. As academics we must commit to decoloniality, and perhaps using post-colonial theory in both our teaching and thinking about ourselves and the work we do is one of the best means by which we can clarify our stance and also make some inroads into the decoloniality of a colonial discipline.

Acknowledgements

This paper is part of the IDEAS project which is funded by a Newton Grant from the Economic and Social Research Council, UK (ES/P002161/1) and the National Research Foundation of South Africa (UTSA160329161196).

References

1. Asher N. Decolonization and education: Locating pedagogy and self at the interstices in global times. In: Colona RS, editor. Postcolonial challenges in education. New York: Peter Lang AG; 2009. p. 67–77.
2. Collard RC, Dempsey J, Sundberg J. A manifesto for abundant futures. Ann Ass Am Geog. 2015;105(2):322–330.
3. Grosfoguel R. The epistemic decolonial turn: Beyond political economic paradigms. Cult Stud. 2007;21:211–223. https://doi.org/10.1080/09502380601162514
4. Mignolo W. Cosmopolitanism and the decolonial option. Stud Philos Educ. 2010;29:111–127.
5. Jazeel T. Mainstreaming geography's decolonial imperative. Trans Inst Br Geogr. 2017;42:334–337. https://doi.org/10.1111/tran.12200
6. Noxolo P. Introduction: Decolonising geographical knowledge in a colonised and recolonising postcolonial world. Area. 2017;49(3):317–319. https://doi.org/10.1111/area.12370
7. Grande S. Red pedagogy: The un-methodology. In: Dennzin NK, Lincoln YS, Smith LT, editors. Handbook of critical and indigenous methodologies. Los Angeles, CA: Sage; 2008. p. 223–253.
8. Bloor D. Anti-Latour. Stud Hist Philos Sci A. 1999;30(1):81–112.
9. Sone EM. Nature in crisis: An ecological construction and conservation of the environment in Swazi oral literature. S Afr J Afr Lang. 2014;34(2):195–205. https://doi.org/10.1080/02572117.2014.997056
10. Battiste M. Animating sites of postcolonial education: Indigenous knowledge and the humanities. CSSE Plenary Address. Saskatoon: University of Saskatchewan; 2004.
11. Stellenbosch University. Department of Geography and Environmental Studies [webpage on the Internet]. c2018 [cited 2018 Jul 14]. Available from: http://www0.sun.ac.za/geography/people/
12. Dalu MT, Gunter AW, Wasserman RJ, Dalu T. Owning the lake, not just the rod: The continuing challenge of'the old boys' in knowledge production. S Afr J Sci. 2018;114(7/8), Art. #a0275, 2 pages. http://dx.doi.org/10.17159/sajs.2018/a0275
13. Carter L. Thinking differently about cultural diversity: Using postcolonial theory to (re) read science education. Sci Educ. 2004;88(6):819–836. https://doi.org/10.1002/sce.20000
14. Mackenzie CG. Black students in 'white' universities: The character and provision of liberal higher education institutions in post-apartheid South Africa. Compare: J Comp Int Educ. 1994;24(1):67–78. https://doi.org/10.1080/0305792940240106
15. Johnson J, Cant G, Howitt R, Peters E. Creating anticolonial geographies: Embracing indigenous knowledges and rights. Geogr Res. 2007;45:117–120. https://doi.org/10.1111/j.1745-5871.2007.00441.x
16. Fairhurst UJ, Davies RJ, Fox RC, Goldschagg P, Ramutsindela M, Bob U, et al. Geography: The state of the discipline in South Africa (2000–2001). S Afr Geog J. 2003;85(2):81–89. https://doi.org/10.1080/03736245.2003.9713787
17. Sithole B. My knowledge, your knowledge, whose knowledge is it? Reflections from a researcher's journey through universities in North and South. In: Adriansen HK, Madsen LM, Jenses S, editors. Higher education and capacity building in Africa: The geography and power of knowledge under changing conditions. London: Routledge; 2016. https://doi.org/10.4324/9781315734620
18. Terefe MT. Ethiopia was colonised [webpage on the Internet]. c2017 [cited 2017 Aug 15]. Available from: http://africanarguments.org/2017/06/21/ethiopia-was-colonised
19. University of Pretoria. Geography, Geoinformatics & Meteorology staff [webpage on the Internet]. c2018 [2018 Jul 14]. Available from: https://www.up.ac.za/en/geography-geoinformatics-and-meteorology/article/46217/staff
20. Mbembe A. Decolonising the university: New directions. Art Human Higher Educ. 2016;15:29–45.

AUTHOR:
Stanley Maphosa[1]

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

CORRESPONDENCE TO:
Stanley Maphosa

EMAIL:
stanley@assaf.org.za

HOW TO CITE:
Maphosa S. The Academy of Science of South Africa and science diplomacy. S Afr J Sci. 2019;115(9/10), Art. #a0314, 3 pages. https://doi.org/10.17159/sajs.2019/a0314

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
ASSAf, National System of Innovation, science advice, young scientists, gender equality

PUBLISHED:
26 September 2019

The Academy of Science of South Africa and science diplomacy

The Academy of Science of South Africa (ASSAf) recently released its second report on the state of Scholarly Publishing in South Africa. Titled *Twelve Years Later: Second ASSAf Report on Research Publishing in and from South Africa*, the report outlines ASSAf's important role in promoting innovation and scholarly activity. However, ASSAf has a substantial range of other strategic goals; here I highlight those that relate to providing scientific advice to national government; playing a critical role in international science diplomacy (including transborder issues in and beyond Africa); and sharing best practice for other academies to engage in bilateral joint committees and action planning.

ASSAf was established in 1994/1995 and became the only national science academy formally recognised by the South African government in 2002, determined by the ASSAf Statute (*Act 67 of 2001*) – one of the earlier, beneficial results of the advent of democracy in South Africa. In order to fulfill its national and international roles, ASSAf adopted in its name the term 'science' in the singular, referring to 'knowledge' and reflecting a common way of enquiring rather than an aggregation of different disciplines. In terms of the Act, the Academy's mission is to:

1. recognise scholarly achievement and excellence in the application of scientific thinking for the benefit of society;
2. mobilise Members to ensure that they are available to contribute their expertise in the service of society;
3. conduct systematic and evidence-based studies on issues of national importance, producing authoritative reports that have significant impact on policymaking;
4. promote the development of an indigenous system of South African research publications, increasing their quality, visibility, accessibility and impact;
5. publish science-focused periodicals that will showcase the best of southern African research to a wide national and international audience;
6. develop productive partnerships with national, regional and international organisations with a view to building our capacity in science and its application within the National System of Innovation;
7. create diversified sources of funding for sustainable functioning and growth of a national Academy; and
8. communicate effectively with relevant stakeholders through various media and fora.

The activities described here are those that give substance to item 6 of the mission. In addition, what follows is supported by, and reinforces, other South African national policy frameworks such as the Ten-Year Innovation Plan, the National Research and Development Strategy, and the National Development Plan. The growing importance of the science content in critical foreign policy issues has required the South African government to pursue a concerted science diplomacy strategy in which ASSAf plays a significant role.

International cooperation

ASSAf's international responsibilities have increased markedly over the years, as its reputation has grown both locally and abroad. ASSAf has contributed to global statements and ensured the participation of its elected Members and other South African scientists in high-level engagements.

To give this effect, ASSAf has a number of strategic partners. The global network of academies, which in 2016 was renamed the InterAcademy Partnership (IAP), with three branches – IAP for Health, IAP for Science and IAP for Research – all based in Trieste, Italy, is one such partner. ASSAf plays a major role in the executive committees of all three branches with a mandate to represent developing nations.

ASSAf also works with the BRICS science academies and is championing the formation of the BRICS network of academies of science. This relation is intended to enhance collaboration between the academies of the five countries under a common framework, and to offer science advice to the BRICS annual summit. ASSAf has also been involved in the formation of the International Network of Government Science Networks (INGSA) Africa Chapter and still supports that Network. Through the G20 country academies of science, which are called the Science 20 (S20), ASSAf engages actively in science diplomacy through statements, reports and other activities that provide science advice at senior levels. ASSAf also vigorously profiles and encourages South African scientists to contribute at international levels as individuals or as technical backup to science diplomacy.

The World Academy of Sciences (TWAS) is another strategic ASSAf partner, and since 2015 ASSAf has hosted the TWAS Southern Africa Regional Partner (TWAS-SAREP) office. The main objective of TWAS-SAREP is to promote the goals of TWAS in the region, with a strong focus on awarding prizes and supporting young scientists. The Regional Partner supports science diplomacy training on the continent and networking among young affiliates through the TWAS Young Affiliates Network (TYAN), and ensures that the Regional Young Scientists Conference is held in a range of African countries – especially those lagging somewhat in science development. In this way, TWAS-SAREP works with national young and senior academies and assists in academy development.

A further strategic partner is the International Science Council (ISC); ASSAf also hosts the ISC Regional Office for Africa (ISC-ROA). There is a close alignment between the goals of ASSAf and ISC-ROA, particularly in terms of Africa-wide collaboration, thematic areas of interest and promoting young scientists, all of which present significant

opportunities for synergistic partnerships. ASSAf is currently strengthening and consolidating South–South relations, reflecting a shift in the balance of the global distribution of power as well as the increasing influence of emerging economies in the multilateral science and innovation system.

Overseas collaboration

Collaboration covers bilateral and multilateral agreements with overseas academies of science, support for overseas strategic partnerships, and engagement with multilateral organisations such as the European Commission – among others. ASSAf brings overseas academies to the continent to support African collaborations, women in science and technology, and young scientists, including the South African Young Academy of Science (SAYAS) and TWAS-SAREP as flagships of ASSAf. It also links overseas academies and partners with other programmes within ASSAf for collaborative activities.

In this respect, ASSAf engages actively with the German National Academy of Sciences – Leopoldina – with which they have undertaken several joint activities. ASSAf also engages with other academies, especially those with which the South African Department of Science and Technology (DST) has ongoing or planned bilateral cooperation, and participates in Joint Committee and Action Planning Meetings with countries that have signed bilateral cooperation with the DST to ensure that academies of science of both countries can provide science advice in jointly selected priority areas. Currently, ASSAf has Memoranda of Understanding (MOUs) with Germany, Austria, China, Belarus and India, and is currently renewing bilateral agreements with Russia.

In addition, ASSAf has worked with the British Academy, the American Association for the Advancement of Science (AAAS), the Royal Society in the UK and other academies on a number of collaborative activities – and has begun working with a network of diplomatic missions and embassies based in Pretoria. Through the Embassy Lecture Series, ASSAf has access to high-profile scientists and policymakers invited to South Africa. Working jointly with visitors' home country embassies, ASSAf hosts lectures by these visitors across South Africa. Lectures have already been held in collaboration with the USA, Switzerland, UK, Italy and New Zealand. Other embassies have expressed interest in this initiative. The Embassy Lecture Series is, among other things, an entry point for ASSAf to work with academies of countries that have embassy missions in South Africa.

On a multilateral level, ASSAf is an implementing partner of the European Union Horizon 2020 Programme and is responsible for reporting on the participation of South African and European scientists in this framework of research funding. ASSAf is also involved with other Regional Networks of Science Academies in Europe, the Asia-Pacific region and Latin America to ensure that South African scientists can participate in activities created by these academies, and to assist ASSAf in developing collaborations with individual member academies. ASSAf also works with the Human Rights Network of Academies. Particular attention will, in the coming year, be paid to the establishment of new links with the BRICS group of academies and the BRICS Network of Science Academies. The network will position itself so as to be able to give science advice to the BRICS Summit.

ASSAf works with Commonwealth Academies of Science with the support of the Royal Society in the UK, and, as mentioned above, with science academy members of the G20 countries – the so-called Science 20 (S20) – to provide strategic foresight and science advice to the G20 Summit through statements produced annually. ASSAf also has links with G20 international offices in Tokyo and Brussels through South Africa's diplomatic missions, which are dedicated to promoting cooperation with Japan and the European Union. ASSAf works closely with the DST-seconded official to the secretariat of the Southern African Development Community (SADC) in Gaborone, Botswana. In multilateral frameworks, South Africa seeks to encourage and support scientific cooperation and the building of trust and relationships among its partners, and to foster consensus on contentious or difficult issues.

African collaborations

African countries face many common development challenges and are bound by the pan-African policies of the African Union Commission. Academies have a major role to play in collectively addressing these matters. South Africa's future is directly linked to that of Africa, and it follows that, through its Africa Strategy, ASSAf continues to support regional and continental academy development processes, awareness-raising and science advocacy, especially for investment in research and development by African governments. ASSAf works with the Network of African Science Academies (NASAC) in strengthening African science academies and establishing new ones, especially within SADC. To this end, ASSAf has been in discussions with Angola, Lesotho, Namibia and Eswatini (formerly Swaziland), and, based on that engagement, Eswatini has established the Kingdom of Eswatini Academy of Science and Angola will soon be establishing its own academy. ASSAf made a presentation to the SADC Ministers responsible for science and technology in Eswatini in June 2017, following which the Ministers committed to the formation of new academies and the supporting of existing academies. ASSAf also assisted in the formation of the Botswana Academy of Science.

Beyond the SADC, ASSAf supported the Rwanda Academy of Science to develop a constitution, thus enabling them to apply for registration as a member of NASAC. ASSAf has MoUs with academies in Benin, Nigeria, Senegal, Uganda and Mauritius and partners with these and other academies on both bilateral and multilateral levels.

It is important to note that ASSAf participates in Joint Committee and Action Planning Meetings for African countries that have Science and Technology bilateral agreements with South Africa. ASSAf recognises Africa Day and uses it as a key opportunity to raise awareness of science on the continent, and also works with the Next Einstein Forum on Africa Science Week.

ASSAf conducts colloquia for scientists in the South African diaspora in other African countries – the largest number being from Zimbabwe, Nigeria and Kenya. In addition, ASSAf collaborates with the Science, Technology and Innovation (STI) hub of New Partnership for African Development (NEPAD), the African Academy of Sciences (AAS) and ISC-ROA in developing a database of African scientists. ASSAf is positioning itself as a critical contributor to the implementation of the Science, Technology and Innovation Strategy for Africa (STISA-2024) and towards achieving the Sustainable Development Goals. Working with TWAS-SAREP in all its activities, ASSAf promotes Science Diplomacy in Africa: in 2018, ASSAf began championing science diplomacy training at home and in the rest of Africa with partners such as AAAS, TWAS and DST. In March 2018, ASSAf hosted the first session of TWAS and AAS Science Diplomacy training in Africa. This training had previously always been held in Trieste. Other training sessions will be held in four other TWAS Regional Partner Offices, beginning with Egypt in 2019. The Academy is working with DST to conceptualise a curriculum for South African Science Diplomacy, to be launched in 2019.

Finally, ASSAf is working to strengthen links within SADC, especially on STI and the Gender Protocol. With TWAS-SAREP and ISC-ROA currently being hosted at ASSAf, and in collaboration with NASAC and other continental science entities, ASSAf ensures that the best African scientists from these organisations are involved in the continental science advice ecosystem.

Young scientists and gender-related activities

Since 2010, ASSAf has hosted an annual Young Scientists' Conference, aligned to the themes of the International Year of the UN or the AU, as part of its commitment to supporting the development of young scientists. ASSAf provides funding to SAYAS for its operational activities in its mission to support the growth of the young academy. ASSAf also assists SAYAS in raising its profile within the Global Young Academy and with other African science academies. In this way, ASSAf creates opportunities for joint activities with SAYAS and ensures that SAYAS Members participate in ASSAf's Standing Committees and study panels.

ASSAf has actively supported the nomination of young scientists for awards and leadership opportunities and acts as the implementing agency for some of these, including the BRICS Young Scientists Conference; TWAS Regional Young Scientist Prize and TWAS Young Affiliates' Young Physician Leaders Programme which is linked to IAP for Health and the

World Health Organization; Lindau Nobel laureate meetings; and AU-TWAS Young Scientists' Prize.

The Organization for Women in Science for the Developing World (OWSD) South Africa National Chapter is hosted by ASSAf, and all gender-related activities in ASSAf have been coordinated through the work of this Chapter. ASSAf has assisted other SADC member states (Eswatini for example) to establish Women in Science, Engineering and Technology (WISET) National Chapters. The WISET National Chapters were introduced by SADC through a charter and every country should have one by mid-2019.

ASSAf has also been the focal point for GenderInSITE in the southern African region since 2015. GenderInSITE is a global initiative that aims to raise the awareness of decision-makers on gender aspects of the Science Innovation Technology and Engineering (SITE) initiative, aimed at both men and women. Through this initiative, ASSAf has had an opportunity to consolidate and strengthen gender-related activities and policy influence across southern Africa. Through GenderInSITE, ASSAf has been working with SADC Member States to produce factsheets for the monitoring of the implementation of the Gender Protocol.

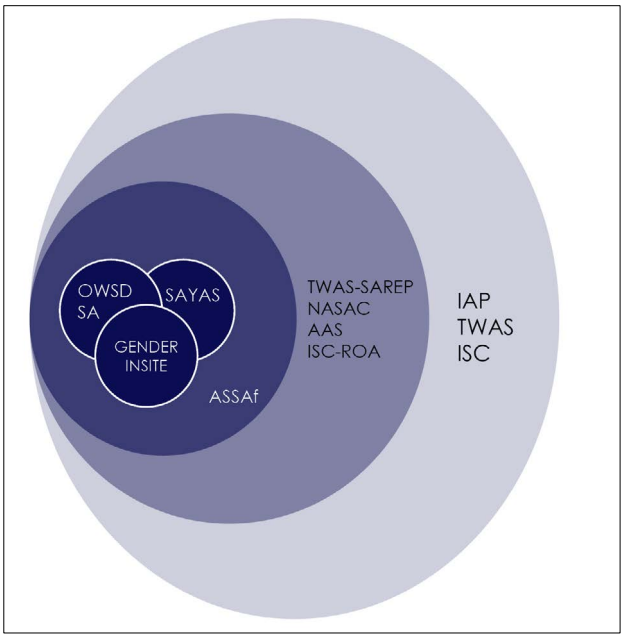
Of course, ASSAf does not have exclusive responsibility for matters related to science diplomacy and works closely on matters of mutual interest with a wide range of science councils and with government. Many of South Africa's national science councils or other public-funded research and technology organisations also have teams dedicated to international cooperation. One of these is the National Research Foundation (NRF), which is responsible for the implementation of international science and technology cooperation agreements. The ASSAf science diplomacy agenda comprises multiple initiatives, but all target strategic national priorities.

ASSAf's priorities for science diplomacy can perhaps best be summarised by emphasising that international scientific cooperation is pursued both as an objective in its own right and as an instrument to attain strategic national and foreign policy objectives. Its international relations often provide some of the resources that ASSAf seeks to support activities that are not funded through the baseline budget. In 2019 alone, the International Liaison Programme raised ZAR15 million in support of science diplomacy interventions. These diplomatic interventions have all played a part in ensuring that 'science for sustainable development' enjoys priority focus in global forums.

Conclusion

One of the current flagship areas for South African science diplomacy is radio astronomy in Africa, specifically the Square Kilometre Array (SKA). ASSAf supports SKA and includes it as an example in science diplomacy training and other engagements in Africa and beyond. Certainly, there is a positive picture of ASSAf's science diplomacy efforts and much to be proud of. Nonetheless, a more detailed analysis might also interrogate the obstacles, dead ends, and frustrations that international cooperation faces, but that will need to wait.

As described above, ASSAf's science diplomacy agenda comprises, and has achieved success in, five areas: developing and maintaining (1) strategic partners; (2) African collaborations; (3) overseas collaborations; (4) young scientists' liaison; and (5) gender equality in science, technology and innovation (Figure 1). The dynamic interfaces between these five components will certainly increase. This is perhaps the largest challenge for ASSAf's future science diplomacy engagements: to have an agenda that is sufficiently focused in order to ensure an optimal investment of resources, but sufficiently flexible to be able to respond to the rapidly changing dynamics of international relations in the 21st century, which, if not driven by science, most certainly will require science-based responses and support.



IAP, The InterAcademy Partnership; TWAS, The World Academy of Sciences; SAREP, Sub-Saharan Africa Regional Partner; ISC, International Science Council; ROA, Regional Office for Africa; NASAC, Network of African Science Academies; AAS, African Academy of Sciences; SAYAS, South African Young Academy of Science; OWSD, Organization for Women in Science for the Developing World

Figure 1: ASSAf's science diplomacy agenda comprises overseas (outer circle) and African (middle circle) collaborations as well as young scientists' liaison and gender equality in science, technology and innovation (inner circles).

AUTHORS:
Lorren K. Haywood[1]
Caradee Y. Wright[2,3]

AFFILIATIONS:
[1]Smart Places, Council for Scientific and Industrial Research, Pretoria, South Africa
[2]Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa
[3]Environment and Health Research Unit, South African Medical Research Council, Pretoria, South Africa

CORRESPONDENCE TO:
Caradee Wright

EMAIL:
cwright@mrc.ac.za

HOW TO CITE:
Haywood LK, Wright CY. Private sector contribution to SDG 3: Health and Well-being – a South African case study. S Afr J Sci. 2019;115(9/10), Art. #6452, 3 pages. https://doi.org/10.17159/sajs.2019/6452

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
sustainable development, global health, public health, environmental health, business

PUBLISHED:
26 September 2019

Private sector contribution to SDG 3: Health and Well-being – a South African case study

Since the Sustainable Development Goals (SDGs) were globally adopted by world leaders in 2015[1], worldwide the private sector has openly embraced them and committed to address the SDG targets[2-4]. The United Nations Global Compact[5] lists 10 principles that enable businesses to adopt socially responsible practices into their strategies. Similarly, the Global Reporting Initiative is an internationally recognised set of guidelines to enable businesses to report on how they address sustainability.[6]

While it is primarily governments' responsibility to provide the enabling environment for SDG implementation, they will not be achieved without private sector involvement.[2] Businesses have a significant role to play as engines of economic growth, employment, finance, technology and innovation. They report their social contributions in annual sustainability and integrated reports. Because a healthy population contributes to a healthy workforce, it is in the interest of business to address human health.[7-9]

SDG 3 – Good Health and Well-being – seeks to ensure health and well-being at every life stage.[10] This goal is broader than the Millennium Development Goals on child/maternal mortality and communicable diseases.[11] SDG 3 includes 13 targets and 26 high-level indicators which address all major health priorities, including reproductive, maternal and child health; communicable, non-communicable and environmental diseases; universal health coverage; and access for all people to safe, effective, quality and affordable medicines and vaccines. Aspects of health are also captured in targets of several other SDGs.[11] Interconnectedness of SDG 3 goals with other SDG goals highlights that population health is a major beneficiary of sustainable growth and central to the achievement of SDGs.[12]

Universal health coverage (SDG Target 3.8) is a unifying platform for making progress on SDG 3[13] and it emphasises the importance of universal access to health-care services. The private sector has provided financial mechanisms and has also provided expertise in enlarging human capacity and conducting research for health care.[14] The contribution that the private sector can make to achieving SDG 3 is being recognised[7,14] and should be reported by businesses. Through their corporate social investment (CSI) programmes, business can leverage capabilities that contribute to universal health coverage[28], including innovative data collection technologies, enhanced disease surveillance, improved supply chain management practices for storage and delivery of essential health supplies.

SDGs provide an opportunity for businesses to support public health through their value chains, communication activities, occupational health and safety practices, and employee benefits. By ensuring that employees have safe working conditions and access to health services, businesses build better relationships with their employees. The purpose of the current study was to understand the extent to which South African businesses explicitly report contributions to SDG global health objectives. This analysis is important because the country faces significant public health issues that cannot be overcome without private sector contribution.

South Africa and its public health related challenges

South Africa faces a quadruple burden of disease[15] and the mortality rate attributed to cardiovascular disease, cancer, diabetes and/or chronic respiratory disease is ranked third worst in sub-Saharan Africa. Mean probability of dying from one or more of these four diseases in 2015 in South Africa (ages 30–70 years) was 26% compared with 21% for all of sub-Saharan Africa.[16] There is concern about the increase in the prevalence of non-communicable diseases.[17] The Global Burden of Disease 2015 SDG Collaborators created an overall health-related SDG index for which South Africa was ranked 134 (out of 188).[18]

Business case for social involvement in public health

Today, it is not enough for a business to only make a profit for shareholders; their licence to operate must include a contribution to social capital.[19] The unprecedented pace of global change presents businesses with new risks, such as effects on ecosystems and population health, and affects their modi operandi. These interconnected relations are pressuring the private sector to appreciate that it is not isolated from the communities and the physical environment in which it operates.[5,20]

A business contributes to community health by ensuring that it safeguards the environment that ensures commercial viability.[20] Moreover, public–private partnerships for public health can, for example, increase access to pharmaceuticals in developing countries as new technologies come to market.[21] Increase in diseases poses one of the greatest threats to the global economy.[22,23] Recent outbreaks such as Ebola demonstrated extraordinary health, economic and security risks associated with infectious disease epidemics.[24] Ebola alone affected 28 639 people and caused 11 316 deaths, leading to USD2.2 billion in lost GDP combined in Guinea, Liberia and Sierra Leone.[24] Such outbreaks threaten not only macroeconomic stability but also food security, human capital development and private sector growth.[8] Investing in global health security and emergency preparedness is a priority in an increasingly globalised world.[7,8] Such investment is critical to health equity, as vulnerable, marginalised people are primary victims. Improving health contributes to economic development and stability.[25] The private sector has resources and expertise to contribute to strengthening health security by protecting their employees and communities in which they operate.[25]

Gathering the evidence

Businesses were drawn from those listed on the Johannesburg Stock Exchange (JSE). The top 100 (out of ~400 listed) were extracted from the JSE website (www.jse.co.za), and of these, 88 were included in the study that was based upon available annual reports over three consecutive financial years (2016–2018). The list of 88

businesses and their JSE-defined sector type (https://www.sharenet.co.za) was exported into Microsoft Excel (Microsoft Office 265 ProPlus) for analysis. In 2019, two researchers searched the Internet for top 100 businesses' annual integrated and/or sustainability reports for 2016, 2017 and 2018. Once these reports were located, they were downloaded and analysed to identify: (1) mentions of SDGs; (2) specific mentions of SDG 3; and, if SDG 3 was mentioned (3) whether it was being applied to communities and society beyond their workforce/immediate customers.

In 2016, 25 businesses specifically mentioned SDGs in either report (Table 1), while only eight reports specifically mentioned SDG 3. The number that reported SDGs and SDG 3 explicitly in 2017 more than doubled (to 18) from that in 2016 and increased again in 2018 (to 20). Businesses stated their contribution to SDG 3 through their health and safety strategies and plans or through CSI initiatives. Where companies discussed their contributions to SDGs and/or SDG 3, this input was mainly through sustainability reports (not a mandatory report as is the case for integrated reports).

Discussion

Even though it is still 'early days' in terms of SDG implementation, we expected more companies to have explicitly linked their contributions to SDG targets and SDG 3 given the UNDC guidance. More of them itemised their SDG contribution in 2017 and 2018 than in 2016 and we expect that this trend will continue.

Of the 88 companies investigated, it was those from the mining sector that currently have the greatest number highlighting their contributions to SDG 3. Companies in the mining sector have to comply with health and safety regulations as a requirement of the *Mine Health and Safety Act (Act No. 29 of 1996)*. Mining health risks include cardiovascular and respiratory diseases, substance abuse and domestic violence. Health and safety programmes are initiated not only for employees but also as a CSI to invest in community health programmes. As health hazards relevant to mining are discussed in SDG 3 Targets 3.3, 3.4 and 3.5, mining companies can align their programmes and reporting. One mining company quantified its data to demonstrate contribution to HIV/Aids and non-communicable disease targets. In this sustainability report, the company detailed relevant SDG 3 targets and also provided dedicated commitment and activities to each target. Another stated exactly how the company contributed to SDG 3 by referring to a memorandum of understanding with the Provincial Department of Health to provide primary health-care services to the communities surrounding their operations.

A beverage business provided insight into its 'healthier world programme' in which it invests in programmes to shift social behaviour around alcohol abuse (SDG 3 Target 3.5). Another company in the same sector noted that although it produced alcoholic drinks, it also had embarked on programmes to educate pregnant women on the impact of alcohol on their unborn children (SDG 3 Target 3.2). There were also programmes to reduce road deaths and injuries from drinking and driving (SDG 3 Target 3.6). But surprisingly, none of the JSE health-related businesses, for example hospital management, long-term care, pharmaceuticals and biotechnology, mentioned SDG 3, despite their potentially critical role.[26]

Generally, businesses that mentioned SDG 3 did so in relation to either their CSI among their own employees, their customers or, more rarely, society by making employee health care available. There were programmes that extended beyond the commercial responsibilities that aimed to improve livelihoods or their employees' families to access health care. For example, a mining business offered free testing for HIV/Aids, counselling and treatment to their employees as well as education awareness programmes among nearby communities. There is a need for businesses to be made aware of the SDGs through the Global Reporting Initiative to encourage them to align their CSI activities to specific SDGs. This would not only benefit the company but would help government track progress towards meeting SDG targets.

Notwithstanding the potential benefit that the business sector can contribute to the achievement of SDG 3, it must be noted that there have been challenging situations and tensions between some CSI programmes and public health.[27] Some CSI activities conducted by the tobacco and alcohol industries were beset with conflicts of interest and systematic biases. Careful evaluation and monitoring, and even possible policy and regulation by independent bodies, may be required to ensure sound, ethical public health and health promotion initiatives by businesses.[27]

In our review of the reports that mentioned being involved in activities relevant to SDG 3, we also considered missed opportunities. We believe there is potential for private health-care providers to work more closely with public providers to expand the reach of health-care services and to close gaps in quality care. Some companies could upgrade hospital and clinic infrastructure and even prioritise access to new services and products via CSI. There were also opportunities where private resources and expertise might be extended to state institutions by introducing innovation or training for health-care workers. And large international businesses could use their transborder reach to tackle global health challenges on the scale required to achieve SDG targets.

Conclusions

Many of South Africa's top JSE-listed companies do not publicly report on their contribution to the SDGs. This is not to say they do not contribute to SDGs or SDG-related public health targets but they are failing to align their activities and/or to report on them in the public domain. Given that the SDGs are guiding principles for humanity's future, it is important that businesses not only report on their activities but also make greater contributions towards achieving the SDGs by working in partnership with government and other organisations.

Table 1: Descriptive findings of the analysis and characteristics of the businesses that explicitly mention the SDGs and/or SDG 3

Characteristic	2016		2017		2018	
Total $N = 88$ $12 =$ Missing	No % (*n*)	Yes % (*n*)	No % (*n*)	Yes % (*n*)	No % (*n*)	Yes % (*n*)
Sustainability Report available online	58 (51)	42 (37)	60 (53)	38 (35)	63 (55)	38 (33)
Integrated Annual Report available online	3 (3)	97 (85)	1 (1)	99 (87)	2 (2)	98 (86)
Either Report explicitly mentions any SDG	72 (63)	28 (25)	58 (51)	42 (37)	77 (68)	23 (20)
Either Report explicitly mentions SDG 3	91 (80)	9 (8)	80 (70)	20 (18)	77 (68)	26 (20)
Number of businesses mentioning SDG 3 by sector:						
Agriculture	1		1		4	
Mining	4		8		7	
Wholesale and trade	0		3		4	
Finance	1		3		3	
Manufacturing	1		1		1	
Telecommunication	1		2		1	

Acknowledgements

C.Y.W. receives research funding support from the South African Medical Research Council and the National Research Foundation (South Africa). L.K.H. receives research funding from a parliamentary grant.

References

1. United Nations. Transforming our world: The 2030 agenda for sustainable development. In: World public sector report 2015: Responsive and accountable governance. New York: United Nations; 2015. p. 73–87. https://doi.org/10.18356/e5a72957-en
2. Global Reporting Initiative/United Nations Global Compact/World Business Council for Sustainable Development SDG Compass. The guide for business action on the SDGs [document on the Internet]. c2015 [cited 2017 Aug 21]. Available from: http://sdgcompass.org/wpcontent/uploads/2016/05/019104_SDG_Compass_Guide_2015_v29.pdf
3. Pricewaterhouse Coopers. Navigating the SDGs: A business guide to engaging with the global goals [document on the Internet]. c2016 [cited 2017 Aug 23]. Available from: https://www.pwc.com/gx/en/sustainability/
4. Business and Sustainable Development Commission. Better business, better world [document on the Internet] c2017 [cited 2017 Aug 17]. Available from: http://report.businesscommission.org/uploads/BetterBiz-BetterWorld_170215_012417.pdf (last accessed17 August 2017)
5. United Nations Global Compact. United Nations Global Compact progress report: Business solutions to sustainable development [document on the Internet] c2017 [cited 2018 Feb 12]. Available from: https://doi.org/10.18356/bb7953d3-en
6. Maubane P, Prinsloo A, Van Rooyen N. Sustainability reporting patterns of companies listed on the Johannesburg Securities Exchange. Public Relat Rev. 2014;40:153–160. https://doi.org/10.1016/j.pubrev.2014.02.014
7. World Economic Forum. How can business improve global health? [webpage on the Internet]. c2015 [cited 2018 Jan 26]. Available from: https://www.weforum.org/agenda/2015/09/how-can-business-improve-global-health/
8. World Economic Forum. The global risks report 2016 [document on the Internet]. c2016 [cited 2018 Jan 26]. Available from: http://www3.weforum.org/docs/GRR/WEF_GRR16.pdf
9. World Economic Forum. Global healthcare: The $300 billion question [webpage on the Internet]. c2017 [cited 2018 Feb 27]. Available from: https://www.weforum.org/agenda/2017/11/the-300-billion-global-health-question
10. United Nations. The Sustainable Development Goals report. New York: United Nations; 2017.
11. Kumar S, Kumar N, Vivekadhish S. Millennium Development Goals (MDGS) to Sustainable Development Goals (SDGS): Addressing unfinished agenda and strengthening sustainable development and partnership. Indian J Community Med. 2016;41:1. https://doi.org/10.4103/0970-0218.170955
12. World Health Organization (WHO). Health in 2015: From MDGs, Millennium Development Goals to SDGs, Sustainable Development Goals. Geneva: WHO; 2015. https://doi.org/10.4324/9781315228068-4
13. Chapman AR. Assessing the universal health coverage target in the Sustainable Development Goals from a human rights perspective. BMC Int Health Hum Rights. 2016;16, Art.#33, 9 pages. https://doi.org/10.1186/s12914-016-0106-y
14. De Wolf AH, Toebes B. Assessing private sector involvement in health care and universal health coverage in light of the right to health. Health Human Rights. 2016;18:79.
15. Pillay-van Wyk V, Msemburi W, Laubscher R, Dorrington RE, Groenewald P, Glass T, et al. Mortality trends and differentials in South Africa from 1997 to 2012: Second National Burden of Disease Study. Lancet Glob Health. 2016;4(9):642–653. https://doi.org/10.1016/s2214-109x(16)30113-9
16. World Health Organization. World health statistics 2017 [webpage on the Internet]. c2017 [cited 2017 Dec 14]. Available from: http://www.who.int/gho/publications/world_health_statistics/2017/en/
17. Day C, Gray A. Health and related indicators. In: Padarath A, Barron P, editors. South African health review 2017. Durban: Health Systems Trust; 2017.
18. Lim SS, Allen K, Bhutta ZA, Dandona L, Forouzanfar MH, Fullman N, et al. Measuring the health-related Sustainable Development Goals in 188 countries: A baseline analysis from the Global Burden of Disease Study 2015. Lancet. 2016;388:1813–1850. https://doi.org/10.1016/S0140-6736(16)31467-2
19. Buhmann K. Public regulators and CSR: The 'social licence to operate' in recent United Nations instruments on business and human rights and the juridification of CSR. J Bus Ethics. 2016;136(4):699–714. https://doi.org/10.1007/s10551-015-2869-9
20. World Economic Forum. Managing the risk and impact of future epidemics: Options for public-private co-operation [document on the Internet]. c2015 [cited 2018 Jan 16]. Available from: http://www3.weforum.org/docs/WEF_Managing_Risk_Epidemics_report_2015.pdf
21. Reich MR. Public-private partnerships for public health. Nat Med. 2002;6:617–620.
22. Reubi D, Herrick C, Brown T. The politics of non-communicable diseases in the Global South. Health Place. 2016;39:179–187. https://doi.org/10.1016/j.healthplace.2015.09.001
23. Islam SMS, Purnat TD, Phuong NTA, Mwingira U, Schacht K, Fröschl G. Non-communicable diseases (NCDs) in developing countries: A symposium report. Global Health. 2014;10:81. https://doi.org/10.1186/s12992-014-0081-9
24. World Bank. Update on the economic impact of the 2014-2015 Ebola epidemic on Liberia, Sierra Leone and Guinea [document on the Internet]. c2015 [cited 2017 Dec 12]. Available from: https://openknowledge.worldbank.org/bitstream/handle/10986/21965/95804.pdf
25. Sturchio JL, Goel A. The private-sector role in public health: Reflections on the new global architecture in health [webpage on the Internet]. c2012 [cited 2018 Nov 20]. Available from: https://csis-prod.s3.amazonaws.com/s3fs-public/legacy_files/files/publication/120131_Sturchio_PrivateSectorRole_Web.pdf
26. Thorsteinsdóttir H, Ovtcharenko N, Kohler JC. Corporate social responsibility to improve access to medicines: The case of Brazil. Global Health. 2017;13:10. https://doi.org/10.1186/s12992-017-0235-7
27. De Vries H. Invited Commentary: Corporate social responsibility and public health: An unwanted marriage. Prev Med. 2016;89:345. https://doi.org/10.1016/j.ypmed.2016.07.003

AUTHORS:
Caradee Y. Wright[1,2]
Matthew Chersich[3]
Angela Mathee[4,5,6]

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]Wits Reproductive Health and HIV Institute, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
[4]Environment and Health Research Unit, South African Medical Research Council, Johannesburg, South Africa
[5]School of Public Health, University of the Witwatersrand, Johannesburg, South Africa
[6]Environmental Health Department, Faculty of Health Sciences, University of Johannesburg, Johannesburg, South Africa

CORRESPONDENCE TO:
Caradee Wright

EMAIL:
cwright@mrc.ac.za

HOW TO CITE:
Wright CY, Chersich M, Mathee A. National Health Insurance and climate change: Planning for South Africa's future. S Afr J Sci. 2019;115(9/10), Art. #5800, 3 pages. https://doi.org/10.17159/sajs.2019/5800

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
weather, environmental health, prevention, health care

PUBLISHED:
26 September 2019

National Health Insurance and climate change: Planning for South Africa's future

Introduction

In South Africa we have witnessed a gripping drought in the Western Cape, devastating heavy rains, fires, flooding and strong winds caused by an upper-air cut-off low in October 2017 in Durban,[1] and regular heatwave and fire risk warnings in several provinces. One foremost cause of weather-related deaths is heat (leading to heatstroke), which is on the rise as towns and cities across South Africa shatter record all-time high temperatures year-on-year.[2] In January 2016, for example, the North-West Health Department reported that 11 people (aged 22–58 years) died over a period of 48 h in the province from heatstroke.[3] Deaths among groups such as the elderly, those with chronic disease and infants rise during heatwaves. These deaths may not be directly related to 'heat shock', but to the additional physiological stress associated with high heat. While there is some uncertainty as to whether these weather events are linked to climate change, the increase in heat levels generally, extreme weather events and current climate predictions for warming temperatures are strongly indicative that this is the case.

By the end of the 21st century, climate change in South Africa is predicted to result in, on average, temperatures 4 °C warmer than they are now.[4] Summers may become longer and warmer. Some areas of the country are projected to get drier while other parts may experience unseasonal rainfall. Even though there is some uncertainty around such climate predictions, preparing and adapting for climate change impacts is imperative.

Climate change will exacerbate the already high burden of disease in the country and create new public health challenges.[5] These challenges include respiratory diseases (from exposure to pollen, mould, smoke and particulate urban air pollution), infectious diseases, disaster-related injury and deaths, and environmental contamination that affects food crops and water supplies. The World Health Organization estimates that 250 000 additional deaths will occur globally each year due to direct and indirect climate change impacts.[6] If South Africa is going to brave the storm of climate change, as well as threats to morbidity and mortality posed by extreme weather events, our primary health care system needs to be robust, resilient, of high quality, and accessible.

South Africa has had some recent environmental health successes.[7] The percentage of households connected to electricity has increased from 77% in 2002 to 84% in 2017. Access to electricity lowers the use of solid fuel[8] combustion which is linked to chronic lung disease, asthma, fire hazards and carcinoma[9]. About 90% of South Africans had access to piped water in 2017 (versus 60% in 2002). These improvements in living conditions shore up resilience of households to climate change. Moreover, they may have contributed to a decline in total mortality from 1481.7/100 000 in 2002 to 1093.6/100 000 in 2015.[10] However, many people still live in unhealthy housing conditions while the public health sector struggles from problems including insufficient funding and inadequate staffing.[11] While access to primary health care has reportedly increased in recent decades, especially in rural areas, quality of care has not always kept pace.[12] In addition, South Africa faces a quadruple burden of disease and a mix of health system challenges.[13] These conditions, in a context of widespread poverty and extreme inequality, limit the chances of optimal adaptation to climate change, but also of implementing a National Health Insurance (NHI) scheme in a cost-effective manner.

In June 2017, the National Department of Health gazetted the White Paper on NHI for implementation by 2025. NHI, or universal health coverage, aims to meet four objectives[14]: (1) improve access to quality health services for all South Africans, irrespective of whether they are employed or not; (2) pool risks and funds so that equity and social solidarity will be achieved through creation of a single fund; (3) procure services on behalf of the entire population, and efficiently mobilise and control key financial resources; and (4) strengthen the under-resourced and strained public health sector to improve health systems performance. NHI aims to provide access to a uniform package of quality and affordable health services to South Africans based on needs, not on socio-economic status.

A fundamental goal of NHI is to eliminate inequality in access to primary health care. The inequality is evident from the fact that 82 out of 100 South Africans make use of public health care (with limited resources, poor management and limited service delivery) and private health care is used by 18 out of 100 South Africans.[15] Private health care is characterised by high and rising costs of health care and thus of medical scheme cover, disempowered and uninformed consumers, ineffective constraints on rising volumes of care, practitioners that are subject to little regulation and failures of accountability.[16] NHI will be implemented through the creation of a single fund that is publicly funded and publicly administered. The cost (in 2017 terms) to run NHI is estimated to be at least ZAR369 billion by 2025.[17]

NHI is anticipated to be the catalyst behind an improvement in primary health care services and infrastructure across the country; however, one of the World Health Organization's prerequisites for successful universal health coverage is an existing efficient health-care system which is lacking in South Africa.[18] Since 2015, initiatives have been taken to enhance the system[19] through ward-based community primary health care workers conducting home visits, mobile clinics servicing schools, and district clinical specialist teams reaching out to improve maternal and child health. Activities have been on a pilot basis and evaluations show mixed results.[19] There is also speculation around exactly how NHI will be implemented and in this regard we contend that, as NHI is being initiated, serious thought needs to be given to how climate change is likely to impact health and universal health coverage so as to plan for and mitigate the likely effects as soon as possible.

Possible direct and indirect impacts of climate change on NHI

Potential direct impacts on NHI and primary health care include extreme weather events on health service infrastructure such as heatwaves affecting functionality of medical equipment, changes to cold-chain requirements for transporting medicine and vaccines, thermal comfort in hospitals, and working conditions, productivity and staff well-being. Hotter conditions may constrain health workers' outreach work which often involves walking long distances for home visits. Flooding may interrupt water and power supplies, impede the ability of staff to get to work, affect the safety of staff and patients at health centres, and also jeopardise access to, or integrity of, systems for maintaining patient records.

The notion of universal health coverage provides the foundation needed to ensure health and well-being for all South Africans in a changing climate. As temperatures increase, the socio-economic effects of climate-related health impacts will ripple through communities, and especially affect children and infants susceptible to heat impacts[20] and child-headed households. High rainfall can lead to increased mould in people's homes, especially in poorly built dwellings with limited ventilation, increasing the likelihood of respiratory diseases and exacerbating conditions such as asthma and tuberculosis.[21] Changes in precipitation can also impact food production and food security, with possible implications for malnutrition, already a concern in South Africa.[22] Climate change health impacts are not mentioned in the White Paper on NHI but this should not suggest that health-care professionals in South Africa are unaware of them.

Ways in which NHI can consider climate change

If South Africans are to benefit from universal health coverage through NHI and simultaneously avoid health risks associated with climate change, then one fundamental NHI principle – *prevention of disease* – needs to be centre stage. This includes a focus on preventing climate-sensitive conditions, i.e. cholera and dengue fever, and pre-empting infection outbreaks. More stringent standards governing heat levels in occupational settings and promoting a set of interventions for outdoor workers are also key actions. A holistic approach founded on the essential tenets of public health is required to shape climate change responses and should include tracking diseases and trends; investigating disease outbreaks; informing policymakers about health impacts; creating partnerships – with industry and faith communities – to implement solutions; and ensuring health-care service provision following disasters.[23] Training of health-care providers on health aspects of climate change and research on optimal adaptation strategies are also needed.

Synergies between NHI principles and climate change exist: there are risks and opportunities to inform preparedness and help mitigate against adverse health impacts (Table 1).[24] Health promotion, preparedness and advocating health protective behaviours, e.g. drinking water in hot conditions, must be among the high priorities. Public awareness campaigns are required for socially isolated and marginalised groups. With a greater emphasis on prevention through behavioural change, especially in relation to environmental factors known to affect health, but also through vaccination, we could aim to reduce clinic and hospital visits during extreme weather events.

NHI is planned specifically to reduce the gap in the standard of health care between rich and poor, thereby reducing susceptibility of vulnerable groups. If this goal is to be achieved in a period of climate change, it will need to be carefully considered through social, spatial and economic lenses. Cities, where air pollution is already a concern, could experience worsening levels of air pollution when temperature rises, which may lead to greater prevalence of asthma, hence health-care facilities in cities may need to be well stocked with the appropriate medication.

Finally, sectors need to be integrated – for example, housing and settlements, labour/occupation and education – in order to cover social

Table 1: Health system building blocks, principle or element of National Health Insurance (NHI) and climate change: Identifying risks and opportunities to prepare and mitigate impacts

Health system building block	NHI principle	Climate change	
		Risk	Opportunity
Leadership / governance	Equity and health care as a public good	• Failure to work across sectors with Health in all policies results in burdens of morbidity and mortality, and cost-inefficiency	• Integrate with Disaster Risk Management, Housing and Settlements, Education etc. • Promote health-based guidelines and standards
Health workforce	Efficiency and effectiveness	• Staff succumbing to extreme weather events etc. • Staff unable to reach places of work during extreme weather events • Lowered staff productivity during heat waves • Need for more environmental health staff, and environmental health compliance checks as mitigation efforts rise	• Government increases numbers of health professionals graduating from universities and returns professionals from abroad who could assist in weather-related extreme events, for example • Climate and health considerations are included in curricula for all cadres of health workers
Access to essential medicines and technologies	Equity Efficiency Appropriateness	Public–private partnerships could be unsuccessful and not maximise opportunities to embrace new medical technologies	• Collaboration with NGOs to address preventive measures • Threats of increased climate-sensitive infectious diseases used to advocate for increased access to medicines
Health information systems	Health care as a public good	• Lack of preparedness in hospital and clinics for health impacts from extreme weather events	• Paper-based data system replaced with a quality electronic data system for monitoring and analysis • Data for engaging in research and modelling. Early-warning systems and better integration with South African Weather Service
Health service delivery	Right to access health care Effectiveness Appropriateness	• Facilities could experience damage to infrastructure, and water and power shortages • Facilities not equipped for thermal comfort • Vulnerable people remain left behind if climate change puts pressure on health systems despite NHI in place • Higher burden on the vulnerable, reducing long-term prospects for productive society • Impacts deepen inequality	• Health-care sector as key messengers for climate change impacts, to educate patients about the impacts of climate change and improve preparedness • NHI could increase hospital readiness and resilience for severe weather events, e.g. back-up generators, operable windows for ventilation • Health-care facilities' emissions reduced by switching to cleaner energy options for powering hospitals and clinics • NHI could reduce inequality and provide support to vulnerable groups

determinants of health in implementing NHI, not only for its success but also to alleviate threats of climate change on health, the need for health-care insurance and health-care delivery.[25,26] Social determinants of health[27] warrant special attention in all NHI promotion campaigns for disease prevention. The case study below illustrates the application of a climate change lens to prevailing social challenges to ensure a cost-effective NHI as the health hazard ramifications of climate change unfold.

Need for climate change considerations in all policies

In March 2018, 5-year-old Lumka Mkhetwa died in a pit toilet at a school in Bizana in the Eastern Cape Province. Her death followed that of young Michael Komape who met a similar fate at a school in Limpopo Province in 2014. Understandably, there was widespread public anger over these entirely preventable deaths and calls for urgent replacement of school pit latrines with 'proper' waterborne sewage systems. In an era of climate change, however, there is need for a longer-term view, and a broader perspective on child health and safety at schools.

Human faeces can be one of the most contaminated substances on earth. The purpose of any sanitation system is to remove faeces from human contact as quickly as possible. Most sanitation systems can achieve this if they are properly located, designed, constructed and maintained. Lumka and Michael died, not because they used a pit latrine, but because they had to use pit latrines that were not designed or appropriate for use by children.

During periods of drought, waterborne toilets constitute a major drain on available water supplies, and under certain circumstances, water shortages may render school toilets entirely dysfunctional. Consequences are potentially dire, with school toilets becoming a health hazard, a site for pest infestations, and a source of outbreaks of diarrhoeal diseases. It is possible that well-constructed pit latrines are preferable to waterborne toilets in areas with water scarcity, itself a rapidly expanding concern. But any school toilet must be designed with the health and safety of young children uppermost in mind; children should never be able to fall into pit latrines, no matter the extent of their natural curiosity, or any other circumstance.

Conclusions

It is evident that climate change will present substantial challenges, especially to the communities expected to benefit most from NHI. The aim of NHI is to deliver preventive, promotive, curative and rehabilitative health-care services with an emphasis on disease prevention and health promotion. It is the last two principles that hold particular synergy in adapting to climate change. By making climate change an integral consideration in planning and development it is possible to deliver an NHI that contributes more effectively to reducing inequalities that are likely to stem from evolving environmental hazards to health associated with climate change.

Acknowledgements

A.M. and C.Y.W. receive research support from the South African Medical Research Council and the National Research Foundation (South Africa). We thank Professors S. Fonn and J. Goudge for providing comments on an early draft of this manuscript.

References

1. Durban storm leaves death and destruction in its wake. News24. 2017 October 10. Available from: https://www.news24.com/SouthAfrica/News/live-massive-storm-hits-durban-causes-flooding-chaos-20171010
2. Johannesburg temperature hits record high as South Africa drought persists. Reuters. 2016 January 07. Available from: https://www.reuters.com/article/us-safrica-weather-idUSKBN0UL1F020160107
3. Heat stroke death toll rises to 11. SABC. 2016 January 10. Available from: https://www.enca.com/south-africa/heatstroke-death-toll-rises-11
4. Engelbrecht F, Adegoke J, Bopape MJ, Naidoo M, Garland R, Thatcher M, et al. Projects of rapidly rising surface temperatures over Africa under low mitigation. Environ Res Lett. 2015;10, 052004, 16 pages. https://doi.org/10.1088/1748-9326/10/8/085004
5. US EPA. Climate change indicators: Understanding the connections between climate change and human health [webpage on the Internet]. c2017 [cited 2018 Sep 27]. Available from: https://www.epa.gov/climate-indicators/understanding-connections-between-climate-change-and-human-health
6. World Health Organization. Climate change and health – Key facts [webpage on the Internet]. c2018 [cited 2019 Aug 23]. Available from: http://www.who.int/news-room/fact-sheets/detail/climate-change-and-health
7. Statistics South Africa. Statistical release P0318: General Household Survey 2017 [document on the Internet]. c2018 [2019 Aug 23]. Available from: http://www.statssa.gov.za/publications/P0318/P03182017.pdf
8. Ye Yuxiang, Koch SF, Zhang J. Determinants of household electricity consumption in South Africa. Energy Econ. 2018,75:120–133. https://doi.org/10.1016/j.eneco.2018.08.005
9. World Health Organization. Household air pollution and health [webpage on the Internet]. c2018 [cited 2019 Aug 23]. Available from: http://www.who.int/en/news-room/fact-sheets/detail/household-air-pollution-and-health
10. World Health Organization. WHO Mortality database [webpage on the Internet]. c2018 [cited 2019 Aug 23]. Available from: http://www.who.int/healthinfo/mortality_data/en/
11. Moyimane MB, Matlala SF, Kekana MP. Experiences of nurses on the critical shortage of medical equipment at a rural district hospital in South Africa: A qualitative study. Pan Afr Med J. 2017;28:100. https://doi.org/10.11604/pamj.2017.28.100.11641
12. Eygelaar JE, Stellenberg EL. Barriers to quality patient care in rural district hospitals. Curationis. 2012;35(1), Art. #36, 8 pages. https://doi.org/10.4102/curationis.v35i1.36
13. Pillay-van Wyk V, Dorrington RE, Bradshaw D. Rapidly changing mortality profiles in South Africa in its nine provinces. S Afr Med J. 2017;107(3):168–169. https://doi.org/10.7196/samj.2017.v107i3.12344
14. HealthMan. National Health Insurance White Paper 2017 Summary [document on the Internet]. c2017 [cited 2019 Aug 23]. Available from: https://www.mm3admin.co.za/documents/docmanager/1E9AEA2C-B58D-4AED-B5A2-96187D705AEE/00126403.pdf
15. Statistics South Africa. Public healthcare: How much per person? [webpage on the Internet]. c2017 [cited 2019 Aug 23]. Available from: http://www.statssa.gov.za/?p=10548
16. Competition Commission South Africa. Health market inquiry provisional findings and recommendations report [document on the Internet]. c2018 [cited 2019 Aug 23]. Available from: http://www.compcom.co.za/wp-content/uploads/2018/07/Health-Market-Inquiry-1.pdf
17. NHI – a healthy way to go? IOL Personal Finance. 2018 March 06. Available from: https://www.iol.co.za/personal-finance/nhi-a-healthy-way-to-go-12651180
18. World Health Organization. What is universal health coverage? Geneva: World Health Organization; 2014. https://doi.org/10.2471/blt.08.060046
19. Choonara S, Eyles J. National Health Insurance: South Africa's most progressive reform to date. S Afr Labour Bull. 2016;40(1):31–34. Available from: http://ir.nrf.ac.za/handle/10907/2069
20. Stanberry LR, Thomson MC, James W. Prioritising the needs of children in a changing climate. PLoS Med. 2018;15(7), e1002627, 4 pages. https://doi.org/10.1371/journal.pmed.1002627
21. Foster N, Vassall A, Cleary S, Cunnama L, Churchyard G, Sinanovic E. The economic burden of TB diagnosis and treatment in South Africa. Soc Sci Med. 2015;130:42–50. https://doi.org/10.1016/j.socscimed.2015.01.046
22. Tumushabe JT. Climate change, food security and sustainable development in Africa. In: Oloruntoba S, Falola T, editors. The Palgrave handbook of African politics, governance and development. New York: Palgrave Macmillan; 2017. p. 853–868. https://doi.org/10.1057/978-1-349-95232-8_53
23. Messery M. Insurer Climate Risk Disclosure Survey report and scorecard: 2016 Findings and recommendations [webpage on the Internet]. c2016 [cited 2019 Aug 2019]. Available from: https://www.ceres.org/resources/reports/insurer-climate-risk-disclosure-survey-report-scorecard
24. Centers for Disease Control and Prevention (CDC). The public health system and the 10 essential public health services [webpage on the Internet]. c2019 [cited 2019 Aug 23]. Available from: https://www.cdc.gov/stltpublichealth/publichealthservices/essentialhealthservices.html
25. National Department of Health. National Health Insurance for South Africa – Towards universal health coverage [document on the Internet]. c2015 [cited 2019 Aug 23]. Available from: https://www.gov.za/sites/default/files/National_Health_Insurance_White_Paper_10Dec2015.pdf
26. Curtis S, Fair A, Wistow J, Val DV, Oven K. Impact of extreme weather events and climate change for health and social care systems. Environ Health. 2017;16(suppl.1):128. https://doi.org/10.1186/s12940-017-0324-3
27. Marmot M. Social determinants of health inequalities. Lancet. 2005;365:1099–1104. https://doi.org/10.1016/s0140-6736(05)71146-6

AUTHORS:
Cheledi Tshehla[1,2]
Caradee Y. Wright[1,3]

AFFILIATIONS:
[1]Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa
[2]South African Weather Service, Pretoria, South Africa
[3]Environment and Health Research Unit, South African Medical Research Council, Pretoria, South Africa

CORRESPONDENCE TO:
Cheledi Tshehla

EMAIL:
Cheledi.Tshehla@weathersa.co.za

HOW TO CITE:
Tshehla C, Wright CY. 15 Years after the *National Environmental Management Air Quality Act*: Is legislation failing to reduce air pollution in South Africa? S Afr J Sci. 2019;115(9/10), Art. #6100, 4 pages. https://doi.org/10.17159/sajs.2019/6100

ARTICLE INCLUDES:
☐ Peer review
☒ Supplementary material

KEYWORDS:
human health, environmental health, regulation, guidelines

PUBLISHED:
26 September 2019

15 Years after the *National Environmental Management Air Quality Act*: Is legislation failing to reduce air pollution in South Africa?

Air pollution is characterised by the presence of chemicals or compounds in the air which are usually not present or are present at levels higher than those considered to be safe for human health.[1] Air pollution is the main cause of environmental effects such as acid rain (formed primarily by nitrogen oxides and sulfur oxides in the atmosphere) which can acidify soil and water bodies leading to a threat on food security; and ground-level ozone which is responsible for destruction of agricultural crops and commercial forests.[2] Air pollution can cause detrimental changes to the quality of life. According to the World Health Organization, air pollution is one of the greatest environmental threats to human health that can lead to increased mortality and morbidity. Pollutants mostly associated with health effects are particulate matter, ozone, sulfur dioxide and nitrogen dioxide.[3]

Efforts have been made locally through the transition in legislation from the *Atmospheric Pollution Prevention Act (APPA) Number 45 of 1965* (focused on air pollution emitters) to the *National Environmental Management Air Quality Act (NEMAQA) No. 39 of 2004* to not only reduce emissions of air pollutants but also to monitor effects of air pollution on the environment. National Ambient Air Quality Standards for pollutants such as particulate matter (PM_{10}), lead (Pb), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO), ozone (O_3), and benzene (C_6H_6) were gazetted in 2009[4], with $PM_{2.5}$ gazetted in June 2012[5]. However, the problems associated with air pollution are far from being solved, particularly with the observed levels of particulate matter and ozone in areas declared as hotspots (Priority Areas) in South Africa. The main sources of particulate matter in these areas have been identified as industry, mining, motor vehicles, and biomass and domestic burning.[6-8] Ground-level ozone is formed as a result of photochemical reactions in the atmosphere in the presence of sunlight.[9] The current approach to implementation of air quality legislation to reduce air pollution may be inadequate considering evidence of negative impacts and risks.[10,11]

Effective management of air quality in South Africa will require sound policy implementation, air quality monitoring and the enforcement of legislation and standards. Cooperation between government departments, economic sectors, research institutions and the public is of great importance in the battle against air pollution. The political buy-in of all spheres of government (municipal, provincial and national) is needed to ensure that environmental issues are at the top of the agenda in every sitting of the legislator to ensure that environmental programmes are allocated enough attention and appropriate resources. A published study has shown that a direct positive effect of democratic institutions on environment quality is higher in developed countries than in developing countries.[12]

The aim of this Commentary is threefold: (1) to provide an overview of the current NEMAQA legislative instruments for air pollution prevention; (2) to consider the current state of NEMAQA implementation approaches; and (3) to reflect on future approaches for effective implementation of NEMAQA and ultimate reduction of air pollution in South Africa.

Air pollution and its management

Air pollutants are solid particles, gases and liquid droplets in the air that can adversely affect ecosystems and human health.[13] Major ambient air pollutants include toxic metals, volatile organic compounds, PM_{10}, $PM_{2.5}$, NOx, SO_2, O_3 and CO[3]. Air pollutants are classified according to the source of emission into two main groups: primary and secondary pollutants. Primary air pollutants are emitted directly into the air from sources. They can have effects both directly and as precursors of secondary air pollutants (such as O_3, NO_3^-, SO_4^{2-}, H_2SO_4) which are formed by chemical reactions in the atmosphere.[14] Air pollutants can be emitted by natural sources such as wildfires, volcanic activities and crustal materials as well as anthropogenic activities such as power plants, smelters, mines, vehicles and domestic wood and coal burning.[15] The distribution of these pollutants is dependent on meteorological conditions.[16]

NEMAQA specifies that to reduce and manage air quality there needs to be:

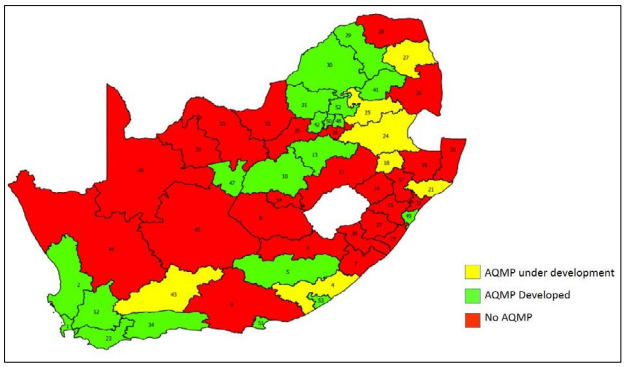
- decentralisation of air quality management among all spheres of government;
- identification and quantification of all sources;
- compliance monitoring and enforcement;
- setting of ambient and emissions standards;
- development of Air Quality Management Plans (AQMPs) by all spheres of government and emissions reductions and management plans by all source emitters;
- access to information and public consultation; and
- norms and standards for air quality monitoring and management.

South Africa's national Department of Environmental Affairs (DEA) is mandated to develop, review and revise systems and procedures for attaining compliance with Air Quality Standards in South Africa. The provincial DEAs must monitor ambient air quality in their provinces as well as the performance of municipalities in implementing the *Air Quality Act*. Local authorities are required, in terms of the *Air Quality Act* (Section 8(a)), to monitor ambient air quality and emissions from point, non-point and mobile sources. Therefore, authorities must study emissions reports from licensed emitters to ensure that they comply with the conditions of their licences. The municipal by-laws should

also be structured in such a way that they address emissions from small industries that are not regulated by atmospheric emission licences.

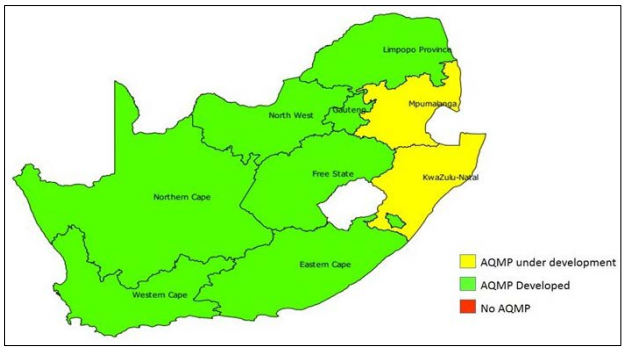
Consideration of progress to date

The evidence collected during the development of the Mpumalanga Provincial AQMP suggests that decentralisation of air quality management in South Africa is not effectively managed, with most local authorities not performing this function due to several constraints (Supplementary tables 1–4). The *Municipal Systems Act No. 32 of 2000* requires local authorities to include their AQMPs in their Integrated Development Plans for resource allocation. However, only a few district municipalities have AQMPs in place, as shown in Figure 1. All provinces have developed their AQMPs with the exception of KwaZulu-Natal and Mpumalanga which are still under development (Figure 2). For the three areas that have been declared Priority Areas (that is, areas where the ambient air quality standards are being, or may be, exceeded, or any other situation which is causing, or may cause, a significant negative impact on air quality in the area) in terms of Section 18(1) of the *Air Quality Act*, processes are underway to implement the Priority Area AQMPs. Given that the status quo has not changed, there could be underlying problems in the implementation of the AQMPs.



Source: South African Air Quality Information System[22]

Figure 1: Current development status of district municipality Air Quality Management Development Plans in South Africa.



Source: South African Air Quality Information System[22]

Figure 2: Current development status of provincial Air Quality Management Development Plans in South Africa.

In terms of identification and quantification of sources of air pollutants, no formal study has been undertaken solely by any of the three spheres of government to identify and apportion sources of air pollutants. In 2016, a source apportionment study was commissioned by the DEA for the Vaal Triangle Airshed Priority Area (VTAPA).[17] This study was commissioned after the 2013 mid-term AQMP review revealed that there was no improvement in air quality in the VTAPA.[18] Given the unchanging status quo on air pollution levels in the Priority Areas, the source apportionment studies are necessary for effective implementation of abatement strategies. Air pollution has no boundaries; therefore, air pollution from the Priority Areas could also be transported to other regions in the country. As such, a directive should also have been issued to all provincial DEAs to commission source apportionment studies in their respective provinces for effective air quality management.

The establishment of the Environmental Management Inspectorate to monitor and enforce compliance with the *National Environmental Management Act* was to ensure that all those undertaking activities that may lead to detrimental effects on the environment are held accountable. However, the regulatory authorities either are not scrutinising the compliance reports from industries to look at the root-cause analyses, not following up on non-compliance or not doing trends analyses of the reports to check consistency. In a number of DEA Implementation Task Team meetings, industries have mentioned that they submit reports to regulatory authorities but their reports are not attended to nor do they receive feedback from authorities. The response from authorities was that there are not enough personnel to study the reports. This situation presents a potential loophole that industries may have identified within the regulatory framework and, as such, provides an opportunity for non-compliance with the minimum emissions standards by industries.

According to the South African Government website, there are 278 municipalities in South Africa.[19] Supplementary table 5 shows the 121 government-managed ambient air quality monitoring stations that have been reporting to the South African Air Quality Information System (SAAQIS) since 2009 when the system was launched. If all the municipalities were implementing Section 8(b) of the *Air Quality Act*, there would be 278 ambient air monitoring stations reporting to SAAQIS (i.e. one station per municipality) to ensure widespread spatial coverage necessary for effective compliance monitoring. The DEA and South African Weather Service have embarked on a project to revamp SAAQIS through the development and implementation of SAAQIS phase III. This improvement has led to live reporting of air quality monitoring stations. The DEA is in the process of reviving stranded municipal air quality monitoring stations to increase the number of stations reporting to SAAQIS. Privately owned stations are also targeted to report live on the SAAQIS system. New features on SAAQIS phase III include a mobile app to view live ambient data and the air quality index, and the ability to download data on the public page.

National norms and standards are required to ensure that there is standardised ambient air quality monitoring and management in the country. But to date, there is no document that outlines the national norms and standards for air quality management in South Africa that can be used as a directive to all regulating authorities and the source emitters. With a total of 164 monitoring stations reporting to SAAQIS, only a few are considered to have credible data that can be used for scientific research to inform policy development. As shown in Supplementary tables 2–4, the reason for non-credible data can be attributed to lack of funding and skilled personnel to perform air quality related functions. This situation is a matter of concern given the non-compliance with the National Ambient Air Quality Standards shown in Supplementary tables 6–8.

There has not been a cost–benefit analysis undertaken by the government since the promulgation of the *Air Quality Act* in 2004 to determine the impact of air pollution on the economy of South Africa. Some industries in South Africa have been applying for the postponement of minimum emissions standards for several years. Furthermore, there has not been projected cost–benefit analyses on how much will be saved by the country if air pollution was reduced to acceptable limits by complying with the minimum emissions standards and ambient air quality standards. Eskom undertook a cost–benefit analysis[20] of the offset (defined as an intervention to counterbalance an adverse environmental impact) project in the Highveld Priority Area. However, the focus of this analysis was on reducing pollution from households and compliance with the 2020 NEMAQA minimum emission standards was not considered. A report[21] by the Centre for Environmental Rights and groundWork indicates that the initiatives undertaken by Eskom to reduce household emissions from coal burning in Zamokuhle Township through air quality offset interventions did not bear positive results because of the high cost of electricity. The high cost of liquid petroleum gas in South Africa is also a negative factor in the proposed Eskom retrofit project and will lead to communities reverting to coal use for space heating and cooking.

In terms of mining, the Department of Mineral Resources is responsible for issuing Atmospheric Emission Licences and granting environmental authorisations. Environmental authorisations are a key tool in effective environmental management, including the management of air quality. However, the Department of Mineral Resources officials are not designated as Air Quality Officers and as such air quality related matters may not be fully explored during the authorisation process. This arrangement makes the management of mining-related air pollution very difficult in South Africa.

There is a gap between science and policy

Supplementary tables 6–8 show the 2018 National Ambient Air Quality Standards exceedance tables for the areas which have been declared pollution hotspots in South Africa: VTAPA declared in 2006, Highveld Priority Area declared in 2007 and Waterberg/Bojanala Priority Area declared in 2012. The VTAPA AQMP identified the main sources of air pollution in the area as biomass burning, domestic fuel burning, mining operations, petrochemical sector, power generation, transportation, waste burning, iron and steel and ferroalloy industries, and smaller industries. The Highveld and Waterberg/Bojanala Priority Area AQMPs identified the major sources of air pollution as residential fuel burning, coal mining, power generation, transport, biomass burning and burning coal mines and smouldering coal dumps, landfills, incinerators, waste treatment works, tyre burning, agricultural dust, and biogenic sources.[7,8] The variation in source categories in Priority Areas clearly shows that these sources will be complex to manage and will require multi-stakeholder partnership in implementation of abatement strategies. It is evident from the exceedance data that there is a problem with particulate matter and ozone in all the areas. However, there have not been any studies commissioned by the DEA to comprehensively identify sources of particulate matter and ozone (except in the VTAPA) and there are no known memoranda of understanding between DEA and research institutions to develop and fund programmes aimed at tackling this research gap.

The *Air Quality Act* requires new Atmospheric Emission Licence applicants to undertake an atmospheric emissions modelling study. Many air quality dispersion models rely on surface meteorological parameters to model air pollution dispersion, particularly in complex terrain. Section 4.2.1 of the draft regulations regarding Air Dispersion Modelling (Notice 1035 of 2012) in the *Air Quality Act* requires site-specific meteorological data for modelling purposes in complex terrain. However, there has not been any collaboration between the DEA and the South African Weather Service to ensure that there is a sufficient number of surface meteorological monitoring stations in remote areas with complex terrain. One such place is the Greater Tubatse Municipality which has several industrial facilities and a complex terrain but no meteorological stations. (A Research Article in this issue reports on air pollution in the Greater Tubatse Municipality). Institutional collaborations between government entities and research institutions may narrow the gap between science and strategic policy development and implementation for successful management of air quality. Air pollution reduction could be achieved by strengthening collaboration between government departments such as DEA and Department of Mineral Resources for better management of pollution from the mining sector; and allocating funding for environmental issues at all spheres of government to be centralised at DEA for better management of air quality. It could also entail developing a cost–benefit study for the implementation of the *Air Quality Act*; and making source apportionment a pre-requisite for the development of air quality management plans by authorities and for all industry applications for postponement of complying with the minimum emission standards by April 2020, and for atmospheric emission licence application for new facilities. The source apportionment and source quantification results will ensure that the contribution of major sources, as well as the impact that results from granting postponements and/or new licences, will be known. Establishing expert panels to identify research programmes aimed at addressing air pollution problems would also be beneficial, as it would ensure that resources are channelled to research studies that are relevant to air quality improvement. Lastly, air pollution programmes should be introduced from the foundation phase of basic education to build a nation that is conscious of and educated about air quality issues.

Conclusions

The *Air Quality Act* was passed in South Africa over 15 years ago, but it is evident that several of its strategic objectives have yet to be met. Even though emissions reduction is implemented by some industries, and there also are efforts by local authorities to develop and implement by-laws to reduce household emissions, the introduction of new small industries, and the failure to effectively reduce pollution from domestic burning, waste burning, biomass burning, vehicle emissions and mining activities within the air pollution hotspots makes it impossible to achieve the desired air pollution reduction. Particulate matter and ozone are two pollutants for which there is non-compliance with the National Ambient Air Quality Standards. Therefore a comprehensive study to look at the major precursors of ozone is necessary to develop abatement strategies for ozone. There is a need to relook at the drivers and factors influencing policy implementation such as political buy-in (by educating politicians on air quality matters) particularly in local authorities and reprioritisation of societal needs, especially with respect to housing and economic development in relation to protection of the environment and human health.

Acknowledgements

C.Y.W. receives research funding support from the South African Medical Research Council and the National Research Foundation (South Africa). C.T. thanks the South African Weather Service for provision of resources, space and time for conducting the research.

References

1. Sharma SB, Jain S, Khirwadka P, Kulkarni S. Effects of air pollution on the environment and human health. Indian J Res Pharm Biotechnol. 2013;1(3):391–396. Available from: http://www.ijrpb.com/issues/Volume%201_Issue%203/ijrpb%201(3)%2020%20page%20391-396.pdf
2. Ashfaq A, Sharma P. Environmental effects of air pollution and application of engineered methods to combat the problem. J Ind Pollut Control. 2013;29(1):19–29. Available from: http://www.icontrolpollution.com/articles/environmental-effects-of-air-pollution-and-application-of-engineered-methods-to-combat-the-problem-.php?aid=45739
3. World Health Organization (WHO). Ambient air pollution: A global assessment of exposure and burden of disease. Geneva: WHO; 2016. Available from: https://apps.who.int/iris/handle/10665/25014
4. South African Department of Environmental Affairs (DEA). National ambient air quality standards [document on the Internet]. c2009 [cited 2019 Feb 04]. Available from: https://www.environment.gov.za/sites/default/files/legislations/nemaqa_airquality_g32816gon1210.pdf
5. Western Cape Department of Environmental Affairs and Development Planning. State of Air Quality Management 2015 [document on the Internet]. c2015 [cited 2019 Feb 04]. Available from: https://www.westerncape.gov.za/eadp/files/atoms/files/State%20Of%20Air%20Quality%20Monitoring%20Report%202015.pdf
6. South African Air Quality Information System (SAAQIS). The Vaal Triangle Air-Shed Priority Area Network AQMP. Pretoria: South African Weather Service; 2009. Available from: http://www.saaqis.org.za/documents/Vaal%20Triangle%20Air-Shed%20Priority%20Area%20(VTAPA)%20AQMP%20Regulations_29-05-2009.pdf
7. South African Air Quality Information System (SAAQIS). Declaration of the Highveld Priority Area. Pretoria: South African Weather Service; 2007. Available from: http://www.saaqis.org.za/documents/HIGHVELD%20PRIORITY%20AREA%20AQMP.pdf
8. South African Department of Environmental Affairs (DEA). Declaration of the Waterberg Bojanala Priority Area. Pretoria: South African Weather Service; 2013. Available from: https://www.environment.gov.za/sites/default/files/gazetted_notices/nemaqa_waterberg_declaration_g35435gen495_0.pdf
9. World Bank Group. Ground-level ozone. Washington DC: World Bank; 1998. Available from: https://www.ifc.org/wps/wcm/connect/dd7c9800488553-e0b0b4f26a6515bb18/HandbookGroundLevelOzone.pdf?MOD=AJPERES

10. Maponya P, Rampedi IT. Impact of air pollution on maize production in the Sasolburg Area, South Africa. J Agric Sci. 2013;5(11):181–188. https://doi.org/10.5539/jas.v5n11p181

11. Keen S, Altieri K. The health benefits of attaining and strengthening air quality standards in Cape Town. Clean Air J. 2016;26(2):22–27. https://doi.org/10.17159/2410-972X/2016/v26n2a9

12. Romualt KS. Democratic institutions and environmental quality: Effects and transmission channels. Paper presented at: EAAE 2011 Congress on Change and Uncertainty Challenges for Agriculture, Food and Natural Resources; 2011 August 30 – September 02; Zurich, Switzerland. Available from: https://ageconsearch.umn.edu/bitstream/120396/2/Somlanare_Kinda_354.pdf

13. Bhatia SC. Environmental pollution and control in chemical process industries. Delhi: Khanna Publishers; 2001; p. 163.

14. Kicinski M, Nawrot TS. Neurobehavioral effects of air pollution in children. In: Costa L, Aschner M, editors. Environmental factors in neurodevelopmental and neurodegenerative disorders. London: Academic Press; 2015. p. 89–105. https://doi.org/10.1016/B978-0-12-800228-5.00005-4

15. Singh OP. Air pollution: Types, sources and abatement. Environment and natural resources: Ecological and economic perspective. New Delhi: Regency Publications; 2015. p. 101–124.

16. Ediagbona TF, Ukpebor EE, Okieimen FE. Correlation of meteorological parameters and dust particles using scatter plot in a rural community. Ife J Sci. 2013;15(3):445–453.

17. South African Air Quality Information System (SAAQIS). [National air quality officer's annual report on air quality management [document on the Internet]. c2015 [cited 2019 Feb 07]. Available from: https://saaqis.environment.gov.za/Pagesfiles/NAQO%27s_Annual%20Report_2015.pdf

18. South African Department of Environmental Affairs (DEA). Vaal Triangle Airshed Priority Area source apportionment study: Preliminary results [document on the Internet]. c2018 [cited 2019 Feb 13]. Available from: http://www.airqualitylekgotla.co.za/assets/2018_5.4-vtapa-source-apportionment-study.pdf

19. South African Government. [Local government]. c2019 [cited 2019 Feb 16]. Available from: https://www.gov.za/about-government/government-system/local-government

20. Eskom Holdings SOC Limited. The execution of a household emission offset pilot study in the Highveld Priority Area, Mpumalanga [document on the Internet]. c2015 [cited 2019 Feb 20]. Available from: http://www.eskom.co.za/AirQuality/Documents/PilotStudySemi-FinalReport.pdf

21. groundWork. Government: Wheels come off the Eskom offset [webpage on the Internet]. c2018 [cited 2019 Feb 22]. Available from: http://www.groundwork.org.za/archives/2018/news20180814-Wheels_come_off_the_Eskom_offset.php

22. South African Air Quality Information System (SAAQIS). AQ Planning [webpage on the Internet]. No date [2019 Jan 23]. Available from: https://saaqis.environment.gov.za//home/text/358).

AUTHORS:
Cheledi Tshehla[1,2]
Caradee Y. Wright[1,3]

AFFILIATIONS:
[1]Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa
[2]South African Weather Service, Pretoria, South Africa
[3]Environment and Health Research Unit, South African Medical Research Council, Pretoria, South Africa

CORRESPONDENCE TO:
Cheledi Tshehla

EMAIL:
Cheledi.Tshehla@weathersa.co.za

DATES:
Received: 28 Mar. 2019
Revised: 14 May 2019
Accepted: 05 June 2019
Published: 26 Sep. 2019

HOW TO CITE:
Tshehla C, Wright CY. Spatial variability of PM_{10}, $PM_{2.5}$ and PM chemical components in an industrialised rural area within a mountainous terrain. S Afr J Sci. 2019;115(9/10), Art. #6174, 10 pages. https://doi.org/10.17159/sajs.2019/6174

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Priscilla Baker

KEYWORDS:
complex terrain, passive samplers, mixing height, air pollution potential

FUNDING:
None

Spatial variability of PM_{10}, $PM_{2.5}$ and PM chemical components in an industrialised rural area within a mountainous terrain

We describe the measurement and spatial variability of particulate matter (PM) chemical composition, PM_{10} and $PM_{2.5}$ in the Greater Tubatse Municipality, South Africa. Monthly samples were collected over 12 months (July 2015 to June 2016) using the inexpensive and easy to operate passive samplers of the University of North Carolina. Sites for sample collection were located at private residences, a church, a hospital and a school. Concentrations of PM_{10}, $PM_{2.5}$ and PM chemical components were determined using computer-controlled scanning electron microscopy with energy-dispersive X-ray spectroscopy. The annual observed concentrations at all sites were below the South African National Ambient Air Quality Standards of 40 μg/m^3 for PM_{10} and 25 μg/m^3 for $PM_{2.5}$. The Cr-rich and CrFe-rich particles showed substantial heterogeneity with high concentrations observed near the chrome smelters, and Si-rich particles were highest near the silicon mine. SiAl-rich particles were highest at sites close to busy roads, while SiAlFe-rich particles were less spatially distributed. The low spatial variability of SiAlFe-rich particles indicates that these elements are mainly found in crustal material. Using the synoptic meteorological parameters of The Air Pollution Model, we were unable to effectively determine correlations between PM_{10} and mixing height, Monin–Obukhov length, air pollution potential, or coefficient of divergence.

Significance:

- We have shown that the use of University of North Carolina passive samplers coupled with computer-controlled scanning electron microscopy is effective in determining the chemical composition of PM.
- The use of passive samplers is a cheap and effective method to collect data in remote areas of South Africa which have limited or no electricity supply.
- Assessment of the spatial distribution of PM and PM chemical components can assist in the development of effective air quality management strategies.

Introduction

Airborne particulate matter (PM) is a term used to describe solid particles or a mixture of solid and liquid droplets suspended in the air.[1] The particle mixture may vary in size distribution, composition and morphology and may be in the form of sulfates, nitrates, ammonium and hydrogen ions, trace elements (including toxic and transition metals), organic material, elemental carbon (or soot) and crustal components.[2,3] PM may originate from either primary or secondary sources. Primary particles are those directly emitted into the atmosphere from sources such as road vehicles, coal burning, industry, windblown soil, dust and sea spray. Secondary particles are particles formed within the atmosphere by chemical reactions or condensation of gases. The major contributors of secondary particles are sulfate and nitrate salts formed from the oxidation of sulfur dioxide and nitrogen oxides, respectively.[4] Ambient PM has long been associated with adverse effects on respiratory, cardiovascular and cardiopulmonary health.[5-7] The severity of such health effects depends largely on the size, concentration and composition of inhaled particles.[8] PM pollution emanating from industrialisation has serious environmental impacts mainly because of the release of toxic substances and trace metals into the atmosphere.[9]

Industrialisation and urbanisation of rural areas can lead to the emission of large amounts of PM and chemical elements into the atmosphere. These emissions result in widespread air pollution problems[10], and these problems have proved to be more regional and complex with time[11]. The Greater Tubatse Municipality (GTM) in South Africa is home to a large number of people and a variety of anthropogenic pollution sources such as chrome smelters, mines (for chrome, silicon and platinum), agricultural operations, biomass combustion, brick manufacturing, vehicles and unpaved roads, which can contribute to PM emissions. Differences in the composition of particles emitted by these sources may lead to spatial heterogeneity in the composition of the atmospheric aerosols. Hence, understanding the spatial variability of PM is of great importance for environmental planning and management purposes by both the industries and governing authorities. Therefore, this study will lay a foundation for developing effective intervention strategies to reduce PM emissions in the GTM. In South Africa, PM is only regulated in two size fractions (PM_{10} and $PM_{2.5}$). However, to date, there are no ambient air quality standards for elemental particles. The list of metals regulated under the *National Environmental Management: Act No. 39 of 2004* should be expanded to include metals such as chromium, iron, arsenic, copper, cobalt, manganese and other metals that have been identified[12] to have the potential to cause environmental health threats.

Apart from air pollution challenges due to anthropogenic activities, South Africa has a varying topography ranging from flat to complex terrain that can have differing effects on the dispersion of air pollutants. The shape of the landscape plays an important role in trapping or dispersing pollutants. Air pollution in mountain valleys tends to be higher in colder months than in warmer months.[13] The distribution of pollutants depends largely on the meteorology and the landscape of the area. Surface heterogeneity plays a major role in the interaction between the atmosphere and the underlying surface, and it affects moist convection, and systematically produces responses in both local circulation and regional climate.[14-18] Complex terrain such as that of the GTM is characterised by high mountains and steep inclinations. In this

type of terrain, the wind flow is very hard to predict. However, the steep slopes give rise to thermally induced circulations like mountain valley breezes which strongly modify the characteristics of synoptic flow.[19-22] The ability of the atmosphere to disperse pollutants depends on the local circulations, mixing height, stability of the atmosphere and wind strength. However, the complex nature of the terrain in the GTM and the lack of electricity supply in some areas of the municipality, makes it impossible to rely only on a network of continuous ambient air pollution monitoring.

A number of methods have been developed over the years to collect and analyse air pollutant samples, using both active and passive techniques. The passive sampling techniques involve non-active means such as gravitational settling to collect air samples onto the substrate. This method of sampling is cheaper than active sampling and allows for the deployment of more samplers to evaluate air pollution spatially.[23] The GTM has only one air quality monitoring station that is not sufficiently well maintained to produce good quality data. As a result, a network of passive samplers was used to determine the spatial variation of $PM_{2.5}$ and PM_{10}, which in future can be used as a baseline for the deployment of active samplers in the area.

Mountain winds

Wind circulations in the free atmosphere above the mountains and valleys are governed by pressure gradients between large circulation systems.[24] The lower troposphere interacts with mountains, valleys and vegetation that in turn alter the circulation patterns. Mountainous terrain has a high degree of topographical variation and land-cover heterogeneity.[25] This variation in topography influences the atmosphere in two ways.[26] The first is in the form of momentum exchange between the atmosphere and the surface that occurs as a result of flow modification by mountains in the form of mountain lee waves, flow channelling and flow blocking.[27] The second effect involves energy exchange between the terrain and the atmosphere. The thermally induced winds depend on the temperature differences along the mountain plains systems and the strength of the synoptic systems and the cloud cover, with weak synoptic systems and cloud-free atmosphere producing more pronounced winds.[20,28] Mountain winds blow parallel to the longitudinal axis of the valley, directed up-valley during daytime and down-valley during night time. The circulation is closed above the mountain ridges by a return current flowing in the reverse direction. The actual development of thermally driven winds is often complicated by the presence of other wind systems developed on different scales.[22,28] Anabatic flows are more temporally limited during wintertime than summertime due to the shorter exposure period to sunlight.[29]

Mixing height

Mixing height (MH) is the height to which relatively vigorous mixing occurs in the lower troposphere. Temperature inversions are most common in mountainous terrain where cool mountain air sweeps down into the valley at night, below the warm, polluted air. This inversion keeps the emitted pollutants close to the ground instead of allowing them to disperse into the atmosphere. A flow of thermal or synoptic origin channelled inside a mountain valley can transport plumes along the valley floor, thus limiting crosswind dispersion. Pollution stagnation in the bottom of the valleys can be favoured by the temperature inversion that develops inside the valley during the night and is destroyed by the growing convective boundary layer in the morning.[30] The thermally induced MH influences the concentration and transport of pollutants[31], and is used in air quality models to determine atmospheric pollutant dispersion[32-34]. However, in mountainous terrain, processes such as MH and mountain slope winds are coupled together[35] to transport air pollutants across mesoscales to synoptic scales[36]. Research by De Wekker and Kossmann[27] has illustrated that the dispersion of pollutants in mountainous terrain does not depend on the boundary layer but rather on the thermally induced mountain slope winds.

Monin–Obukhov length

The Monin–Obukhov (MO) similarity theory has been applied in air pollution modelling for determining the dispersion of air pollutants. The MO measures the stability of the atmosphere, with stable atmospheric conditions favouring higher pollutant concentrations and unstable conditions allowing the dispersion of pollutants and hence lowering pollutant concentrations.[37] However, the MO is restricted to horizontal homogeneous terrains where there are no sudden roughness changes (such as in forested area, hilly or mountainous terrain) to modify the velocity profile and turbulent transport of heat and momentum.[38] Figueroa-Esspinoza and Salles[38] and Grisogono et al.[39] reported that MO theory is unable to account for the transport of pollutants in mountain valleys because the flow dynamics of the valleys are governed by anabatic and katabatic flows. These flows are generated by the mountain slopes and are normally decoupled from the synoptic flows above.

Ventilation coefficient

Gross[40] defined the ventilation coefficient (VC) as the product of the MH and the average wind speed, which can also be defined as a measure of the volume rate of horizontal transport of air within the MH per unit distance normal to the wind. Iyer and Raj[41] describe the VC as a measure of the atmospheric condition that gives an indication of the air quality and air pollution potential. When the coefficient is higher, it is an indication that the atmosphere is able to disperse air pollutants effectively, resulting in a better state of air quality, whereas low ventilation indicates poor pollutant dispersion resulting in high pollution levels. The VC varies diurnally during summer and winter with high coefficients observed in the late afternoon and low values in the early mornings. Winter coefficients are also lower than those in summer due to low MH and reduced wind speeds in winter,[42,43] and the influence of the dominant anti-cyclones that are experienced over southern Africa during the winter months.

Air pollution potential

Gross[40] and Nath and Patil[44] describe air pollution potential (APP) as the measure of the inability of the atmosphere to adequately dilute and disperse pollutants emitted into it. The APP depends on meteorological conditions such as the MH, wind speed, atmospheric stability and solar radiation.[45] Once the pollutants are emitted into the atmosphere, their transportation is dependent on the mean wind speed which carries the pollutants away from the source to their sinks, and their convective mixing is dependent on the vertical temperature gradient.[44] The higher values of APP indicate that the atmosphere is unfavourable for the dilution and dispersion of pollutants[46] and indicate high concentrations of observed pollutants at the receiving environment. The low values of APP indicate that the atmosphere is conducive for the dispersion of pollutants which will result in low concentrations on the receiving environment.[44] The APP can be used as a management tool for siting of ambient air quality monitoring stations and for land-use planning in the development of new residential areas and zoning of new industrial sites.

The aim of this work was to determine the spatial variability of PM_{10}, $PM_{2.5}$ and PM chemical composition. Further analysis of the MO theory, MH, VC and the atmospheric pollution potential was performed to determine whether these factors have any influence on the PM_{10} concentrations in the study area.

Methods

Study area

Sampling of PM was undertaken in a rural area of the GTM in Limpopo Province, South Africa (Figure 1). The main towns in the area are Steelpoort and Burgersfort which are sustained through economic activities such as mining and smelting of chromium ores. Furthermore, there are agricultural and forestry activities and transportation that also add to the economic activities in the area. Most of the households in the area are dependent on wood burning for space heating and cooking. The GTM has a complex terrain with high mountains and steep inclinations. The elevation of the surface area is approximately 740 m above sea level with the surrounding mountains extending to a height of approximately 1200–1900 m above sea level. The area is located in the subtropical climate zone where the maximum and minimum average temperatures are 35 °C and 18 °C, respectively in summer, and 22 °C and 4 °C, respectively in winter.[46] The annual rainfall for the area ranges between 500 mm and 600 mm.[47]

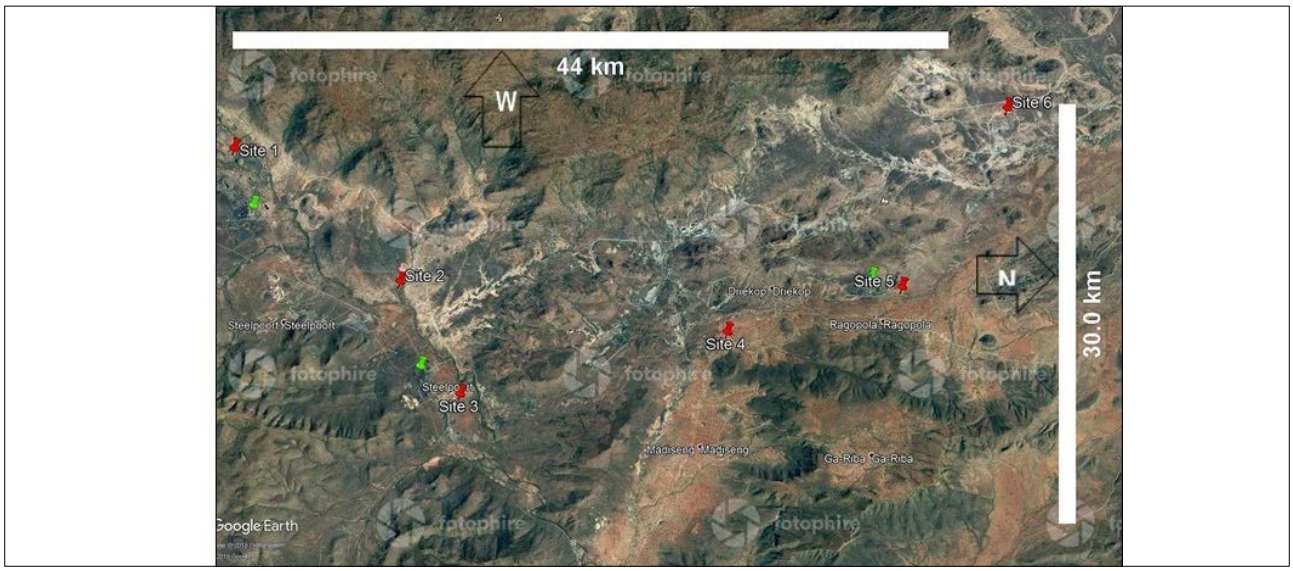


Figure 1: Google Earth map of the study area showing passive sampler locations (indicated by red pins) and smelters (indicated by green pins).

Site selection

The locations of the monitoring sites were selected to optimise spatial sampling for exposure assessment. A sequential sampling technique[48,49] was used to design an optimal sampling network of six sites in the GTM. This technique is based on extended knowledge of the area to be sampled and factors controlling the distribution of pollutants. These factors could be the terrain and various phenomena like meteorological conditions and the chemistry of pollutants.[50] The number of sites selected was influenced by budgetary constraints due to costs associated with laboratory analysis of samples. The sites were located at private residences, a church, a hospital and a school, to ensure a secure area with easy access for site visits.

Sampling and sample analysis

The University of North Carolina passive samplers designed by Wagner and Leith[23] and housed in a protective shelter designed by Ott et al.[51] were deployed at six sites for PM sampling. Ott et al.[51] designed the shelter to shield the passive sampler from precipitation and to minimise the influence of wind speed on particle deposition.[52] The samplers consist of a scanning electron microscopy stub, a collection substrate and a protective mesh cap.[53] The samplers were deployed for a period of ±30 days from July 2015 to June 2016, except for the months of August–September and September–November for which they were deployed for a period of >35 days. The longer sampling periods were selected to ensure that there was sufficient particle loading on the samplers.[52]

The $PM_{2.5}$ and PM_{10} concentrations and the elemental composition of individual particles deposited on the passive sampler were determined by computer-controlled scanning electron microscopy with energy-dispersive X-ray spectroscopy (CCSEM-EDS). Before sample analysis by photoemission electron microscopy (according to the method of Hopke and Casuccio[54]), the samples were coated with a thin layer of graphitic carbon under vacuum to bleed off the charges induced by the electron beam in the SEM. The photoemission electron microscopy was operated at 20 kV.[52] We used the method of Lagudu et al.[53] to determine the chemical composition of PM using CCSEM analysis. Briefly, CCSEM scans the collection substrate of the SEM stub for individual particles and provides fluoresced X-ray spectra and an image of each particle. The method involves rastering the electron beam over the sample while monitoring the resultant backscattered signal. At each point, the image intensity is compared to a pre-set threshold level. Once a coordinate is reached at which the signal is above the threshold level, the electron beam is driven across the particle in a pre-set pattern to determine the size of the particle. Upon measurement of the particle size, the elemental composition of the particle is then determined by collection of characteristic X-rays using EDS techniques. Individual particles characterised during CCSEM analysis are then grouped into particle classes based on their elemental composition. The individual particle masses are finally calculated by multiplying the assigned density of the particle by its volume. Each particle is assigned a density based on common oxide in proportion to the elements present as determined by the EDS analyses.[53] The particle classes obtained from the analysis include carbon-rich (C-rich), chromium-rich (Cr-rich), iron-rich (Fe-rich), iron/chromium-rich (FeCr-rich), silicon-rich (Si-rich), calcium-rich (Ca-rich), silicon/aluminium/iron-rich (SiAlFe-rich), silicon/magnesium-rich (SiMg-rich) and silicon/aluminium-rich (SiAl-rich).

Data analysis

The coefficient of divergence (COD) was used to characterise the spatial variation of PM_{10}, $PM_{2.5}$ and PM chemical components. The COD is defined as:

$$\mathrm{COD} = \sqrt{1/p \sum_{i=1}^{p} \left(\frac{x_{ij} - x_{ik}}{x_{ij} + x_{ik}}\right)^2} \qquad \text{Equation 1}$$

where x_{ij} and x_{ik} are the concentration for sampling interval *i* at sites *j* and *k,* respectively, and *p* is the number of sampling intervals. In terms of spatial distribution, a COD of 0 means that there are no differences between the observed concentrations at the two sites, while a value approaching 1 indicates that the two sampling sites are different.[53,55] Graphical analysis was also used in determining spatial variation. The inverse distance weighted (IDW) interpolation within the mapping software (ArcMap version 10.0) was applied to the annual and monthly concentrations of PM_{10}, $PM_{2.5}$ and the PM chemical components given that the number of sites was restricted by the cost. When data are sparse, the underlying assumptions about the variation among samples may differ and the use of a spatial interpolation method and parameters may become critical.[56,57] The performance of the spatial interpolation method is better when the sample density is higher.[58-60] However, the accuracy of regression modelling is not really dependent on the sampling density, but rather on how well the data are sampled and how significant the correlation is between the primary variable and secondary variable(s).[61] To predict a value for any unmeasured location, IDW uses the measured values surrounding the prediction location. The measured values closest to the prediction location have a greater influence on the predicted value than those farther away. IDW assumes that each measured point has a local influence that diminishes with distance. It gives greater weight to points closest to the prediction location, and the weight diminishes as a function of distance.[62]

APP calculation

The Air Pollution Model (TAPM) was used in the calculation of parameters needed to determine the APP. The dynamic parameters that were calculated included the MO length, wind velocity, planetary boundary layer height and turbulence parameters. The APP was determined according to the method of Swart[63] using Equation 2:

$$P(APP) = P(|\bar{V}|)P(H)P(L) \quad \text{Equation 2}$$

where P(APP) is the air pollution potential index, $P(|\bar{V}|)$ is the wind speed, P(H) is the planetary boundary layer and P(L) is the atmospheric stability. The APP index for a specific area can be classified as being favourable, moderate or unfavourable depending on the conditions set out for the parameters that are the driving force behind the APP calculation, as shown in Table 1.

Table 1: Parameters and limits for air pollution potential (APP) calculation

Parameter	Unfavourable	Moderate	Favourable
Wind speed	0–2 m/s	2–5 m/s	>5 m/s
Mixing height	0–400 m	400–1000 m	>1000 m
Monin–Obukhov length	0 to 200 m	>1000 m	0 to -200 m

In this study, APP, MO, MH and VC values were calculated and correlated with the PM measurement values collected during the sampling campaign.

Results and discussion

Figure 2 shows the annual concentrations of PM_{10}, $PM_{2.5}$ and PM chemical components. The annual concentrations of PM_{10} were 38.11 μg/cm^3 at Site 3, 31.28 μg/cm^3 at Site 2, 31.02 μg/cm^3 at Site 1, 24.65 μg/cm^3 at Site 5, 24.10 μg/cm^3 at Site 4, and 20.98 μg/cm^3 at Site 6. Annual PM_{10} concentrations were below the South African National Ambient Air Quality Standard of 40 μg/cm^3. The $PM_{2.5}$ concentrations are on average lower than the concentrations of SiAl-rich and SiAlFe-rich particles. This finding can be attributed to the fact that some $PM_{2.5}$ particles may have evaporated during the 3–5 week period during which the samplers were deployed in the field. The Fe-rich particles were the least abundant, with an annual average below 1 μg/cm^3 across all sites. The C-rich particles had the same signature as $PM_{2.5}$, with the lowest concentration being around Site 6. The highest concentrations for Ca-rich particles were observed around sampling Site 6 which is located about 1.6 km west-northwest of Marula Platinum Mine, with the lowest concentrations around Site 4 and Site 5. The highest Si-rich, SiMg-rich, SiAl-rich and SiAlFe-rich particle concentrations were observed around sampling Site 3 which is about 1.7 km from the Samancor chrome smelter and 1.9 km from the silicone mine, with the lowest concentrations being observed at Site 6. The highest observed concentrations for FeCr-rich particles were at sampling Sites 3 and 5, and Site 5 is about 2 km from the ASA chrome smelter. The Cr-rich particles were highest at Site 1 and Site 3; Site 1 is about 2.5 km from the Glencore chrome smelter. The annual concentrations of Cr-rich particles across all the sites were above the 0.11 μg/m^3 annual limit set by the New Zealand Ministry of Environment.[64] The highest PM_{10} concentrations were measured during the winter months (May–July) except at Site 6 (Mashegoane) where the highest concentrations were observed during the month of November when there was soil tillage in preparation for crop sowing just before the rainy season. SiAl-rich and SiAlFe-rich particles were the most abundant particles with Fe-rich particles being less abundant. Si-rich, Cr-rich and CrFe-rich particles were more abundant closer to their sources.

Spatial variation

The annual spatial concentration map was generated using geographic information system software (Figure 3). The number of sampling sites was limited due to budgetary constraints, so IDW was used because it does not require a threshold for number of points. The choice of the IDW statistical method proved to be useful as it was able to predict the spatial variation of PM_{10}, $PM_{2.5}$ and PM chemical components in the study area. This output is very important for cash-strapped local authorities that are tasked with the responsibility of managing air quality in their jurisdiction because they can perform this analysis with limited resources. The maps in Figure 3 indicate that there is a distinct spatial heterogeneity in the study area with variability in both low and elevated concentrations being observed at different sites for PM_{10}, $PM_{2.5}$ and PM chemical components. This difference can be attributed to the vast distribution of sources in the area. The highest concentrations for annual PM_{10}, Cr-rich, Fe-rich, Si-rich, SiAl-rich, SiAlFe-rich and SiMg-rich particles were observed around Site 3, which is about 1.7 km from the Samancor chrome smelter and 1.9 km from the Silicone mine, which are located south-southeasterly of the sampling site. The highest Fe-rich, Si-rich and SiMg-rich concentrations were sparsely distributed. Lowest concentrations for the same particles were observed around Site 6. The highest concentrations for $PM_{2.5}$ and C-rich particles were observed around Site 1, and they have similar distribution patterns. The lowest concentrations of PM_{10}, $PM_{2.5}$ and PM chemical components were observed around Site 6, extending to Site 5 and Site 4. The only exception was SiAlFe-rich particles for which the lowest concentrations were observed to the northeast (Site 6) and southeast (Site 1) of the study area. FeCr-rich particles showed highest concentrations closer to the smelters around Site 3 and Site 5.

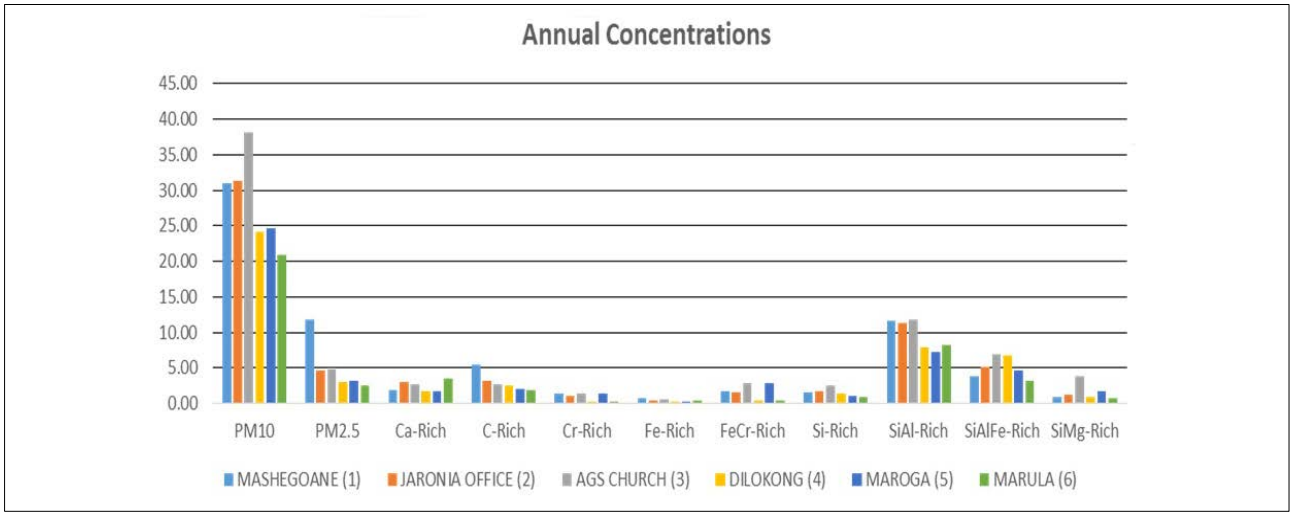


Figure 2: Annual concentrations (μg/cm^3) of PM_{10}, $PM_{2.5}$ and PM chemical components (site number in parentheses).

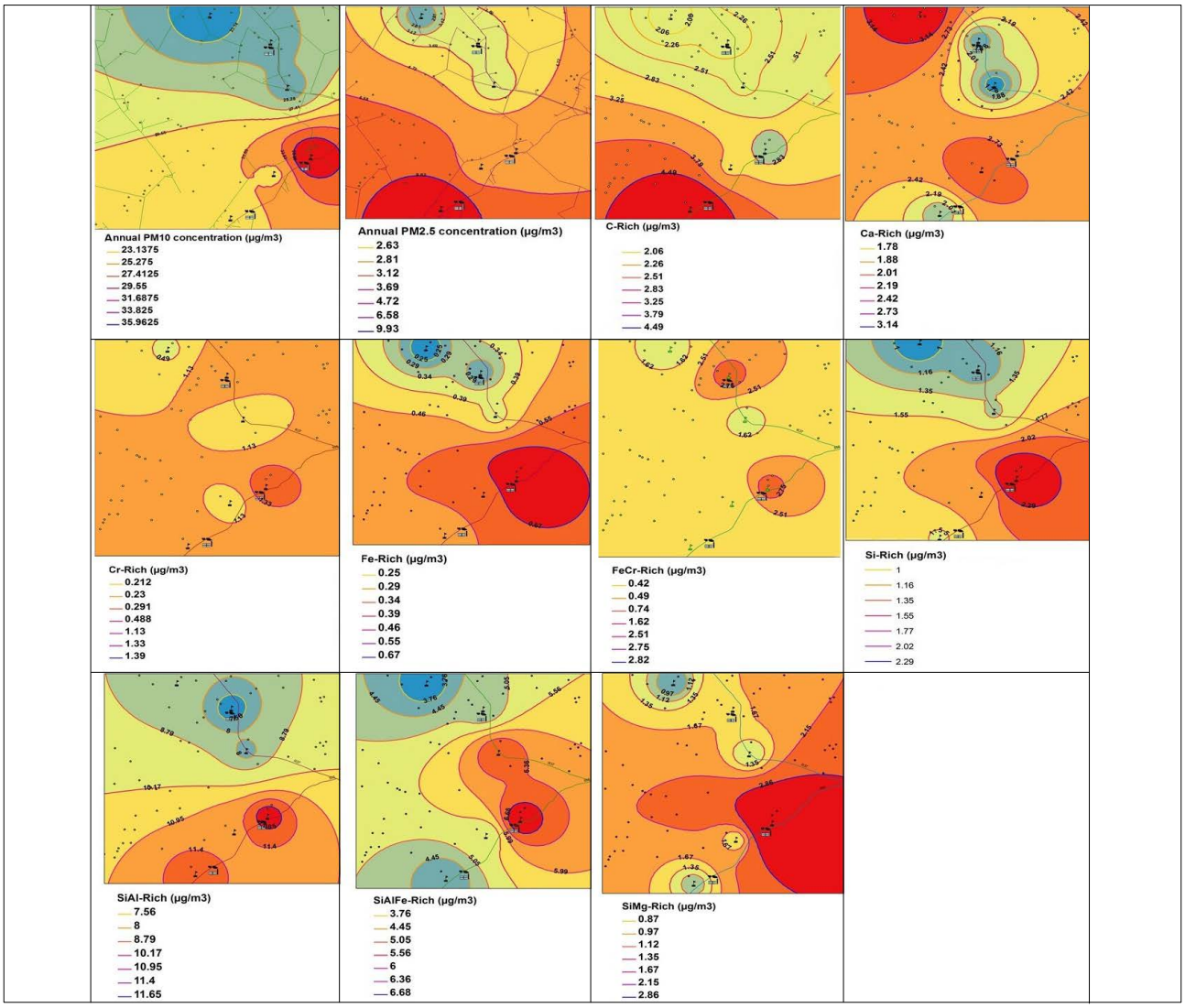


Figure 3: Maps of annual spatial variation for PM_{10}, $PM_{2.5}$, and C-rich, Ca-rich, Cr-rich, Fe-rich, FeCr-rich, Si-rich, SiAl-rich, SiAlFe-rich and SiMg-rich particles. High concentrations are shown in dark red and low concentrations are depicted in blue.

The highest Ca-rich concentrations were observed around Site 6, which is located in an area with black cotton soil. However, because the sites were not equally spaced in the study area, other methods such as the COD and *r* were used to confirm and validate the results of the spatial analysis determined using the geographic information system.

The COD values were calculated (Table 2) to characterise the spatial heterogeneity of PM_{10}, $PM_{2.5}$ and PM chemical components.

Table 2: Results of coefficient of divergence (COD) analysis for Greater Tubatse Municipality

Species	COD
PM_{10}	0.24
$PM_{2.5}$	0.29
C	0.25
Ca	0.38
Cr	0.6
Fe	0.41
FeCr	0.59
Si	0.35
SiAl	0.3
SiAlFe	0.28
SiMg	0.42

COD values higher than 0.2 indicate spatial heterogeneity, while COD values less than 0.1 indicate homogeneity of concentrations. All components in the study area had COD values greater than 0.2, which is an indication that there was a heterogeneous relation observed between the sites in the study area, and is in agreement with the observations in Figure 3. The lowest heterogeneity values for COD ranged from 0.24 to 0.4 and were observed for PM_{10} (0.24), C-rich (0.25), SiAlFe-rich (0.28), $PM_{2.5}$ (0.29), SiAl-rich (0.3), Si-rich (0.35) and Ca-rich (0.38) particles. The moderate to highest COD values were observed for Fe-rich (0.41), SiMg-rich (0.42), FeCr-rich (0.59) and Cr-rich (0.6) particles. The highest COD values were observed for sites located in the vicinity of point source emitters, which indicates that the communities residing in the vicinity of these point sources are more vulnerable to the exposure of these particles than those living further downwind.

Influence of APP, MO, MH and VC on the distribution of PM_{10}

The annual influence of APP, MH, MO and VC on the distribution of PM_{10} concentrations is shown in Figure 4a–d. The highest annual PM_{10} concentrations are centred on Site 3 and distributed more to the east of the sampling site. The highest values for APP (Figure 4a) are centred to the south of the study area around Site 1, which is in contrast to the high APP values which are an indication of low dilution and poor dispersion of concentrations. The distribution of high concentrations to the east of Site 3 suggests that these concentrations move over the mountain

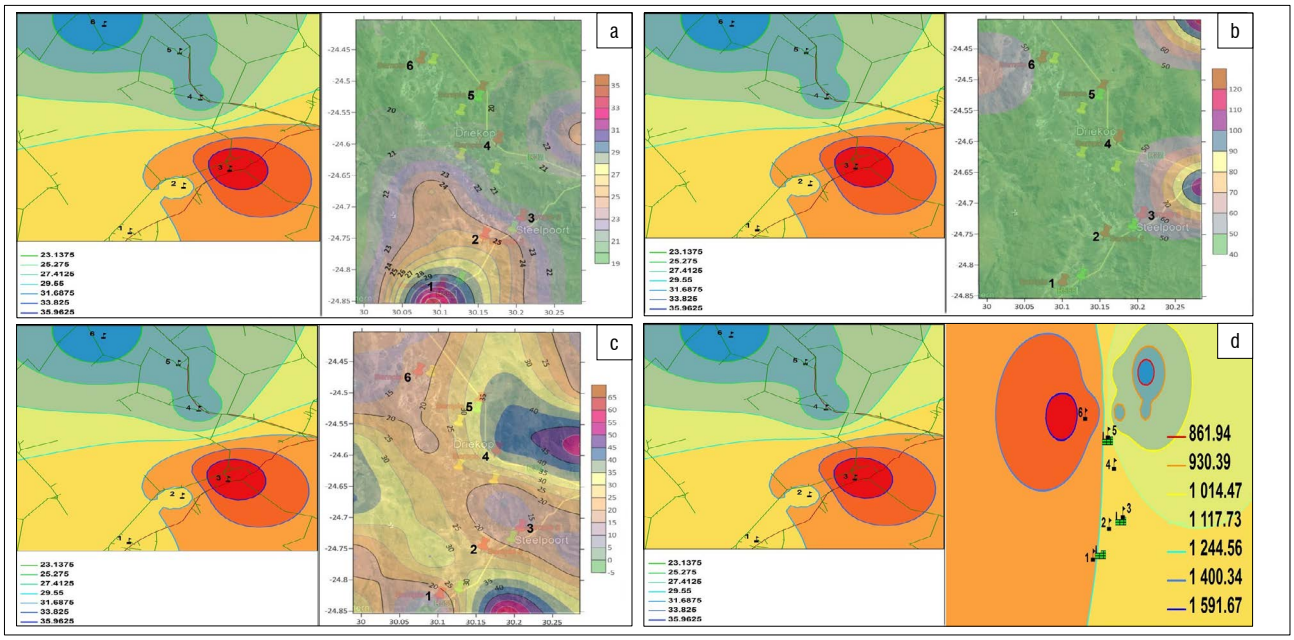


Figure 4: Spatial distribution of PM_{10} and (a) air pollution potential, (b) mixing height, (c) Monin–Obukhov length and (d) ventilation coefficient.

slope to the east of Site 3, which is an indication that mountain winds may be responsible for this flow pattern. The lowest concentrations are centred on Site 6 and extend to Site 5 and Site 4. The lowest APP values are also encountered in the same area as that of the low PM_{10} concentrations which is in contrast to the APP definition. Site 1 and Site 2 have moderate PM_{10} concentrations, which could be attributed to the fact that the area from Site 4 to Site 6 is within a broader mountain valley floor base, compared to the area from Site 1 to Site 3 which has a narrow mountain valley floor base.

Figure 4b shows the comparison between PM_{10} and MH. There is no correlation between PM_{10} and MH. The highest MH was observed to the northeast of Site 3 which is supposed to be an area of low PM_{10} concentrations; however, high PM_{10} concentrations were observed in this region of high MH. The lowest PM_{10} concentrations were observed where there was generally low MH, which is in contrast to the notion that low MH values are associated with poor dilution and dispersion resulting in accumulation of pollutants.

The relationship between PM_{10} and MO is shown in Figure 4c. The area indicated by light green is an area with MO values below 0 and indicates favourable conditions for pollution dispersion. However, the lowest PM_{10} concentrations were observed in an area with moderate stability values, with high PM_{10} concentrations observed in an area of moderate MO. Therefore, MO was unable to correctly indicate the locations of high and low PM_{10} concentrations. This anomaly between PM_{10} concentrations and MO in a complex terrain is because MO is dependent on horizontal wind flows and local equilibrium.[39,65] However, these conditions do not hold in a complex terrain.[39] The MO was derived from synoptic circulations and showed stable conditions in areas where high and low PM_{10} concentrations were observed. The model's inability to account for discontinuities in steep terrain suggests that the PM_{10} concentrations within the valley floor were influenced by the thermal circulations within the valley, with upslope winds due to thermal heating favouring low PM_{10} concentrations and downwind flows due to thermal cooling leading to stagnation and a possible increase in PM_{10} concentrations. However, this hypothesis needs to be further tested in future studies with continuous ambient monitoring in the GTM.

Figure 4d shows the relationship between PM_{10} and the VC. The VC shows moderate to high values to the west of the study area with moderate to low values spreading to the east of the study area. The highest VC values are observed around Site 6 with moderate values across all sites and low values observed to the northeast of Site 6. The observations show that there is a slight correlation between low PM_{10} concentrations and high VC values, and poor correlation between high PM_{10} concentrations and VC values.

The seasonal influence of APP, MH, MO and VC on the distribution of PM_{10} concentrations is shown in Figure 5 and Figure 6, for winter and summer, respectively. The highest PM_{10} concentrations during the winter month of July were observed around sampling Site 5 which is located to the northeast of ASA chrome smelter. The lowest PM_{10} concentrations were observed around Site 1 and Site 6, with moderate concentrations distributed across Site 2, Site 3 and Site 4. The highest APP values were concentrated around Site 1. Moderate APP values were observed to the northeast of Site 5 which is where high PM_{10} concentrations were observed, and to the east of the study area. Low APP values were observed in areas with moderate PM_{10} concentrations. The winter APP was unable to clearly identify areas with high and low PM_{10} concentrations. The highest winter MH was observed to the southeast of the study area with moderate values spreading from southwest to northeast of Site 5. All other sites are located in regions with low MH, which is in contrast to the expected relation between MH and the expected dispersion ability of the atmosphere. The most favourable areas ($MO \leq 0$) for the dispersion of pollutants are indicated in Figures 5 and 6 by light green around Site 1 and Site 2. These areas are where the lowest PM_{10} concentrations were observed. The most stable MO values were spatially distributed across Site 2, Site 3 and Site 4, which are areas where the highest pollution was expected. However, the highest concentrations were observed in an area of moderate MO values. This finding is an indication that the MO cannot clearly identify areas of high PM_{10} concentrations in winter. Strong ventilation (VC) was observed to the west of the mountain valley and weak ventilation to the east of the mountain valley with moderate VC observed within the valley floor. This indicates that the winds within the valley were decoupled from winds outside the valley, and as a result, the VC cannot adequately predict the dispersion of pollutants in the study area.

During the summer month of December (Figure 6), the highest PM_{10} concentrations were distributed around Site 1 and lowest concentrations observed around Site 4, Site 5 and Site 6, with moderate concentrations observed around Site 2 and Site 3. The high APP was distributed around Site 1 with moderate values distributed across Site 2, Site 3 and Site 4, and lower values around Site 5 and Site 6. The PM_{10} concentrations are in agreement with the observed APP for all sites except Site 4 which is supposed to lie within a similar APP to that of Site 5 and Site 6. High MH values were observed to the southeast of the study area with

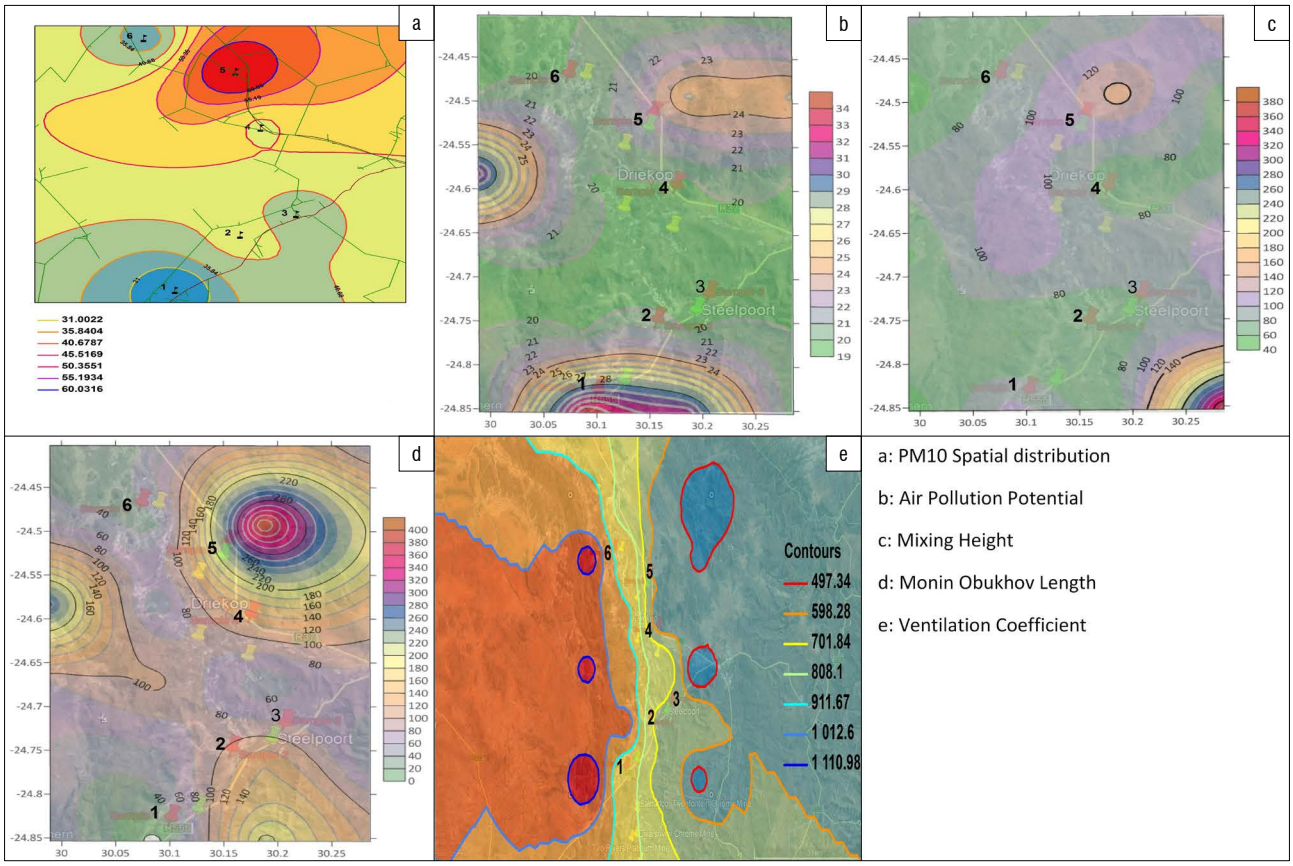


Figure 5: Influence of air pollution potential (APP), mixing height (MH), Monin–Obukhov length (MO) and ventilation coefficient on PM_{10} concentrations in winter.

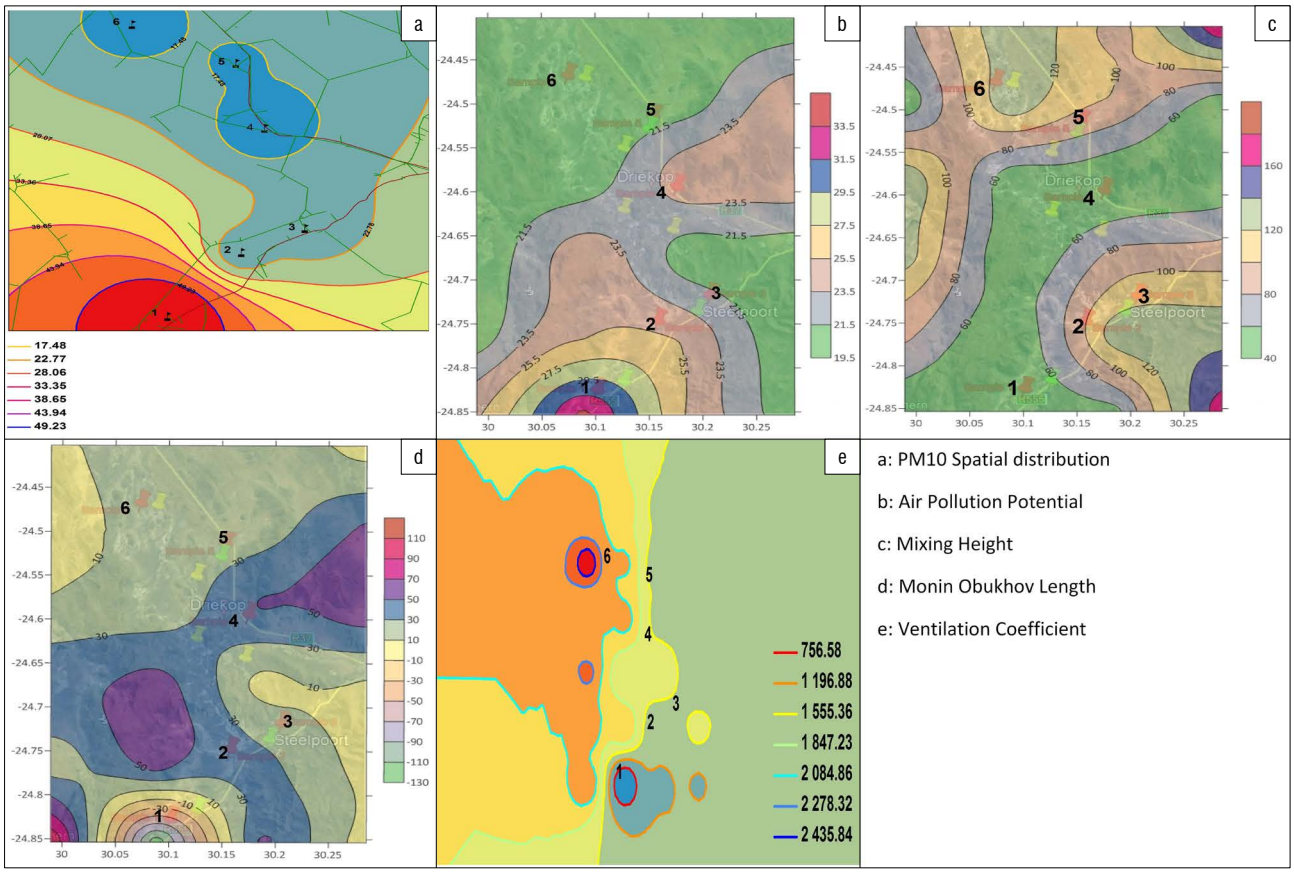


Figure 6: Influence of air pollution potential (APP), mixing height (MH), Monin–Obukhov length (MO) and ventilation coefficient on PM_{10} concentrations in summer.

moderate values distributed across Site 2, Site 3, Site 5 and Site 6. The lowest MH values were observed across Site 1 and Site 4. The PM_{10} concentration was therefore expected to be highest around Site 1 and Site 4 in accordance with the definition of MH, with moderate PM_{10} expected across all other sites. However, only results from Site 1 were in agreement with the observed MH. The most favourable conditions (with respect to MO) for the dispersal of pollutants were observed around Site 1, with unfavourable conditions observed around Site 2 and Site 4, and moderate conditions around Site 3, Site 5 and Site 6. The lowest VC values were observed around Site 1 and the highest VC values were observed around Site 6, with moderate values distributed across the remaining sites. Therefore, the VC was able to predict the areas for high PM_{10} concentrations (Site 1) and low PM_{10} concentrations (Site 6). Site 1 and Site 6 are located in the valley openings with Site 1 being in an area with a narrow valley opening and Site 2 being in an area with a wide valley opening. The strength of the valley flows depends on the valley volume. Wind speeds are often larger near the valley head where valley volume is small and the pressure gradient is high relative to distance from ridge top to ridge top. Wind speed weakens near the valley opening where the valley volume is larger and the pressure gradient is low relative to the distance between ridge tops.[66] Therefore, wind erosion may have played a major role in the observed high PM_{10} concentrations at Site 1 and, similarly, calm conditions may have been responsible for the observed low PM_{10} concentrations at Site 6. However, the VC did not have the same influence on the other sites which are situated in the middle of the valley floor. The reason could be that the TAPM model inputs terrain following coordinate systems and was unable to account for discontinuities in the steep terrain of the study area.

Conclusion

The University of North Carolina passive samplers coupled with CCSEM_EDS were used to determine spatial heterogeneity of PM chemical components. The concentrations of $PM_{2.5}$, PM_{10} and PM chemical components were spatially heterogeneous with high heterogeneity observed near the industrial sources for FeCr-rich and Cr-rich particles and Si-containing particles. The COD values also showed that the highest heterogeneity was observed near the industrial sources. Findings showed little or no correlation between PM_{10} and the meteorological parameters MH, MO length, APP and COD.

The findings highlight a very important point: passive samplers can be used (particularly in developing world contexts) as a substitute to more expensive continuous samplers to determine the spatial variation of PMs and their chemical components for effective environmental planning. The IDW interpolation within the mapping software (ArcMap version 10.0) was able to predict the spatial variation of PM_{10}, $PM_{2.5}$ and PM chemical components that indicated the existence of different conditions within the air shed, and therefore this variation may require different control strategies to mitigate the impacts of pollution within the air shed. The second finding was that synoptic winds used by the TAMP model were unsuccessful in determining the influence of APP, MO, MH and VC on the distribution of PM_{10} concentrations in a complex terrain. This finding clearly indicates that these parameters are dependent largely on winds generated by temperature changes and mountain slopes in mountainous terrain. However, the VC was able to predict the areas for high PM_{10} concentrations at the valley openings where the VC is influenced by the impact of pressure gradient on the wind strength. Therefore, for future analysis of the behaviour of pollutants in a complex terrain, a network of meteorological station balloon soundings within the valley floor and adjacent slopes needs to be set up in order to capture the actual meteorological parameters that influence the behaviour of air pollution. The ambient air quality should be monitored continuously to verify the findings of this study.

Acknowledgements

C.Y.W. receives research funding support from the South African Medical Research Council and the National Research Foundation (South Africa). C.T. thanks the South African Weather Service for provision of resources, space and time for conducting this research.

Authors' contributions

C.T. conceptualised the study, collected the samples, performed the GIS analysis, interpreted the data, wrote the initial draft of the manuscript and the revised version of the manuscript. C.Y.W. provided critical feedback and helped to shape the manuscript.

References

1. Gautam S, Prusty BK, Patra AK. Pollution due to particulate matter from mining activities. Reciklaža i održivi razvoj. 2012;5:53–58. Available from: https://www.rsd.tfbor.bg.ac.rs/download/arhiva_radova/2012/6_SnehaGAUTAM_ST.pdf
2. Seinfeld JH, Pandis SN. Atmospheric chemistry and physics from air pollution to climate change. New York: John Wiley and Sons; 1998.
3. Khlystov A, edited by Wittig B, Davidson C. Quality assurance project plan for Pittsburgh Air Quality Study (PAQS) [document on the Internet]. c2001 [cited 2018 Jun 15]. Available from: https://www3.epa.gov/ttnamti1/files/ambient/super/pittqapp.pdf
4. Quality of Urban Air Review Group (AQUARG). Airborne particulate matter in the United Kingdom. Third report of the Quality of Urban Air Review Group [document on the Internet]. c1996 [cited 2018 Jun 10]. Available from: https://uk-air.defra.gov.uk/assets/documents/reports/empire/quarg/quarg_11.pdf
5. Pope CA, Dockery W. Health effects of fine particulate air pollution: Lines that connect. J Air Waste Manag Assoc. 2006;56:709–742. ttps://doi.org/10.1080/10473289.2006.10464485
6. Brook RD, Rajagopalan S, Pope CA, Brook JR, Bhatnagar A, Diez-Roux AV, et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation. 2010;121(21):2331–2378. https://doi.org/10.1161/CIR.0b013e3181dbece1
7. World Health Organization (WHO). Health effects of particulate matter: Policy implications for countries in eastern Europe: Policy implications for countries in eastern Europe, Caucasus and central Asia. Copenhagen: WHO; 2013. Available from: http://www.euro.who.int/__data/assets/pdf_file/0006/189051/Health-effects-of-particulate-matter-final-Eng.pdf
8. Hinds WC. Aerosol technology, properties, behavior, and measurements of airborne particles. 2nd ed. New York: John Wiley and Sons; 1999.
9. Shah MH, Shaheen N, Jaffar M, Khalique A, Tariq SR, Manzoor S. Spatial variations in selected metal contents and particle size distribution in an urban and rural atmosphere of Islamabad, Pakistan. J Environ Manage. 2006;78:128–137. https://doi.org/10.1016/j.jenvman.2005.04.011
10. Wang Y, Eliot MN, Koutrakis P, Gryparis A, Schwartz JD, Coull BA. Ambient air pollution and depressive symptoms in older adults: Results from the MOBILIZE Boston study. Environ Health Perspect. 2014;122(6):553–558. https://doi.org/10.1289/ehp.1205909
11. Chan CK, Yao X. Air pollution in mega cities in China. Atmos Environ. 2008;42:1–42. https://doi.org/10.1016/j.atmosenv.2007.09.003
12. Environmental Protection Agency (EPA). Technology Transfer Network (TTN). National air toxics assessments [webpage on the Internet]. c2010 [cited 2018 Jun 14]. Available from: https://www.epa.gov/national-air-toxics-assessment
13. Ahrens CD. Meteorology today. 9th ed. Belmont, CA: Thomas Brooks/Cole; 2007. Available from: https://epdf.tips/meteorology-today-9th-edition.html
14. Courault D, Drobinski P, Brunet Y, Lacarrere P, Talbot C. Impact of surface heterogeneity on a buoyancy-driven convective boundary layer in light winds. Bound Layer Meteorol. 2007;124:383–403. https://doi.org/10.1007/s10546-007-9172-y
15. Reen BP, Stauffer DR, Davis KJ, Desai AR. A case study on the effects of heterogeneous soil moisture on mesoscale boundary layer structure in the southern Great Plains, USA. Part II: Mesoscale modelling. Bound Layer Meteorol. 2006;120:275–314. https://doi.org/10.1007/s10546-006-9056-6
16. Yuan X, Xie Z, Zheng J, Tian X, Yang Z. Effects of water table dynamics on regional climate: A case 10 study over East Asian Monsoon Area. J Geophys Res. 2008;113, D21112, 16 pages. https://doi.org/10.1029/2008JD010180
17. Yates DN, Chen F, Nagai H. Land surface heterogeneity in the Cooperative Atmosphere Surface Exchange Study (C 5 ASES-97). Part II: Analysis of spatial heterogeneity and its scaling. J Hydrometeorol. 2003;4:219–234. https://doi.org/10.1175/1525-7541(2003)4<219:LSHITC>2.0.CO;2
18. Zhang N, Williams Q, Liu H. Effects of land-surface heterogeneity on numerical simulations of mesoscale atmospheric boundary layer processes. Theoret Appl Climatol. 2010;102:307–317. https://doi.org/10.1007/s00704-010-0268-9

19. Atkinson BW. Mesoscale atmospheric circulation. London: Academic Press; 1981.

20. Whiteman CD. Observations of thermally developed wind systems in mountainous terrain. In: Blumen W, editor. Atmospheric processes over complex terrain. Meteorological Monographs vol. 23. Boston, MA: American Meteorological Society; 1990. p. 5–42. https://doi.org/10.1007/978-1-935704-25-6_2

21. Durran DR. Mountain waves and downslope winds. In: Blumen W, editor. Atmospheric processes over complex terrain. Meteorological Monographs vol. 23. Boston, MA: American Meteorological Society; 1990. p. 59–81. https://doi.org/10.1007/978-1-935704-25-6_4

22. Whiteman CD, Doran JC. The relationship between overlying synoptic-scale flows and winds within a valley. J Appl Meteorol. 1993;32:1669–1682. https://doi.org/10.1175/1520-0450(1993)032<1669:TRBOSS>2.0.CO;2

23. Wagner J, Leith D. Passive aerosol sampler. Part I: Principle of operation. Aerosol Sci Technol. 2001;34(2):186–192. https://doi.org/10.1080/027868-201300034808

24. Jiménez PA, González-Rouco JF, Montávez JP, Navarro E, García-Bustamante E, Valero F. Surface wind regionalization in complex terrain. J Appl Meteorol Climatol. 2007;47:308–324. https://doi.org/10.1175/2007JAMC1483.1

25. Helgason W, Pomeroy JW. Characteristics of the near surface boundary layer within a mountain valley during winter. J Appl Meteorol Climatol. 2012;51:583–597. https://doi.org/10.1175/JAMC-D-11-058.1

26. Geiger R. Das Klima der bodennahen Luftschicht**.** Ein Lehrbuch der Mikroklimatologie. 4. neubearbeitete und erweiterte Auflage mit 281 Abb., 646 S [The climate of ground-level air-layer. A textbook of microclimatology 4. New and expanded edition with 281 Abb.,646 S]. Brunswick: Verlag Friedrich Vieweg & Sohn; 1961. German. https://doi.org/10.1002/jpln.19620960108

27. De Wekker SFJ, Kossmann M. Convective boundary layer heights over mountainous terrain: A review of concepts. Front Earth Sci. 2015;3(77):1–22. https://doi.org/10.3389/feart.2015.00077

28. Zardi D, Whiteman D. Diurnal mountain wind systems. In: Chow F, De Wekker S, Snyder B, editors. Mountain weather research and forecasting: Recent progress and current challenges.: Dordrecht: Springer; 2013. p. 35–119. https://doi.org/10.1007/978-94-007-4098-3_2

29. Bianco L, Djalaova IV, King CW, Wilczak JM. Diurnal evolution and annual variability of boundary-layer height and its correlation to other meteorological variables in California's Central Valley. Bound Layer Meteorol. 2011;140:491–511. https://doi.org/10.1007/s10546-011-9622-4

30. Working Group 4 (WG4). COST 710: Pre-processing of meteorological data for dispersion models. Wind flow models over complex terrain for dispersion calculations [document on the Internet]. c1997 [cited 2018 Mar 11]. Available from: https://www2.dmu.dk/atmosphericenvironment/cost/docs/COST710-4.pdf

31. Coulter RL. A comparison of three methods for measuring mixing-layer height. J Appl Meteorol. 1979;8:1495–1499. https://doi.org/10.1175/1520-0450(1979)018<1495:ACOTMF>2.0.CO;2

32. Saurez MJ, Arakawa A, Randall DA. The parameterization of the planetary boundary layer in the UCLA general circulation model: Formulation and results. Mon Weather Rev. 1983;111:2224–2243. https://doi.org/10.1175/1520-0493(1983)111<2224:TPOTPB>2.0.CO;2

33. Wesely ML, Cook DR, Hart RL, Speer RE. Measurements and parameterization of particulate sulfur dry deposition over grass. J Geophys Res. 1985;90:2131–2143. https://doi.org/10.1029/JD090iD01p02131

34. Konor CS, Boezio GC, Mechoso CR, Arakawa A. Parametrization of PBL process in an atmospheric general circulation model: Description and preliminary assessment. Mon Weather Rev. 2009;37:1061–1082. https://doi.org/10.1175/2008MWR2464.1

35. Rotach A, Gohm MN, Lang D, Leukauf I, Stiperski, Wagner JS. On the vertical exchange of heat, mass and momentum over complex, mountainous terrain. Front Earth Sci. 2015;3:76. Available from: https://www.frontiersin.org/articles/10.3389/feart.2015.00076/full

36. Steyn DG, De Wekker SFJ, Kossmann M, Martilli A. Boundary layers and air quality in mountainous terrain. In: Chow FK, De Wekker SFJ, Snyder B, editors. Mountain weather research and forecasting: Recent progress and current challenges. Berlin: Springer; 2013. p. 261–289. https://doi.org/10.1007/978-94-007-4098-3_5

37. Faiz A, Sinha K, Walsh M, Varma A. Automotive air pollution: Issues and options for developing countries. Washington DC: Infrastructure and Urban Development, World Bank; 1990. Available from: http://documents.worldbank.org/curated/en/743671468739209497/pdf/multi-page.pdf

38. Figueroa-Esspinoza B, Salles P. Local Monin–Obukhov similarity in heterogeneous terrain. Atmos Sci Lett. 2014;15:299–306. https://doi.org/10.1002/asl2.503

39. Grisogono B, Kraljevic L, Jericevic A. Notes and correspondence the low-level katabatic jet height versus Monin-Obukhov height. Q J R Meteorol Soc. 2007;133:2133–2136. https://doi.org/10.1002/qj.190

40. Gross E. The National Air Pollution Potential Forecast Program: ESSA technical memorandum WBTM NMC 47 [document on the Internet]. c1970 [cited 2018 Mar 03]. Available from: https://apps.dtic.mil/dtic/tr/fulltext/u2/714568.pdf

41. Iyer US, Raj P. Ventilation coefficient trends in the recent decades over four major Indian metropolitan cities. J Earth System Sci. 2013;122:537–549. https://doi.org/10.1007/s12040-013-0270-6

42. Devara PCS, Raj PE. Lidar measurements of aerosols in the tropical atmosphere. Adv Atmos Sci. 1993;10:365–378. https://doi.org/10.1007/BF02658142

43. Krishnan P, Kunhikrishnan PK. Temporal variations of ventilation coefficient at a tropical Indian station using UHF wind profiler. Curr Sci. 2004;86(3):447–451. Available from: https://www.currentscience.ac.in/Downloads/article_id_086_03_0447_0451_0.pdf

44. Nath P, Patil RS. Climatological analysis to determine air pollution potential for different zones in India. In: Brebbia CA, Power H, Longhurst JWS, editors. Air pollution VIII [document on the Internet]. c2000 [cited 2018 Apr 15]. Available from: https://www.witpress.com/Secure/elibrary/papers/AIR00/AIR-00048FU.pdf

45. Anil Kumar KG. Air pollution climatology of Cochin for pollution management and abatement planning. Mausam. 1999;50:383–390. Available from: http://metnet.imd.gov.in/mausamdocs/15046.pdf

46. Schulze BR. Climate of South Africa. Part 8: General survey WB28. Pretoria: SA Weather Bureau; 1986.

47. South African Department of Water Affairs and Forestry (DWAF). Water Services Planning Reference Framework, Sekhukhune District Municipality [document on the Internet]. c2005 [cited 2018 Jun 13]. Available from: http://www.sekhukhunedistrict.gov.za/sdm-admin/documents/Sekhukhune%202011-12%20IDP.pdf

48. Goovaerts P. Kriging vs stochastic simulation for risk analysis in soil contamination. In: Soares A, Gomez-Hernandez J, Froidevaux R, editors. geo-ENV I – Geostatistics for environmental applications. Dordrecht: Kluwer Academic Publishers; 1997. p. 247–258. https://doi.org/10.1007/978-94-017-1675-8_21

49. Van Groenigen JW, Stein A, Zuurbier R. Optimization of environmental sampling using interactive GIS. Soil Tech. 1997;10:83–97. https://doi.org/10.1016/S0933-3630(96)00122-5

50. Fraczek W, Bytnerowicz A, Legge A. Optimizing a monitoring network for assessing ambient air quality in the Athabasca oil sands region of Alberta, Canada. Alpine Space – Man Environ. 2009;48:127–142. Available from: https://www.researchgate.net/publication/313609727_Optimizing_a_monitoring_network_for_assessing_ambient_air_quality_in_the_Athabasca_oil_sands_region_of_Alberta_Canada

51. Ott DK, Cyrs W, Peters TM. Passive measurement of coarse particulate matter, PM10-2.5. J Aerosol Sci. 2008;39(2):156–167. https://doi.org/10.1016/j.jaerosci.2007.11.002

52. Sawvel EJ, Willis R, West RR, Casuccio GS, Norris G, Kumar N, et al. Passive sampling to capture the spatial variability of coarse particles by composition in Cleveland, OH. Atmos Environ. 2015;105:61–69. https://doi.org/10.1016/j.atmosenv.2015.01.030

53. Lagudu URK, Raja S, Hopke PK, Chalupa DC, Utell MJ, Casuccio G, et al. Heterogeneity of coarse particles in an urban area. Environ Sci Technol. 2011;45:3288–3296. https://doi.org/10.1021/es103831w

54. Hopke PK, Casuccio G. Scanning electron microscopy. In: Receptor modeling for air quality management vol. 7. Amsterdam: Elsevier; 1991; p. 149–212. https://doi.org/10.1016/S0922-3487(08)70129-1

55. Wongphatarakul V, Friedlander SK, Pinto JP. A comparative study of $PM_{2.5}$ ambient aerosol chemical databases. Environ Sci Technol. 1998;32:3926–3934. https://doi.org/10.1021/es9800582

56. Burrough PA, McDonnell RA. Principles of Geographical Information Systems. New York: Oxford; 1998.

57. Hartkamp AD, White JW, Hoogenboom G. Interfacing GIS with agronomic modeling: A review. Agron J. 1999;91:917–928. https://doi.org/10.2134/agronj1999.915761x

58. Webber D, Englund E. Evaluation and comparison of spatial interpolators. Math Geol. 1992;24(4):381–391. https://doi.org/10.1007/BF00891270

59. Isaaks H, Srivastava MR. An introduction to applied geostatistics. New York: Oxford University Press; 1989. p. 561. Available from: https://www.scribd.com/document/266290280/An-Introduction-to-Applied-GEOostatistics-Isaaks-and-Srivastava-1989-Oxford

60. Stahl K, Moore RD, Floyer JA, Asplin MG, McKendry IG. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. Agric Forest Meteorol. 2006;139:224–236. https://doi.org/10.1016/j.agrformet.2006.07.004

61. Hengl T. A practical guide to geostatistical mapping of environmental variables. EUR 229404 EN. Luxembourg: Office for Official Publications of the European Communities; 2007. Available from: https://www.lu.lv/materiali/biblioteka/es/pilnieteksti/vide/A%20Practical%20Guide%20to%20Geostatistical%20Mapping%20of%20Environmental%20Variables.pdf

62. Li L, Losser T, Yorke C, Piltner R. Fast inverse distance weighting-based spatiotemporal interpolation: A web-based application of interpolating daily fine particulate matter $PM_{2.5}$ in the contiguous U.S. using parallel programming and k-d tree. Int J Environ Res Public Health. 2004;11:9101–9141. https://doi.org/10.3390/ijerph110909101

63. Swart A. An analysis of the air dispersion potential over Uubvlei, Oranjemund, Namibia [dissertation]. Pretoria: University of Pretoria; 2015. Available from: https://repository.up.ac.za/bitstream/handle/2263/60862/Swart_Assessment_2017.pdf?sequence=1&isAllowed=y

64. Fisher G, Metcalfe J. Good practice guide for assessing discharges to air from industry. Wellington: NZ Ministry for the Environment; 2016. https://doi.org/10.13140/RG.2.1.3803.4162

65. Monin AS, Obukhov AM. Basic laws of turbulent mixing in the surface layer of the atmosphere [translation]. Tr Akad Nauk SSSR Geophiz Inst. 1954;24(151):163–187. Available from: https://mcnaughty.com/keith/papers/Monin_and_Obukhov_1954.pdf

66. Whiteman CD. Mountain meteorology: Fundamentals and applications. New York: Oxford University Press; 2000. p. 355. Available from: https://www.amazon.com/Mountain-Meteorology-Fundamentals-David-Whiteman/dp/0195132718#reader_0195132718

AUTHORS:
Dawn D. Mahlobo[1]
Thando Ndarana[2]
Stefan W. Grab[1]
Francois A. Engelbrecht[3]

AFFILIATIONS:
[1]School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand, Johannesburg, South Africa
[2]Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa
[3]Global Change Institute, University of the Witwatersrand, Johannesburg, South Africa

CORRESPONDENCE TO:
Dawn Mahlobo

EMAIL:
ddmahlobo@gmail.com

DATES:
Received: 31 Oct. 2018
Revised: 19 Feb. 2019
Accepted: 30 June 2019
Published: 26 Sep. 2019

HOW TO CITE:
Mahlobo DD, Ndarana T, Grab SW, Engelbrecht FA. Granger causality of the local Hadley cell and large-scale cloud cover over South Africa. S Afr J Sci. 2019;115(9/10), Art. #5724, 10 pages. https://doi.org/10.17159/sajs.2019/5724

ARTICLE INCLUDES:
☒ Peer review
☒ Supplementary material

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

EDITORS:
Nicolas Beukes
Yali Woyessa

KEYWORDS:
Hadley circulation, mass flux, total cloud cover, solar energy potential

FUNDING:
National Research Foundation (South Africa)

Granger causality of the local Hadley cell and large-scale cloud cover over South Africa

This study demonstrates that Hadley cell dynamics could be used as a proxy to determine cloud cover and thus solar energy potential over South Africa. Granger causality was used to investigate causal interactions between the Hadley cell and cloud cover for the period 1980–2015, and such links were established. Areas of strong causality are found over the northwestern parts of South Africa. Moreover, weak causality from cloud cover to the Hadley cell does exist, with vertical velocity being the main variable responsible for this causality, which hence indirectly links cloud cover to Hadley cell causality.

Significance:

- Hadley cell dynamics may be used to identify regions of cloudlessness over South Africa.
- Hadley cell dynamics may further be used as a proxy for cloud cover towards understanding the solar energy potential in South Africa within the context of climate variability and change.

Introduction

The South African government has identified a number of renewable energy options to inform the country's energy mix on the 2030 horizon and beyond.[1] One of these options is solar energy, which depends on the ability of incoming short wave radiation to penetrate through the atmosphere to the ground, where solar energy conversion technologies are located. In this study, we demonstrate that Hadley cell dynamics could be used as a proxy to determine cloud cover and thus the solar energy potential over South Africa. The abundance in solar radiation and sunshine over South Africa is because the country is located in the subtropical belt – exactly where the Hadley cell descending branch is located. Solar radiation reaching the surface has a strong longitudinal gradient, which exists in relation to the longitudinal gradients in cloud cover, rainfall and thus sunshine duration.[2] Interactions between the Hadley cell and cloud cover may therefore be key in understanding and anticipating the potential of solar power as a renewable energy source for South Africa.

The Hadley circulation is a key component of the global circulation and accounts for ascending motion in the tropics and descending air in the subtropics. Weather and climate in both the tropics and subtropics are thus strongly influenced by the Hadley circulation.[3] Southern Africa, with its location in the subtropics, is on average under the influence of the descending branch of the Hadley cell.[3] This influence is the main reason why southern Africa is in general a relatively dry and warm region; in fact, much of southern Africa is semi-arid.[4] Winters in South Africa are generally dry with clear skies over the interior, due to the dominance of the subtropical high-pressure belt (the surface manifestation of the descending branch of the Hadley cell) during these months.[5] It is only the most southern parts of southern Africa, namely the southwestern Cape and Cape south coast of South Africa, that receive substantial amounts of rainfall during winter.[3] During wet summers, the non-divergent part of the Hadley circulation causes a southward shift in the Indian Ocean cyclonic cell, resulting in surplus water vapour transport across southern Africa from the north.[6] During dry summers, the spatial extent of the tropical western Indian Ocean anticyclone decreases, thereby leading to the reduction of water vapour transport from the southeast, whilst the descending branch of the Hadley cell strengthens over southern Africa.[7] Both the long-term climate and inter-annual climate variability over southern Africa are thus strongly controlled by Hadley cell circulation dynamics, with pronounced implications for the region's agriculture, water security and biodiversity. Moreover, Hadley cell dynamics have also been linked to variations in sunshine duration across southern Africa, which may be important within the context of a growing renewable energy sector in the region.[8]

The meridional shifts of the Hadley cell are linked to the seasonal migration of the Inter-Tropical Convergence Zone (ITCZ).[9] Tropical disturbances occur in summer over southern Africa as the ITCZ propagates southwards to approximately 17°S.[3] Cloud bands associated with most of the late summer rainfall over the region link the ITCZ to the north[3] with a westerly wave to the south. These tropical-extratropical cloud bands are defined as regions of elongated cloudiness that start in the tropics and extend southeastwards into the mid-latitudes.[6] The cloud bands export moisture and heat from the tropics to the middle latitudes, through which the deep Hadley cell overturning is replaced during periods of weakened ITCZ activity.[10] The Hadley cell overturning is associated with convective intensity over southern Africa, anticyclonic ridging and moist air inflows from the Indian Ocean.[9]

Climate change over southern Africa may be anticipated to be closely linked to changes in the dynamics of the regional Hadley cell. In fact, the mean positions of storm tracks, high and low pressure systems, jet streams and associated precipitation patterns are all projected to change in response to the expansion of the Hadley cell in a warmer world.[5] Over southern Africa, the strengthening and expanding subtropical high-pressure belt is, under climate change, projected to contribute to the southward displacement (or blocking) of frontal systems bringing rainfall over southern Africa.[11] Moreover, a general strengthening of the descending branch of the Hadley cell in summer has been postulated as a key reason why southern Africa is projected to become generally drier under global warming.[3,11]

Large-scale cloudiness over South Africa is linked to the dynamics of the Hadley circulation over the country.[8] Whether changes in cloudiness are directly caused by changes in the Hadley circulation or whether cloudiness may impact on the Hadley cell via feedback processes, remains to be rigorously investigated. To address this gap, causality between changes in the Hadley cell and cloud cover need to be established. Thus, the aim of this study was to use Granger

causality[12] to investigate causal links between the Hadley cell and large-scale cloud cover over South Africa. Understanding such links is important in understanding the rainfall anomalies and solar energy potential in the region, within the context of climate variability and change.

Data and methods

Data

The study covers the period 1980 to 2015. Data used in this study were extracted from the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim Reanalysis (ERA-Interim).[13] The ERA-Interim has a fine resolution and is sufficient enough to provide appropriate diagnostics of the Hadley circulation over South Africa. Meridional velocity from 1000 hPa to 10 hPa was used to calculate the zonally averaged mass stream function (see Supplementary appendix 1). The zonally asymmetric Hadley cell time series was calculated using the mass flux,[14] as outlined in Supplementary appendix 2. The vertical velocity at 500 hPa (which is the level of maximum upward vertical motion) was used to calculate the Hadley cell diagnostics (Supplementary appendix 2). Data are on a horizontal resolution of 0.75°x 0.75° on 37 pressure levels.[13,15] Total cloud cover was also extracted from the ERA-Interim data set. To account for both short- and long-term effects of the Hadley cell on total cloud cover and vice versa, daily time steps were used. The total cloud cover data for the area 18–34°S and 15–34°E was divided into low and high cloudiness years. The high and low cloudiness years for December–January–February (DJF) and June–July–August (JJA) are shown in Table 1.

Table 1: Years of low and high cloudiness over the area 18–34°S and 15–34°E

	Season	
	December–February	**June–August**
Low cloudiness	1986, 1987, 1990, 1992, 1993, 1995, 2004	1980, 1984, 1986, 1987, 1988, 1990, 1992, 1994, 1995, 1996, 1997, 1999, 2000, 2006, 2008, 2009
High cloudiness	1982, 1983, 1991, 1997, 2001, 2004, 2007, 2010, 2011, 2012, 2013, 2014	1982, 1985, 1991, 1993, 1998, 2002, 2004, 2005, 2007, 2009, 2010, 2012, 2013, 2014

Methods

Several studies have used linear and lag correlation statistics to help establish the links between time series in climate science.[16,17] However, it is a challenge to identify the direction of causality from such methods. Causality studies between climate variables have also been undertaken through Bayesian network inference[18,19] and Granger causality.[20-23] The two frameworks were compared to each other using biological data, from which it was established that the Bayesian network performs better for shorter temporal data sets, while for longer data sets, Granger causality seems to perform better.[24] One remarkable feature of Granger causality is that it has a decomposition property, which is not present in the Bayesian network inference.[24] This feature enables one to establish the best frequency at which causality may be established between two time series. Thus, Granger causality seems to be the best method for testing the direction and strength of causality between two time series. We therefore introduce the notion of Granger causality to establish the causal relationships between the Hadley cell and cloud cover.

Granger causality can be defined as variable Y Granger causing variable X, if X can be predicted better by using the past values of Y, more than the past values of X itself. This definition, when applied to the study, means that the cloud cover has a Granger causal relationship to the Hadley cell, if past values of cloud cover could be used to help predict the Hadley cell. Granger causality analysis tests for both the presence and direction of causality.[25] Granger causality was initially designed and mainly applied to econometrix data, yet several studies have applied Granger causality to the atmospheric sciences.[19-22] The main challenge in employing Granger causality to climate data is the fact that climate systems are highly non-linear.[12] Studies have employed non-linear Granger causality to overcome such challenges.[26] However, it has been shown that using average data (e.g. seasonal averages) can produce near-linear relationships between climate variables,[23] and hence reasonable estimates of causal links can be obtained from a linear model.

Granger causality studies using climate data include a causality study investigating southern and northern hemisphere temperatures[27], and a Granger causality study between the North Atlantic Oscillation and Atlantic sea surface temperatures at a seasonal scale[20]. Although most climate studies make use of a bivariate system when investigating causality between two variables[20,22,26], bivariate systems have problems of spurious causality and of non-causality due to omission of a relevant variable[22]. These problems can be solved by introducing an auxiliary third variable in the analysis.[22,28] An alternative method for shorter time series is cross validation.[26] Other techniques used to test a direct Granger causality of Y on X include ex conditional Granger causality[29] and partial Granger causality[30]. However, studies have not yet employed Granger causality to the Hadley cell and cloud cover. We thus tested this interaction using a four-step procedure including unit root testing and differencing, selecting the appropriate model for the time series data, and testing for Granger causality.

Unit root testing

The fundamental issue in testing for causality between variables is to use a suitable time series that is stationary or does not contain unit roots. Stationarity in a time series is defined as one with a statistical process (mean or standard deviation) that does not change over time, whereas a non-stationary time series may lead to false causality results.[31] The most common way of testing for stationarity is through the augmented Dickey–Fuller test which uses estimates from an augmented autoregression as follows:

$$\Delta y^{n}_{t=1} = \beta_1 + a y_{t-1} + y\sum \Delta y_{t-1} + \varepsilon_t \,, \qquad \text{Equation 1}$$

where y_t represents all variables (in the natural logarithmic form) at time t, Δ is the first difference operator, β_1 is a constant, and n is the optimal lag length on the dependent variable. The test for a unit root is conducted on the coefficient of y_{t-1} in the regression model. The null and alternative hypotheses are represented by (H_0) and (H_1), respectively. The null hypothesis states that data need to be differenced to make it stationary, while the alternative hypothesis states that data are stationary and do not need to be differenced. To check for the existence of a unit root in variable y_t, we use: $H_0 := 0$ versus $H_1 : < 0$. The coefficient should be significantly different from zero (less than zero) for the hypothesis that y contains a unit root to be rejected. Rejection of the null hypothesis indicates stationarity in the series.

Differencing

The Granger causality test is ideal for stationary time series; however, if any of the time series in question are not stationary or if there is a root, then the series should be temporally differenced (Equation 2). For a variable y depending on another variable x (i.e. $y = f(x)$), and for a set of n points on an equi-spaced grid, the first derivative with respect to time t, f'_t at $i = 1$,.....n, the backward difference will be given by:

$$\Delta y_t \equiv y_{(t-1)} - y_t \qquad \text{Equation 2}$$

Model selection

Vector autoregressions (VAR) are often used in climate science to estimate the maximum lags used for testing data for Granger causality.[32-34] The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are used to find the optimal maximum lag.[35] These two information criteria follow a general form, which consists of the log-likelihood estimates as well as the penalty functions for the parameters in the model. Model estimates with the least information criteria is the best fitting model. The general information criteria are given as:

$$IC(p) = ln|\sum\hat{}\,| + C_T \frac{p}{T}, \qquad \text{Equation 3}$$

where $\sum\hat{}$ is an approximation of the residual covariance matrix associated with the fitted VAR(p) model, C_T is a deterministic penalty term, T denotes the number of observations used for estimation and p denotes the lag order. The definition of the penalty term C_T differs according to the choice of information criterion used:

for AIC,

$$C_T = 2k^2, \quad \text{Equation 4}$$

and for BIC,

$$C_T = k^2, \quad \text{Equation 5}$$

where k is the number of equations in the VAR model.

Granger causality

The presence and direction of Granger causality between each grid point of the Hadley cell and cloud cover is tested by means of VAR, indicated by Equations 6 and 7:

$$X_t = \alpha_1 + \sum_{i-1}^{S} \beta_i X_{t-i} + \sum_{i-1}^{S} \gamma_{1i} Y_{t-i} + e_{1t} \quad \text{Equation 6}$$

and

$$Y_t = \alpha_2 + \sum_{i-1}^{S} \beta_{2i} X_{t-i} + \sum_{i-1}^{S} \gamma_{2i} Y_{t-i} + e_{2t} \quad \text{Equation 7}$$

where α, β, and γ are regression coefficients, e is error term, and s is lag length, which is determined by using AIC and BIC (Equation 3). The structural VAR consisting of present values of the Hadley cell and cloud cover as functions of the lagged values of the dependent variables, and the present and lagged values of the independent variables, is the basis for the derivation of Equations 6 and 7. Structural VAR is based on the fact that for each season, present values of both the Hadley cell and cloud cover depend not only on present values of the other variable, but also on the history of the other variable. The direction of causal order is thus determined by estimating the restricted forms of Equations 6 or 7, and by eliminating the causal variable. For example, to determine whether the Hadley cell Granger causes cloud cover, we estimate a restricted form of Equation 6 in which cloud cover is eliminated. The restricted version of Equation 6 is thus:

$$X_t = \alpha_1' + \sum_{i-1}^{S} \beta_{1i}' X_{t-i} + e_{1t} \quad \text{Equation 8}$$

Conversely, it can also be determined whether the Hadley cell can be Granger caused by cloud cover, by estimating a restricted version of Equation 7 where the lagged values of the Hadley cell are omitted. We further tested whether the restricted model is statistically significantly different from the unrestricted model, as per Equation 9:

$$\omega = \frac{\left(\frac{RSS_r - RSS_u}{S}\right)}{\left(\frac{RSS_u}{T-k}\right)}, \quad \text{Equation 9}$$

where RSS is the sum of the residuals squared; the subscripts r and u refer to the restricted and unrestricted versions of Equations 6 or 7, respectively; T is the number of observations; k is the number of regressors in the unrestricted version of the equation; and s is the number of coefficients restricted to zero in Equation 8. The test statistic can be evaluated against an F distribution with s and T-k degrees of freedom in the numerator and denominator, respectively, in order to evaluate the null hypothesis that the cloud cover does not Granger cause the Hadley cell.

In this study, Equations 6 and 7 are used to analyse links between the Hadley cell and total cloud cover for each season (DJF and JJA) separately. The restrictive assumption is then constructed, such that the coefficients vary with the lag lengths and seasons. Equations 6 and 7 are then modified to cater for each season. As daily time steps were used to construct time series for both DJF and JJA, a lag length of one thus indicates the previous day, a lag length of two implies two previous days, and so on. It is recognised that the detection of Granger causality does not necessarily imply a physical causal mechanism between the two fields. Conclusions about the presence and direction of causality depend on the validity of the statistical models. A challenge of using Granger causality between two variables is that Granger causal implication estimates may be biased by the omission of relevant variables (e.g. vertical motion due to weather systems may contribute to total cloud cover, and its omission may have bearing on the Granger causal statistic) that are in fact the causal variables. Notwithstanding such limitations, the causality test is more reliable than lagged correlation statistics because the latter shows only the interaction between two variables, and may not indicate the presence and direction of causality.

Results and discussion

Climatology of the Hadley cell and total cloud cover

The Hadley cell is usually defined in terms of the zonally averaged stream function.[5,10] The zonally averaged stream function (defined in Supplementary appendix 1) is displayed in Figure 1, together with the vertical cross section of the divergent circulation in the meridional plane displayed by means of wind vectors. The clockwise direction of the mass stream function is indicated by blue shades; similarly the anticlockwise direction is illustrated by red shades of the stream function. The Hadley circulation consists of two branches: an ascending branch equatorward and a descending branch poleward. On the descending branch of the Hadley cell, the wind vectors advocate downward motion. Similarly, vertical ascent is evident on the ascending branch of the Hadley cell. The seasonal strength of the Hadley circulation is also evident from the stream function as well as the divergent circulation.

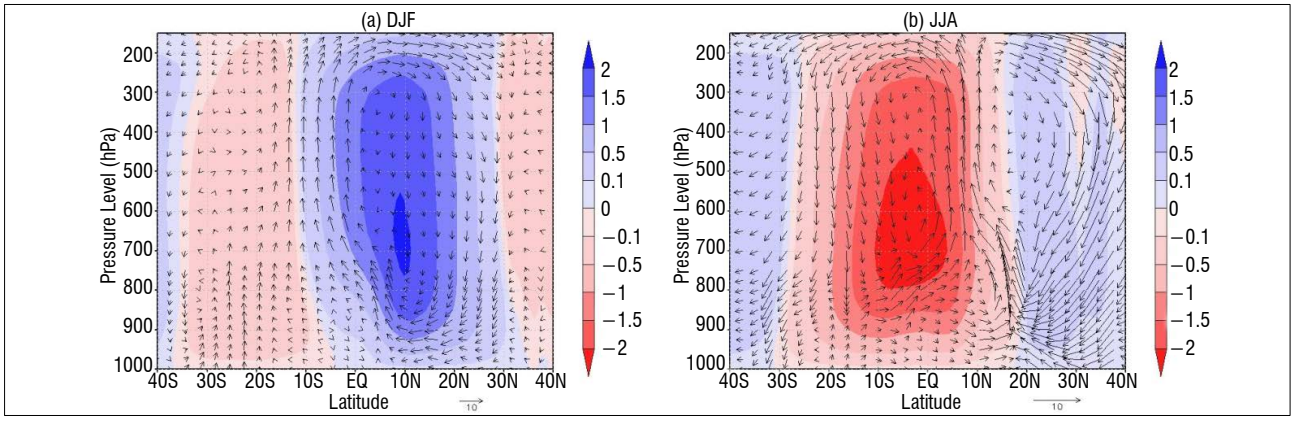


Figure 1: Climatology of zonally averaged stream function with clockwise direction (red shades) and anticlockwise direction (blue shades) and divergent circulation on a meridional plane (wind vectors) for (a) December–February (DJF) and (b) June–August (JJA). The contour intervals for the stream function are 10^{11}.

In DJF, the Hadley circulation is positioned more southwards; its descending branch is weak and has a wider gradient (Figure 1a). In JJA, the Hadley cell is positioned more northwards; it is more pronounced and stronger, as indicated by the tight gradient of the stream function (Figure 1b).

The meridional migration of the Hadley cell also displays intraseasonal variability as demonstrated by composite anomalies of the mass stream function for low and high cloudiness years during DJF (Figure 2a and Figure 2b) and JJA (Figure 2c and Figure 2d). For the low cloudiness years, the mass stream function anomalies have reversed sign compared to the climatological mass stream function shown in Figure 1a and 1b. This means that for the low cloudiness years (Table 1) during DJF (Figure 2a) and JJA (Figure 2c), the Hadley cell is positioned towards the equator and its width is narrower than in climatology.[36] The equatorward displacement of the Hadley cell is linked to subsidence over the subtropical regions of the southern hemisphere. In high cloudiness years for both DJF (Figure 2b) and JJA (Figure 2d), the mass stream function anomalies have the same sign as the climatological mass stream function.[36] In both seasons, the statistical significance covers a wider area for the high cloudiness years than the low cloudiness years, implying that the Hadley cell extends more polewards than normal, enhancing cloud formation over the southern hemisphere subtropics. While the zonally averaged stream function gives a general overview of the Hadley cell, it is easy to miss regional features of the Hadley circulation due to zonal averaging. Therefore, to examine the behaviour of the Hadley cell over South Africa, a zonally asymmetric diagnostic of the Hadley cell is necessary.

The zonally asymmetric Hadley cell is represented by the meridional mass flux at 500 hPa, as indicated in Supplementary appendix 2. The climatology of the zonally asymmetric Hadley cell and total cloud cover for DJF is provided in Figures 3 and 4, respectively. For both the DJF and JJA seasons (Figure 3 and Figure 4), the downward mass flux is dominant over the country, as indicated by negative mass flux values. Simultaneously, lower cloud cover is evident throughout the country. Both mass flux and cloud cover are indicative of subtropical weather. Vertical velocity plays an integral role in both the Hadley cell and total cloud cover. Negative vertical velocity is associated with vertical uplift and cloud formation, whereas positive vertical velocity values indicate subsidence and limited cloud development.

Negative vertical velocity is evident in DJF (Figure 3 and Figure 4), but very close to zero, which means that even with some vertical uplift, it is lower because the mean vertical motion over South Africa is downward. The eastern escarpment also indicates negative values of vertical velocity. High negative vertical velocity values are confined to areas of positive mass flux over the northern parts of the subcontinent. However, over South Africa, negative velocity is confined to the western interior of South Africa and the eastern escarpment. These areas are characterised by the downward mass flux (Figure 3) and relatively higher values of cloud cover (Figure 4), which means the downward mass flux is not responsible for the negative vertical velocity, but is rather the effect of weather systems that dominate over interior regions of South Africa during summer months, while uplift is due to orography (over the eastern escarpment) that contributes to negative vertical velocity and cloud cover. In JJA, a semi-permanent high pressure system dominates the country, resulting in a strong negative mass flux (Figure 5), as well as limited cloud development over the country (Figure 6). Subsidence due to the Hadley cell could be the main cause for the lack of upward vertical motion (negative vertical velocity), and could mean that limited cloud development results from a lack of upward vertical motion. These results demonstrate the role that vertical velocity in both the Hadley cell and cloud cover exhibit. Therefore, the effect of vertical velocity should thus be kept in mind when interpreting causality between the Hadley cell and cloud cover.

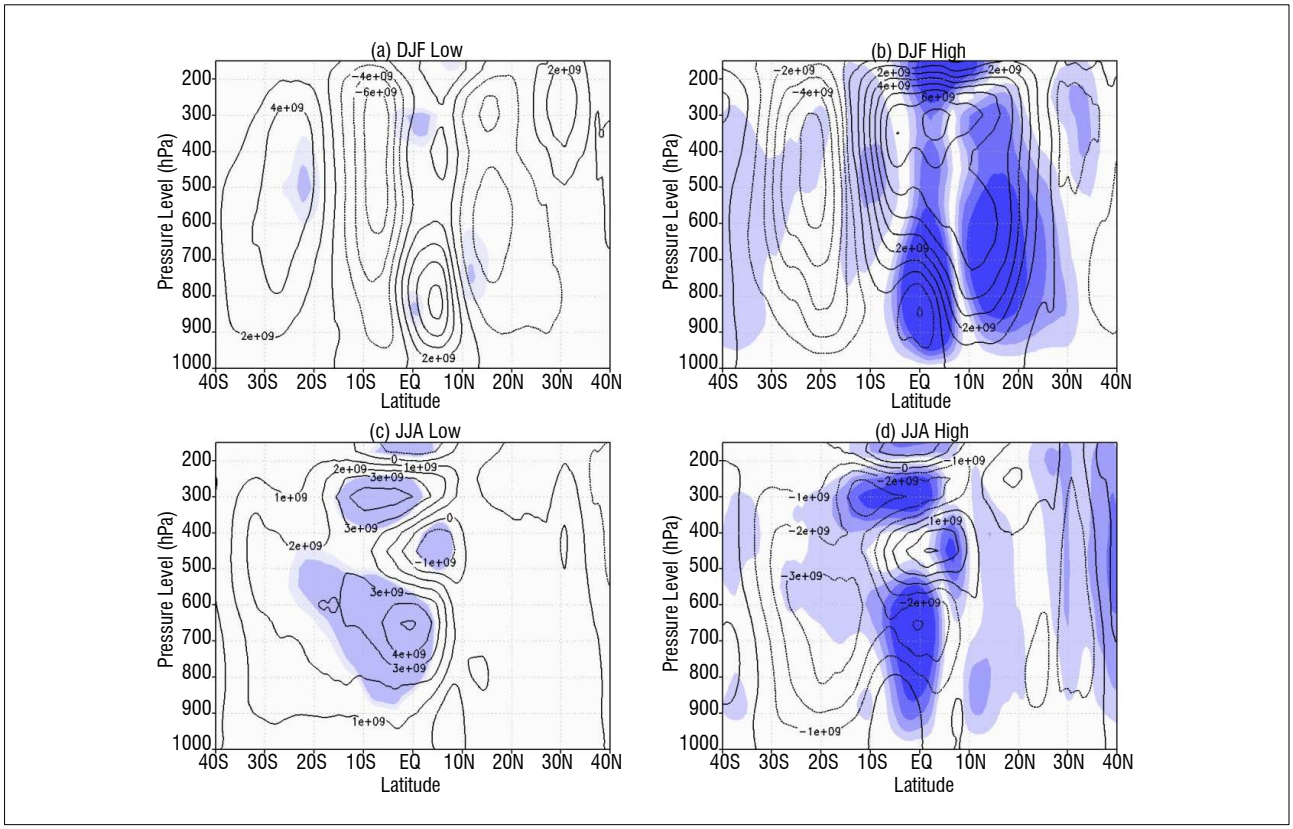


Figure 2: The mass stream function composite anomalies (contour plot) and statistical significance (blue shaded plot) for low cloudiness in (a) December–February (DJF), (c) June–August (JJA) and high cloudiness in (b) DJF and (d) JJA.

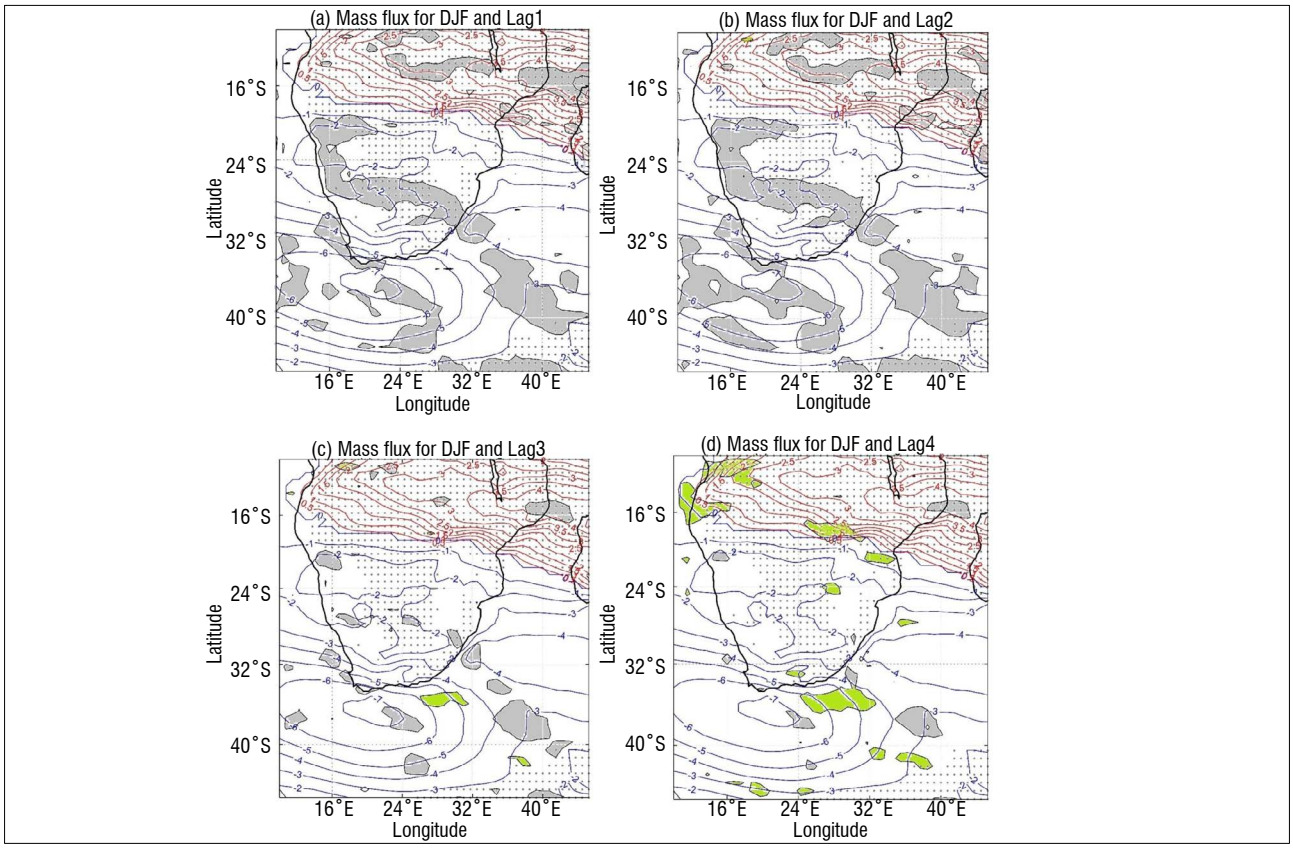


Figure 3: Seasonal mean of (a–d) the mass flux (m_ϕ) in December–February (DJF); units are kg/m²s. Negative (positive) values of the mass flux are represented by the blue (red) contours. The contours are plotted in 10^3 kg/m²s intervals. Correlation coefficients for the Hadley cell and total cloud cover relation in DJF for (a) Lag 1, (b) Lag 2, (c) Lag 3 and (d) Lag 4. Grey shading denotes correlation coefficients greater than 0.3 and green shading indicates correlation coefficients less than -0.3. Regions of upward vertical velocity (ω) in Pa/s are represented by grey dots. Where there are no grey dots, there is downward motion (i.e. $\omega > 0$).

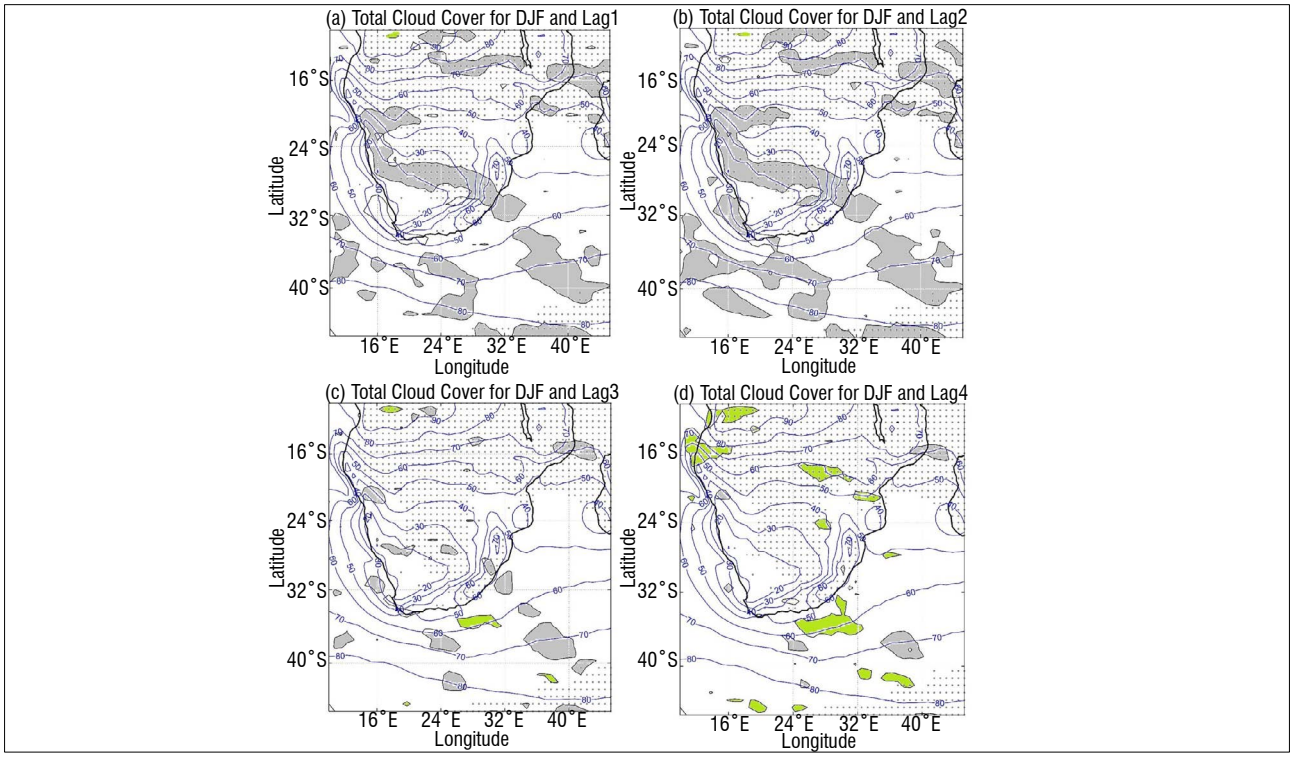


Figure 4: Total cloud cover area fraction (a–d) presented as percentage (contours) in December–February (DJF). Correlation coefficients for the Hadley cell and total cloud cover relation in DJF for (a) Lag 1, (b) Lag 2, (c) Lag 3 and (d) Lag 4. Grey shading denotes correlation coefficients greater than 0.3 and green shading indicates correlation coefficients less than -0.3. Regions of upward vertical velocity (ω) in Pa/s are represented by grey dots. Where there are no grey dots, there is downward motion (i.e. $\omega > 0$).

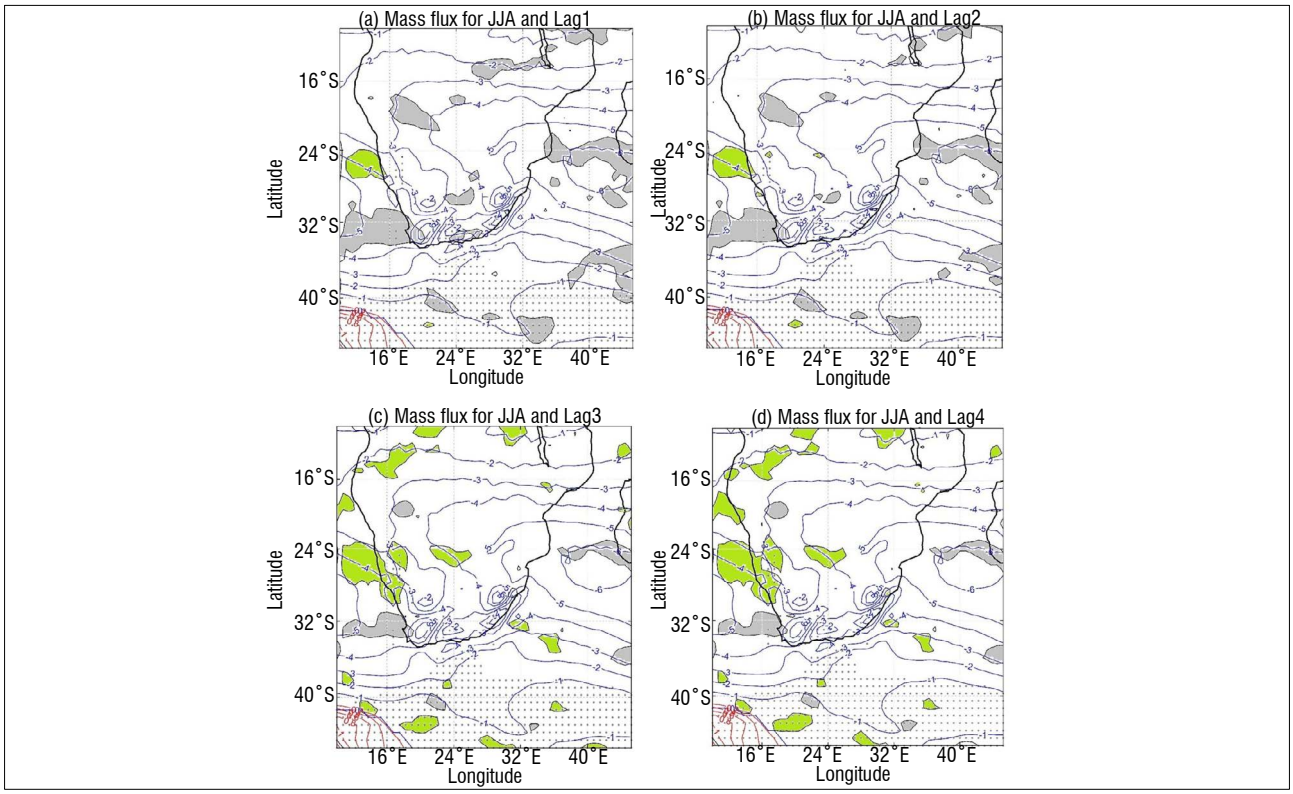


Figure 5: Seasonal mean of (a–d) the mass flux (mϕ) in June–August (JJA); units are kg/m^{2}s. Negative (positive) values of the mass flux are represented by the blue (red) contours. The contours are plotted in 10^3 kg/m^{2}s intervals. Correlation coefficients for the Hadley cell and total cloud cover relation in JJA for (a) Lag 1, (b) Lag 2, (c) Lag 3 and (d) Lag 4. Grey shading denotes correlation coefficients greater than 0.3 and green shading indicates correlation coefficients less than -0.3. Regions of upward vertical velocity (ω) in Pa/s are represented by grey dots. Where there are no grey dots, there is downward motion (i.e. $\omega > 0$).

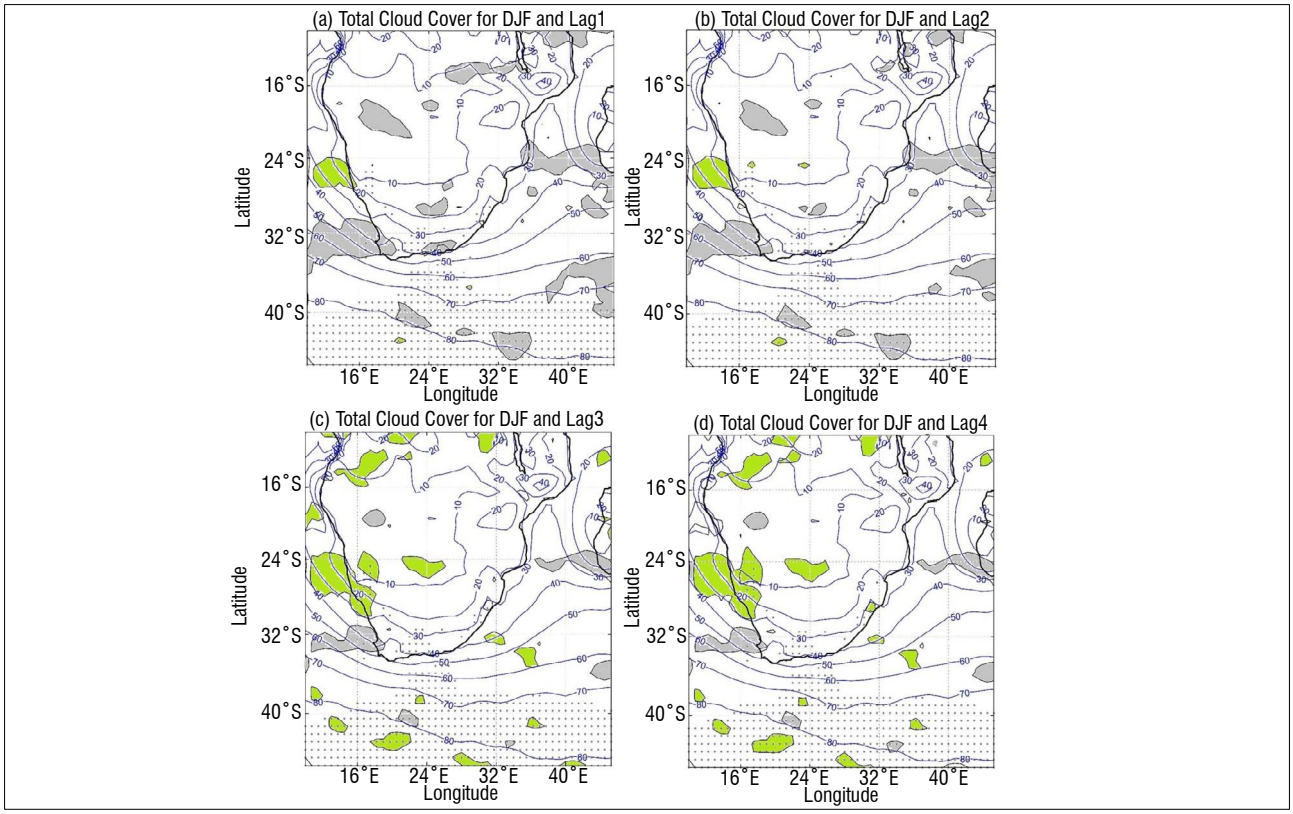


Figure 6: Total cloud cover area fraction (a–d) presented as percentage (contours) in June–August (JJA). Correlation coefficients for the Hadley cell and total cloud cover relation in JJA for (a) Lag 1, (b) Lag 2, (c) Lag 3 and (d) Lag 4. Grey shading denotes correlation coefficients greater than 0.3 and green shading indicates correlation coefficients less than -0.3. Regions of upward vertical velocity (ω) in Pa/s are represented by grey dots. Where there are no grey dots, there is downward motion (i.e. $\omega > 0$).

Stationarity

Based on the results of the augmented Dickey–Fuller test (Table 2), the unit root hypothesis could not be rejected for both time series; however, when the backward difference was performed for both time series, the unit hypothesis was rejected. This result suggests that both Hadley cell and total cloud cover are not stationary and might be considered stationary at the first differenced level. After attaining stationarity of the time series, the best model was further investigated by identifying the maximum lags that could be used in the VAR model. Most climate studies have used a lag length of 4 as the optimal length for causality analysis. In this study, we established an optimal lag length by using the AIC and BIC for DJF and JJA. Results from the two information criterion tests favoured a lag length of 2 as the maximum lag for fitting the VAR model (Figure 7a and 7b). This means that the previous 2 days are critical for predicting cloud cover when using the Hadley cell as the predictor, and vice versa.

Table 2: The Dickey–Fuller test for Hadley cell (HC), cloud cover (CC), and first differenced (DF) Hadley cell and cloud cover

Time series	DF value	*p*-value	Decision
HC	0	0.1	Do not reject
CC	0	0.1	Do not reject
HC'	-5.327	0.0001	Reject
CC'	-4.865	0.0001	Reject

Correlations

A number of studies make use of linear and lag correlations to study associations between two time series in climate science.[16,17] Lagged correlations between the Hadley cell and total cloud cover for DJF and JJA are averaged over South Africa (Figure 7c and Figure 7d). In the first week (1 to 7 days), relatively higher correlations (0.56 for DJF; 0.39 to 0.516 for JJA) were established. A substantial decrease in lag lengths then follows from the second week onwards. The decrease in lagged correlations from the second week onwards could mean that longer timescales (e.g. weeks to months) might have detrimental effects on the prediction of total cloud cover using the Hadley cell. In the first week, the second day (day 2) gave the highest correlation (0.56 and 0.516) for both DJF and JJA.

To determine zonal gradients of lag correlations between the Hadley cell and total cloud cover for DJF and JJA, a zonally asymmetric analysis for lag lengths of 1 to 4 days was performed, as informed by Figures 3–6. Positive correlations for lag lengths of 1 to 2 days over South Africa are indicated for the northwestern parts of South Africa and Namibia in DJF (Figures 3a, 3b, 4a and 4b). The northwestern parts of South Africa are the most arid in the country, and mostly dominated by the Hadley cell. The eastern parts of South Africa also show positive correlations for lag lengths of 1 to 2 days. No strong correlation is established over the southern interior of South Africa in DJF, which is attributed to the fact that cloud cover over these areas is caused by weather systems that are prevalent over South Africa, rather than the Hadley cell. Lag lengths of 2 days followed by a lag length of 1 day has a stronger positive correlation than lag lengths of 3 and 4 days, as indicated in Figure 3c, Figure 3d, Figure 4c and Figure 4d.

In JJA, a positive correlation between the Hadley cell and cloud cover for lag lengths of 1 and 2 days has been established over the interior of South Africa, as shown in Figure 5a and Figure 5b. This is attributed to the fact that a high pressure system is dominant over the interior of South Africa in JJA.[3] Subsidence leading to limited cloud development is hence expected over these areas. A strong positive correlation between the Hadley cell and cloud cover is recorded over the Atlantic Ocean, west of South Africa. From the four lag lengths investigated, the lag length of 1 showed the strongest correlation. No discernible correlation is established over South Africa for lag lengths of 3 to 4 days (Figure 5c, Figure 5d, Figure 6c and Figure 6d).

Granger causality

To determine whether variability in cloud cover is caused by the Hadley cell, the bivariate model indicated in Equations 6 and 7 was used. The F-statistic and critical values of the bivariate model were established at each grid point. The presence of Granger causality is denoted by values of the F-statistic being greater than the critical value and, conversely, values of the F-statistic lower than the critical value show the absence of Granger causality. The strength of Granger causality from the Hadley cell to total cloud cover (denoted as HC-CC) and total cloud cover to the Hadley cell (denoted as CC-HC) is indicated in Figures 8 and 9 for DJF and JJA, respectively. The presence of Granger causality is established from the HC-CC interaction over South Africa. In DJF, the western and northwestern regions show higher values of HC-CC causality. However, maximum causality is located over the far northeastern parts of the subcontinent between latitudes 15°S to 10.5°S.

The presence of CC-HC Granger causality has also been established over South Africa, but is limited to only the eastern parts of the country, while causality is absent over the western and southeastern parts. Some places

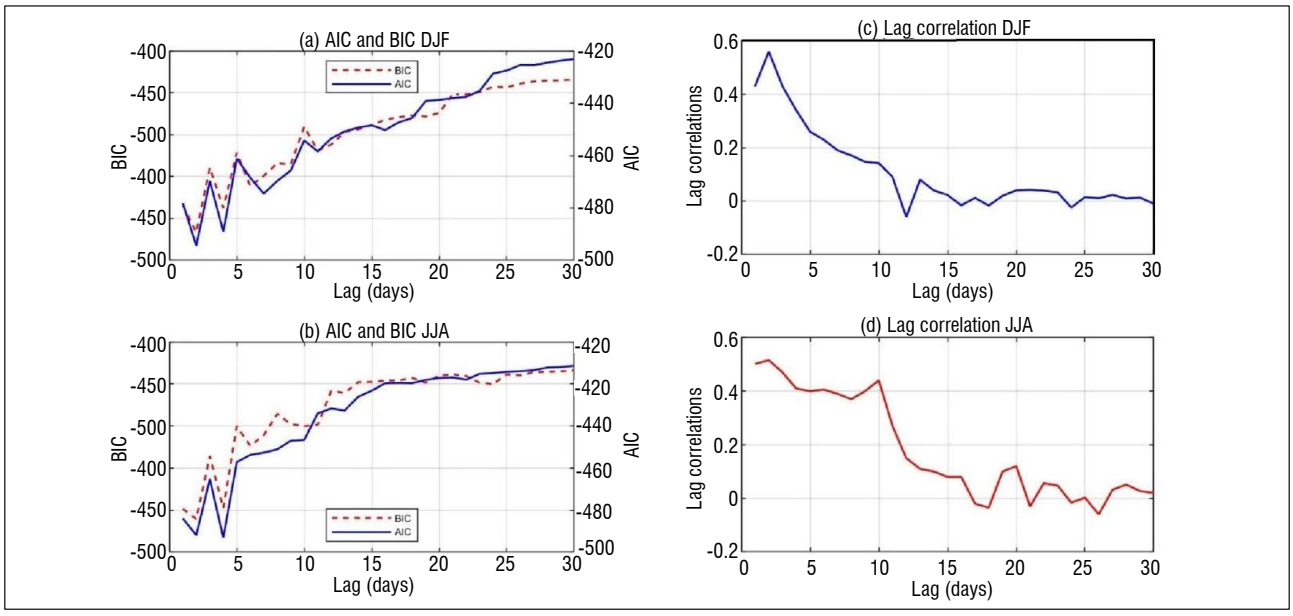


Figure 7: (a,b) Akaike information criterion (AIC, blue solid line) and Bayesian information criterion (BIC, red dashed line) values for (a) December–February (DJF) and (b) June–August (JJA); and (c,d) lag correlations for (c) DJF and (d) JJA.

over Mozambique and Zimbabwe indicate very weak CC-HC causality. In JJA, the HC-CC causality is stronger than the CC-HC causality, and the maximum causality values are situated over the northern interior of South Africa. No causality has been established over the southwestern parts of South Africa. One notable feature is that the HC-CC maximum causality has moved further northwards in JJA than it was in DJF, which could be linked with the northwards migration of the ITCZ during austral winter. Although the CC-HC causality is weak over most parts of South Africa, especially in JJA, the southwestern parts advocate an absence of Granger causality.

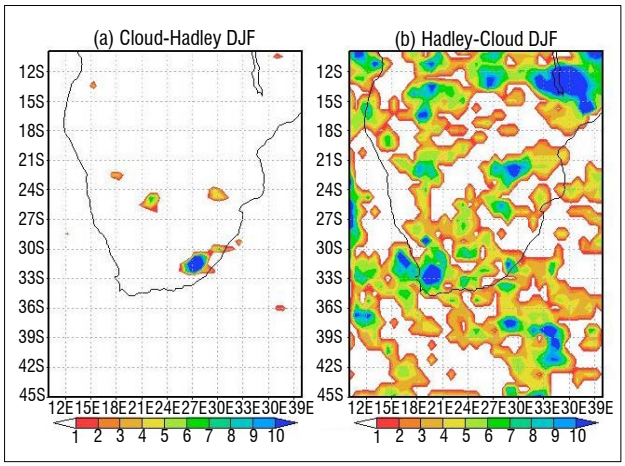


Figure 8: (a) Granger causality of total cloud cover to the Hadley cell; and (b) Hadley cell to total cloud cover in December–February (DJF).

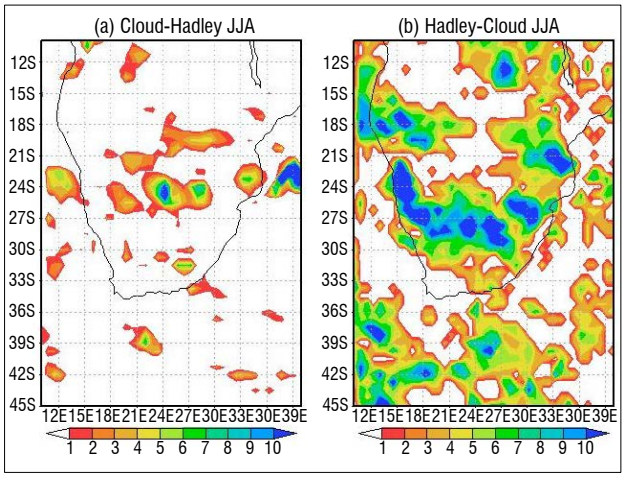


Figure 9: (a) Granger causality of total cloud cover to the Hadley cell; and (b) Hadley cell to total cloud cover in June–August (JJA).

As previously mentioned,[20,22,26] the main challenge with Granger causality in a bivariate model is the exclusion of other variables that may affect the causing variable, thus leading to false causality of the variables being tested. This could mean that, although it is weak, the CC-HC causality may be false as a result of the omission of a third variable that is affecting the causality result. For example, cloud cover is a function of vertical velocity, which in turn may affect the Hadley cell. The effect of vertical velocity due to other weather systems has an indirect impact on the causality of both the Hadley cell and cloud cover. Secondly, the structure of the CC-HC causality in DJF is oriented in a northwesterly to southeasterly direction, resembling the structure of tropical temperature troughs. This orientation could mean that the vertical velocity due to tropical temperature troughs, which are a summer phenomenon over southern Africa,[37] is responsible for the causality from the cloud cover to the Hadley cell, leading to CC-HC false causality. In JJA, the HC-CC causality is stronger than the CC-HC causality. The HC-CC stronger causality is greater over the interior and northwards of South Africa. The strong HC-CC causality could be attributed to the fact that the Hadley cell informs limited cloud development over the subtropics and during JJA (Figure 9a). The far northern and southwestern parts of South Africa show relatively weaker HC-CC causality. Cloud cover over the southwestern parts is greater because of frontal systems, and not entirely due to the Hadley cell, and hence produces weaker causality.

To illustrate the effect of vertical velocity causality on both the Hadley cell and cloud cover, Granger causality was analysed between vertical velocity and cloud cover (Figure 10). The omega to cloud cover interaction is denoted as OMEGA-CC. Causality was established for both DJF and JJA. In DJF, the OMEGA-CC causality is located predominantly over the eastern to western escarpment of South Africa (Figure 10a), which means that vertical velocity due to topography (orographic uplift) is the main cause of cloud cover along the escarpment regions of South Africa. The western and adjacent oceanic regions show maximum OMEGA-CC causality. Strong OMEGA-CC causality is also evident between Mozambique and Zimbabwe. In JJA, most of the South African interior shows very weak OMEGA-CC causality; however, the northern and northwestern parts of South Africa show strong OMEGA-CC causality (Figure 10b). Figure 11 depicts Granger causality of vertical velocity to the Hadley cell which we identify as OMEGA-HC.

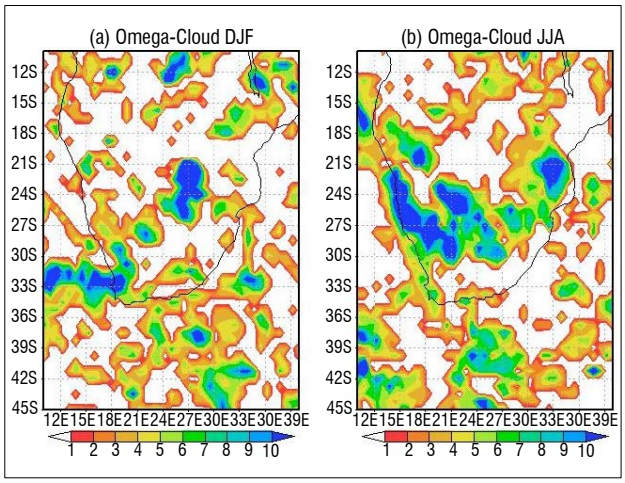


Figure 10: Granger causality of omega to cloud cover in (a) December–February (DJF) and (b) June–August (JJA).

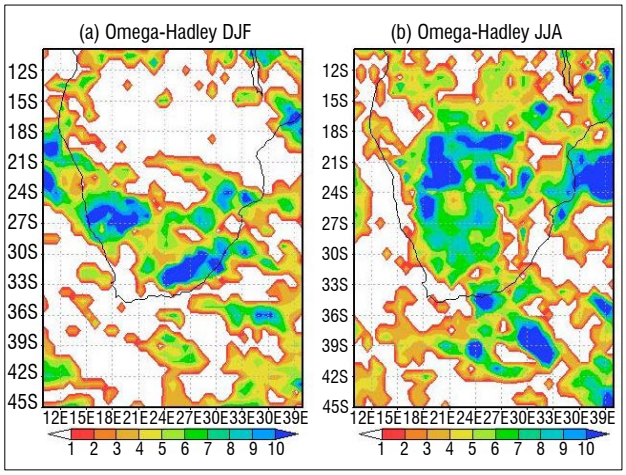


Figure 11: Granger causality of omega to the Hadley cell in (a) December–February (DJF) and (b) June–August (JJA).

The OMEGA-HC causality is further southwards and weaker in DJF and further northwards and stronger in JJA, which may be linked with the northwards-southwards seasonal migration of the ITCZ. In DJF, OMEGA-HC causality is strong over the summer region of South Africa and stronger along the southeastern parts. In JJA, strong OMEGA-HC causality is located over the interior of South Africa where the subtropical high pressure cell is located. The southwestern parts of South Africa show an absence

of OMEGA-HC causality for both DJF and JJA, which may be attributed to weather over the southwestern parts being Mediterranean in nature, and mostly owing to southerly frontal systems, rather than the Hadley cell.

Conclusions

We used Granger causality time series modelling to quantitatively diagnose the influence of the seasonal Hadley cell on cloud cover and vice versa. Bivariate vector autoregressive time series models for the period 1980 to 2015 were fitted to the seasonal Hadley cell and total cloud cover. According to the AIC and BIC statistics, the optimum lag used in the Hadley cell – cloud cover bivariate model was 2 days. The highest correlation coefficient of the lag correlations from averaged data was also 2 days. The lag correlations calculated from each grid point indicate lag lengths of 1–2 days for strongest correlation. The strength of these correlations is seasonally dependent and suggests linkage between the meridional shift of the Hadley cell and the seasonal migration of the ITCZ.[10] The meridional shift of the Hadley cell is seasonal, but with intraseasonal variability as informed by the seasonal climatology of the mass stream function as well as from composite anomalies of the mass stream function during high and low cloudiness years. The lag correlation between the Hadley cell and cloud cover is stronger in DJF than in JJA. This difference may be due to synoptic systems (e.g tropical temperate troughs and westerly waves) leading to cloud formation, which dominate in DJF. Such systems have life spans of one to several days.[37,38] It should also be noted that typically 2 days before the development of tropical temperate troughs over South Africa, there is a migration of easterly moisture from Botswana to the subtropical region of South Africa, which is an example of the strong 2-day lag correlation in DJF.[38]

We further established that the Hadley cell can Granger cause cloud cover for both DJF and JJA. This causality is associated mostly with limited cloud development. For example, causality is found over the interior of South Africa, where the subtropical high pressure cell dominates. In DJF, the western parts of South Africa experience limited cloud development, and this area is mostly where causality has been established. Weak causality from cloud cover to the Hadley cell has also been established. This case is typical of spurious causality in bivariate systems caused by the exclusion of a third variable in the Granger causality analysis. Further analysis has shown that vertical velocity is the main variable that could have led to the causality of cloud cover to the Hadley cell. Finally, although we have shown that Granger causality is a useful tool in establishing causality from the Hadley cell to cloud cover, to obtain the best outputs, the scientific background of the variables being analysed should be well understood before correct interpretations of the causality results can be made.

Formally establishing the causality of anomalously strong regional Hadley cell circulation in terms of limited cloud development over southern Africa is of profound importance within the context of future climate change over the region. Southern Africa has been assessed to be likely to become generally drier under low mitigation climate change futures.[11,39-41] The underlying circulation changes in the model projections of climate change that led to this assessment have in fact been shown to include the frequent occurrence of mid-level anti-cyclones and subsidence[11,39] over the region, which implies a strengthening of the descending branch of the Hadley cell. Increasingly limited cloud development over southern Africa in a warming world, in response to changes in the regional Hadley cell, may thus be an important factor driving the relatively high rate of observed and projected temperature increases over southern Africa.[42] Increasing limited cloud development may further enhance southern Africa's solar potential within the context of a growing renewable energy sector, although such changes will be largely detrimental to agriculture and may further compromise water security in the region.

Acknowledgements

This work was financially supported by the National Research Foundation of South Africa (grant no. 102215). We are grateful to Mr Chang-Hyan Park for his valuable input prior to the submission of this manuscript and to the anonymous reviewers for their input in improving the manuscript.

Authors' contributions

D.D.M. conceptualised the study, analysed the data and wrote the paper. T.N., S.W.G. and F.A.E. conceptualised the study, and reviewed and approved the final manuscript.

References

1. South African Department of Energy (DoE). Integrated resource plan for electricity (IRP) 2010-2030. Updated report. Pretoria: DoE; 2013.
2. Loeb NG, Schuster GL. An observational study of the relationship between cloud, aerosol and meteorology in broken low-level cloud conditions. J Geophys Res. 2008;113(D14), D14214, 9 pages. https://doi.org/10.1029/2007JD009763
3. Tyson PD, Preston-Whyte RA. The weather and climate of southern Africa. 2nd ed. Cape Town: Oxford University Press; 2000.
4. Engelbrecht CJ, Engelbrecht FA. Shifts in Köppen-Geiger climate zones over southern Africa in relation to key global temperature goals. Theor Appl Climatol. 2016;123:247–261. https://doi.org/10.1007/s00704-014-1354-1
5. Seidel DJ, Fu Q, Randel WJ, Reichler TJ. Widening of the tropical belt in a changing climate. Nat Geosci. 2008;1:21–24. https://doi.org/10.1038/ngeo.2007.38
6. d'Abreton PC, Lindesay JA. Water vapour transport over southern Africa during wet and dry early and late summer months. Int J Climatol. 1993;13:151–170. https://doi.org/10.1002/joc.3370130203
7. Tennant WJ, Reason CJC. Associations between the global energy cycle and regional rainfall in South Africa and southwest Australia. J Clim. 2005;18:3032–3047. https://doi.org/10.1175/JCLI3464.1
8. Mahlobo DD, Ndarana T, Grab S, Engelbreht F. Integrated climatology and trends in the subtropical Hadley cell, sunshine duration and cloud cover over South Africa. Int J Climatol. 2018;39:1805–1821. https://doi.org/10.1002/joc.5917
9. Manabe S, Holloway JL. The seasonal variation of the hydrological cycle as simulated by a global model of the atmosphere. J Geophys Res. 1975;80:1617–1649. https://doi.org/10.1029/JC080i012p01617
10. Kang SM, Lu J. Expansion of the Hadley cell under global warming: Winter versus summer. J Clim. 2012;25:8387–8393. https://doi.org/10.1175/JCLI-D-12-00323.1
11. Engelbrecht FA, McGregor JL, Engelbrecht CJ. Dynamics of the conformal-cubic atmospheric model projected climate-change signal over southern Africa. Int J Climatol. 2009;29:1013–1033. https://doi.org/10.1002/joc.1742
12. Granger CWJ. Investigating causal relations by econometric models and cross spectral methods. Econometrica.1969;37:424–438. https://doi.org/10.2307/1912791
13. Dee DP, Uppala SM, Simmo AJ, Berrisford P, Poli P, Kobayashi S, et al. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Q J Royal Meteorol Soc. 2011;37:553–597. https://doi.org/10.1002/qj.828
14. Schwendike J, Govekar P, Reeder MJ, Wardle R, Berry GJ, Jakob C. Local partitioning of the overturning circulation in the tropics and the connection to the Hadley and Walker circulations. J Geophys Res. 2014;119:1322–1339. https://doi.org/10.1002/2013JD020742.
15. Nguyen H, Evans A, Lucas C, Smith I, Timbal B. The Hadley circulation in reanalyses: Climatology, variability, and change. J Clim. 2013;26:3357–3376. https://doi.org/10.1175/JCLI-D-12-00224.1.
16. Nemani R, Keeling CD, Hashimoto H, Jolly WM, Piper SC, Tucker CJ, et al. Climate driven increases in global terrestrial net primary production from 1982 to 1999. Science. 2003;300:1560–1563. https://doi.org/10.1126/science.1082750
17. Wu D, Zhao X, Liang S, Zhou T, Huang K, Tang B, et al. Time-lag effects of global vegetation responses to climate change. Glob Change Biol. 2015;21:3520–3531. https://doi.org/10.1111/gcb.12945
18. Abramson B, Brown J, Edwards W, Murphy M, Winkler R. A Bayesian system for forecasting severe weather. Int J Forecast.1996;12:57–71. https://doi.org/10.1016/0169-2070(95)00664-8
19. Cofino A, Cano R, Sordo C, Gutierrez JM. Bayesian networks for probabilistic weather prediction. In: Proceedings of the 15th European Conference on Artifical Intelligence. Amsterdam: IOS Press; 2002. p. 695–700.

20. Wang W, Anderson BT, Kaufmann RK, Myneni RB. The relation between the North Atlantic Oscillation and SSTs in the North Atlantic Basin. J Clim. 2004;17:4752–4759. https://doi.org/10.1175/JCLI-3186.1

21. Elsner JB. Evidence in support of the climate change Atlantic hurricane hypothesis. Geophys Res Lett. 2006;33, L16705, 3 pages. https://doi.org/10.1029/2006GL026869

22. Mosedale TJ, Stephenson DB, Collins M, Mills TC. Granger causality of coupled climate processes: Ocean feedback on the North Atlantic Oscillation. J Clim. 2006;19:1182–1194. https://doi.org/10.1175/JCLI3653.1

23. Attanasio A, Pasini A, Triacca U. Granger causality analyses for climatic attribution. Atmos Clim Sci. 2013;3:515–522. https://doi.org/10.4236/acs.2013.34054

24. Zou C, Feng J. Granger causality vs. dynamic bayesian network inference: A comparative study. BMC Bioinform. 2009;10, Art. #122, 17 pages. https://doi.org/10.1186/1471-2105-10-122

25. Diks C, Mudelsee M. Redundancies in the earth's climatological time series. Phys Lett. 2000;275:407–414. https://doi.org/10.1016/S0375-9601(00)-00613-7

26. Papagiannopoulou C, Miralles DG, Decubber S, Demuzere M, Verhoest NEC, Dorigo WA, et al. A non-linear Granger-causality framework to investigate climate–vegetation dynamics. Geosci Model Dev. 2017;10:1945–1960. https://doi.org/10.5194/gmd-10-1945-2017

27. Kaufmann RK, Stern DI. Evidence for human influence on climate from hemispheric temperature relations. Nature.1997;388:39–44. https://doi.org/10.1038/40332

28. Geweke J. Measurement of linear dependence and feedback between multiple time series. J Am Stat Assoc.1982;77:304–313. https://doi.org/10.2307/2287238

29. Barnett L, Seth AK. The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. J Neurosci Methods. 2014;223:50–68. https://doi.org/10.1016/j.jneumeth.2013.10.018

30. Guo S, Seth AK, Kendrick KM, Zhoud C, Feng J. Partial Granger causality – eliminating exogenous inputs and latent variables. 2008;172:79–93. https://doi.org/10.1016/j.jneumeth.2008.04.011

31. Sims CA, Stock JH, Watson MW. Inference in linear time series models with some unit roots. Econometrica.1990;58:113–144. https://doi.org/10.2307/2938337

32. Triacca U. On the use of Granger causality to investigate the human influence on climate. Theor Appl Climatol. 2001;69:137–138. https://doi.org/10.1007/s007040170019

33. Attanasio A, Pasini A, Triacca U. A contribution to attribution of recent global warming by out-of-sample Granger causality analysis. Atmos Sci Lett. 2012;13:67–72. https://doi.org/10.1002/asl.365

34. Pasini A, Triacca U, Attanasio A. Evidence of recent causal decoupling between solar radiation and global temperature. Environ Res Lett. 2012;7, Art. #034020, 6 pages. https://doi.org/10.1088/1748-9326/7/3/034020

35. He W. Model selection for cointegrated relationships in small sample [MSc dissertation]. Port Elizaberth: Nelson Mandela Metropolitan University; 2008.

36. Guo Y-Pe, LI J-Ping. Impact of ENSO events on the interannual variability of Hadley circulation extents in boreal winter. Adv Clim Change Res. 2016;7:46–53. https://doi.org/10.1016/j.accre.2016.05.001

37. Hart NCG, Reason CJC, Fauchereau N. Tropical extratropical interactions over southern Africa: Three cases of heavy summer season rainfall. Mon Weather Rev. 2010;138:2608–2623. https://doi.org/10.1175/2010MWR3070.1

38. Singleton AT, Reason CJC. Variability in the characteristics of cut-off low pressure systems over subtropical southern Africa. Int J Climatol. 2007;27:295–310. https://doi.org/10.1002/joc.1399

39. Christensen JH, Hewitson B, Busuioc A, Chen A, Gao X, Held I, et al. Regional climate projections. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, et al., editors. Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press; 2007.

40. Niang I, Ruppel OC, Abdrabo MA, Essel A, Lennard C, Padgham J, et al. Africa. In: Barros VR, Field CB, Dokken DJ, Mastrandrea MD, Mach KJ, Bilir TE, et al., editors. Climate change 2014: Impacts, adaptation, and vulnerability. Part B: Regional aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press; 2014.

41. Hoegh-Guldberg O, Jacob D, Taylor M, Bindi M, Brown S, Camilloni I, et al. Chapter 3: Impacts of 1.5°C global warming on natural and human systems. In: Masson-Delmotte V, Zhai P, Pörtner H-O, Roberts D, Skea J, Shukla PR, et al., editors. Global warming of 1.5°C: An IPCC special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change. Intergovernmental Panel on Climate Change; 2018.

42. Engelbrecht F, Adegoke J, Bopape M, Naidoo M, Garland R, Thatcher M, et al. Projections of rapidly rising surface temperatures over Africa under low mitigation. Environ Res Lett. 2015;10, Art. #085004, 16 pages. https://doi.org/10.1088/1748-9326/10/8/085004

AUTHORS:
Rakhee Lakhraj-Govender[1]
Stefan W. Grab[2]

AFFILIATIONS:
[1]School of Environmental, Water and Earth Sciences, Tshwane University of Technology, Pretoria, South Africa
[2]School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand, Johannesburg, South Africa

CORRESPONDENCE TO:
Rakhee Lakhraj-Govender

EMAIL:
rakhee.lg@gmail.com

DATES:
Received: 12 Feb. 2019
Revised: 06 May 2019
Accepted: 07 May 2019
Published: 26 Sep. 2019

HOW TO CITE:
Lakhraj-Govender R, Grab SW. Rainfall and river flow trends for the Western Cape Province, South Africa. S Afr J Sci. 2019;115(9/10), Art. #6028, 6 pages. https://doi.org/10.17159/sajs.2019/6028

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITORS:
Nicolas Beukes
Yali Woyessa

KEYWORDS:
Mann–Kendall, water resource management, decreasing trends, spatio-temporal distribution

FUNDING:
None

Rainfall and river flow trends for the Western Cape Province, South Africa

Climate change has the potential to alter the spatio-temporal distribution of rainfall, subsequently affecting the supply and demand of water resources. In a water-stressed country such as South Africa, this effect has significant consequences. To this end, we investigated annual and winter rainfall and river flow trends for the Western Cape Province over two periods: 1987–2017 and 1960–2017. Annual rainfall for the most recent 30-year period shows decreasing trends, with the largest magnitude of decrease at the SA Astronomical Observatory rainfall station (-54.38 mm/decade). With the exception of the significant decreasing winter rainfall trend at Langewens (-34.88 mm/decade), the trends vary between stations for the period 1960–2017. For the period 1987–2017, statistically significant decreasing winter trends were found at four of the seven stations, and range from -6.8 mm/decade at Cape Columbine to -34.88 mm/decade at Langewens. Similarly, the magnitudes of decreasing winter river flow at Bree@Ceres and Berg@Franschoek are greater for the more recent 30-year period than for 1960–2017. Correlation coefficients for Vilij@Voeliv rainfall and four river flow stations (Berg@Franschoek, Bree@Ceres, Wit River@Drosterkloof and Little Berg@Nieuwkloof) are stronger for shorter periods (i.e. 1987–2017 and 2007–2017) than that for the longer period, 1960–2017. The Intergovernmental Panel on Climate Change emphasises the importance of studies to assist with model prediction uncertainties. To this end, our study expands the understanding of regional hydrological responses to rainfall change in the water stressed region of the Western Cape Province.

Significance:

- Historical trend analyses provide the basis for future rainfall and river flow projections and also improve our understanding of hydrological responses to rainfall change, which is important for water resource planning and management in light of recent rainfall shortages experienced in the Western Cape region.
- Anthropogenic and natural influences on river systems must be taken into account when assessing the impact of rainfall changes on river flow because these factors ultimately affect water resource management.
- We identified decreasing winter rainfall trends across the Western Cape Province for the most recent three decades.

Introduction

Anthropogenic activities such as an increasing human population, economic development and urbanisation, place enormous pressure on water resources across the globe.[1] Climate change significantly impacts the supply and demand of such water resources.[2] The Intergovernmental Panel on Climate Change provides robust evidence that current climate change is projected to decrease renewable surface water.[3] One of the most significant consequences of climate change is the change in magnitude and frequency of precipitation.[4] Rising temperatures increase evaporation levels, spatio-temporal run-off distribution and water resource storage.[5] It is anticipated that global warming will intensify the hydrological cycle, resulting in extreme high/low river flows, depending on the region.[6] In a global analysis of 200 rivers, significant trends were recorded for one third of these rivers – 45 decreasing and 19 increasing.[1]

The Intergovernmental Panel on Climate Change reports (2014) indicate that studies examining hydrological responses to climate change in Africa are limited because of inadequate observational data.[3] However, there has been some work assessing spatio-temporal rainfall variability across South Africa.[7,8,] While statistically significant decreases in rainfall were recorded for the central and northeastern regions of South Africa, mixed signals were reported at different stations in some regions, over the period 1960–2010.[7] For the period 1921–2015, increasing annual rainfall trends were reported at the majority of stations across the western and southern interior regions, while decreasing rainfall trends were recorded across the northern and northeastern regions of South Africa.[8]

Studies investigating the relationship between long-term river flow and rainfall trends have been limited in South Africa. For instance, decreasing rainfall trends were measured at four of six stations over the Luvuvhu River catchment, for the period 1920/1921–2005/2006.[9] For the same period, two streamflow stations showed increasing trends and two decreasing trends. For the upper Limpopo River catchment (i.e. Crocodile, Marico, Mahalapse and Lotsane Rivers), annual rainfall and evaporative losses were used to calculate the summer surface water balance for the period 1959–2014.[10] Here too, decreasing river flow and surface water balance trends were measured.[10] Given that no similar studies have been undertaken to assess the situation for the increasingly water-stressed Western Cape Province, our primary aim was to determine rainfall and river flow trends for this region, and to investigate the relationship between historical rainfall and river flow using Mann–Kendall, Sen's slope estimator and descriptive statistics.

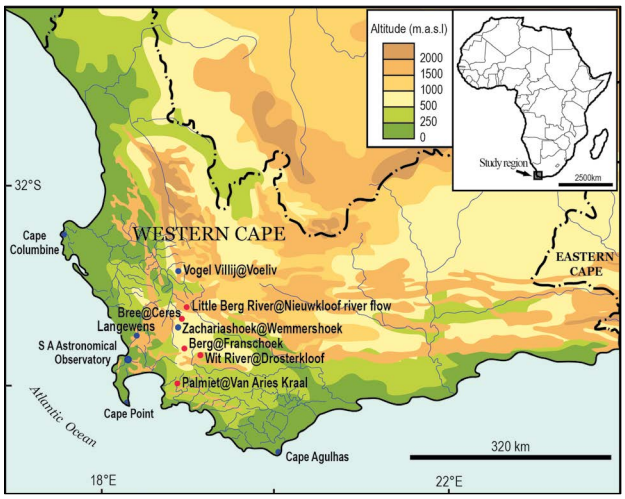
Methodology

Study area and data

South Africa is a water-stressed country with high spatio-temporal rainfall variability.[11] The Western Cape Province of South Africa has a warm temperate Mediterranean climate, with rainfall predominating during austral winter and early spring.[12] Mean annual rainfall varies across the Western Cape, with mountainous regions receiving up to ca 3000 mm of rain, while low-lying regions (40 m) receive less than 200 mm.[12] The Western Cape region is well known

for its history of water shortages.[13] The province recently (2014–2017) experienced the worst water deficiencies in 113 years.[13] A recent study focusing on drought monitoring indicators and 3-, 6- and 12-month Standardized Precipitation Indexes for the Western Cape Province, concluded that the 2015/2016 drought was part of a natural cycle.[13]

To illustrate the spatial variability of rainfall, monthly rainfall data provided by the South African Weather Service for the stations Cape Agulhas, Cape Columbine, Cape Point, Langewens and SA Astronomical Observatory, were used to determine annual and austral winter rainfall trends for the Western Cape region. The remaining rainfall data (i.e. that from Vogel Vilij@Voeliv and Zachariashoek@Wemmershoek stations) were obtained from the South African Department of Water and Sanitation (from the Internet) and used to determine the relationship with river flow stations located downstream (i.e. Bree@Ceres, Berg@Franschoek, Wits River@Drosterkloof, Palmiet@van Aries Kraal and Little Berg@Nieuwkloof). Stations were selected based on the availability of data (stations with data for the entire period of study, i.e. 1960–2017), quality (screening of extremely low/high monthly values for obvious errors) and completeness of data records (does not exceed 5% missing data), and location within the winter rainfall region (Figure 1; Table 1).



Blue dots, rainfall stations; red dots, river flow stations

Figure 1: Map showing the Western Cape rainfall stations from which data were sourced in this study.

To determine the influence of rainfall changes on river flow, rainfall stations located upstream of the river flow gauges were selected to compare rates of change with time. Annual and winter rainfall and river flow trends were determined using the non-parametric modified Mann–Kendall statistical test and quantified with Sen's slope estimator.[14-16] Data for the winter season were analysed given that the selected stations are located in a winter rainfall region. Rainfall for June, July and August were used to calculate the winter rainfall average, as 48% of the mean annual rainfall for this region falls during these winter months.[17] Mann–Kendall is calculated as[16]:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sign(x_i - x_j), \qquad \text{Equation 1}$$

where S is the statistic, n is the number of data points and x_i and x_j are sequential data points.

$$sign(x_i - x_j) = \begin{cases} +1, \text{ if } (x_i - x_j) > 0 \\ 0, \text{ if } (x_i - x_j) = 0 \\ -1, \text{ if } (x_i - x_j) < 0 \end{cases} \qquad \text{Equation 2}$$

The variance of S is calculated as:

$$Var(S) = \frac{n(n-1)(2n+5) - \sum_{i=1}^{m} t_i (t_i - 1)(2t_i + 5)}{18}, \qquad \text{Equation 3}$$

where t_i denotes number of ties with i tied values and m is the number of tied groups of values. Then the standard normal variate Z is computed as[16]:

$$Z = \begin{cases} \frac{S-1}{\sqrt{Var(S)}} & S > 0 \\ 0 & S = 0 \\ \frac{S+1}{\sqrt{Var(S)}} & S < 0 \end{cases} \qquad \text{Equation 4}$$

The trend is considered significant when $Z > 1.96$ (or < -1.96) for a selected 5% significance level ($\alpha = 0.05$). The trend is quantified with Sen's slope estimator[16], which is assumed to be linear according to Equation 5:

$$f(t) = Qt + B, \qquad \text{Equation 5}$$

where t is the time index, Q is the slope and B is the constant.[16]

The Pearson correlation coefficient was used to determine the relationship between annual rainfall and river flow for the periods 1960–2017, 1987–2017 and 2007–2017. Descriptive statistics (standard deviation and coefficient of variation) were also calculated to compare rainfall and river flow.

Table 1: List of stations located in the Western Cape Province from which data were sourced for this study

Station name	Latitude	Longitude	Altitude (masl)	Station type
Vogel Villij@Voeliv	-33.34	19.04	683	Rainfall
Zachariashoek@ Wemmershoek	-33.70	19.01	1735	Rainfall
Cape Columbine	-32.48	18.29	62	Rainfall
Cape Point	-34.20	18.28	228	Rainfall
SA Astronomical Observatory	-33.56	18.28	15	Rainfall
Langewens	-33.27	18.25	179	Rainfall
Cape Agulhas	-34.49	20.01	11	Rainfall
Bree@Ceres	-33.38	19.30	269	River flow
Berg@Franschoek	-33.89	19.08	315	River flow
Wit River@Drosterkloof	-33.97	19.15	838	River flow
Little Berg@Nieuwkloof	-33.36	19.08	1021	River flow
Palmiet@van Aries Kraal	-34.20	18.98	315	River flow

Results

Rainfall trends

Here we present the results for the annual and winter rainfall trend analyses, with Mann–Kendall and Sen's slope. The majority of results for the Mann–Kendall statistic and Sen's slope estimator are in agreement. For the period 1960–2017, no statistically significant annual rainfall trends are observed at any of the stations in this study. For the period 1987–2017, a statistically significant annual decreasing trend of 54.38 mm/decade was recorded at the SA Astronomical Observatory (Table 2; Figure 2). Notably, the Cape Point station recorded 85 mm of rainfall on 13 November 2013, which corresponds with the flash flood experienced in parts of the Western Cape on that day. This flash flood subsequently affected the annual rainfall trend at Cape Point, which shows a decline in rainfall from 2013 (741.80 mm) to 2017 (325 mm). Rainfall decreased by 279 mm at Cape Agulhas over the 5-year period (2013–2017) for the most recent 30-year period. This may be owing to coastal influences, as this station is located at a distance from the other stations (see Figure 1). The Vogel Vilij@Voeliv and Zachariashoek@Wemmershoek rainfall stations are particularly important stations because of their location within the Western Cape's Water Supply System Dams. Although these stations do not show significant monotonic trends, it is worth noting that the Zachariashoek@Wemmershoek station recorded the highest annual rainfall in 2013 (1753 mm) and lowest values during three consecutive years: 2015 (674 mm), 2016 (918 mm) and 2017 (668 mm), over the period 1960–2017. Similarly, the lowest annual rainfall at Vogel Vilij@Voeliv station was measured for the years

2015 (282 mm), 2016 (520 mm) and 2017 (294 mm) over the period 1960–2017.

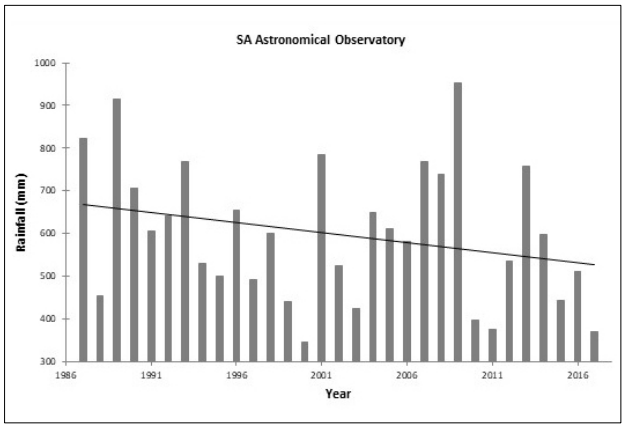


Figure 2: Statistically significant decreasing annual rainfall at the SA Astronomical Observatory station over the period 1987–2017.

During winter, only one of the seven stations (Langewens) show a significant decreasing rainfall trend for the period 1960–2017. However, for the more recent 30-year period (1987–2017), five of the seven stations show significant decreasing rainfall trends (Table 2). Notably, the majority of winter rainfall trends during the period 1960–2017 have become significant for the period 1987–2017 (Table 2). Statistically significant decreasing winter rainfall trends with the greatest magnitudes were identified at Langewens (34.80 mm/decade), SA Astronomical Observatory (-12.20 mm/decade) and Cape Columbine (-22.33 mm/decade) for the most recent 30-year period (Table 2). This finding may reflect the intensifying impact of climate change on rainfall patterns during recent decades. Winter rainfall decreased significantly at Cape Columbine, Cape Agulhas, Cape Point, SA Astronomical Observatory and Langewens, with an overall average decrease of 13.53 mm/decade for the period 1987–2017.

River flow trends

At the hydrological gauge Wits River@Drosterkloof, annual river flow shows a significant increasing trend for the period 1960–2017, which changed to a significant decreasing trend of the same magnitude (-0.24 m^3/s) for the more recent period (1987–2017) (Table 3). While the annual Bree@Ceres river flow displays a significant decreasing trend (-0.1 m^3/s) for the period 1960–2017, no trend is recorded for the more recent 30-year period. Notably the p-value is 0.07, indicating that the rate of decrease is insignificant. Similarly, decreases (-0.03 m^3/s) in annual river flow recorded at Berg@Franschoek for the period 1987–2017 are insignificant ($p=0.09$).

During winter, statistically decreasing river flow trends are observed at Bree@Ceres, while river flow at Berg@Franschoek significantly increased (0.13 m^3/s) for the period 1960–2017. The magnitude of the decrease in river flow at Bree@Ceres is greater for the more recent decades (1987–2017), compared to that for the period 1960–2017. The average winter river flow decrease of 0.52 m^3/s (Bree@Ceres and Berg@Franschoek) equates to 5.2% of the average winter river flow decrease over the most recent 30-year period. River flow during winter also changed at the Wit River@Drosterkloof station when comparing the two periods of study, showing an increase for the period 1987–2017. No significant winter or annual river flow trends are evident at the Palmiet@van Aries Kraal flow gauge (Table 3).

Table 2: Rainfall trends across the Western Cape Province over the periods 1960–2017 and 1987–2017, determined using Mann–Kendall (*Z*) and Sen's slope estimator (*Q*)

Station name	Annual rainfall trends (mm/decade)				Rainfall trends during winter (mm/decade)			
	1960–2017		1987–2017		1960–2017		1987–2017	
	Q	*Z*	*Q*	*Z*	*Q*	*Z*	*Q*	*Z*
Vogel Vilij@Voeliv	**12.26**	**2.05**	**-26.17**	-2.43	9.11	4.48	**-6.75**	-1.61
Zachariashoek@Wemmershoek	4.36	0.37	-3.39	-1.90	**14.37**	3.37	5.59	-0.27
Cape Columbine	1.11	0.21	**-31.12**	**-5.06**	1.05	1.10	**-6.80**	**-2.49**
Cape Agulhas	-5.75	-1.54	**-28.00**	-3.84	-4.36	-1.85	**22.00**	**-4.82**
Cape Point	-1.20	- 0.31	**14.86**	1.55	-0.86	-0.28	**0.13**	1.85
SA Astronomical Observatory	1.88	0.27	**54.38**	**-3.76**	**4.98**	2.14	**-12.20**	**-1.50**
Langewens	-4.44	-1.58	**-38.62**	-6.94	**-9.00**	**-4.19**	**-34.88**	**-11.13**

Note: Z is the Mann–Kendall statistic; the sign indicates an increasing (+) or decreasing (-) trend. The trend is deemed significant when Z>1.96 for the selected 5% level of significance, $\alpha=0.05$. All values are rounded off to two decimal places. Sen's slope estimator (Q) *is the magnitude of the trend. Values in bold indicate a significant trend.*

Table 3: River flow trends (m^3/s) for the Western Cape Province over the periods 1960–2017 and 1987–2017, determined using Mann–Kendall (*Z*) and Sen's slope estimator (*Q*)

River flow stations	Annual river flow trends				River flow trends during winter			
	1960–2017		1987–2017		1960–2017		1987–2017	
	Q	*Z*	*Q*	*Z*	*Q*	*Z*	*Q*	*Z*
Bree@Ceres	-0.10*	**-2.19**	-0.50	-1.22	-0.13	-1.10	**-0.92**	**-2.99**
Berg@Franschoek	0.03	0.98	-0.03	-1.66	0.07*	**3.01**	**-0.13**	**-2.44**
Little Berg River@Nieuwkloof	0.03	0.15	0.33	-1.19	-0.03	-0.04	-0.82	-0.88
Wit River@Drosterkloof	0.24*	**2.02**	-0.22	0.12	0.65*	**6.69**	0.22	0.51
Palmiet@van Aries Kraal	-0.10	-0.83	-0.28	-0.84	-0.25	-0.77	-0.10	-1.20

Note: Z is the Mann–Kendall statistic; the sign indicates an increasing (+) or decreasing (-) trend. The trend is deemed significant when Z>1.96 for the selected 5% level of significance, $\alpha=0.05$. All values are rounded off to two decimal places. Sen's slope estimator (Q) *is the magnitude of the trend. *Statistically significant with* p<0.05. *Values in bold indicate a significant trend.*

Relationship between rainfall and river flow

To illustrate the impact of annual rainfall on annual river flow for the region, correlation coefficients were determined for the periods 1960–2017, 1987–2017 and 2007–2017. Correlation coefficients were determined only for rainfall stations (Vogel Vilij@Voeliv and Zachariashoek@ Wemmershoek) located upstream of the river flow stations (Table 4). Zachariashoek@Wemmershoek rainfall data were correlated with Berg@ Franschoek and Palmiet@van Ariel Kraal river flow data, while Vogel Vallij@Voeliv rainfall data were correlated with river flow data from Bree@ Ceres, Berg@Franschoek, Palmiet@van Aries Kraal and Wit River@ Drosterkloof stations. Annual river flow is positively related to annual rainfall for the period 1960–2017 (see Figure 3 for an example).

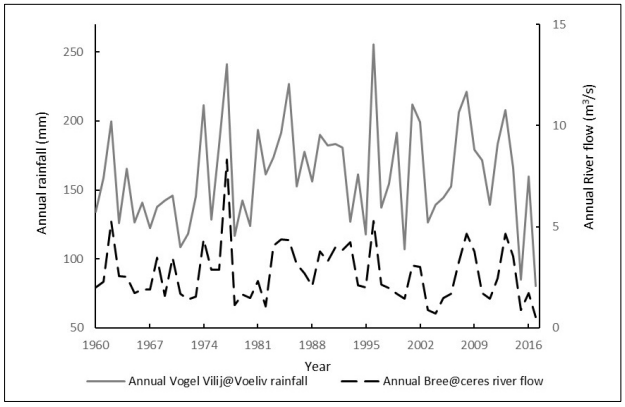


Figure 3: Annual Vogel Vilij@Voeliv rainfall and Bree@Ceres river flow over the period 1960–2017.

The highest correlation coefficient (0.95) obtained is for Vogel Vilij@ Voeliv rainfall station and Little Berg@Nieuwkloof river flow station over the period 2007–2017, likely given their close proximity to each other. The relationship between rainfall at Vilij@Voeliv station with that for several river flow stations (Berg@Franschoek, Bree@Ceres, Wit River@Drosterkloof and Little Berg@Nieuwkloof), is stronger for shorter monitoring periods (i.e. 1987–2017 and 2007–2017) than for the long-term period 1960–2017 (Table 5). The strongest correlations are for rainfall at Vogel Vilij@Voeliv station and river flow at the Little Berg@Nieuwkloof (0.95), Berg@Franschoek (0.87) and Bree@Ceres (0.85) stations for the most recent 10 years (2007–2017). Similarly, correlation coefficients for rainfall at Zachariashoek@Wemmershoek station and river flow at Wit River@ Drosterkloof and Palmiet@van Aries Kraal stations are stronger for recent, shorter periods than for the long-term period (1960–2017). This finding suggests that there have been variable extractive or absorptive pressures on river flow during different periods (i.e. 1960–2017, 1987–2017 and 2007–2017), and tentatively points toward anthropogenic influences (such as water extraction or storage) on such river flow.

Coefficient of variation results show that the largest annual river flow variability is at Bree@Ceres (53%). The second largest annual river flow variability occurred at Little Berg River@Nieuwkloof (51%), followed by Berg@Franschoek (47%), Palmiet@van Aries Kraal (41%) and Wit River@Drosterkloof (36%) (Table 5). However, annual rainfall variability was smaller in magnitude (i.e. 23% at Vogel Vallij@Voeliv and 28% at Zachariashoek@Wemmershoek) than was river flow. We used coefficient of variation to compare annual variability of rainfall with river flow: river flow varied between 36% and 53%, which is almost double that for the rainfall stations (23% to 28%). This finding suggests that a large portion of river flow variability may not be accounted for by rainfall variability alone, but rather that other factors (such as extraction rates or evaporation rates from river systems) have played an (increasingly) important role in annual river flow variability through time.

Discussion and conclusion

Spatio-temporal availability of water resources is governed by rainfall variability, which impacts hydrology.[18] The study of long-term rainfall trends and the potential impact on river flow is important for water resource management planning, which is essential for agriculture, industry and anthropogenic uses. The main aim of this study was to establish annual

Table 4: Pearson correlation coefficients (*r*) for annual rainfall and river flow over the periods 1960–2017, 1987–2017 and 2007–2017

Rainfall station	River flow gauge	Correlation coefficient for the periods:					
		1960–2017		1987–2017		2007–2017	
		r	*p*-value	*r*	*p*-value	*r*	*p*-value
Vogel Vilij@Voeliv	Berg@Franschoek	**0.66**	0.000	**0.74**	0.000	**0.87**	0.001
Vogel Vilij@Voeliv	Bree@Ceres	**0.75**	0.000	**0.76**	0.000	**0.87**	0.001
Vogel Vilij@Voeliv	Wit River @Drosterkloof	0.52	0.000	**0.68**	0.000	**0.76**	0.006
Vogel Vilij@Voeliv	Little Berg@Nieuwkloof	**0.84**	0.000	**0.84**	0.000	**0.95**	0.000
Zachariashoek@Wemmershoek	Berg@Franschoek	**0.71**	0.000	**0.65**	0.000	0.61	0.047
Zachariashoek@Wemmershoek	Wit River @Drosterkloof	**0.62**	0.000	**0.73**	0.000	**0.74**	0.009
Zachariashoek@Wemmershoek	Palmiet@van Aries Kraal	**0.74**	0.000	**0.81**	0.000	**0.82**	0.000

Zachariashoek@Wemmershoek rainfall station is located upstream of Berg@Franschoek, Wit River @Drosterkloof and Palmiet@van Aries Kraal. Vogel Vallij@Voeliv rainfall station is located upstream of Bree@Ceres, Berg@Franschoek and Wit River@Drosterkloof.
Values in bold are significant.

Table 5: Descriptive statistics of annual river flow and rainfall stations

River flow stations	Mean (m^3/s)	Standard deviation (m^3/s)	Maximum (m^3/s)	Minimum (m^3/s)	Coefficient of variation (%)
Bree@Ceres	2.69	1.43	8.32	0.53	53
Berg@Franschoek	0.70	0.33	1.90	0,17	47
Little Berg River@Nieuwkloof	2.03	1.04	5.48	0.20	51
Wit River@Drosterkloof	3.87	1.41	7.58	0.63	36
Palmiet@van Aries Kraal	2.47	1.01	5.05	0.61	41
Rainfall stations					
Vogel Vallij@Voeliv	160.58	37.42	255.54	80.23	23
Zachariashoek@Wemmershoek	78.33	21.95	146.06	42.34	28

and winter rainfall and river flow trends and to investigate the relationship between historical rainfall and river flow trends over the Western Cape region. We established these trends using Mann–Kendall and Sen's slope nonparametric statistical analysis for the periods 1987–2017 and 1960–2017.

The magnitude and significance of the annual rainfall trends was determined, at least in part, by the selected period of study (1960–2017 vs 1987–2017). While no trends are shown at four of the seven stations for the period 1960–2017, annual rainfall for the more recent 30-year period showed decreasing trends at six stations across the Western Cape Province. The largest decreasing annual rainfall trend (54.38 mm/decade), compared to trends at remaining stations in this study, occurred at the SA Astronomical Observatory station. Previous work has not adequately presented annual rainfall trends for the Western Cape region because continuous data for the period of analysis (1921–2015) were not available.[8] For the aforementioned study, only two inland stations were used in the Western Cape region, and thus the region was not well represented. It seems that the last 7 years have contributed toward the rainfall decrease observed in our study, because a previous study investigating the period 1960–2010 reported no significant or consistent annual rainfall trends for the Western Cape region.[7] However, a further study has demonstrated that stations located close to the west coast have reported increasing numbers of rain days.[11] Such differences highlight the changing nature of analysis outcomes depending on the temporal framework.

We have found statistically significant decreasing trends at four of the seven rainfall stations during winter over the period 1987–2017. Similarly, a decreasing winter rainfall trend was found across Cape Town over the period 1979–2017 using Mann–Kendall and Sen slope estimator.[18] The aforementioned study demonstrates that the 2015–2017 drought was a consequence of the poleward shift of the southern hemisphere moisture corridor and displacement of large-scale synoptic features, such as the jet-stream and South Atlantic storm tracks.[18] This poleward shift of moisture corridors, resulting in rainfall shortage, is attributed to the strong influence of Southern Annular Mode in conjunction with expansion of semi-permanent subtropical anticyclones in the South Atlantic and South Indian Oceans.[18]

Various correlation coefficients between annual rainfall and river flow for different periods (i.e. 1960–2017, 1987–2017 and 2007–2017) indicate that other factors play a role in the declining river flow trends, possibly both anthropogenic (such as extraction of water for consumption and agricultural expansion) or natural (such as an increase in evaporation, stronger and/or more frequent ENSO phases) factors. According to the 'State of Rivers' report for the Berg River system, South African rivers provide water to farmers and rural communities for crops and livestock.[19] The 'State of Rivers' report for the Breede River catchment indicates that water is used for intensive irrigation of orchard crops and vineyards for wine and table grapes along the Breede and Palmiet River catchments.[20] Water from the Breede River catchment is also transferred to Theewaterskloof Dam, the largest dam supplying water to Cape Town residents.[20] Instances in which rainfall decreased but river flow increased (e.g. Wit River@Drosterkloof) may be attributed to run-off or discharge into the river from nearby wineries that produce large quantities of wastewater.[21] The higher correlation coefficients recorded for the shorter and more recent periods (1987–2017 and 2007–2017) suggest that river flow is affected by the aforementioned anthropogenic factors at different points in time. The coefficient of variation for annual river flow being almost double that for rainfall, also suggests that river flow variability may not be entirely controlled by rainfall variability, again suggesting the influence of other (likely anthropogenic) factors. Previous studies for the Cobres River basin in southern Portugal also concluded that such results have important implications for water resource management.[22]

In conclusion, regional climate influences on river flow in the Western Cape region are highly complex because river flow is influenced by rainfall changes (affecting the frequency of flash floods and drought), variations in evaporation rates as a result of temperature variations, and increased water extraction for consumption. Population growth and agricultural and industrial development in South Africa have placed increased pressure on water resources. Future work is required to assess extraction rates and changes in evaporation rates associated with rising temperatures on river flow. Historical trend analyses provide the basis for future rainfall and river flow projections, which then have important implications for water resource planning and management.

Acknowledgements

We thank the South African Weather Service for the rainfall data and the South African Department of Water and Sanitation for hydrological data, obtained from the Internet. We appreciate the constructive inputs provided by the anonymous reviewers.

Authors' contributions

This paper is based on work conducted by R.L-G. for a Master of Science degree. R.L-G. was responsible for data collection, data analyses, methodology and write-up. S.W.G. was responsible for conceptualisation, writing revisions and student supervision.

References

1. Jiménez Cisneros BE, Oki T, Arnell NW, Benito G, Cogley JG, Döll P, et al. Freshwater resources. In: Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, et al., editors. Climate change 2014: Impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK/New York: Cambridge University Press; 2014. p. 229–270. https://doi.org/10.1017/cbo9781107415379.008
2. New M. Climate change and water resources in the southwestern Cape, South Africa. S Afr J Sci. 2002;98:369–376. https://hdl.handle.net/10520/EJC97508
3. Niang I, Ruppel OC, Abdrabo MA, Essel A, Lennard C, Padgham J, et al. Africa. In: Barros VR, Field CB, Dokken DJ, Mastrandrea, MD, Mach, KJ, Bilir TE, et al., editors. Climate change 2014: Impacts, adaptation, and vulnerability. Part B: Regional aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK/New York: Cambridge University Press; 2014. p. 1199–1265. https://doi.org/10.1017/cbo9781107415379.025
4. Brekke LD, Kiang JE, Olsen JR, Pulwarty RS, Raff DA, Turnipseed DP, et al. Climate change and water resources management – A federal perspective. USA Geological Survey Circular. 2009;1331, 65 pages. https://doi.org/10.3133/cir1331
5. Panthi J, Dhahal P, Shrestha ML, Aryal S, Krakauer NY, Pradhanang SM, et al. Spatial and temporal variability of rainfall in the Gandaki River Basin of Nepal Himalaya. Climate. 2015;3:210–226. https://doi.org/10.3390/cli3010210
6. Asadieh B, Krakauer NY. Global change in streamflow extremes under climate change over the 21st century. Hydrol Earth Syst Sci. 2015;21:5863–5874. https://doi.org/10.5194/hess-21-5863-2017
7. MacKellar N, New M, Jack C. Observed and modelled trends in rainfall and temperature for South Africa: 1960–2010. S Afr J Sci. 2014;110(7/8), Art. #2013-0353, 13 pages. https://doi.org/10.1590/sajs.2014/20130353
8. Kruger AC, Nxumalo MP. Historical rainfall trends in South Africa. Water SA. 2017;43:285–297. https://doi.org/10.4314/wsa.v43i2.12
9. Odiyo JO, Makungo R, Nkuna TR. Long-term changes and variability in rainfall and streamflow in Luvivhu river catchment, South Africa. S Afr J Sci. 2015;111, Art. #2014-0169, 9 pages. https://doi.org/10.17159/sajs.2015/20140169
10. Jury M. Climate influences on upper Limpopo river flow. Water SA. 2016;42:63–71. https://doi.org/10.4314/wsa.v42i1.08
11. Naidoo S, Rodda N, Stenstrom TA, Schmidt S, Dent M, Bux F, et al. Water security in South Africa: Perceptions and municipal obligations, governance and water reuse. Water SA. 2016;42:456–465. https://doi.org/10.4314/wsa.v42i3.11
12. South African Department of Water Affairs and Forestry (DWAF). The assessment of water availability in the Berg Catchment (WMA 19) by means of water resource related models. Report no. 5: Update of catchment hydrology. Volume 2: Upper Breede River. Pretoria: National Water Resource Planning, DWAF; 2007. Available from: http://www.dwa.gov.za/Documents/Other/WMA/19/Reports/Rep5-Vol2-Upper%20Breede%20River%20Hydrology.pdf
13. Botai CM, Botai JO, De Wit JP, Ncongwane KP, Adeola AM. Drought characteristics over the Western Cape Province, South Africa. Water. 2017;9:876–892. https://doi.org/10.3390/w9110876

14. Mann HB. Nonparametric tests against trend. Econometrica. 1945;13:245–259. https://doi.org/10.2307/1907187

15. Sen PK. Estimates of the regression coefficient based on Kendall's Tau. J Am Stat Assoc. 1968;63:1379–1389. https://doi.org/10.1080/01621459.1968.10480934

16. Novotny EV, Stefan HG. Stream flow in Minnesota: Indicator of climate change. J Hydrol. 2007;334:319–333. https://doi.org/10.1016/j.jhydrol.2006.10.011

17. Du Plessis JA, Schloms B. An investigation into the evidence of seasonal rainfall pattern shifts in the Western Cape, South Africa. J S Afr Inst Civ Eng. 2017;59:47–55. https://doi.org/10.17159/2309-8775/2017/v59n4a5

18. Sousa PM, Blamey R, Reason CJC, Ramos AM, Trigo RM. The 'Day Zero' Cape Town drought and the poleward migration of moisture corridors. Environ Res Lett. 2018;13:1-11. https://doi.org/10.1088/1748-9326/aaebc7

19. River Health Programme. State-of-Rivers report: Berg River System. Pretoria: Department of Water Affairs and Forestry; 2004.

20. River Health Programme. State-of-Rivers report: Rivers of the Breede Water Management Area. Pretoria: Department of Water Affairs and Forestry; 2011.

21. Howell CL, Myburgh PA, Lategan EL, Hoffman JE. Seasonal variation in composition of winery wastewater in the Breede River Valley with respect to classical water quality parameters. S Afr J Enol Vitic. 2015;37:31–38. https://doi.org/10.21548/37-1-756

22. Da Silva RM, Santos CAG, Moreira M, Corte-Real J, Silva VCL, Medeiros IC. Rainfall and river flow trends using Mann Kendall and Sen's slope estimator statistical tests in the Cobres River Basin. Nat Hazards. 2015;77:1205–1221. https://doi.org/10.1007/s11069-015-1644-7

AUTHORS:
Nicky Allsopp[1]
Jasper A. Slingsby[1,2]
Karen J. Esler[3]

AFFILIATIONS:
[1]South African Environmental Observation Network (SAEON) Fynbos Node, Cape Town, South Africa
[2]Centre for Statistics in Ecology, Environment and Conservation, Department of Biological Sciences, University of Cape Town, Cape Town, South Africa
[3]Centre for Invasion Biology, Department of Conservation Ecology and Entomology, Stellenbosch University, Stellenbosch, South Africa

CORRESPONDENCE TO:
Nicky Allsopp

EMAIL:
allsopp@saeon.ac.za

DATES:
Received: 20 Dec. 2018
Revised: 11 Mar. 2019
Accepted: 30 May 2019
Published: 26 Sep. 2019

HOW TO CITE:
Allsopp N, Slingsby JA, Esler KJ. Identifying research questions for the conservation of the Cape Floristic Region. S Afr J Sci. 2019;115(9/10), Art. #5889, 8 pages. https://doi.org/10.17159/sajs.2019/5889

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
John Butler-Adam

KEYWORDS:
Fynbos biome, environmental sustainability, Mediterranean-type ecosystem, IPBES Conceptual Framework

FUNDING:
None

Identifying research questions for the conservation of the Cape Floristic Region

We conducted a survey among people working in the nature conservation community in an implementation, research or policy capacity to identify research questions that they felt were important for ensuring the conservation of the Cape Floristic Region. Following an inductive process, 361 submitted questions were narrowed to 34 questions in seven themes: (1) effective conservation management; (2) detecting and understanding change: monitoring, indicators and thresholds; (3) improving governance and action for effective conservation; (4) making the case that biodiversity supports critical ecosystem services; (5) making biodiversity a shared concern; (6) securing sustainable funding for biodiversity conservation; and (7) prioritising research. The final questions were evaluated against the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services Conceptual Framework to test whether the questions addressed elements identified by this Framework as those essential to ensure that conservation contributes to a positive future for the Cape Floristic Region. We found that all elements in this Framework received attention from the collective group of questions. This finding suggests that the conservation community we approached recognises implicitly that research in multiple disciplines as well as interdisciplinary approaches are required to address societal, governance and biological issues in a changing environment in order to secure the conservation of the Cape Floristic Region. Because the majority of people responding to this survey had a background in the natural sciences, a challenge to tackling some of the questions lies in developing integrative approaches that will accommodate different disciplines and their epistemologies.

Significance:

- We present a hierarchical compendium of research questions to generate the knowledge required to conserve the Cape Floristic Region as a social-ecological system.
- The conservation community of the Cape Floristic Region collectively recognises that effective conservation management needs to be supported by knowledge of ecosystems, factors that impact them and context appropriate conservation approaches. In addition, knowledge to develop effective governance and institutions, sustainable funding and broader societal participation in conservation are also identified.
- The questions reflect the elements and linkages of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services Conceptual Framework, suggesting that the questions presented follow global prerogatives for developing a sustainable future.
- The range and complexity of knowledge gaps presented suggest the need for a broader research agenda that includes the social sciences and humanities to address conservation in the Cape Floristic Region.

Introduction

Globally, initiatives such as the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) and the Sustainable Development Goals (SDGs) emphasise the need for wise management of the natural environment, because its decline will impact human well-being and ultimately our future on this planet. Similarly, the World Economic Forum has increasingly highlighted environmental risks as threatening ecosystem services.[1] Environmental pressures with global or local impact are threatening systems such as the Cape Floristic Region (CFR), a globally unique biodiversity hotspot and conservation priority.[2] Effective conservation of natural systems and countering of anthropogenic drivers of change that threaten the environment, biodiversity and ecosystem services requires guidance from well-grounded research. Conservation research, in turn, needs to be prioritised by stakeholders more broadly than the research community alone.[3]

Several features set the CFR apart from other globally important conservation areas. The dominant vegetation is a shrubland, generally on low nutrient soils, in a winter rainfall regime, with 68% of the over 9000 plant species present endemic to this region.[4] Stochastic fire cycles drive many processes from evolution to biotic interactions.[5] Hence, many environmental mitigation schemes promulgated at global levels (e.g. the Bonn Challenge's forest-themed restoration) may be unsuitable in this unique ecosystem.

In this study, which formed part of a larger study to identify research priorities for global Mediterranean-type ecosystems[26], we adapted the approach of Sutherland et al.[3,6] by canvassing widely in relevant communities for their research priorities for conservation. Although ours is not the first attempt to collate conservation research priorities in the CFR, it differs from Steyn et al.'s[7] in that it was not developed as a funding strategy for biophysical conservation research nor as an expert review of research directions.[8]

We present a summary of the conservation research questions provided by the community of practitioners and scholars in public and private sectors in the CFR and evaluate their research questions within local and global contexts.

The questions address conservation of biodiversity, which is recognised as underpinning many of the SDGs directly or indirectly. We wanted to know whether the questions related to the CFR are reflective of the kind of knowledge required by current global initiatives, such as the SDGs, to ensure an environmentally sustainable and societally equitable future, by assessing how many elements and linkages these questions addressed in the IPBES Conceptual Framework for connecting people and nature.[9]

Methods

This project formed part of an initiative of researchers from the five Mediterranean-type ecosystems, associated with the Society for Conservation Biology Europe Section and the International Society of Mediterranean Ecologists, to identify the 100 priority questions that, if answered, would have a high probability of increasing the success of actions targeted at the conservation of biological diversity in the five Mediterranean-type ecosystems of the world.[26]

Identifying stakeholders and soliciting questions

Following ethical clearance from Stellenbosch University for working with human subjects (SU-HDS-000323), the questionnaire (Data set 1)[10], in the form of an online Google Form, was distributed by email to people associated with conservation in the CFR through implementation or research. Each recipient was asked to provide up to 10 questions which, if answered, would, in their opinion, have a high probability of increasing the success of actions targeted at the conservation of biological diversity in the CFR. We did not explicitly request, nor prohibit, the sharing of the email, so some respondents may have been additional to our distribution list (see below). Respondents submitted their questions anonymously online, and responded to additional questions aimed to solicit a profile of educational, work and sector characteristics of respondents (Data set 2[10]). Respondents received at least one reminder by email.

The broader CFR conservation community is small and well networked.[11] We selected potential respondents on the basis of key sectors in conservation and key people within these sectors or organisations (decision-makers, public and private conservation practitioners, and researchers working at government policy, conservation or research agencies, non-governmental organisations, consultancies and universities). Generally, these were people that we knew personally, had met at meetings, who held relevant positions in key organisations, or who had attended the annual Fynbos Forum (a conservation research, practice and policy conference) in the last five years. Our biases were towards people who had worked in conservation-related fields (i.e. not students). We had difficulty identifying people in the business world associated with conservation (e.g. those involved in environmental responsibility programmes) and recognise this as a gap.

Processing responses

While respondents were asked to assign their questions to predetermined categories for the global project (a deductive approach), we chose to derive the summary questions for the CFR following an inductive approach, clustering the submitted questions until generalised themes emerged.

All three authors jointly reduced the original 361 questions (362 after splitting compound questions and eliminating submissions that could not be turned into questions)(Data set 3)[10] to 34 summarised questions and clustered these into seven themes. We then revisited the original questions and extracted more specific questions (126) which added further context to the summarised questions.

We chose this approach over that taken by Sutherland et al.[3,6] because we felt that it captured the array of questions and topics posed by respondents more fully than an elimination of questions to select 100 original questions favoured by a committee.

We concede that there are opportunities for bias in whichever approach is taken, but in our approach with fewer original questions to manage, it was possible to better preserve the intentions of the original questions. Our approach also allowed us to include the essence of poorly articulated questions on an equal footing to grammatically well-constructed and scientifically nuanced questions, as we wanted to provide a platform for a broad cross-section of active participants in different spheres of conservation irrespective of their written English fluency. We were also able to explore poorly constructed questions which yielded yes/no type answers for their underlying research requirements.

We assessed the conservation scope of the questions to provide an additional verification that the clustering process correctly emphasised general themes and topics of the submitted questions in terms of what aspects of conservation they addressed. This was done by counting how many times words (or the core of words e.g. implement or implementation) or terms appeared in the submitted questions. These terms were clustered into topics (Data set 4).[10]

Finallly, we assessed each of the final 34 questions against the IPBES Framework[9] to see which elements and linkages of the IPBES Framework it addressed. For example, for the question 'How effective are restoration interventions in restoring biodiversity and ecosystem function?', restoration is seen as falling into the element 'Direct drivers' that, if successful, will influence 'Nature' which in turn affects processes that influence 'Nature's benefits to people'. We scored how many times the elements and linkages were addressed by the 34 summary questions.

Results

Respondent profile

We sent the questionnaire to 176 people (114 men, 62 women) and 53 (30%) responded (26 men, 23 women and 4 undisclosed)(Data set 2).[10] Respondents provided, on average, 6.8 questions each for a total of 361 individual questions. Of those who responded, 17 were employed in research, 16 in government conservation entities, 10 in environmental non-governmental organisations, 7 were consultants, and the balance of 3 were in other employment. From the original pool of solicited people, researchers were less likely to respond (24% responded) than people in government conservation entities, environmental non-governmental organisations or consultants (average response rate 35%). The average age of respondents was 45 years (range 29–63 years), average years of experience in broad conservation was 16.8 years (range 1–39 years) and average length of employment in their current capacity was 9.4 years (range 2 months to 35 years). In terms of qualifications, 22 held doctoral degrees, 19 master's, 5 honours, 1 bachelor's, and 2 post-school diplomas (4 were undisclosed). The majority of respondents had studied the biological (n=24) or conservation (n=13) or environmental (n=5) sciences. Among this group, eight were from other disciplines: two each from the humanities and education, while horticulture, agriculture, business management and energy studies had single representatives and three did not disclose their studies.

Conservation perspectives were evenly distributed across plants, animals and society; the scale most focused on was landscape or ecosystem (Figure 1).

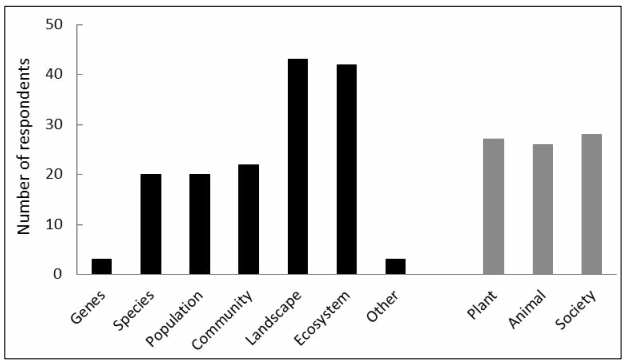


Figure 1: Respondents were asked at what ecological scale (black bars) they focused their efforts and whether this focus was predominantly on plants, animals or society (grey bars). Respondents (n=53) could choose more than one focus area in each category.

The synthesised questions

We developed a hierarchical classification of the 34 summary questions under seven themes (Table 1). Further elaboration of the themes and derived questions was provided by 126 sub-questions. This structure provides a means of directing people to the area of their interest and a more accessible way of presenting research questions, thus allowing readers to more readily identify their areas of interest and proceed from the more general to the specific depending on their objectives.

Table 1: Research questions were arrived at through an inductive analysis of 361 questions submitted by members of the broader Cape Floristic Region (CFR) conservation community (Data set 3[10]). Main questions are organised under themes and further clarification is provided in the right-hand column where relevant. The number of original questions from respondents contributing to a theme or research question is indicated in brackets.

Theme A: Effective conservation management (143 original questions)	
A1. What fundamental knowledge of CFR evolution (speciation, phylogeny, diversification, endemism) and ecology (life-history traits, biotic and abiotic interactions and processes driving community assemblages) do we still need for effective conservation planning and management? (21 original questions)	Can vegetation types serve as surrogates of other biota for conservation planning? What knowledge of breeding systems, dispersal, recruitment and other critical life-history traits, influence of soils and interactions along ecotones, do we need for effective conservation? What additional taxonomic knowledge and understanding of phylogenetics do we need for effective conservation?
A2. How effective is the current conservation estate for protecting biodiversity into the future? (7 original questions)	What are indicators of effective conservation inclusive of all landscape elements, freshwater systems and terrestrial? Are all components adequately mapped to determine their conservation status? How effective is the conservation estate at conserving all components of biodiversity?
A3. To what extent are conservation management objectives such as ecosystem service or biodiversity conservation being achieved outside formal protected areas? (6 original questions)	What levels of conservation are being achieved on formal and informal private protected areas, under stewardship agreements, and on land not under conservation management?
A4. How do we prioritise ecosystems for conservation management or restoration? (7 original questions)	
A5. What approaches to landscape planning and management can be developed that will enable us to optimise and sustain ecosystem services and biodiversity in transformed and production landscapes in a changing world? (35 original questions)	What are the impacts of various (established and emerging) land uses on biodiversity and ecosystem services? How do we avoid harmful impacts of infrastructure development? How can we optimise infrastructure developments for biodiversity e.g. power lines, rail and road reserves as biodiversity corridors? What green (engineered) infrastructure approaches can be incorporated? How do we manage agricultural systems for more ecologically sustainable outcomes? How can we integrate mixed use landscapes for better conservation? How effective are land sparing versus land sharing approaches? How can urban planning contribute to ecological sustainability?
A6. How effective are restoration interventions in restoring biodiversity and ecosystem function? (7 original questions)	What restoration methods work for farm lands? How do we restore abiotic components? How do we restore following invasive species? Are there innovative technologies for restoration? What degradation processes are irreversible? When is partial restoration acceptable?
A7. What spatial configurations are most effective for conservation? (6 original questions)	How effective are corridors (including agriculture, restored lands, riparian zones) and buffers around sensitive systems, among other spatial configurations? What spatial surrogates for evolutionary processes actually maintain and sustain these processes?
A8. What are the thresholds beyond which fragmented or otherwise impacted ecosystems become dysfunctional in maintaining biodiversity and ecosystem services? (12 original questions)	What is the carrying capacity of wildlife on land units? How effective at conservation are fragments with respect to altered fire regime (exclusion or escalation)? Are there minimum areas below which extinction debt is accrued? What is the genetic integrity on fragments? Are tops of mountains facing similar issues to fragments?
A9. How can the design and management of current protected areas be optimised to anticipate and adapt to likely range shifts in plant and animal species due to climate change? (3 original questions)	How do we mitigate or avoid future losses with current planning and management?
A10. How do we develop a nuanced and responsive approach to integrated invasive species management in a changing world? (22 original questions)	What biocontrol options exist for all invasive species, and for which different stages of the life cycle? How do we detect emerging weeds and prevent their establishment? How do we integrate all levels of policy and management in approaches? How do we ensure effective and sustainable approaches?
A11. How do we manage fire in a responsive manner in a changing world? (10 original questions)	What are different optimal fire return intervals for different vegetation units? What are the thresholds of potential concern for vegetation fires in terms of return interval and spatial cover? What role can fire play in restoration, especially integrated into alien plant clearing? How does one manage wildfire at the urban interface?
A12. Are there effective ex-situ approaches to species conservation outside their natural habitat? (7 original questions)	Is assisted migration a feasible option? What are the implications for gene pools? Is seed banking an option? How effective is search and rescue?

Table 1: Continued.

Theme B: Detecting and understanding change: Monitoring, indicators and thresholds (87 original questions)	
B1. How do global change drivers impact biodiversity, from species to ecosystem function, including across boundaries? (33 original questions)	What is the impact of a range of drivers (colonialism, loss of herbivores and predators, urbanisation, agriculture, fragmentation, mining, commercial bees, wild harvesting, tourism, recreational activities, invasive species, N_2-fixing invasive species, water abstraction from rivers and aquifers, CO_2 and climate change) on biodiversity? What emerging impacts arise from interactions between CO_2 and climate change and other drivers of change? What components of biodiversity are resilient to global change? Which ones are susceptible to climate change? Does groundwater buffer surface water temperatures? What can we learn from the impacts of global change on biodiversity across Mediterranean-type ecosystems?
B2. How is global change affecting ecosystem processes? (17 original questions)	How are ecosystem processes such as fire regime (intensity and frequency), fog precipitation, species interactions and aquatic processes being affected by global change?
B3. What impacts does global change have on ecosystem services and livelihoods dependent on biodiversity and what adaptation or mitigation will be required? (4 original questions)	How can humans adapt to reduced natural resources and ecosystem services impacted by climate change? What have we learnt from past extreme events (e.g. floods and droughts) that will aid us to understand and manage future impacts?
B4. What are the critical interactions and thresholds for (irreversible) biodiversity change? (13 original questions)	What indicators are robust for detecting thresholds of potential concern or vulnerability timeously for remedial action to be implemented? Which species show common responses to climate change? What are the climate thresholds for recruitment, growth, survival, reproduction, and pollinator and predation interactions? What elements are resilient to change?
B5. How do we best monitor for biodiversity conservation in a variable and changing world? (17 original questions)	How do we determine baselines in highly specious systems which are spatially and temporally variable and subject to stochastic disturbance events? Are current vegetation maps and vegetation units good enough for monitoring? How do we determine which elements are essential for ecological assessment (wetlands, rivers, animals etc.)? How do we deal with special communities (e.g. seeps with locally endemic species) nested in vegetation units? What interactions between units (e.g. feedbacks between marine and terrestrial systems) may be critical to these systems and need assessment?
B6. How do we determine sustainable guidelines for harvesting or consumptive use of indigenous species? (3 original questions)	What levels of harvesting of wild resources are sustainable? Are there life-history traits that can indicate suitability for harvesting? What are knock-on effects of harvesting? What alternatives to wild harvesting would work?
Theme C: Improving governance and action for effective conservation (48 original questions)	
C1. What institutional and governance structures are most effective at conserving biodiversity? (3 original questions)	What institutional arrangements promote sustainable development across all sectors?
C2. How can legislation and policy become more effective in conserving biodiversity? (12 original questions)	Is existing environmental legislation effective? How can prosecution under environmental acts become more effective? Can regulatory processes be simplified while improving effectiveness? How does environmental legislation affect behaviour? How can environmental impact assessments become more effective in guiding sustainable development and reducing biodiversity loss (beyond species of special concern)? How can environmental legislation be devolved to the most relevant and effective level?
C3. In a developing world context and in an environment of limited resources, how are ecosystem-based concerns best integrated into environmental decision-making? (8 original questions)	What strategies are effective in bringing environmental concerns into planning and development? What developments impact negatively on environmental sustainability? What compromises and trade-offs are acceptable in conservation in the face of poverty and inequity? What developments negatively affect nature-based livelihoods?
C4. What biodiversity research outputs do we need to influence conservation management and decision-making at different scales and across different sectors? (22 original questions)	How best can biodiversity concerns be mainstreamed in municipal Spatial Development Frameworks and other planning tools? What tools best support conservation objectives in planning? What mechanisms would support the effective translation of scientific knowledge for management and decision-making? How do we communicate climate change science for effective and efficient conservation management and decision-making at different scales and across different sectors? How can climate change research be made scale effective?
C5. How do we translate international conservation strategies for implementation at the local level? (2 original questions)	

Table 1: Continued.

Theme D: Making the case that biodiversity supports critical ecosystem services (19 original questions)	
D1. What is the evidence base for links between healthy biodiversity and ecosystem services? (7 original questions)	What impact do invasive species (both plants and animals) have on ecosystem services? What biodiversity-friendly practices would improve ecosystem service delivery? Do restored ecosystems deliver ecosystem services effectively? What is the value of biodiversity-based ecosystem services?
D2. How does biodiversity contribute to human well-being? (2 original questions)	What role does nature play in the mental, physical, emotional, social and spiritual well-being of citizens?
D3. What aspects of biodiversity support ecosystem function and resilience? (5 original questions)	Do CFR species offer redundancy that supports resilience? Can we increase the resilience of systems? What role do mammals play in supporting resilience?
D4. How do we communicate the evidence base for links between healthy biodiversity and ecosystem services effectively to society, managers and decision-makers? (5 original questions)	How is biodiversity information incorporated into non-biological curricula/disciplines? What are government officials' levels of understanding of ecosystem services? How can we convince communities and governments to invest in ecosystem services? Who should take up messaging around biodiversity and ecosystem services? How can scientists communicate better? How can knowledge be effectively communicated to conservation managers?
Theme E: Making biodiversity a shared concern (44 original questions)	
E1. How can we accommodate different world views to provide motivation for conservation? (14 original questions)	How do people's world view, value systems and generational differences influence their acceptance of conservation messages? Which values, perceptions or world views act as barriers to conservation behaviour? What is the potential for heritage protection (e.g. archaeological sites, cultural heritage) to complement biodiversity conservation?
E2. How do we obtain people's support and action for biodiversity conservation? (20 original questions)	How can people be supported to become more environmentally aware? How do we develop local ecological literacy at schools and in society? What has worked in breaking down barriers to biodiversity-friendly behaviour? How can conservationists change their messaging to be more effective with other groups e.g. business? What aspects or actions by the biodiversity sector alienate the general public and decision-makers? How can science best be translated to make biodiversity messages accepted? Does the use of flagship species work in promoting biodiversity behaviour? Can we take advantage of natural disasters for public engagement on environmental issues? How can individual citizens best be motivated to take responsibility for environmental concerns in their realm of interest? What knowledge tools that promote conservation can support decision-making by society around everyday aspects of their lives? What roles can ordinary citizens play to enhance biodiversity conservation? How can we expand the areas in which people can get involved in conservation action? How can citizen science be expanded?
E3. What incentives and enablers can be used to promote conservation behaviour among different sectors of society? (10 original questions)	What influences people's conservation behaviour? What regulatory regimes are effective in influencing conservation behaviour? What enablers or incentives ensure that people, including land managers, take ownership of conservation action on land? What disincentivises conservation behaviour?
Theme F: Securing sustainable funding for biodiversity conservation (21 original questions)	
F1. How do we promote investment in ecosystem services in a sustainable way? (9 original questions)	How do we build public support for funding restoration and protection of ecosystem service? How much will mismanagement of ecosystem services cost? How do we assess the contribution of biodiversity to ecosystem services? How do we make the case for future savings through current investment in ecosystem services?
F2. What effective mechanisms can be used to fund conservation organisations? (8 original questions)	How can we counter approaches that require self-funding of conservation: e.g. ecotourism as the sole funder of conservation? What are the best public–private partnership models to attract investment in conservation? What additional value-adding skills can private sector stakeholders provide? What skills and resources would improve financial management of conservation? What impact will global change have on current revenue gaining activities e.g. ecotourism?
F3. How can we improve the benefits to livelihoods of investments in protected areas and biodiversity conservation? (4 original questions)	How can using a social ecological systems thinking approach in conservation bring better benefits to people? How can conservation be achieved with other land-use options? How can the conservation and livelihood outcomes of public conservation orientated poverty-relief programmes be improved?
Theme G: Prioritising research (1 original question)	
G1. How do we effectively communicate conservation research needs to ensure they are taken up by researchers? (1 original question)	

The scope covered

An assessment of the words used by respondents in their questions shows a strong focus on the environment, its characterisation, properties and the processes that regulate them (Table 2, Data set 4).[10]

Words associated with conservation action were also frequently mentioned. Primarily, the interest was focused on vegetation, with freshwater systems also predominant. The dominance of vegetation as the focus of conservation effort by the group submitting questions is emphasised by the fact that terms associated with various types of fauna were only mentioned 38 times, while terms associated with vegetation types were recorded 210 times. However, the focus was not exclusively on biodiversity as terms associated with the human dimension and covering social, economic or governance aspects were mentioned 354 times (Table 2). Terms associated with tracking change and monitoring impact of interventions (monitoring and assessment in Table 2) were mentioned 139 times. Direct drivers of change were mentioned predominantly in the context of land use (both urban and agricultural), followed by invasive species and climate change.

Table 2: Clustering of terms used in respondents' submitted questions (see Data set 4)[10]

Topics	Number of times terms associated with topics are mentioned
Biodiversity conservation	
Ecosystem properties and processes	756
Conservation action	488
Vegetation	210
Monitoring and assessment	139
Freshwater systems	77
Animals	38
Soil systems	27
Marine	3
Human dimensions	
People, their conservation values and behaviour	142
Economic considerations	125
Governance considerations	87
Drivers of change	
Land use	135
Invasive species	48
Climate change	39

Meeting current global environmental challenges

An assessment of how reflective the 34 CFR conservation research questions (Table 1) are of the elements and pathways in the IPBES Conceptual Framework is depicted in Figure 2. There are strong emphases on how direct drivers affect nature; how nature provides benefits to people and how this affects quality of life. Many questions also emphasised the role of institutions and governance.

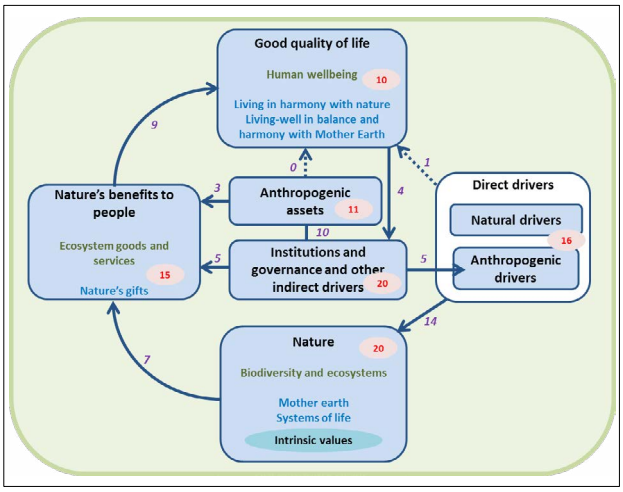


Figure 2: A depiction of how the Cape Floristic Region conservation questions address the core elements and linkages of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) Conceptual Framework. Red numbers indicate the number of questions addressing an element and the purple italic numbers indicate the number of questions addressing the linkages. One question may address more than one element or linkage. The IPBES Framework is redrawn from Diaz et al.[9] where further explanations of the elements and linkages are described.

Discussion

We are fortunate to have documentation of early efforts to prioritise conservation research in the CFR.[8,12] The first on formal record was concerned with diminishing streamflow in the 1930s.[13] In response, the Jonkershoek Small Catchment Experiments were established to understand the effect of fynbos versus plantation pines on water delivery.[13] In the current study, questions associated with the management of freshwater systems and control of invasive alien species remain high priorities.

In 1945, Wicht and other prominent biologists under the auspices of the Royal Society of South Africa investigated the causes of the 'degradation' of Cape vegetation.[12] Enduring themes from that period and that remain of concern today are fire management, catchment hydrology, invasive alien species and land-use conversion. Issues that have receded as conservation priorities are grazing management and soil erosion.[14] However, an emerging concern is the impact of stocking natural areas with wild herbivores, many extralimital. Conservation for the sake of providing recreational and educational opportunities for city dwellers was touched on in the Wicht report,[12] but the need for research on these topics and on urban ecology is more recent. Governance, institutional arrangements and interactions between wider society and conservation practitioners are more urgent in the current era, reflecting the nexus between conservation and the broader society that informs the SDGs and IPBES.

A further assessment of research needs and funding priorities for the CFR identified six themes.[7] Steyn et al.'s[7] first theme (Discovering and understanding the Cape Floristic Region's biodiversity) is similar to A1 in the current study (Table 1) but their subsequent topics (Ecosystem health and services, Fragmentation, Climate change, Alien Invasives, and Freshwater systems) speak more narrowly to specialist biodiversity researchers than to implementers of research findings. We consider land use as a driver of change more broadly, and issues around social, governance and conservation management which emerged strongly in our study as requiring research in their own right, are dealt with by Steyn et al.[7] as communicating with and influencing people under each theme but are not elaborated on as research topics.

While questions around fundamental biodiversity knowledge (A1) and impacts of global change on biodiversity (B1) are dominant, most of the other questions in our Themes A and B focus on research to inform the

design, management and monitoring of conservation measures. Over a third of all the questions (Themes C–F) deal with developing knowledge around topics which do not focus directly on biodiversity, including governance, promoting conservation through ecosystem services, communicating and eliciting conservation action from broader society and how to fund biodiversity conservation sustainably. A 2014 synthesis of CFR research suggested a focus predominantly on biodiversity components.[5] While in respect of invasive alien species, South African research is largely focused on ecological processes and impacts, with social and applied research under-represented.[15]

The use of words and terms in our respondents' questions suggests independently that the derived questions reflected the intentions of the original questions. Vocabulary focused on understanding ecosystems and transferring this knowledge into conservation action, while recognising that governance, institutions and wider society are important elements. One lone, but somewhat relevant question in Theme G, is how researchers might be encouraged to pursue conservation relevant research. This question is pertinent because the majority of questions we solicited came from people working broadly in conservation implementation and not from researchers. Thus the research agenda is driven mainly by conservation practitioners and policymakers. Moreover, many of the questions require research in the social sciences and humanities and these groups were poorly represented among our respondents. The challenge of integrating social sciences and humanities effectively in conservation research, and not merely delegating them to service provision roles, is well recognised.[16-18]

There is general congruence between Sutherland et al's.[3] global questions and our CFR questions, although 23 of the global questions were not addressed for the CFR (e.g. marine systems, polar ice or permafrost, and sea level rise). In contrast, while gaps in fundamental ecological knowledge are explicitly addressed by CFR participants, Sutherland et al. mention this only in their preamble[3] to their Ecosystem Management and Restoration section.

The shrublands of the CFR have the highest number of threatened ecosystems[19] and the highest density of taxa of conservation concern in South Africa[20]. Soils are of very low organic content and wildfire regularly destroys above-ground biomass as an intrinsic process that sustains fynbos; so this system is unlikely to support ecosystem-based carbon capture projects to mitigate climate change. This highlights the need for locally relevant approaches to manage biodiversity in shrublands even when they run counter to the global focus on forestation as a mechanism for capturing carbon. Global conservation topics that are also not germane to the CFR are nanotechnology, GMOs and climate change associated animal vectors.

In contrast to the global questions,[3] the CFR questions more frequently dealt with the nuances of governance and human values, perceptions and behaviour associated with conservation (and how these can be influenced). These are probably best addressed at the local level and reflect heightened awareness for research on these topics as a consequence of initiatives such as the Cape Action Plan for People and the Environment (C.A.P.E.).[21] The IPBES Conceptual Framework[9] also strongly emphasises the interlinkages between these components, suggesting the local questions reflect concerns that are globally recognised as important for biodiversity conservation.

Historically, concern by the public for conservation in the CFR was vested in resources such as game and wildflowers that wealthy or propertied people wanted to protect from the competing demands made by poor people for livelihood support.[22] In the modern era, understanding the social context of conservation, and of using conservation to provide pathways out of poverty is pronounced. While acknowledging the interdependence between people and the environment, as reflected in the IPBES Conceptual Framework,[9] some questions also addressed how development prerogatives for poverty alleviation could be met while ensuring a sustainable future. Similar questions have been raised around the SDGs.[23]

The capacity of the current questions to populate most of the elements and major linkages of the IPBES Conceptual Framework[9] suggests that our respondents are well aware that ensuring conservation of the CFR requires knowledge that includes people and their institutions as well as biodiversity. The alignment of the questions with the IPBES Conceptual Framework suggests that a large sector of the general conservation community in the CFR regard conservation in the context of global sustainability and connectedness to society, mirroring the role that conservation can play in attaining the SDGs. Despite this, the original questions overwhelmingly reflect a positivist approach of delivering evidence based on natural science methods of research, not surprising given the background of the majority of respondents in this study.

The CFR questions reflect that an interdisciplinary and intersectoral approach is called for that addresses dynamic social-ecological systems changing through time as a consequence of global and local pressures. The concept that biodiversity conservation is about pristine areas is also eroding with matters being raised around urban and production landscapes.

Our results are unlikely to be affected by low survey response rates or non-response bias. Sheehan[24] showed a decline in response rates to email solicited surveys from rates over 50% in the early days of email to an average response of 33% for the period 1996–2000. In this context, and given that there is a negative relationship between response rates and long surveys such as ours,[25] we consider our response rate of 30% (53 respondents) to be high. Variation in response rates by occupation was low, and researchers, who showed a slightly lower response rate, still made up the largest respondent group. We feel that our online approach made our survey accessible to a broader range of respondents from different sectors and the anonymity of respondents facilitated submissions on a broader range of topics than alternatives like workshopping.

We conclude that the questions provided by the conservation sector in the CFR show a willingness on the part of this community to work towards conservation in a societal context that will support, for example, the SDGs. However, we recognise that inclusion of multiple disciplinary and societal perspectives is missing from this analysis. Such a broadening of perspectives is likely to change the nature of the questions presented here as plurality of knowledge is brought to bear on this broadly natural science approach to conservation knowledge generation.[16-19]

Acknowledgements

This work was carried out in the scope of the 'Priority questions for Mediterranean biodiversity conservation' initiative, promoted by the Society for Conservation Biology and the International Society of Mediterranean Ecologists. We thank the leaders of this project, Francisco Moreira and Pedro Beja, and the research team for providing a stimulating environment in which to conduct this research. We are grateful to the anonymous respondents who contributed their priority research questions, reflective of careful thought and considerable time investment.

Authors' contributions

This project emerged from a project proposed by Francisco Moreira and Pedro Beja, University of Porto, to a global group of Mediterranean-type ecosystem scientists at a Society of Conservation Biology meeting attended by K.J.E. who invited J.A.S. and N.A. to make up the South African team. Henceforth, all the authors worked as equal partners in developing the project approach based on the international team's proposal, developing an online questionnaire and inviting respondents. K.J.E. obtained ethical clearance from Stellenbosch University. All authors worked together to inductively reduce the submitted questions to a shorter list of priority questions presented in this manuscript. Individually, the authors worked on co-agreed aspects of the analysis of submitted questions that overlapped with the global project and which informed our analyses but are not necessarily presented in this paper. N.A. wrote the first draft of the paper. All authors participated in further revisions. N.A., with advice from K.J.E. and J.A.S., prepared the data for archiving.

References

1. World Economic Forum (WEF). The global risks report 2018. 14th ed. Geneva: WEF; 2019. Available from: http://www3.weforum.org/docs/WEF_Global_Risks_Report_2019.pdf

2. Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GAB, Kent J. Biodiversity hotspots for conservation priorities. Nature. 2000;430:853–858. https://doi.org/10.1038/35002501

3. Sutherland WJ, Adams, WM, Aronson RB, Aveling R, Blackburn TM, Broad S, et al. One hundred questions of importance to the conservation of global biological diversity. Conserv Biol. 2009;23:557–567. http://dx.doi.org/10.1111/j.1523-173 9.2009.01212.x

4. Manning J, Goldblatt P. Plants of the Greater Cape Floristic Region 1. The Core Cape flora. Strelitzia. 2012;29:1–853.

5. Allsopp N, Colville JF, Verboom GA, editors. Fynbos: Ecology, evolution, and conservation of a megadiverse region. Oxford: Oxford University Press; 2014. https://doi.org/10.1093/acprof:oso/9780199679584.001.0001

6. Sutherland WJ, Armstrong-Brown S, Armsworth PR, Brereton T, Brickland J, Campbell CD, et al. The identification of one hundred ecological questions of high policy relevance in the UK. J Appl Ecol. 2006;43:617–627. https://doi.org/10.1111/j.1365-2664.2006.01188.x

7. Steyn L, Wilson T, Kirkwood D, Krug C. Research strategy and funding priorities for the Cape Floristic Region [document on the Internet]. In: Data conservation research priorities for the Cape Floristic Region [data set]. c2007 [cited 2019 Mar 11]. Available from: https://doi.org/10.6084/m9.figshare.7484903

8. Van Wilgen BW, Carruthers J, Cowling RM, Esler KJ, Forsyth AT, Gaertner M, et al. Ecological research and conservation management in the Cape Floristic Region between 1945 and 2015: History, current understanding and future challenges. Trans Roy Soc S Afr. 2016;71(3):207–303. http://dx.doi.org/10.1080/0035919X.2016.1225607

9. Diaz S, Sebsebe D, Carabuas J, Joly C, Lonsdale M, Ash N, et al. IPBES Conceptual Framework – connecting nature and people. Curr Opin Env Sust. 2015;14:1–16. https://doi.org/10.1016/j.cosust.2014.11.002

10. Allsopp N, Slingsby JA, Esler KJ. Data conservation research priorities for the Cape Floristic Region [data set]. In: figshare. https://doi.org/10.6084/m9.figshare.7484903

11. Gelderblom C, Wood J. The Fynbos Forum, its impacts and history [webpage on the Internet]. c2018 [cited 2019 Mar 11]. Available from: https://fynbosforum.org.za/index.php/documents-archive/home-page/200-the-fynbos-forum-its-impacts-and-history/file

12. Wicht CL. Report of the Committee on the Preservation of the Vegetation of the South Western Cape. Cape Town: Royal Society of South Africa; 1945.

13. Bennett BM, Kruger FJ. Ecology, forestry and the debate over exotic trees in South Africa. J Hist Geogr. 2013;42:100–109. https://doi.org/10.1016/j.jhg.2013.06.004

14. Raimondo D, Von Staden L, Foden W, Victor JE, Helme NA, Turner RC, et al. editors. Red list of South African plants. Strelitzia. 2009;25:1–668.

15. Abrahams B, Sitas N, Esler KJ. Exploring integrative research in the context of invasive alien plant management. S Afr J Sci. 2019;115(3/4), Art. #4263, 11 pages. https://doi.org/10.17159/sajs.2019/4263

16. Dunkley R, Baker S, Constant N, Sanderson-Bellamy A. Enabling the IPBES conceptual framework to work across knowledge boundaries. Int Environ Agreements. 2018;6:779–799. https://doi.org/10.1007/s10784-018-9415-z

17. Fish R, Church A, Winter M. Conceptualising cultural ecosystem services: A novel framework for research and critical engagement. Ecosyst Serv. 2016;21:208–217. http://dx.doi.org/10.1016/j.ecoser.2016.09.002

18. Vadrot ABM, Rankovic A, Lapeyre R, Aubert P-M, Laurans Y. Why are social sciences and humanities needed in the works of IPBES? A systematic review of the literature. Innovation. 2018;31(suppl1):S78–S100. https://doi.org/10.1080/13511610.2018.1443799

19. RSA. National list of ecosystems that are threatened and in need of protection. National Environmental Management: Biodiversity Act (Act 10 of 2004). Government Gazette. 2011;34809. Available from: http://biodiversityadvisor.sanbi.org/wp-content/uploads/2012/10/20111209-National-Gazette-No-34809-of-09-December-2011-Volume-558.pdf

20. SANBI. Statistics: Red list of South African plants [database on the Internet]. Version 2017.1 [cited 2018 Nov 30]. Available from: http://redlist.sanbi.org/stats.php

21. Ashwell A, Sandwith T, Barnett M, Parker A, Wisani F. Fynbos Fynmense: People making biodiversity work. SANBI Biodiversity Series 4. Pretoria: South African Biodiversity Institute; 2006. Available from: https://www.sanbi.org/wp-content/uploads/2018/03/biodiversityseries4.pdf

22. Van Sittert L. Making the Cape Floral Kingdom: The discovery and defence of indigenous flora at the Cape ca.1890–1939. Landsc Res. 2003;28:113–129. https://doi.org/10.1080/01426390306529

23. International Council for Science (ICSU), International Social Science Council (ISSC). Review of the Sustainable Development Goals: The science perspective. Paris: ICSU; 2015. Available from: https://council.science/publications/review-of-targets-for-the-sustainable-development-goals-the-science-perspective-2015

24. Sheehan KB. E-mail survey response rates: A review. J Comput Mediat Commun. 2001;6(2). https://doi.org/10.1111/j.1083-6101.2001.tb00117.x

25. Fan W, Yan Z. Factors affecting response rates of the web survey: A systematic review. Comput Human Behav. 2010;26:132–139. https://doi.org/10.1016/j.chb.2009.10.015

26. Moreira F, Allsopp N, Esler KJ, Wardell-Johnson G, Ancillotto L, Arianoutso M, et al. Priority questions for biodiversity conservation in the Mediterranean biome: Heterogeneous perspectives across continents and stakeholders. Conserv Sci Pract. 2019; e118, 11 pages. https://doi.org/10.1111/csp2.118

AUTHORS:
Kokeli P. Ryano[1,4]
Karen L. van Niekerk[2]
Sarah Wurz[2,3]
Christopher S. Henshilwood[2,4]

AFFILIATIONS:
[1]Department of History, School of Humanities, University of Dodoma, Dodoma, Tanzania
[2]SFF Centre for Early Sapiens Behaviour (SapienCE), Department of Archaeology, History, Cultural Studies and Religion, University of Bergen, Bergen, Norway
[3]School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand, Johannesburg, South Africa
[4]Evolutionary Studies Institute, University of the Witwatersrand, Johannesburg, South Africa

CORRESPONDENCE TO:
Kokeli Ryano

EMAIL:
kpryano@gmail.com

DATES:
Received: 11 Sep. 2018
Revised: 03 Mar. 2019
Accepted: 16 May 2019
Published: 26 Sep. 2019

HOW TO CITE:
Ryano KP, Van Niekerk KL, Wurz S, Henshilwood CS. Shellfish exploitation during the Oakhurst at Klipdrift Cave, southern Cape, South Africa. S Afr J Sci. 2019;115(9/10), Art. #5578, 9 pages. https://doi.org/10.17159/sajs.2019/5578

ARTICLE INCLUDES:
☒ Peer review
☒ Supplementary material

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

EDITOR:
Maryna Steyn

KEYWORDS:
shell-fishing, predation pressure, palaeoclimate, environment, Terminal Pleistocene, coastal subsistence

FUNDING:
Research Council of Norway, University of Bergen, University of the Witwatersrand, National Research Foundation (South Africa), Palaeontological Scientific Trust (PAST), European Research Council

Shellfish exploitation during the Oakhurst at Klipdrift Cave, southern Cape, South Africa

Klipdrift Cave in the southern Cape, South Africa, provides new insights into shellfish harvesting during the Later Stone Age (14–9 ka) period associated with the Oakhurst techno-complex. Two shellfish species dominate: *Turbo sarmaticus* and *Dinoplax gigas.* An abrupt shift in the relative frequencies of these species occurs in the middle of the sequence with *T. sarmaticus* almost completely replacing *D. gigas*. The shift in dominant species is likely due to environmental change caused by fluctuating sea levels rather than change in sea surface temperatures. The shellfish assemblage shows that local coastal habitats at Klipdrift Cave were somewhat different from those of contemporaneous sites in the southern Cape. Although the shellfish specimens are smaller at Klipdrift Cave than those from Middle Stone Age localities such as Blombos Cave, there is no robust indication that larger human populations at Klipdrift Cave during the Oakhurst period might have caused this change in size. Environmental or ecological factors could have restricted shellfish growth rates as some experimental works have suggested, but this possibility also remains to be further explored.

Significance:

- The dominance of *D. gigas* and *T. sarmaticus* at Klipdrift Cave is surprising, as it indicates a local habitat slightly different from other similar sites during the Oakhurst period.
- The shift in dominance from *D. gigas* to *T. sarmaticus* indicates changing climatic and environmental conditions during the Oakhurst period, 14–7 ka.

Introduction

Early evidence for the exploitation of shellfish for subsistence traces back to at least 164 ka during the Middle Stone Age (MSA) in South Africa[1], and by 100–60 ka shellfish were systematically and intensively exploited at a handful of sites[2-5]. Evidence for the use of shellfish for purposes other than food, such as making containers and ornaments, appears from 100 ka to 75 ka in the southern Cape.[6-8] It is, however, possible that many older sites containing shellfish remains were destroyed by the Marine Isotope Stage 5e sea level transgression.[1] Further, there is little evidence for shellfish exploitation between 50 ka and 14 ka, mainly because of the paucity of coastal sites from this time period.

Evidence for shell-fishing re-appears at around 14 ka in the southern Cape,[9] at the end of the period associated with the Robberg techno-complex. Shellfish become more abundant in sites during the subsequent period linked to the Oakhurst techno-complex, around 14–7 ka, although sites from this period with shellfish are still relatively uncommon.[10,11] The most abundant evidence for intensive shellfish exploitation comes from the 'megamidden' period, between 3 ka and 2 ka, from the West coast, which is dotted with extensive open shell middens.[12]

Here we present new data on shellfish exploitation during the Oakhurst period from a recently excavated Later Stone Age (LSA) site – Klipdrift Cave (KDC) – situated in the southern Cape. We compare our data to those from contemporary sites in the region by examining the role that shellfish played in the subsistence of these coastal dwellers. Climatic and predation pressure hypotheses are examined against these data. We investigate three main issues: the nature of marine shellfish exploitation; the intensity of shellfish collection; and the extent to which climate and environmental conditions can be deduced from shellfish remains.

The Oakhurst period in context

The term Oakhurst techno-complex is used here for sites that typically date to ~14–7 ka[10,13] and that follow the Robberg techno-complex, although regional variants occur across southern Africa[14]. Furthermore, the transition between entities such as the Robberg and Oakhurst in the Cape region was not synchronous.[11] The Oakhurst is characterised by non-microlithic and bladelet poor lithic assemblages dominated by unstandardised flakes, and frequent use of coarser grained lithic raw materials.[15-17] Formal tools are rare and consist mainly of medium to large scrapers. Non-lithic artefacts associated with the Oakhurst include worked bone, ostrich eggshell beads and ornaments as well as worked marine shell and beads.[10] The Robberg period is characterised by relatively few shellfish remains, but during the Oakhurst, shellfish, fish, marine mammals and seabirds are present in the deposits. Initially the Oakhurst people also hunted large-to-medium-sized game such as eland and warthog, but these prey were later replaced by browsers, likely reflecting a change towards a woodier habitat.[18]

Shellfish remains during the Oakhurst Complex on the southern Cape coast

Of the eight securely dated Oakhurst sites (Figure 1a) in the southern Cape[10], four – Nelson Bay Cave (NBC)[9], Matjes River Shelter (MRS)[19], Byneskranskop 1 (BNK 1)[20], and Oakhurst Shelter[21] – contain significant shellfish remains, probably because they were close to the coast. At inland sites such as Boomplaas Cave (BPA) and Kangkara Cave,[15] Wilton Large Rock Shelter[22] and Melkhoutboom Cave[23], shellfish are rare, and, when present, may have predominantly been used for non-subsistence purposes such as the manufacture of ornamental and decorative items.

The shellfish from three of the four sites have been described to varying degrees. The shellfish from Oakhurst Shelter were listed to species level but not quantified. The impression is that the shellfish samples were not retained during excavation.[24] It is, however, evident that the white sand mussel (*Donax serra*) is the most common species in the Oakhurst levels.[21] Therefore, NBC, MRS and BNK 1 are described in more detail below.

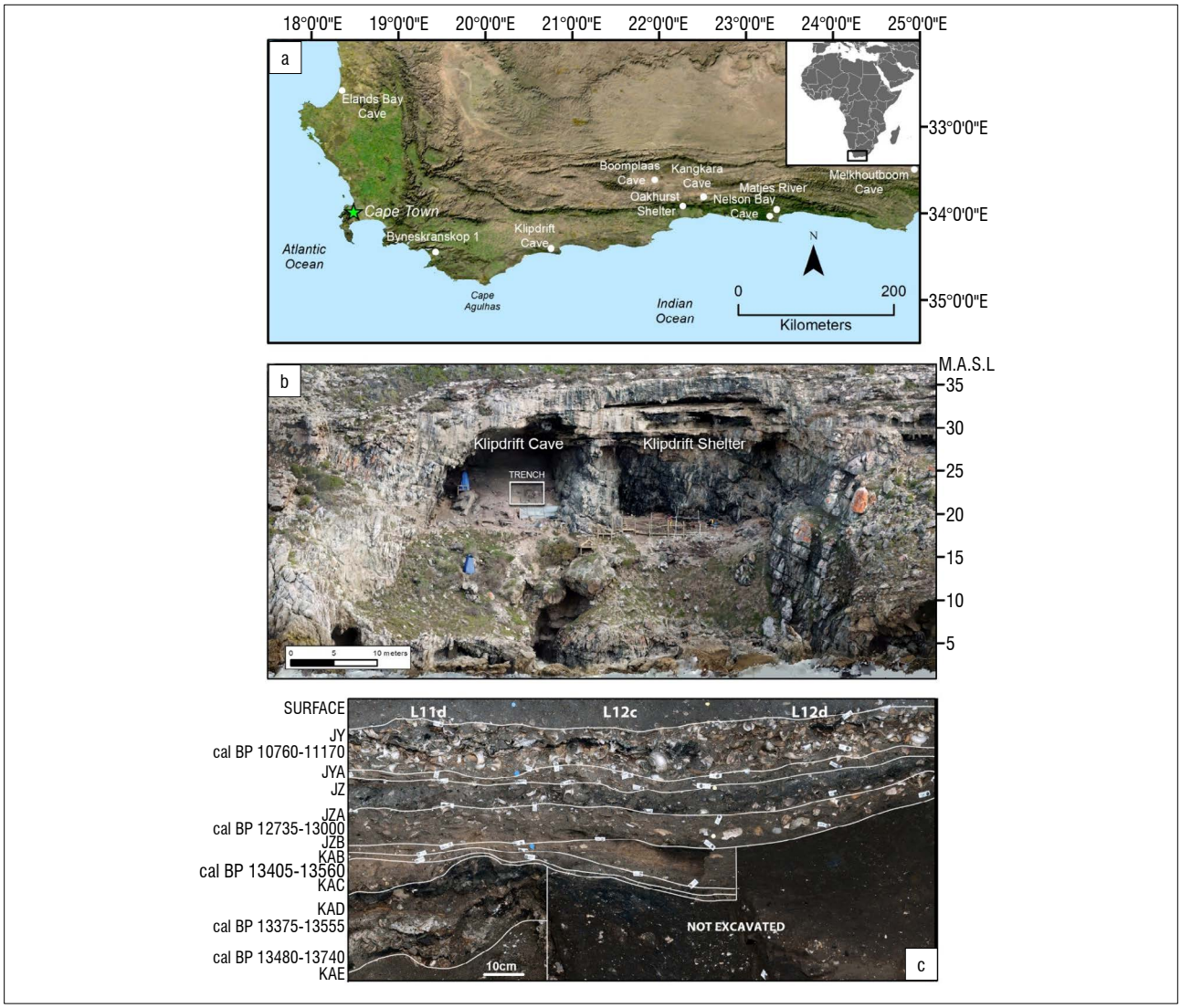


Figure 1: (a) Sites containing material from the Oakhurst Complex in the southwestern Cape, (b) Klipdrift Complex (Klipdrift Cave is the western section and Klipdrift Shelter is to the east), and (c) part of the Klipdrift Cave natural stratigraphic profile and associated accelerator mass spectrometry dates (dates provided by Beta Analytics).

NBC and MRS are situated approximately 10 km apart with similar rocky shores and sandy beaches in close vicinity. The most common species in the Oakhurst levels are the bivalves *Perna perna* (brown mussel), *D. serra* and *Choromytilus meridionalis* (black mussel). At both sites, there is an inverse correlation in the frequency of these species over time.[9,19] At NBC, *C. meridionalis* is replaced by *P. perna* around 10 ka whilst *D. serra* frequencies increase at the same time.[9,11] At MRS, *P. perna* increases sharply relative to *C. meridionalis* by 9.5 ka[13], but, unlike at NBC, *D. serra* is most common in the earlier layers, dated to 10 ka, and decreases as *P. perna* increases[19]. These changes in species representation over time have been interpreted as indicative of changing environmental conditions.[9,19]

At BNK, shellfish are present in the Oakhurst layers in relatively low frequencies; the total minimum number of individuals (MNI) is 310. The frequencies gradually increase from the oldest 13 ka to the youngest 7 ka Oakhurst layers, but only become significantly abundant in the overlying layers attributed to the Wilton period.[11,20] The low numbers probably reflect an increased distance from the shore prior to the Wilton. Numbers of *C. meridionalis* and *P. perna* are negligible, although the former outnumbers the latter in all the Oakhurst layers. *D. serra* is the most common species in all but the youngest Oakhurst layer, dated to 7 ka, where *Turbo sarmaticus* (alikreukel) become more common. Although the authors listed non-modified *D. serra* under the general shellfish table, they were not convinced that they were brought in primarily as food, but suggest that they had been used as artefacts, as some perforated fragments were found.[20] Modified *D. serra* shells have not been reported from NBC and MRS.

Shellfish as a proxy for human population sizes

It has been argued that the sizes of gastropods, and variation in shellfish species, from archaeological sites may provide an estimate of the extent and intensity of harvesting of these animals as food, which in turn can be used as a reflection of human population size.[4,25,26] The reduced size of gastropods in LSA sites relative to those from MSA localities in southern Africa has been used to imply intensified collection due to higher human population size during the LSA.[4,25,27] In addition, MSA assemblages tend to contain a smaller range of mostly larger species, and the few smaller species tend to occur in low numbers relative to LSA assemblages.[26] As such, gastropod size and species abundance have been used to argue for smaller human population sizes during the MSA.[4,25,27] Another argument is that the reduced gastropod size could be due to environmental factors affecting shell growth rates, particularly as non-food species are also reduced in size in the LSA compared to MSA.[28] It is also interesting that data on *Nerita tessellata* (tessellated nerite snail) from the Caribbean island of Nevis indicate that shellfish size increased despite intensive exploitation by human populations between 890 and 1440 AD.[29]

Site background

KDC (34°27.0963'S 20°43.4582'E) is located on the coast in the De Hoop Nature Reserve on Portion 20 of farm 516, Swellendam District in the southern Cape, South Africa (Figure 1a and 1b). It forms part of the Klipdrift Complex where Klipdrift Shelter, a Howiesons Poort locality, is also found.[5] KDC was first excavated in 2010 and subsequently in 2011. The surface of the deposit in KDC is truncated possibly by mid-Holocene higher sea levels, but excavations over an area of 2.75 m^2 revealed horizontal in-situ depositional layers with exceptional preservation of bone, shell, charcoal and ostrich eggshell. The lithics are assigned to the Oakhurst techno-complex, and the top and bottom of the excavated sequence have been dated to approximately 10 ka and 13 ka, respectively (Figure 1c).[13]

Materials and methods

All the shellfish remains retained in the 3-mm sieve from layers JY through KAE, which form 85% of the total excavated volume (0.93 m^3) at KDC, were analysed. Whole shellfish <2 cm are not considered to be food items[30] but rather animals that landed up in the site incidentally, for example through attachment to bigger shells, and are not included in this analysis.

The methods and techniques adopted for analysing marine shellfish from KDC involve species identification, determining the MNI, weighing the shells, and measurements of the maximum 'length' of *T. sarmaticus* opercula and limpet shells. Both MNI and weight are used as rare species may be underrepresented when only MNI is used[31] and further because post-depositional damage can affect the MNI counts[32].

MNI counts for *T. sarmaticus* are derived from counting apices and opercula and the highest value is considered the MNI. The weight for *T. sarmaticus* species given here includes the opercula and shell weights. The MNI values for other gastropods are calculated by counting the apices. For *Dinoplax gigas* (the giant chiton), the front, middle (the number of middle valves divided by six), and rear valves were counted separately. The greatest total for the three categories was taken as the MNI. Left and right hinges of bivalves were counted separately, and the highest value taken as the MNI.

Results

Species composition

Eleven mollusc species with a total MNI of 5330 were identified from 197.69 kg of shellfish remains (Table 1). Both *Diloma sinensis* and *D. tigirina* (periwinkles) are present, but, as the apices are usually separate from the identifiable body whorl, the shell weights and MNI have been combined for these two species and listed as *Diloma* spp. No shell fragments of *D. variegata* were found, and it is therefore assumed that only the former two species are represented by the apices of this family at KDC. All the species identified occur in the southern Cape today,[33] and no cold-water indicator species (e.g. *Cymbula granatina*, granite limpet) are present.

All species listed in Table 1 are edible, and most were presumably collected primarily as food. It is possible that the white mussels (*D. serrra*) were first eaten, and some shells subsequently used for other purposes, as 14 of the valves have ~10 mm circular perforations near the centre. The angular surf clam, *Scissodesma spengleri*, present in small numbers throughout the sequence, may also have been used for purposes other than food.[34] This species occurs subtidally in the deeper surf zone, and is therefore difficult to collect live but specimens do wash up after storms.[35] The KDC specimens do not appear waterworn, but, according to G. Branch (2013, written communication, May 15), these washed up shells are seldom damaged or waterworn. Thus, it is not clear whether these specimens were collected dead or alive. Some of the valves have what appears to be retouch on the ventral side, and might have functioned as a sort of scraper, but this possibility needs further investigation. Incidental, non-food species consist mostly of barnacle fragments and juvenile limpets.

Shellfish exploitation through time

T. sarmaticus and *D. gigas* are the most frequently occurring species throughout the assemblage (Tables 1 and 2), contributing over 93% in terms of MNI and 95% in terms of weight. All other species combined contribute <4% in terms of weight, and 7% in terms of MNI to the total assemblage (Table 2). *P. perna*, *Haliotis midae* (abalone), *Haliotis spadicea* (Venus ear, a small abalone), *Scutellastra longicosta* (long-spined limpet) and *S. spengleri* (surf clam) occur in negligible numbers. *Cymbula oculus* (goat's eye limpet), *Diloma* spp. and *Burnupena cincta* (whelk) occur in slightly higher numbers than the aforementioned, but still at very low frequencies relative to *D. gigas* and *T. sarmaticus*.

D. gigas is the most abundant species in the site, both in terms of weight and MNI (Table 1). There is an inverse relationship in frequency between *D. gigas* and *T. sarmaticus* through time, with the former being most abundant in the lower part of the sequence (layers KAE–JZB), and the latter in the upper four layers (JZA–JY). On a much smaller scale, the frequencies of *B. cincta*, *C. oculus* and *Diloma* spp. follow a similar

Table 1: Klipdrift Cave: Minimum number of individuals (MNI) and weight (g) of shellfish from the Oakhurst layers

Species	Layer																			
	JY		JYA		JZ		JZA		JZB		KAB		KAC		KAD		KAE		TOTAL	
	MNI	g	MNI	g	MNI	g	MNI	g	MNI	g	MNI	g	MNI	g	MNI	g	MNI	g	MNI	g
Turbo sarmaticus	279	16 477	167	4959	280	19 605	196	15 889	14	534	22	735	17	520	44	2 168	32	2280	1 051	63 167
Dinoplax gigas	53	906	83	1598	137	2843	85	1896	25	397	107	1999	132	3261	1771	75 142	1527	39 509	3920	127 551
Diloma spp.	3	145	10	405	63	1330	13	453	1	7	1	8	1	37	3	29	1	8	96	2422
Cymbula oculus	33	449	31	691	46	901	2	69	1	2	1	7	1	4	1	4	1	4	117	2131
Scutellastra longicosta			2	11	1	7	1	20											4	38
Perna perna	4	10	1	1	1	1			1	1	1	3	1	4					9	20
Haliotis midae	1	3	1	2			1	5			1	6							4	16
Haliotis spadicea							1	2											1	2
Burnupena cincta	1	1			1	1	1	7	1	16	6	72	14	171	24	328	17	428	65	1024
Donax serra	7	163	4	30	9	128	8	76	1	27	10	158	3	103	5	130	2	117	49	932
Scissodesma spengleri	1	5	1	1	1	5	1	6	1	4	1	31	1	55	6	208	1	73	14	388
Total	**382**	**18 159**	**300**	**7698**	**539**	**24 821**	**309**	**18 423**	**45**	**988**	**150**	**3019**	**170**	**4155**	**1854**	**78 009**	**1581**	**42 419**	**5330**	**197 691**

Table 2: Klipdrift Cave: Relative abundance (%) of each species per layer based on minimum number of individuals (MNI) and weight (g)

Species	Layer																		Total	
	JY		JYA		JZ		JZA		JZB		KAB		KAC		KAD		KAE			
	MNI	g	MNI	g	MNI	g	MNI	g	MNI	g	MNI	g	MNI	g	MNI	g	MNI	g	MNI	g
Turbo sarmaticus	73.0	90.7	55.7	64.4	51.9	79.0	63.4	86.2	31.1	54.0	14.7	24.3	10.0	12.5	2.4	2.8	2.0	5.4	19.7	32.0
Dinoplax gigas	13.9	5.0	27.7	20.8	25.4	11.5	27.5	10.3	55.6	40.2	71.3	66.2	77.6	78.5	95.5	96.3	96.6	93.1	73.5	64.5
Diloma spp.	0.8	0.8	3.3	5.3	11.7	5.4	4.2	2.5	2.2	0.7	0.7	0.3	0.6	0.9	0.2	0.0	0.1	0.0	1.8	1.2
Cymbula oculus	8.6	2.5	10.3	9.0	8.5	3.6	0.6	0.4	2.2	0.2	0.7	0.2	0.6	0.1	0.1	0.0	0.1	0.0	2.2	1.1
Scutellastra longicosta	0.0	0.0	0.7	0.1	0.2	0.0	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
Perna perna	1.0	0.1	0.3	0.0	0.2	0.0	0.0	0.0	2.2	0.1	0.7	0.1	0.6	0.1	0.0	0.0	0.0	0.0	0.2	0.0
Haliotis midae	0.3	0.0	0.3	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.7	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
Haliotis spadicea	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Burnupena cincta	0.3	0.0	0.0	0.0	0.2	0.0	0.3	0.0	2.2	1.6	4.0	2.4	8.2	4.1	1.3	0.4	1.1	1.0	1.2	0.5
Donax serra	1.8	0.9	1.3	0.4	1.7	0.5	2.6	0.4	2.2	2.7	6.7	5.2	1.8	2.5	0.3	0.2	0.1	0.3	0.9	0.5
Scissodesma spengleri	0.3	0.0	0.3	0.0	0.2	0.0	0.3	0.0	2.2	0.4	0.7	1.0	0.6	1.3	0.3	0.3	0.1	0.2	0.3	0.2
Total	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**

pattern: *B. cincta* is most common in the lower layers associated with *D. gigas* and all but disappears in the upper layers, whereas the relative frequencies of *C. oculus* and *Diloma* spp. increase in the upper layers (Figure 2). *D. serra* is present in all layers in low numbers, but its relative frequency is highest in the same layers where *B. cincta* is most common.

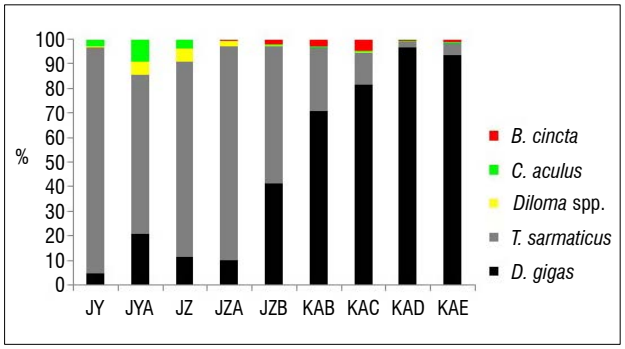


Figure 2: Relative abundance (%) per layer of the five most common shellfish species, based on weight, from Klipdrift Cave.

P. perna are absent from the lowermost two layers and layer JZA, and constitute only between 0.2% and 2.2% of MNI in the layers in which they do occur (Table 2). *H. midae* and *H. spadicea* occur in negligible quantities, but it is notable that they are only present in layers above JZB, except for a few fragments in KAB (Table 2).

KDC contains a limited number of species ($n=11$) relative to the other sites, particularly MRS ($n=20$). Some shellfish species, such as the *Aulacomya atra* (ribbed mussel), *C. compressa* (kelp limpet) and *C. granatina*, are restricted to one site (MRS). Limpets are rare at KDC and BNK 1 and more common at NBC and MRS. *S. spengleri* is present only at KDC and BNK 1 (Supplementary table 1).

Shellfish density through time

Shellfish densities at KDC are very high in the three uppermost and two lowermost layers of the sequence (Figure 3). Densities are the lowest between layers KAC and JZA. Layer KAD has the highest shell density (~374 kg/m^3) and JZB the lowest, at ~28 kg/m^3. Jerardino[36] cautions against using density measures particularly when making inter-site subsistence comparisons. However, as KDC is a 'closed' cave (as opposed to open air sites), intra-site density comparisons are less likely to be significantly problematic, although deposition rates may have differed between layers.

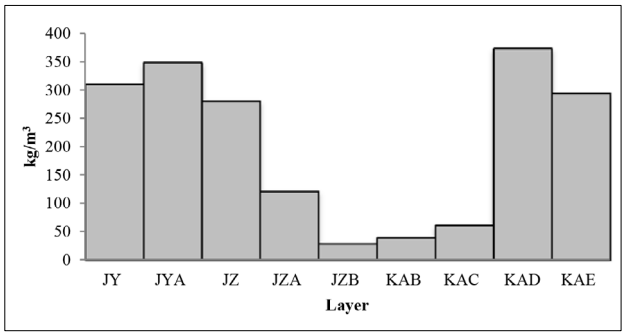


Figure 3: Density of shellfish per layer (kg/m^3) in Klipdrift Cave.

Shellfish size

Turbo sarmaticus opercula

The southern Cape species most frequently used for size measurements are various limpets and the opercula of *T. sarmaticus*. The latter are used as a proxy for shell size as shells tend to be fragmented in archaeological assemblages.[37] Descriptive statistics for *T. sarmaticus* opercula from KDC are given in Table 3 and summarised using box plots (Figure 4).

Table 3: Klipdrift Cave: *Turbo sarmaticus* opercula descriptive statistics per layer (mm)

Layer	*n*	Minimum	Maximum	Mean	Median	s.d.
JY	94	18	46	32.4	33	6.23
JYA	50	10	46	31.7	32	6.69
JZ	114	20	44	33.8	34	5.61
JZA	52	17	48	35.4	36.5	6.43
JZB	7	26	41	34.7	36	5.19
KAB	7	13	36	28.3	30	7.23
KAC	5	14	38	29.8	33	9.83
KAD	13	24	47	37.1	39	6.55
KAE	6	16	45	32.3	35	11.13

Opercula lengths range between 10 mm and 48 mm through the sequence. The median value is highest in layer KAD, at 39 mm, and lowest in KAB (30 mm) (Table 3). The data for opercula are also presented using box plots (Figure 4). This figure shows some variations within the sequence.

For example, larger individuals are apparent in layers KAD and JZA while KAB and JYA have proportionally smaller individuals.

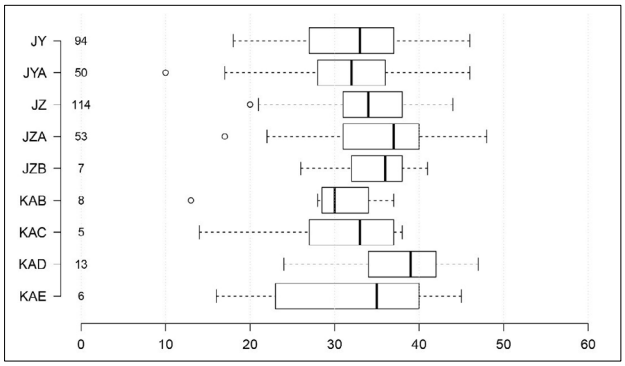


Figure 4: Box plots summarising the lengths of opercula of *Turbo sarmaticus* from Klipdrift Cave through time. Centre lines show the medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles; outliers are represented by dots. Number of sample points per layer is given on the left. (Constructed using BoxPlotR at http://shiny.chemgrid.org/boxplotr/).

KDC measurements were compared with those from other LSA localities (Blombosfontein [BBF], Blombos Cave [BBC], and NBC) and MSA sites (Klasies River [KR], Klipdrift Shelter [KDS], and BBC) from the southern Cape (Supplementary table 2; Supplementary figure 1). BBF 2 specimens are slightly larger than those from both KDC and NBC sites while there is a progressive decrease in size of *Turbo* opercula for younger LSA sites. The difference in size of specimens between Oakhurst and post-Oakhurst assemblages such as BBF 3 is significant (Supplementary figure 1). The general trend is that specimens from MSA sites are larger than those from LSA ones; for example, there is a significant difference between BBC M2 and KDC (Supplementary figure 1).

Cymbula oculus shells

The most common limpet species present at KDC is *C. oculus*. As whole (measurable) *C. oculus* shells were rare ($n=17$), the measurements were combined for all layers (Supplementary table 3). While the sample size is small, we include the data here as a contribution to the available information on shellfish size patterns during the Oakhurst period in the southern Cape. As with the *T. sarmaticus* measurements, *C. oculus* measurements from KDC were plotted against those from other MSA and LSA sites in the southern Cape (Supplementary figure 2). *C. oculus* specimens from KDC are smaller in size than those from NBC Oakhurst layers (Supplementary table 3) using the median index (66.5 mm against 62 mm). *C. oculus* specimens from MSA sites such as KR and BBC are larger than those from KDC by at least 7 mm (Supplementary figure 2).

Discussion

We address two main issues here: first, whether the change in shellfish species over time at KDC could be attributed to climate, environment and/or human choices; and second, whether human agents were responsible for the differences in the shellfish sizes observed at KDC.

Change of shellfish species composition at Klipdrift Cave

The KDC data present two clear patterns of exploitation: (1) the dominance of *D. gigas* in the lower layers, with AMS dates centring on 14 ka and 13 ka and (2) the high frequency of *T. sarmaticus* in the upper layers, from layer JZA (Figures 2 and 3). The question we explore here is whether this shift in the presence of species is related to changes in sea surface temperatures, habitat change or deliberate human choice.

Shellfish species composition has often been used as an indicator of sea surface temperatures.[2,9,19] However, only a few species are effective temperature indicators. These species include *C. granatina* and *A. atra*[35,38] that occur on the west coast and are indicative of cool temperatures. The (non-food) species such as *Cellana radiata capensis* and *Alaba pinnae* indicate warmer waters[38] but these species do not occur at KDC. The species conventionally used as temperature indicators – *C. meridionalis* for 'mostly cool' temperatures and *P. perna* and *C. oculus* for 'mostly warm' sea surface temperatures – cannot be regarded as reliable proxies for sea surface temperatures.[38] An experimental study has suggested that *C. meridionalis* do not thrive in temperatures above 18 °C[39], but *C. meridionalis* can co-exist with *P. perna* in the south coast surviving in sea surface temperatures above 20 °C[40]. This supports the Langejans et al.[38] suggestion that *C. meridionalis* and *P. perna* may not be reliable temperature indicators. Furthermore, there are minor differences in habitat preferences between *C. meridionalis* and *P. perna* that may cause them to co-exist spatially separated in the same locality. *C. meridionalis*, for example, occurs on rocks on the low shore that are associated with sand while *P. perna* occurs on the high shore on rocks which are not usually covered by sand.[40]

As most of the species that occur at KDC thrive in both warm and cool sea temperatures (Table 4), it is difficult to infer sea surface temperatures at the times of occupation. However, the absence of *C. granatina*, a more reliable cold-water indicator species[35,38] at KDC, NBC and MRS suggests that sea surface temperatures in the southern Cape coast were mildly warm during the Oakhurst period.

Table 4: Shellfish species present at Klipdrift Cave and sea surface temperature[38]

Species	Sea surface temperature
Dinoplax gigas	Warm and cool
Turbo sarmaticus	Warm and cool
Diloma sinensis	Warm and cool
Diloma tigrina	Warm and cool
Cymbula oculus	Mostly warm
Perna perna	Mostly warm
Haliotis midae	Warm and cool
Haliotis spadicea	Warm and cool
Burnupena cincta	Warm and cool
Donax serra	Warm and cool
Scissodesma spengleri	Mostly warm

Species representation can also reflect past habitats. In this regard it is surprising that *C. meridionalis* is not present, even in the lower KDC sequence, as the dominant presence of *D. gigas* suggests sand inundated rocky shores – a habitat that is attractive to *C. meridionalis*. This species is also present at the other Oakhurst sites mentioned. The low incidence of sessile mussels such as *P. perna* at KDC (Tables 1 and 2) is also unusual as they are typically common in LSA sites of the southern Cape coast such as NBC[9], MRS[19] and the BBF sites (dating from ca 6 ka to 0.5 ka)[32].

The fluctuating presence of *D. gigas* and *T. sarmaticus* at KDC may be due to changes in the habitat best suited to each species over time. The shift from the dominance of the more sand-tolerant species, *D. gigas*, in the lower layers (12 ka) to the dominance of *T. sarmaticus* in the upper layers may suggest scouring out of sand in the later period. *T. sarmaticus* would have thrived in a habitat with more exposed rocks and less sand. The near absence of sessile mussels at KDC may suggest a sheltered sandy bay in front of the cave at times – an environment not favoured by these species.[41] The slight increase in sessile mussels towards the top of the sequence could indicate a change to rockier shores and rock pools that would also have attracted *T. sarmaticus* and limpets such as *C. oculus*.[41]

The subtle changes in coastal morphology suggested by the shifting dominance of the species may have been a result of rising sea levels and is less likely due to changing sea surface temperatures. The rising sea levels signalled the transition from the Last Glaciation towards the Holocene epoch. The complete absence of *H. midae* and the low incidence of *P. perna* and *C. oculus* (species that do not tolerate overly

sandy environments) in the ca 10 ka layers[42] support a scenario of a sandy dominated marine environment around this time.

A final scenario to consider for the change of species composition at KDC is whether this change relates to human choice, acknowledging that it is complicated to discriminate between changes resulting from human choice and those from the environment.[28,43] Although the dominant species at KDC prefer slightly different habitats, it is common for them to, at times, occur in close vicinity, suggesting that both could have been available for collection during gathering events.

One possible indication that human choice was responsible for the difference in representation through time is the size of *T. sarmaticus* in the lower layers. If the coastal zones were newly colonised by this species in the lower layers, then one would expect the population to consist of smaller animals, migrating from crowded subtidal populations, not the larger ones that tend to stay in the lower subtidal areas.[33,44] While *Turbo* opercula measurements in the lower layers at KDC (Table 3) indicate a relatively small average size, there are some large individuals present, particularly in layers KAE and KAD (Figure 4). The presence of such large individuals suggests that a mature *Turbo* population was present and available and could be tentative evidence that people actively chose to collect *D. gigas* in the older levels, despite the availability of good-sized *T. sarmaticus*. However, it seems more convincing at present to suggest that the shifting dominance of *D. gigas* and *T. sarmaticus* at KDC was caused by habitat change rather than human preference. A similar scenario is suggested for the MSA site of KDS at the same locality, where *D. gigas* replaces *T. sarmaticus* and *H. midae* in the upper layers.[5] This argument may be tested by future research when refined palaeoenvironmental reconstructions of KDC become available. In the instance of *C. meridionalis* and *P. perna*, it is unlikely that people would discriminate between the former and the latter when collecting, as they are presumably the same in terms of size and taste.[9] Thus, the absence of *C. meridionalis* and the rarity of *P. perna* may be most likely explained by environmental factors rather than human decision to not collect them.

Cause(s) of size reduction

Shellfish sizes in archaeological sites have been linked to human population sizes and the intensity of harvesting. Here, the comparison of shellfish size between MSA and LSA sites in the southern Cape is discussed only for *T. sarmaticus* opercula and *C. oculus* shell measurements. Given the relative rarity of *C. oculus* at KDC, their overall small size (Supplementary table 3) is unlikely caused by human predation pressure. There is no criterion established for comparing *D. gigas* sizes although they are numerous at KDC[45], KDS and BBC MSA[38,46] and are also present at MRS, between 9.6 ka and 7 ka[19]. Comparing their size through time may be a subject for later research.

As detailed above, KDC *T. sarmaticus* opercula are smaller in size than those from the MSA of KDS, BBC and KR sites, but larger than most post-Oakhurst assemblages from BBC, KR and BBF (Supplementary figure 1). The few measurable *C. oculus* at KDC are also smaller than those from MSA sites and more like those from the Oakhurst layers at NBC.

Although most analysts of the southern Cape coast molluscs[3,4,25,27] argue that intensive harvesting of shellfish due to larger human population is a leading cause of reduction in average size of marine molluscs through time, others[28,43,47,48] question this proposition. Jerardino et al.[43] and Sealy and Galimberti[28], for example, point out that *C. meridionalis* in MSA and LSA occurrences are similar in size while limpets (e.g. *C. oculus* and *S. argenvillei*) and turban shells (e.g. *T. sarmaticus*) are smaller in the LSA. Klein and colleagues[4,25,27] believe that the lack of significant difference in *C. meridionalis* sizes between MSA and LSA sites on the west coast is due to this species' rapid colonisation relative to that of slower growing gastropods or turban shells such as *T. sarmaticus*.

It may be significant that non-food shellfish such as *Nassarius kraussianus* (tick shell) from the MSA at BBC are significantly larger than those from the LSA levels at the same site and at Die Kelders.[28] It is unlikely that the reduced sizes of *N. kraussianus* in LSA contexts can be attributed to intensive collection as they were not that intensively collected.[28] Hence, the differences in sizes of shellfish, especially limpets and the turban shells in the MSA and LSA sites, may have been caused by a combination of both natural and human factors.[28]

Non-human factors that affect the shellfish growth rates include sea surface temperature and turbidity, salinity, topography, wave action, desiccation, shellfish population densities and food supply.[28,44,47-49] Oceanic productivity or the production of organic matter by phytoplankton, generates food for marine life such as shellfish.[50] Productivity changed over time and it is known that the primary productivity of the Subantarctic Ocean changed over the last 70 ka with marked algal production at ca 58.8, 53, 46 and 38.5 ka.[51] Oceanic productivity data for the southern Cape coast are not available, but productivity may have been influenced by the Subantarctic Ocean.[28] Variations in oceanic primary productivity affect the food chain and, in turn, may affect size and distributions of shellfish species.[28] The growth of *T. sarmaticus* is affected not only by lack of food but also by its quality.[44] Changes in oceanic productivity may have resulted in changes in the availability and the quality of food on the southern Cape coast, although this supposition remains to be firmly established.

In the case of the KDC data, to test whether increased predation led to a decrease in size, we predicted size reduction in *T. sarmaticus* from the older to the younger layers, when exploitation of this species intensified. This prediction is based on the premise that a present but unexploited Turbo community will contain many large specimens.[52,53] It has also been hypothesised that humans tend to target the largest specimens first when gathering shellfish and the smaller ones may be collected later and thus the overall size distribution would become skewed.[3,4,25,27,54-56] If shellfish collectors were intentionally seeking out a species at KDC, one would expect the initial assemblage to contain the largest specimens, and a gradual reduction in size through time as increased predation leads to fewer large specimens being available.[3,4,27,56]

Opercula measurements show that sizes decrease from layer JZA upwards, and the difference in size between JZA and the uppermost two layers, JYA and JY, is statistically significant.[13] Thus, the decrease in size through time of *T. sarmaticus* opercula at KDC, especially after JZA, may support a scenario of intensive exploitation leading to reduction in size. This decrease coincides with an increase in shellfish densities, which could be because of more intensive harvesting or occupation intensity at this time. However, this does not explain why the KDC opercula are smaller than those in MSA contexts. It is possible that *T. sarmaticus* at KDC had slower growth rates than during the MSA due to not yet established environmental factors. Although *T. sarmaticus* were rare in the lower part of the sequence at KDC, when presumably little exploitation occurred, they are still smaller than MSA ones, which suggests that environmental conditions affected their growth rates.

Although there were probably larger human populations during the LSA[25], non-human factors could also have impacted shellfish size[43]. Reduced size of shellfish may also be a function of more frequent harvesting by smaller groups.[43] Until all the causal factors are carefully weaved together, larger human population as the only driver of shellfish size reduction is untimely.[43]

The environment and the Oakhurst subsistence economy

Shellfish remains are rare during the Robberg, a period that precedes the Oakhurst (e.g. at NBC[57]), as sea levels were lower during the Last Glacial Maximum. Shellfish become abundant again from the Oakhurst period and thereafter. The increase in shellfish subsistence during the Oakhurst period coincides with the rise of sea levels. The sea level transgression after 14.5 ka[58] brought the coastline very close to the present-day Oakhurst sites on the southern Cape coast. The shellfish species exploited at KDC differ somewhat from those at MRS and NBC, and changes through time are evident at all three sites. At KDC, for example (Figure 2), there is a change from the dominance of *D. gigas* to *T. sarmaticus* in the sequence after/around 12 ka (from layer JZA), the period that coincides with the driest environment in the sequence as suggested by ostrich eggshell isotopes.[45] At MRS and NBC, *P. perna* replaces *C. meridionalis* at about 10 ka. These changes are likely a result of changes in local habitat conditions through time and site-specific shores. The isotopic data from

ostrich eggshells also indicate maximum aridity in the sequence in layer JZA.[45] Furthermore, shellfish density is lowest in layer JZB (Figure 3), which also shows a decrease in lithic production. When these trends are compared to the temperature data for the terrestrial sequence of pollen at Wonderkrater[59] (Thackeray and Scott's[59] Figure 2 reproduced here as Figure 5), it is clear that layer JZB, where the lowest density of shellfish is recorded, coincides with the time when temperatures were probably the lowest during the Younger Dryas. The Younger Dryas event might indeed have had a cooling effect over environments in southern Africa, but the influence of this effect may have varied regionally as highlighted by Fitchett and colleagues[60]. On the other hand, the lowest densities of both shellfish and lithic artefacts at layer JZB may also suggest a low-occupation period at the cave, but this argument needs to be supported by other faunal data.

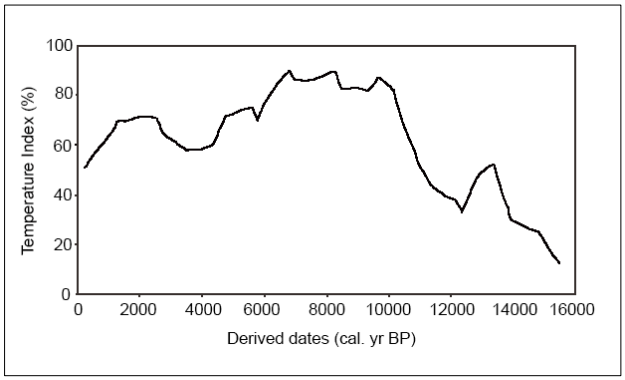


Figure 5: Temperature index for Wonderkrater within the past 16 000 years (calibrated dates) showing layers KAE, JZB and JY at KDC. Layer JZB, for which the shellfish density is the lowest, coincides with the time when temperatures might have been relatively low within the Younger Dryas. Figure modified after Thackeray and Scott's[59] Figure 2, with permission.

The change in shellfish composition probably reflects habitat change that involved removal of sand from the rocks after or around 11 ka due to increasing sea levels. The decrease of the sand mussels at MRS and NBC in the upper layers also supports the shrinking of sandy shores at this time. It is noteworthy that the shellfish subsistence practices track the change in environment much more closely than the lithic technology does at KDC.[13] The lithic technology remains stable.

Conclusions

As there are only a few sites in the southern Cape with exceptionally preserved shellfish remains, this paper broadens our understanding and provides new data on shell-fishing during the Oakhurst period. There are inter- and intra-site variations in the shellfish species exploited – perhaps because of subtle habitat changes.

Of the 11 mollusc species that occur at KDC, 2 are dominant: *D. gigas* is abundant in the lower layers while *T. sarmaticus* is more numerous upwards. The density of shellfish at KDC is lower in the middle of the sequence, which may be due to sea level regression and/or less intensive occupation of the site at that time. The shift from the dominance of *D. gigas* to *T. sarmaticus* may have been caused by rising sea levels, resulting in environmental changes from sand-covered rocky shores prior to 11 ka and more exposed rocks thereafter. The absence of *C. meridionalis* at KDC, which is present at both NBC[9] and MRS[19] during the Oakhurst, and the rarity of *P. perna*, may be due to unsuitable habitats for these species at KDC, and not related to sea surface temperatures. Effective sea surface temperature indicator shellfish species are not present at KDC but the absence of cold temperature species suggests that sea surface temperatures were relatively warm. The terrestrial climate during the Oakhurst at KDC was most likely warm and arid.

The sizes of *T. sarmaticus* opercula and *C. oculus* from the Oakhurst levels at KDC and NBC are smaller than those from MSA sites. The opercula are larger than in the post-Oakhurst LSA sites. A combination of factors may explain this scenario better than a single cause.

Acknowledgements

Partial funding for this research was provided to C.S.H., K.L.v.N. and S.W. by the Research Council of Norway through its Centres of Excellence funding scheme, Centre for Early Sapiens Behaviour (SapienCE) (project number 262618). C.S.H. was also funded through a South African National Research Foundation Research Chair (SARChI) at the University of the Witwatersrand and by the University of Bergen, Norway. S.W.'s research was further supported by the National Research Foundation of South Africa (grant no. 98826). Additional funding was provided to C.S.H. by the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 249587 and by the Palaeontological Scientific Trust (PAST), Johannesburg, South Africa. Opinions expressed and conclusions arrived at are those of the authors and not those of the funding agencies. Petro Keene and Samantha Mienies assisted during analysis while Richard Klein generously supplied data and created Supplementary figures 1 and 2 of this work. Magnus Haaland and Ole Frederik Unhammer kindly supplied some of the images in Figure 1.

Authors' contributions

K.P.R. contributed to data analysis, the methodology and writing of the initial and revised drafts. K.L.v.N. contributed to data collection and analysis, the methodology and writing of the revised draft and co-directed excavations at the Klipdrift Cave site. S.W. was involved in data collection, data analysis and methodology, student supervision and writing revisions. C.S.H. conceptualised the research project, acquired funding, and provided project leadership and management and student supervision.

References

1. Marean CW, Bar-Matthews M, Bernatchez J, Fisher E, Goldberg P, Herries AIR, et al. Early human use of marine resources and pigment in South Africa during the Middle Pleistocene. Nature. 2007;449:905–908. https://doi.org/10.1038/nature06204
2. Thackeray JF. Molluscan fauna from Klasies River, South Africa. S Afr Archaeol Bull. 1988;43(147):27–32. https://doi.org/10.2307/3887610
3. Klein RG, Avery G, Cruz-Uribe K, Halkett D, Parkington JE, Steele T, et al. The Ysterfontein 1 Middle Stone Age site, South Africa, and early human exploitation of coastal resources. Proc Natl Acad Sci USA. 2004;101(16):5708–5715. https://doi.org/10.1073/pnas.0400528101
4. Avery G, Halkett D, Orton J, Steele T, Tusenius M, Klein RG. The Ysterfontein 1 Middle Stone Age rock shelter and the evolution of coastal foraging. Goodwin Ser. 2008;10:66–89.
5. Henshilwood CS, Van Niekerk KL, Wurz S, Delagnes A, Armitage SJ, Rifkin RF, et al. Klipdrift Shelter, southern Cape, South Africa: Preliminary report on the Howiesons Poort layers. J Archaeol Sci. 2014;45:284–303. https://doi.org/10.1016/j.jas.2014.01.033
6. Henshilwood CS, d'Errico F, Vanhaeren M, Van Niekerk K, Jacobs Z. Middle Stone Age shell beads from South Africa. Science. 2004;304:404. https://doi.org/10.1126/science.1095905
7. d'Errico F, Vanhaeren M, Wadley L. Possible shell beads from the Middle Stone Age layers of Sibudu Cave, South Africa. J Archaeol Sci. 2008;35:2675–2685. https://doi.org/10.1016/j.jas.2008.04.023
8. Henshilwood CS, d'Errico F, Van Niekerk KL, Coquinot Y, Jacobs Z, Lauritzen SE, et al. A 100,000-year-old ochre-processing workshop at Blombos Cave, South Africa. Science. 2011;334:219–222. https://doi.org/10.1126/science.1211535
9. Klein RG. Preliminary report on the July through September 1970 excavations at Nelson Bay Cave, Plettenberg Bay (Cape Province, South Africa). Palaeoecol Afr. 1972;6:177–208.
10. Lombard M, Wadley L, Deacon J, Wurz S, Parsons I, Mohapi M, et al. South African and Lesotho Stone Age sequence updated (I). S Afr Archaeol Bull. 2012;67(195):120–144.
11. Loftus E, Sealy J, Lee-Thorp J. New radiocarbon dates and Bayesian models for Nelson Bay Cave and Byneskranskop 1: Implications for the South African later stone age sequence. Radiocarbon. 2016;58:365–381. https://doi.org/10.1017/RDC.2016.12

12. Jerardino A. Mega shell middens and hunter-gatherer resource intensification along the West Coast of South Africa. In: Bailey GN, Hardy K, Abdoulaye C, editors. Shell energy: Mollusc shells as coast resources. Oxford: Oxbow Books; 2013. p. 207–228.

13. Ryano KP, Wurz S, Van Niekerk KL, Henshilwood CS. The technology of the early Oakhurst lithic techno-complex from Klipdrift Cave, southern Cape, South Africa. Afr Archaeol Rev. 2017;34(1):93–119. https://doi.org/10.1007/s10437-017-9246-2

14. Mitchell P. The archaeology of southern Africa. Cambridge: Cambridge University Press; 2002.

15. Deacon J. The Later Stone Age of southernmost Africa. Cambridge Monographs in African Archaeology. Oxford: BAR Publishing; 1984.

16. Wadley L. The Early Holocene layers of Rose Cottage Cave, eastern Free State: Technology, spatial patterns and environment. S Afr Archaeol Bull. 2000;55:18–31. https://doi.org/10.2307/3888889

17. Wurz S. Human evolution, archaeology and the South African landscape during the last 100 000 years. In: Knight J, Rogerson CM, editors. The geography of South Africa – Contemporary changes and new directions. World Regional Geography Book Series. Cham: Springer; 2019. p. 125–132. https://doi.org/10.1007/978-3-319-94974-1_13

18. Klein RG. The late Quaternary mammalian fauna of Nelson Bay Cave (Cape Province, South Africa): Its implications for megafaunal extinctions and environmental and cultural change. Quat Res. 1972;2(2):135–142. https://doi.org/10.1016/0033-5894(72)90034-8

19. Döckel W. Re-investigation of the Matjes River Rock Shelter. Stellenbosch: Stellenbosch University; 1998.

20. Schweitzer FR, Wilson ML. Byneskranskop 1: A late Quaternary living site in the southern Cape Province, South Africa. Ann S Afr Mus. 1982;88:1–203.

21. Goodwin AJH. Archaeology of the Oakhurst Shelter. Part 1: Course of the excavation. Trans R Soc S Afr. 1938;25:229–247. https://doi.org/10.1080/00359193709519751

22. Deacon J. Wilton: An assessment after fifty years. S Afr Archaeol Bull. 1972;10:10–48. https://doi.org/10.2307/3888813

23. Deacon HJ. Where hunters gathered: A study of Holocene Stone Age people in the eastern Cape. Cape Town: South African Archaeological Society; 1976.

24. Schrire C. Oakhurst: Re-examination and vindication. S Afr Archaeol Bull. 1962;17:181–195. https://doi.org/10.2307/3887544

25. Klein RG, Steele TE. Archaeological shellfish size and later human evolution in Africa. Proc Natl Acad Sci USA. 2013;110:10910–10915. https://doi.org/10.1073/pnas.1304750110

26. Klein RG, Bird DW. Shellfishing and human evolution. J Anthropol Archaeol. 2016;44(Part B):198–205.

27. Steele TE, Klein RG. Mollusk and tortoise size as proxies for stone age population density in South Africa: Implications for the evolution of human cultural capacity. Munibe Antropologia-Arkeologia. 2005;57(1):221–237.

28. Sealy J, Galimberti M. Shellfishing and the interpretation of shellfish sizes in the Middle and Later Stone Ages of South Africa. In: Bicho NF, Haws JA, Davis L, editors. Trekking the shore: Changing coastlines and the antiquity of coastal settlements. New York: Springer; 2011. p. 405–419. https://doi.org/10.1007/978-1-4419-8219-3_17

29. Giovas CM, Fitzpatrick SM, Stone J. Intensifying collection and size increase tessellated nerite snail (*Nerita tessellata*) at the Coconut Walk site, Nevis, northern Lesser Antilles, AD 890–1440. J Archaeol Sci. 2013;40:4024–4038. https://doi.org/10.1016/j.jas.2013.05.008

30. Buchanan WF. Sea shells ashore: A study of the role of shellfish in prehistoric diet and lifestyle at Eland's Bay, southwestern Cape, South Africa. Cape Town: University of Cape Town; 1986.

31. Langejans GHJ, Dusseldorp GL, Van Niekerk K, Henshilwood CS. A hazy shade of winter: late Pleistocene environments and behavioural adaptations at Blombos Cave, South Africa. In: Runge J, editor. New studies on former and recent landscape changes in Africa. New York: CRC Press; 2013. p. 19–51.

32. Henshilwood CS. Holocene prehistory of the southern Cape, South Africa. Cambridge Monographs in African Archaeology. Oxford: BAR Publishing; 2008.

33. Kilburn R, Rippey E. Sea shells of southern Africa. Johannesburg: Macmillan South Africa; 1982.

34. Jerardino A, Marean CW. Shellfish gathering, marine paleoecology and modern human behavior: Perspectives from Cave PP13B, Pinnacle Point, South Africa. J Hum Evol. 2010;59:412–424. https://doi.org/10.1016/j.jhevol.2010.07.003

35. Branch G, Branch M. The living shores of southern Africa. Cape Town: Struik; 1981.

36. Jerardino A. Shell density as proxy for reconstructing prehistoric aquatic resource exploitation: Perspectives from southern Africa. J Archaeol Sci Rep. 2015;6:637–644.

37. McLachlan A, Lombard HW. Growth and production in exploited and unexploited populations of a rocky shore gastropod, *Turbo sarmaticus*. The Veliger. 1981;23:221–229.

38. Langejans GH, Van Niekerk KL, Dusseldorp GL, Thackeray JF. Middle Stone Age shellfish exploitation: Potential indications for mass collecting and resource intensification at Blombos Cave and Klasies River, South Africa. Quat Int. 2012;270:80–94. https://doi.org/10.1016/j.quaint.2011.09.003

39. Clark BC, Griffiths CL. Ecological energetic of mussels *Choromytilus meridionalis* under simulated intertidal rock pool condition. J Exp Mar Biol Ecol. 1990;137:63–77. https://doi.org/10.1016/0022-0981(90)90060-P

40. Marshall DJ, McQuaid CD. Differential physiological and behavioural responses of the intertidal mussels, *Choromytilus meridionalis* (Kr.) and *Perna perna* L., to exposure to hypoxia and air: A basis for spatial separation. J Exp Mar Biol Ecol. 1993;171:225–237. https://doi.org/10.1016/0022-0981(93)90005-9

41. Bustamante RH, Branch GM. Large scale patterns and trophic structure of southern African rocky shores: The roles of geographic variation and wave exposure. J Biogeogr. 1996;23(3):339–351. https://doi.org/10.1046/j.1365-2699.1996.00026.x

42. Zardi GI, Nicastro KR, McQuaid CD, Erlandsson J. Sand and wave induced mortality in indigenous (*Perna perna*) and invasive (*Mytilus galloprovincialis)* mussels. Mar Biol. 2008;153:853–858. https://doi.org/10.1007/s00227-007-0857-z

43. Jerardino A, Branch GM, Navarro R. Human impact on precolonial West Coast marine environments of South Africa. In: Erlandson JM, Rick TC, editors. Human impacts on marine environments. Berkeley, CA: University of California Press; 2008. p. 279–296.

44. Foster GG, Hodgson AN, Balarin M. Effect of diet on growth rate and reproductive fitness of *Turbo sarmaticus* (Mollusca: Vetigastropoda: Turbinidae). Mar Biol. 1999;134:307–315. https://doi.org/10.1007/s002270050548

45. Ryano KP. The Later Stone Age in the southern Cape, South Africa, during the Terminal Pleistocene/Early Holocene with a focus on Klipdrift Cave [PhD thesis]. Johannesburg: University of the Witwatersrand; 2014.

46. Henshilwood CS, Sealy JC, Yates R, Cruz-Uribe K, Goldberg P, Grine FE, et al. Klein Blombos Cave, southern Cape, South Africa: Preliminary report on the 1992-1999 excavations of the Middle Stone Age levels. J Archaeol Sci. 2001;28:421–448. https://doi.org/10.1006/jasc.2000.0638

47. Jerardino A. Large shell middens in Lamberts Bay, South Africa: A case of hunter-gatherer resource intensification. J Archaeol Sci. 2010;37:2291–2302. https://doi.org/10.1016/j.jas.2010.04.002

48. Jerardino A. Prehistoric exploitation of marine resources in southern Africa with particular reference to shellfish gathering: Opportunities and continuities. Pyrenae. 2010;1(41):7–52.

49. McQuaid CD, Branch GM. Influence of sea temperature, substratum and wave exposure on rock intertidal communities: An analysis of faunal and floral biomass. Mar Ecol. 1984;19:145–151. https://doi.org/10.3354/meps019145

50. Sigman DM, Hain MP. The biological productivity of the ocean. Nature Education Knowledge. 2012;3(6):4.

51. Sachs JP, Anderson RF. Increased productivity in the subantarctic ocean during Heinrich events. Nature. 2005;434:1118. https://doi.org/10.1038/nature03544

52. Foster GG, Hodgson AN. Intertidal population structure of the edible mollusc *Turbo sarmaticus* (Vetigastropoda) at an unexploited and exploited sites along the coast of the Eastern Cape Province, South Africa. Afr Zool. 2000;35(2):173–183. https://doi.org/10.1080/15627020.2000.11657088

53. Proudfoot L, Kaehler S, McGarry DK, Uppink PA, Aereboe M, Morris KM. Exploitation status of infralittoral abalone (*Haliotis midae*) and alikreukel (*Turbo sarmaticus*) in the southern section of the Eastern Cape coast, South Africa. S Afr J Sci. 2006;102:162–168.

54. Halkett D, Hart T, Yates R, Volman TP, Parkington JE, Orton J, et al. First excavation of intact Middle Stone Age layers at Ysterfontein, Western Cape Province, South Africa: Implications for Middle Stone Age ecology. J Archaeol Sci. 2003;30:955–971. https://doi.org/10.1016/S0305-4403(02)00273-X

55. Parkington J. Middens and moderns: Shellfishing and the Middle Stone Age of the Western Cape. S Afr J Sci. 2003;99:243–247.

56. Tonner TWW. Later Stone Age shellfishing behaviour at Dunefield Midden (Western Cape, South Africa). J Archaeol Sci. 2005;32:1390–1407. https://doi.org/10.1016/j.jas.2005.03.017

57. Deacon J. Changing patterns in the Late Pleistocene/Early Holocene prehistory of southern Africa as seen from the Nelson Bay Cave stone artifact sequence. Quat Res. 1978;10:84–111. https://doi.org/10.1016/0033-5894-(78)90015-7

58. Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, et al. The Last Glacial Maximum. Science. 2009;325:710–714. https://doi.org/10.1126/science.1172873

59. Thackeray JF, Scott L. The Younger Dryas in the Wonderkrater sequence, South Africa? Ann Transv Mus. 2006;43:111–112.

60. Fitchett JM, Grab SW, Bamford MK, Mackay AW. Late Quaternary research southern Africa: Progress, challenges and future trajectories. Trans R Soc S Afr. 2017;72(3):280–293. https://doi.org/10.1080/0035919X.2017.1297966

AUTHORS:
Suranie Horn[1]
Rialet Pieters[1]
Thomas Bøhn[2]

AFFILIATIONS:
[1]Unit for Environmental Sciences and Management, North-West University, Potchefstroom, South Africa
[2]Institute of Marine Research, Tromsø, Norway

CORRESPONDENCE TO:
Suranie Horn

EMAIL:
suranie.prinsloo@nwu.ac.za

DATES:
Received: 28 Jan. 2019
Revised: 10 Apr. 2019
Accepted: 14 May 2019
Published: 26 Sep. 2019

HOW TO CITE:
Horn S, Pieters R, Bøhn T. A first assessment of glyphosate, 2,4-D and Cry proteins in surface water of South Africa. S Afr J Sci. 2019;115(9/10), Art. #5988, 7 pages. https://doi.org/10.17159/sajs.2019/5988

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
Teresa Coutinho

KEYWORDS:
GMO, Roundup, 2,4-dichlorophenoxyacetic acid, Cry1Ab, ELISA, mixtures

FUNDING:
National Research Foundation (South Africa)

A first assessment of glyphosate, 2,4-D and Cry proteins in surface water of South Africa

Agriculture plays a vital role in the South African economy, as well as in the production of maize for food. Genetically modified maize is transformed to encode for crystalline (Cry) proteins found in *Bacillus thuringiensis* (Bt) and is referred to as Bt maize. Ingestion of specific Cry proteins causes the death of target insects that cause harm to maize plants. Bt crops, along with herbicides such as glyphosate and 2,4-dichlorophenoxyacetic acid (2,4-D), are widely adopted as part of the South African farming regime that aims to increase crop yield and reduce costs of production. As chemical compounds used in agriculture often end up in water sources, their presence should be monitored. There are many such monitoring programmes worldwide, but not in South Africa. We screened surface water sources in a maize-dominated agricultural area in the North West Province in South Africa for the presence of Cry1Ab, glyphosate and 2,4-D using enzyme-linked immunosorbent assays (ELISAs). Cry1Ab was not detected at any site; glyphosate was below the limit of detection at most of the sites but one sample had quantifiable traces of glyphosate; and 2,4-D was detected at all the sites. The concentrations of 2,4-D exceeded those for drinking water according to European guidelines, thus highlighting the need for regular monitoring of these compounds. Many people depend on untreated water resources, which may be contaminated by toxic agricultural chemicals. This report is the first on levels of these target compounds in South African water systems.

Significance:

- This report is the first on the presence of glyphosate, 2,4-D and Cry1Ab in the South African aquatic environment.
- Concentrations of 2,4-D in South African surface waters exceed the European guideline for drinking water, indicating a risk to people using these water sources.
- These preliminary results highlight the need to regularly monitor for the presence of glyphosate, 2,4-D and Cry1Ab in water resources in South Africa.

Introduction

In a water-scarce country such as South Africa, water contaminated with chemicals is of even greater concern for residents dependent on untreated surface and groundwater resources because less water causes these compounds to concentrate. One sector of the economy that inadvertently contributes to water pollution is agriculture. A large portion of the South African economy is driven by the agricultural sector; maize is grown on 2.8 million hectares, with the Free State, Mpumalanga and North West Provinces accounting for approximately 84% of total maize production in the country.[1] Moreover, maize serves as the staple food for the majority of South Africans. Therefore, meeting the basic needs of the population relies on successful agriculture.[2]

Globally, there have been major advances in the agricultural sector over the past 40 years which have increased crop yield and reduced pesticide use.[3] The genes that encode for crystal (Cry) proteins, which are produced by *Bacillus thuringiensis* (Bt), have been incorporated into maize, thereby creating genetically modified (GM) crops. Ingestion of these proteins can be lethal for specific insect groups; for example, ingestion of Cry1Ab toxin is lethal for lepidopterans. In South Africa, Cry1Ab maize has been used with success against the stem borer *Busseola fusca*.[4] However, resistance evolution by target pests threatens the sustainability of Bt maize in Africa[5], in part because of unique challenges, such as a lack of refugia where healthy and susceptible insects can be produced[6].

Cry proteins are considered to be environmentally benign with little or no effects on non-target organisms.[7] However, studies on Cry in aquatic ecosystems have been scarce and recent reports indicate negative effects in mussels, some insects and other invertebrates like *Daphnia magna*.[8] Cry1Ab proteins are not commonly found in water sources but the *Cry1Ab* transgene was detected in river water as far as 82 km away from an area intensively cultivated with Bt maize in Canada.[9] When Cry1Ab occurs in the aquatic system, it readily partitions to clay and organic materials.[10]

Another genetic modification of maize makes plants tolerant to the herbicide glyphosate (the active ingredient in Roundup®). These herbicide-tolerant crops are referred to as Roundup-ready maize and can be sprayed with glyphosate-based herbicides in larger quantities and during the entire period of the growing season without causing damage to the crops.[11]

Glyphosate [N-(phosphonomethyl)glycine] is the most used herbicide in the world.[12] It is a broad-spectrum, non-selective, post-emergent herbicide used for weed and vegetation control. Glyphosate is known to rapidly degrade and strongly adsorb to the soil.[13] Glyphosate's mechanism of action is to inhibit the enzyme 5-enolpyruvyl-shikimate-3-phosphate synthase of the shikimate pathway. The shikimate (shikimic acid) pathway is responsible for the biosynthesis of folates and aromatic amino acids (phenylalanine, tyrosine and tryptophan) in plants, bacteria, fungi, algae and some protozoan parasites.[14] Glyphosate is known to be non-toxic to animals and has a low ecotoxicological potential.[15] However, recent evidence of more profound toxicological effects has made the use of glyphosate (Roundup products) more controversial.[16] Moreover, glyphosate has been classified as a probable human carcinogen by the

International Agency for Research on Cancer[17], but not by the European Food Safety Authority[18].

Insufficient crop management has led to glyphosate-resistant weeds.[19] To address the tolerance of weeds towards glyphosate, farmers use herbicides with different mechanisms of action.[20] One of the herbicides used in South Africa, against which fewer weeds have developed resistance, is 2,4-dichloro-phenoxyacetic acid (2,4-D).[21,22] 2,4-D is a post-emergent auxin herbicide and has been used for selective control of broadleaf weeds.

South Africa is the biggest user of pesticides in sub-Saharan Africa and has more than 500 registered active ingredients.[23] The use of herbicides on GM maize – of which 80% is the Roundup-ready version – has increased drastically over past years, and further increases are expected to occur in the next few years.[22] Glyphosate-based herbicides are the most used herbicides in South Africa, with an estimated 23 million litres sold in 2012. The amount of herbicides used in South Africa (with a maize production of 2 million ha) is far less than that by the top producers such as the USA (40 million ha maize production), Brazil (13 million ha maize production) and China (7 million ha maize production).[24] Generally, pesticides are developed to target specific pests and to be immobile. However, run-off, leaching and spray drift occur and spread the compounds into unintended sections of the environment, and to water sources. These compounds generally occur at low concentrations and it is assumed that they would not have detrimental effects on non-target organisms. However, exposure to low levels of pesticides poses a chronic risk to human health, including endocrine disruption, immune impacts, neurotoxicity, genotoxicity, carcinogenesis and mutagenicity.[25]

This report is the first on the presence of the herbicides glyphosate and 2,4-D as well as Cry proteins in water sources in South Africa. In this study, the aforementioned herbicides were applied to GM maize expressing Cry1Ab proteins on two farms in South Africa. Because this was a screening survey, further studies are needed to determine how these contaminants reach the water; how long after application they remain in the aquatic environment; and how their concentrations change within and between seasons. These compounds are not regularly monitored in South Africa. However, South Africa has a target water quality guideline level for 2,4-D of 20 μg/L of water used for livestock.[26] The persistence of glyphosate, 2,4-D and Cry proteins in the environment and their toxicity are still under scientific discussion worldwide. To the best of our knowledge there are no data published on environmental concentrations of these compounds for South Africa.

Materials and methods

Study area

The sampling sites were located on two farms in close proximity to the Renoster and Vaal Rivers in South Africa. Farm A is in the Free State Province and Farm B is on the border between the North West and Free State Provinces (Figure 1). Fields on Farm A were planted with Bt and Roundup-ready maize and those on Farm B were planted with Roundup-ready maize only. Farm A employed rainfed farming practices whereas Farm B used an irrigation system. On both farms, the pesticide spraying regime consisted of pre-emergent Roundup® and post-emergent Roundup® as well as 2,4-D. It was assumed that the farmers applied the herbicides according to the manufacturer's guidelines. Climatic conditions, such as rainfall, are one of the mechanisms that move these compounds from the point of application to water sources. Rainfall during the month of the sampling periods was 10–25 mm for the pre-herbicide application (October 2014), 100–200 mm for the post-herbicide application (November 2014) and 50–100 mm after the harvest (March 2015).[27]

Sampling

Water was sampled at different intervals during the planting season of 2014/2015 (October–May): (1) pre- and (2) post-herbicide application, as well as (3) after the harvest (Table 1). Water was sampled on Farm A from the Renoster River (A1) and from a dam on the farm (A2) and on Farm B from the Vaal River (B1), from an inflow dam on the farm where water is recycled from run-off after rainfall and irrigation (B2) and used again for irrigation, and from a dam on the farm used for recreational activities (B3). Surface water at a 30-cm depth was sampled in 250-mL high-density polyethylene bottles (Nalgene™, Rochester, NY, USA), protected from UV radiation and kept at 4 °C during transportation.

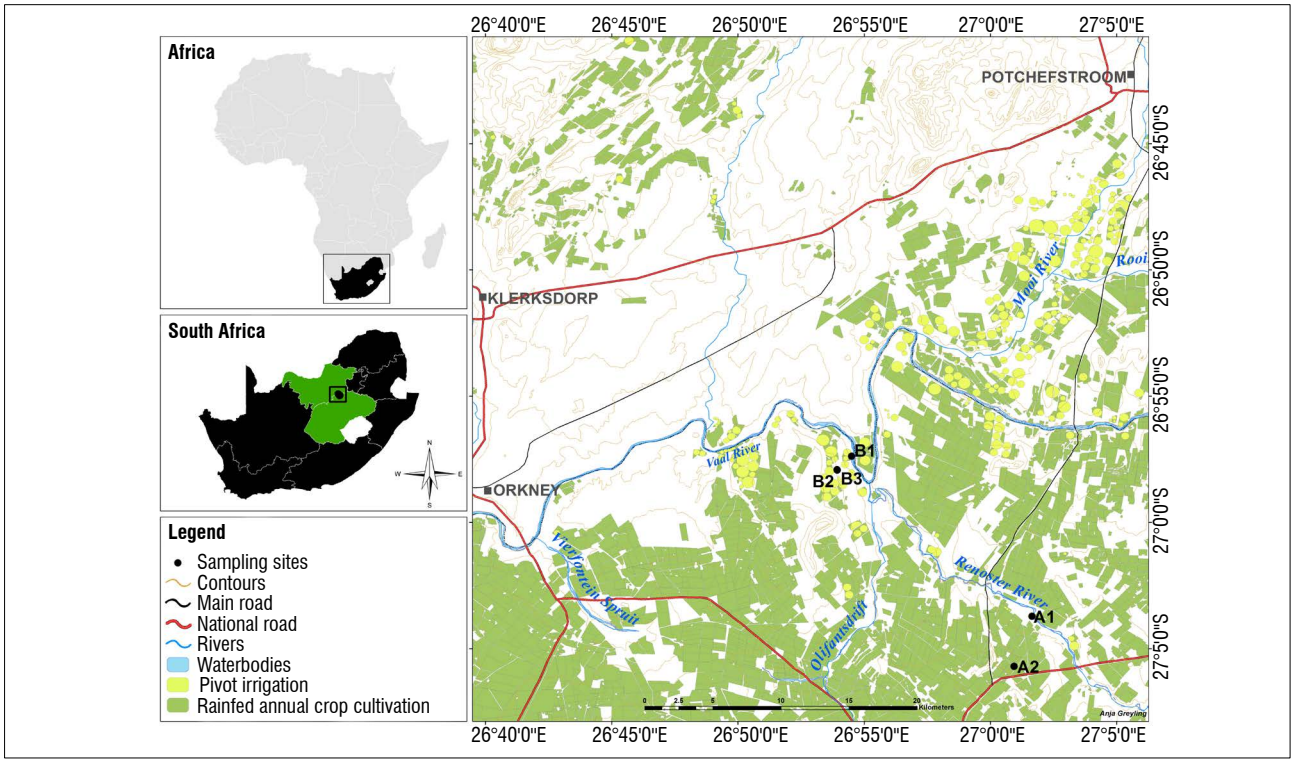


Figure 1: Map of the sampling sites situated on Farms A and B. A1: Renoster River; A2: water from a farm dam; B1: Vaal River; B2: inflow dam; B3: water from a farm dam.

Table 1: Concentrations of the target compounds from various water sources after three different sampling events

Site	Sampling event	Sampling date	Cry1Ab (μg/L)	Glyphosate (μg/L)	2,4-D (μg/L)
Farm A					
River (A1)	Before planting	6 October 2014	<LOD	<LOD	<LOQ
	After spraying	26 November 2014	<LOD	<LOD	0.93±0.08
	End of season	9 March 2015	<LOD	<LOD	<LOQ
Dam (A2)	Before planting	6 October 2014	<LOD	<LOD	<LOQ
	After spraying	26 November 2014	<LOD	<LOD	0.72±0.02
	End of season	9 March 2015	<LOD	<LOD	0.72±0.07
Farm B					
River (B1)	Before planting	6 October 2014	<LOD	<LOD	<LOQ
	After spraying	26 November 2014	<LOD	<LOD	1.02±0.03
	End of season	9 March 2015	<LOD	<LOD	0.96±0.16
Inflow (B2)	Before planting	6 October 2014	<LOD	<LOD	0.83±0.10
	After spraying	26 November 2014	<LOD	0.42±0.04	1.08±0.04
	End of season	9 March 2015	<LOD	<LOD	0.99±0.03
Dam (B3)	Before planting	6 October 2014	<LOD	<LOD	0.74±0.02
	After spraying	26 November 2014	<LOD	<LOQ	0.90±0.08
	End of season	9 March 2015	<LOD	<LOD	0.92±0.08

LOD, limit of detection; LOQ, limit of quantification

Concentrating Cry1Ab proteins from water samples

Each water sample was concentrated using an Amicon® ultracentrifugation tube (Millipore, Billerica, MA, USA) with a 30 000 molecular mass cut-off membrane. In short, a 15 mL aliquot of the sample was centrifuged at 870 *g* for 30 min. The eluent was discarded and another 15 mL was added and again centrifuged at 870 *g* for 30 min. The Amicon® tubes were subjected to a third centrifugation cycle whereafter the Cry proteins were rinsed off the membrane with 1 mL phosphate-buffered saline and Tween assay buffer. This concentrate of the samples was stored at 4 °C and quantified within 24 h.

Enzyme-linked immunosorbent assays

Over the past few years, enzyme-linked immunosorbent assays (ELISAs) have demonstrated results comparable with those of instrumental analytical methods for the quantification of contaminants in water sources. ELISA assays are therefore reliable and good substitutes for screening and monitoring such systems.[28]

Cry1Ab

The commercially available ELISA kit used for quantification of Cry1Ab in the water samples was obtained from Envirologix (Portland, ME, USA) (QualiPlate Kit for Cry1Ab/Cry1Ac Cat # AP003CRBS). The kit does not include a reference standard with a known concentration; the package insert advises that, if the kit is to be used for quantification purposes, a reference standard should be obtained from elsewhere. Lyophilised activated Cry1Ab toxin prepared from Cry1Ab protoxin was acquired from Marianne Pusztai-Carey at the Department of Biochemistry, Case Western University (Cleveland, OH, USA).[29] The lyophilised protein was re-suspended in 10 mM CAPS buffer at pH 10.5 at a concentration of 100 μg/mL and frozen at -80 °C until use.[30] The quantification of the Cry1Ab protein was determined by including two independent 12-point standard curves ranging from 0 to 3.5 μg/L. The samples, blanks and calibrators (Cry1Ab) were loaded in triplicate on the 96-well-microtitre plate pre-coated with antibodies specific for Cry1Ab/Ac and containing Cry1Ab/Ac enzyme conjugate. The plates were left to incubate for 2 h and washed four times with 300 μL wash buffer. A substrate was then added, resulting in a blue colour produced by the hydrolysis of hydrogen peroxide by peroxidase. After 20 min, the stop solution containing 1 N HCl was added and the optical density was measured at 450 nm and 650 nm (reference) using a multimode microplate reader (TriStar LB 941, Berthold, Bad Wildbad, Germany).[31]

Glyphosate

Glyphosate was quantified through the use of the Abraxis ELISA kit (PN 500086; Warminster, PA, USA). The method was performed according to the manufacturer's instructions. A six-point calibration curve that ranged from 0 to 4 μg/L was used to quantify the levels of glyphosate in the sample. In short, the samples, blanks and standards were derivatised and loaded into a 96-well plate coated with antibodies. A glyphosate antibody solution was added and the plates were incubated for 30 min. After incubation, the enzyme conjugate solution was added and the second incubation time was 60 min. Thereafter, the plate was washed three times with 250 μL wash buffer. A colour solution was added and after 30 min incubation, the stop solution was added. Absorbance was measured at 450 nm.[28,32]

2,4-D

To determine the levels of 2,4-D in the surface water, an ELISA specifically for 2,4-D (PN 54003A, Abraxis, Warminster, PA, USA) was employed. The 7-point calibration curve ranged from 0 to 80 μg/L. The water samples, standards and blanks were added to the wells on the test plate. The enzyme conjugate and antibody solution followed shortly after and the plate was incubated for 60 min. After the incubation period, the plates were washed three times using 250 μL wash buffer. After the washing step, a colour substrate was added and incubated for 30 min, after which a stop solution was added and absorbance was read at 450 nm.

Quality control

All samples were quantified in triplicate using ELISAs specific for each target compound. The mean absorbance values were calculated and the coefficient of variation was determined for each sample, requiring a

coefficient of variation of <20%. The limit of detection (LOD) and limit of quantification (LOQ) were determined using a regression analysis of the calibration curves where LOD=3S_b/b and LOQ=10S_b/b with S_b=slope uncertainty and b=slope (Table 2).[33] The concentrations of glyphosate, 2,4-D and Cry1Ab were determined against the linear regression line of the calibration curve, with a correlation coefficient (R^2) as close as possible to 1.

Table 2: Limit of detection (LOD) and limit of quantification (LOQ) values for each of the target compounds

		2,4-D	Glyphosate	Cry1Ab
LOD	**(μg/L)**	0.2	0.2	0.1
LOQ	**(μg/L)**	0.7	0.4	0.5

Results and discussion

Concentration of the compounds in water sources

Cry1Ab

Although the water samples were concentrated 30 times, there were no detectable levels of Cry1Ab proteins in any of the water samples. It is well known that Cry1Ab proteins degrade quickly in water sources, and this was corroborated by the results of the current study (Table 1). Cry1Ab proteins break down when exposed to high temperatures (24–33 °C), thus resulting in microbial degradation. Soil type influences adsorption, making these proteins more persistent, but also decreasing their extractability. Cry1Ab has high conformational stability and retains its activity when absorbed to polar, charged surfaces in soils, which is important when assessing its potential adverse effects in agricultural systems.[34] There is a lack of evidence on the bioactivity and potential health risks of Cry1Ab fragments that may be present in the environment.

In contrast to our results, Tank et al.[35] detected Cry1Ab proteins in 23% of 215 water samples taken from streams near agricultural fields 6 months after harvest. They reported a mean concentration of 14 ng/L and a maximum of 32 ng/L. Whiting et al.[36] detected no Cry1Ab in groundwater samples, but found concentrations of 129 ng/L in run-off water between maize fields. The same research group also analysed soil and run-off sediment, but in contrast to the high levels in water, a maximum mean concentration of only 9 ng/g was detected in soil during the pollination stage of the maize plants. Cry1Ab was detected in run-off water from a non-Bt maize field with levels from below LOD to 42 ng/L, whilst higher levels (maximum concentration of 130 ng/L) were detected from a Bt maize field.[37] It should be noted that the concentrations of Cry1Ab detected in other studies were below the LOD of the current study. The ELISA method used could therefore have missed the presence of Cry1Ab at lower levels. The presence of Cry1Ab proteins in water, although at low levels, highlights the importance of investigating the potential long-term effects of these proteins on non-target organisms.

Glyphosate

The levels of glyphosate were below the LOD at most of the sites (Table 1). The water sampled from the dam (B3) on Farm B had traces of glyphosate with levels between LOD and LOQ after the spraying event. Glyphosate levels of 0.42 μg/L were detected at the in-flow dam on Farm B (B2) after the spraying event. These levels decreased to <LOD at the end of the season (Table 1). Glyphosate is very water soluble and has been found in various water sources around the world, but it also degrades quickly, which can be the reason for low detection. Some studies ascribe the lower than detection limit levels of glyphosate and its quick metabolising capability to its main metabolite aminomethylphosphonic acid (AMPA).[38,39] AMPA was, however, not quantified within the scope of this study. Glyphosate concentrations are also highly influenced by precipitation and can change from year to year.[40]

In contrast to the current study, in other studies from all over the world, glyphosate has been detected in water sources. Sanchís et al.[41] analysed 140 groundwater samples from Spain and found quantifiable levels for 41% of the samples. The mean concentration of glyphosate in Sanchís et al.'s study was 200 ng/L and the maximum concentration was 2.5 μg/L. Glyphosate concentrations of 663 ng/L were found in the Nottawasaga River watershed in Canada.[42] According to Smith et al.[43], 45 μg/L of glyphosate was detected in well water at the Massey Drive substation in the USA 7 weeks after spraying. This station is built on a limestone bed that has high permeability, thus emphasising that glyphosate is very mobile in water sources. In the USA, glyphosate was detected in a stream and wastewater treatment plant effluent samples in a study by Kolpin et al.[44] The maximum concentration they reported was 2.2 μg/L. Also in the USA, an extensive study by Battaglin et al.[39] reported glyphosate levels for different environmental matrices: 73 μg/L in streams; 2.03 μg/L in groundwater; 427 μg/L in ditches and drains; 3.08 μg/L in large rivers; 1 μg/L in soil water; 301 μg/L in wetlands, lakes, and ponds; 2.5 μg/L in precipitation; 476 μg/L in soil and sediment; and 0.3 μg/L in wastewater treatment outfall. It is evident that glyphosate ends up in water sources.

2,4-D

Most of the samples in the current study contained quantifiable levels of 2,4-D with a minimum of 0.72 μg/L and a maximum of 1.08 μg/L. Before planting, the concentrations of 2,4-D were below the LOD in both river samples and the dam on Farm A. It was also detected at low quantifiable levels before planting in both dams on Farm B. The highest concentration was detected after the spraying event and decreased towards the end of the season (Table 1).

According to Wilson et al.[45], 2,4-D amine salts and 2,4-D esters are very mobile but they are not persistent under most environmental conditions. 2,4-D does not adsorb to the soil but readily moves into water resources – a finding confirmed by Mountassif et al.[46] who reported that 91.7% of the applied 2,4-D eventually ends up in water, thus explaining the high levels detected in various countries.

Hernandez et al.[47] detected 0.05 μg/L 2,4-D in Lake Chapala, Mexico, which is an order of magnitude lower than the levels found in the current study. The concentrations of 2,4-D found in our study are in the same range as those in two European studies: Rodil et al.[48] detected levels of 0.062–0.2 μg/L 2,4-D in drinking and surface water in Spain and Tsaboula et al.[49] reported 1.16 μg/L 2,4-D in the Pinios River Basin, Greece. A few US studies by Serrano and DeLorenzo[50], Ensminger et al.[51] and Wijnja et al.[52], reported 2,4-D levels in surface water, urban run-off, a freshwater pond and Kushiwah Creek, Charleston, of 0.1 μg/L to 11.5 μg/L. Rodil et al.[48] reported 2,4-D detected in drinking and surface water in Spain at concentrations ranging between 62 ng/L and 207 ng/L. The estimated recent environmental concentrations of 2,4-D in US water sources ranged from 4 μg/L to 24 μg/L.[53] These concentrations are much higher than the levels obtained in the current study.

The Canadian guideline for the maximum residue limit (MRL) for any pesticide in drinking water is 280 μg/L, and for freshwater aquatic life is 65 μg/L.[54] In the USA, the MRL for pesticides in drinking water is 700 μg/L[54] and the maximum contaminant level – specifically for 2,4-D – is 70 μg/L[55]. In the European Union (EU), the MRL for pesticides in drinking water is less than 0.1 μg/L[54] – a level exceeded by the 2,4-D concentrations found in the current study (Figure 2). Some of the levels of 2,4-D were an order of magnitude higher than the EU guideline (Figure 2), which could mean possible effects on human health. A Canadian study found a significantly increased risk of cancer (non-Hodgkins' disease) in men exposed to 2,4-D.[56] Some studies reported that 2,4-D could reduce growth rates, induce reproductive problems, and produce changes in appearance or behaviour, or could cause death of non-target species, including plants, animals and microorganisms.[57] In contrast, other studies examined the systemic toxicity, developmental neurotoxicity, developmental immunotoxicity, reproductive toxicity, endocrine modulation and thyroid effects in humans, and found that 2,4-D is unlikely to pose a significant health risk.[58,59] The debate on the safety of herbicides continues as there may be unknown long-term effects on human health and the environment.[60]

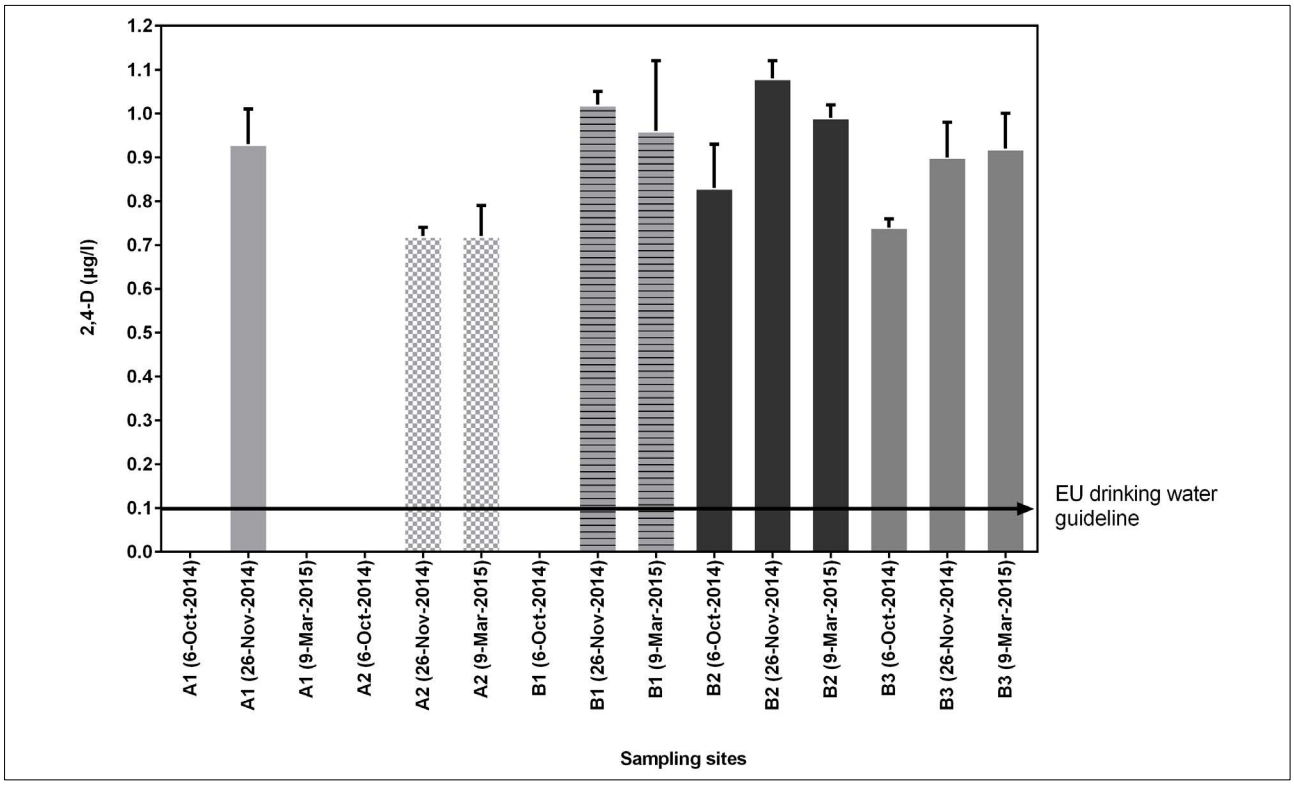


Figure 2: Concentrations of 2,4-D (μg/L) found across all sites and exceedance of the EU drinking water guideline (0.1 μg/L).

Conclusion

South Africa relies on agriculture to supply food to the majority of its people and is the 10th largest maize producer in the world. Both small-scale subsistence farming and modern agriculture are important in the country and both sectors use transgenic insect toxins and may experience development of tolerance to herbicides. Modern agriculture increases food production but may involve excessive use of herbicides and toxins for pest control. Ideally, herbicidal compounds are developed to have a specific mechanism or mode of action to avoid toxic effects in non-target organisms. However, non-target effects need to be investigated and the risk assessed for each chemical substance in use. The first step is to monitor and determine whether herbicides and agricultural toxins used by farmers can be found in the environment. To our knowledge, this has not been done previously for Cry1Ab toxin, glyphosate and 2,4-D in South Africa, although these are dominant agrochemicals in modern South African agriculture. Thus, this report is the first investigation of the presence and concentrations of these substances in water sources in South Africa.

As Cry1Ab, glyphosate and 2,4-D are highly mobile once released into the environment, increased use will elevate the levels in the environment. We did not find Cry1Ab proteins at quantifiable levels and only one sample contained glyphosate. 2,4-D was present at quantifiable levels in more than 70% of the samples and all of these concentrations exceeded the EU guideline for drinking water. Recently, research has revealed adverse health effects of Cry1Ab, glyphosate and 2,4-D exposure to non-target organisms. These effects could also influence biodiversity; therefore, water sources should be monitored to ensure both healthy aquatic ecosystems as well as safe drinking water.

Recommendations

From the results of this first survey conducted over a single maize growing season it is recommended that follow-up studies be done which include more sampling locations across larger geographical regions in South Africa. Also, monitoring should be performed over longer periods to cover variability over seasons and between years. We recommend the use of ELISAs as a screening tool followed by confirmation of positive results using other analytical methods.

Acknowledgements

We thank Anja Greyling for creating the map and acknowledge the South African Department of Agriculture, Forestry and Fisheries as the source of the map data. We thank the National Research Foundation (NRF) of South Africa for funding (grant numbers 106242 and 103487). Opinions expressed and conclusions derived are those of the authors and are not necessarily to be attributed to the NRF.

Authors' contributions

S.H. was responsbile for conceptualisation; data collection; sample analysis; data analysis and validation; and writing of the initial draft. R.P. contributed to the conceptualisation and was responsible for sample collection; student supervision; funding; and writing revisions. T.B. contributed to the conceptualisation and was responsible for project leadership; funding; and writing revisions.

References

1. Department of Agriculture, Forestry and Fisheries (DAFF). Maize market value chain profile 2010-2011. Pretoria: DAFF; 2011.
2. Jury MR. Economic impacts of climate variability in South Africa and development of resource prediction models. J Appl Meteorol. 2002;41(1):46–55. https://doi.org/10.1175/1520-0450(2002)041<0046:EIOCVI>2.0.CO;2
3. Gouse M, Pray CE, Kirsten J, Schimmelpfennig D. A GM subsistence crop in Africa: The case of Bt white maize in South Africa. Int J Biotechnol. 2005;7(1/2/3):84–94.
4. Kruger M, Van Rensburg JBJ, Van Den Berg J. Perspective on the development of stem borer resistance to Bt maize and refuge compliance at the Vaalharts irrigation scheme in South Africa. Crop Prot. 2009;28(8):684–689. https://doi.org/10.1016/j.cropro.2009.04.001
5. Van den Berg J, Hilbeck A, Bøhn T. Pest resistance to Cry1Ab Bt maize: Field resistance, contributing factors and lessons from South Africa. Crop Prot. 2013;54:154–160. https://doi.org/10.1016/j.cropro.2013.08.010
6. Van den Berg J. Insect resistance management in Bt maize: Wild host plants of stem borers do not serve as refuges in Africa. J Econ Entomol. 2017;110(1):221–229.

7. Soberón M, Gill SS, Bravo A. Signaling versus punching hole: How do *Bacillus thuringiensis* toxins kill insect midgut cells? Cell Mol Life Sci. 2009;66(8):1337–1349. https://doi.org/10.1007/s00018-008-8330-9

8. Venter HJ, Bøhn T. Interactions between Bt crops and aquatic ecosystems: A review. Environ Toxicol Chem. 2016;35(12):2891–2902. https://doi.org/10.1002/etc.3583

9. Douville M, Gagné F, Blaise C, André C. Occurrence and persistence of *Bacillus thuringiensis* (Bt) and transgenic Bt corn cry1Ab gene from an aquatic environment. Ecotoxicol Environ Saf. 2007;66(2):195–203. https://doi.org/10.1016/j.ecoenv.2006.01.002

10. Strain KE, Whiting SA, Lydy MJ. Laboratory and field validation of a Cry1Ab protein quantitation method for water. Talanta. 2014;128:109–116. https://doi.org/10.1016/j.talanta.2014.04.036

11. Benbrook CM. Impacts of genetically engineered crops on pesticide use in the U.S. – The first sixteen years. Environ Sci Eur. 2012;24(24):1–13. https://doi.org/10.1186/2190-4715-24-24

12. Dai P, Hu P, Tang J, Li Y, Li C. Effect of glyphosate on reproductive organs in male rat. Acta Histochem. 2016;118(5):519–526. https://doi.org/10.1016/j.acthis.2016.05.009

13. Simonsen L, Fomsgaard IS, Svensmark B, Spliid NH. Fate and availability of glyphosate and AMPA in agricultural soil. J Environ Sci Heal B. 2008;43(5):365–375. https://doi.org/10.1080/03601230802062000

14. Vivancos PD, Driscoll SP, Bulman CA, Ying L, Emami K, Treumann A, et al. Perturbations of amino acid metabolism associated with glyphosate-dependent inhibition of shikimic acid metabolism affect cellular redox homeostasis and alter the abundance of proteins involved in photosynthesis and photorespiration. Plant Physiol. 2011;157(1):256–268. https://doi.org/10.1104/pp.111.181024

15. Williams GM, Kroes R, Munro IC. Safety evaluation and risk assessment of the herbicide Roundup and its active ingredient, glyphosate, for humans. Regul Toxicol Pharmacol. 2000;31(2 Pt 1):117–165. https://doi.org/10.1006/rtph.1999.1371

16. Cuhra M, Bohn T, Cuhra P. Glyphosate: Too much of a good thing? Front Environ Sci. 2016;4(28):1–14. https://doi.org/10.3389/fenvs.2016.00028

17. International Agency for Research on Cancer (IARC). IARC Monographs Volume 112: Evaluation of five organophosphate insecticides and herbicides. Environ Heal. 2015;112(1):425–433.

18. European Food Safety Authority. Conclusion on the peer review of the pesticide risk assessment of the active substance glyphosate. EFSA J. 2015;13(11), Art. #4302, 107 pages.

19. Shaner DL, Lindenmeyer RB, Ostlie MH. What have the mechanisms of resistance to glyphosate taught us? Pest Manag Sci. 2012;68(1):3–9. https://doi.org/10.1002/ps.2261

20. Chahal PS, Aulakh JS, Rosenbaum K, Jhala AJ. Growth stage affects dose response of selected glyphosate-resistant weeds to premix of 2,4-D choline and glyphosate (Enlist Duo™ Herbicide). J Agric Sci. 2015;7(11):1–10. https://doi.org/10.5539/jas.v7n11p1

21. Landrigan PJ, Benbrook CM. GMOs, herbicides, and public health. N Engl J Med. 2015;373(8):693–695. https://doi.org/10.1056/NEJMp1505660

22. Benbrook CM. Trends in glyphosate herbicide use in the United States and globally. Environ Sci Eur. 2016;28(3):1–15. https://doi.org/10.1186/s12302-016-0070-0

23. Quinn L, De Vos BJ, Fernandes-Whaley M, Roos C, Bouwman H, Kylin H, et al. Pesticide use in South Africa: One of the largest importers of pesticides in Africa. In: Stoytcheva M, editor. Pesticides in the modern world – Pesticides use and management. Rijeka: InTech; 2011. p. 49. https://doi.org/10.5772/950

24. Food & Agriculture Organization of the United Nations Statistics Division (FAOSTAT). Crops and livestock products [webpage on the Internet]. c2016 [cited 2019 Apr 10]. Available from: http://www.fao.org/faostat/en/#data/TP/visualize

25. Brown LR, Robinson DE, Nurse RE, Swanton CJ, Sikkema PH. Soybean response to simulated dicamba/diflufenzopyr drift followed by postemergence herbicides. Crop Prot. 2009;28(6):539–542. https://doi.org/10.1016/j.cropro.2009.02.004

26. Department of Water Affairs and Forestry (DWAF). South African water quality guidelines. Volume 1: Domestic use. Pretoria: DWAF; 1996. Available from: http://www.dwa.gov.za/iwqs/wq_guide/Pol_saWQguideFRESH_vol1_Domesticuse.PDF 27.

27. South African Weather Service (SAWS). Historical rain maps [webpage on the Internet]. No date [cited 2019 Jan 24]. Available from: www.weathersa.co.za/climate/historical-rain-maps

28. Szekacs A, Mortl M, Darvas B. Monitoring pesticide residues in surface and ground water in Hungary: Surveys in 1990-2015. J Chem. 2015;2015:1–15. https://doi.org/10.1155/2015/717948

29. Pusztai-Carey M, Carey PR, Lessard T, Yaguchi M. Isolation, quantitation and purification of insecticidal proteins from *Bacillus thuringiensis*. U.S. patent no. 5,356,788. 1994. October, 18, 1994.

30. Tank JL, Rosi-Marshall EJ, Royer TV, Whiles MR, Griffiths NA, Frauendorf TC, et al. Occurrence of maize detritus and a transgenic insecticidal protein (Cry1Ab) within the stream network of an agricultural landscape. Proc Natl Acad Sci USA. 2010;107(41):17645–17650. https://doi.org/10.1073/pnas.1006925107

31. Strain KE, Whiting SA, Lydy MJ. Laboratory and field validation of a Cry1Ab protein quantitation method for water. Talanta. 2014;128:109–116.

32. Mörtl M, Németh G, Juracsek J, Darvas B, Kamp L, Rubio F, et al. Determination of glyphosate residues in Hungarian water samples by immunoassay. Microchem J. 2013;107:143–151. https://doi.org/10.1016/j.microc.2012.05.021

33. Schoeman C, Mashiane M, Dlamini M, Okonkwo OJ. Quantification of selected antiretroviral drugs in a wastewater treatment works in South Africa using GC-TOFMS. J Chromatogr Sep Tech. 2015;6(4):1–7.

34. Madliger M, Gasser CA, Schwarzenbach P, Sander M. Adsorption of transgenic insecticidal Cry1Ab protein to silica particles. Effects on transport and bioactivity. Environ Sci Technol. 2011;45:4377–4384. https://doi.org/10.1021/es200022q

35. Tank JL, Rosi-Marshall EJ, Royer TV, Whiles MR, Griffiths NA, Frauendorf TC, et al. Occurrence of maize detritus and a transgenic insecticidal protein (Cry1Ab) within the stream network of an agricultural landscape. Proc Natl Acad Sci USA. 2010;107(41):17645–17650. https://doi.org/10.1073/pnas.1006925107

36. Whiting SA, Strain KE, Campbell LA, Young BG, Lydy MJ. A multi-year field study to evaluate the environmental fate and agronomic effects of insecticide mixtures. Sci Total Environ. 2014;497–498:534–542. https://doi.org/10.1016/j.scitotenv.2014.07.115

37. Strain KE, Lydy MJ. The fate and transport of the Cry1Ab protein in an agricultural field and laboratory aquatic microcosms. Chemosphere. 2015;132:94–100. https://doi.org/10.1016/j.chemosphere.2015.03.005

38. Reddy KN, Rimando AM, Duke SO. Aminomethylphosphonic acid, a metabolite of glyphosate, causes injury in glyphosate-treated, glyphosate-resistant soybean. J Agric Food Chem. 2004;52(16):5139–5143. https://doi.org/10.1021/jf049605v

39. Battaglin WA, Meyer MT, Kuivila KM, Dietze JE. Glyphosate and its degradation product AMPA occur frequently and widely in U.S. soils, surface water, groundwater, and precipitation. J Am Water Resour Assoc. 2014;50(2):275–290. https://doi.org/10.1111/jawr.12159

40. Myers JP, Antoniou MN, Blumberg B, Carroll L, Colborn T, Everett LG, et al. Concerns over use of glyphosate-based herbicides and risks associated with exposures: A consensus statement. Environ Health. 2016;15(1):1–13. https://doi.org/10.1186/s12940-016-0117-0

41. Sanchís J, Kantiani L, Llorca M, Rubio F, Ginebreda A, Fraile J, et al. Determination of glyphosate in groundwater samples using an ultrasensitive immunoassay and confirmation by on-line solid-phase extraction followed by liquid chromatography coupled to tandem mass spectrometry. Anal Bioanal Chem. 2012;402(7):2335–2345. https://doi.org/10.1007/s00216-011-5541-y

42. Van Stempvoort DR, Spoelstra J, Senger ND, Brown SJ, Post R, Struger J. Glyphosate residues in rural groundwater, Nottawasaga river watershed, Ontario, Canada. Pest Manag Sci. 2016;72(10):1862–1872. https://doi.org/10.1002/ps.4218

43. Smith NJ, Martin RC, St. Croix RG. Levels of the herbicide glyphosate in well water. Bull Environ Contam Toxicol. 1996;57(5):759–765. https://doi.org/10.1007/s001289900254

44. Kolpin DW, Thurman EM, Lee EA, Meyer MT, Furlong ET, Glassmeyer ST. Urban contributions of glyphosate and its degradate AMPA to streams in the United States. Sci Total Environ. 2006;354(2–3):191–197. https://doi.org/10.1016/j.scitotenv.2005.01.028

45. Wilson RD, Geronimo J, Armbruster JA. 2,4-D dissipation in field soils after applications of 2,4-D dimethylamine salt and 2,4-D 2-ethylhexyl ester. Environ Toxicol Chem. 1997;16(6):1239–1246. https://doi.org/10.1002/etc.5620160620

46. Mountassif D, Kabine M, Mounchid K, Mounaji K, Latruffe N, El Kebbaj MS. Biochemical and histological alterations of cellular metabolism from jerboa (*Jaculus orientalis*) by 2,4-dichlorophenoxyacetic acid: Effects on d-3-hydroxybutyrate dehydrogenase. Pestic Biochem Physiol. 2008;90(2):87–96. https://doi.org/10.1016/j.pestbp.2007.08.001

47. Hernandez AA, Silva MR, Moya CA. Compuestos organo-persistentes y daño genético en núcleos hepáticos de Goodea atripinnis del Lago de Chapala [Organo-persistent compounds and genetic damage in liver nuclei of Goodea atripinnis in Lake Chapala]. Scientia-CUCBA. 2011;13:1–8. Spanish.

48. Rodil R, Quintana JB, Concha-Graña E, López-Mahía P, Muniategui-Lorenzo S, Prada-Rodríguez D. Emerging pollutants in sewage, surface and drinking water in Galicia (NW Spain). Chemosphere. 2012;86(10):1040–1049. https://doi.org/10.1016/j.chemosphere.2011.11.053

49. Tsaboula A, Papadakis EN, Vryzas Z, Kotopoulou A, Kintzikoglou K, Papadopoulou-Mourkidou E. Environmental and human risk hierarchy of pesticides: A prioritization method, based on monitoring, hazard assessment and environmental fate. Environ Int. 2016;91:78–93. https://doi.org/10.1016/j.chemosphere.2011.11.053

50. Serrano L, DeLorenzo ME. Water quality and restoration in a coastal subdivision stormwater pond. J Environ Manage. 2008;88(1):43–52. https://doi.org/10.1016/j.jenvman.2007.01.025

51. Ensminger MP, Budd R, Kelley KC, Goh KS. Pesticide occurrence and aquatic benchmark exceedances in urban surface waters and sediments in three urban areas of California, USA, 2008-2011. Environ Monit Assess. 2013;185(5):3697–3710. https://doi.org/10.1007/s10661-012-2821-8

52. Wijnja H, Doherty JJ, Safie SA. Changes in pesticide occurrence in suburban surface waters in Massachusetts, USA, 1999-2010. Bull Environ Contam Toxicol. 2014;93(2):228–232. https://doi.org/10.1007/s00128-014-1251-4

53. Atamaniuk TM, Kubrak OI, Storey KB, Lushchak VI. Oxidative stress as a mechanism for toxicity of 2,4-dichlorophenoxyacetic acid (2,4-D): Studies with goldfish gills. Ecotoxicology. 2013;22(10):1498–1508. https://doi.org/10.1007/s10646-013-1136-z

54. Rubio F, Veldhuis LJ, Clegg BS, Fleeker JR, Hall JC. Comparison of a direct ELISA and an HPLC method for glyphosate determinations in water. J Agric Food Chem. 2003;51(3):691–696. https://doi.org/10.1021/jf020761g

55. United States Environmental Protection Agency (USEPA). National primary drinking water regulations: Organic chemicals [webpage on the Internet]. c2018 [2019 Apr 09]. Available from: https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations

56. McDuffie HH, Pahwa P, McLaughlin JR, Spinelli JJ, Fincham S. Non-Hodgkin's lymphoma and specific pesticides exposures in men: Cross-Canada study of pesticides and health. Cancer Epidemiol Biomarkers Prev. 2001;10:1155.

57. Gervais JA, Luukinen B, Buhl K, Stone D. 2,4-D Technical fact sheet [document on the Internet]. c2008 [2018 Jul 15]. Available from: http://npic.orst.edu/factsheets/2,4-DTech.pdf

58. Marty MS, Neal BH, Zablotny CL, Yano BL, Andrus AK, Woolhiser MR, et al. An F1-extended one-generation reproductive toxicity study in Crl: CD(SD) rats with 2,4-dichlorophenoxyacetic acid. Toxicol Sci. 2013;136(2):527–547. https://doi.org/10.1093/toxsci/kft213

59. Peterson MA, McMaster SA, Riechers DE, Skelton J, Stahlman PW. 2,4-D past, present, and future: A review. Weed Technol. 2016;30(02):303–345. https://doi.org/10.1614/WT-D-15-00131.1

60. Green JM. The benefits of herbicide-resistant crops. Pest Manag Sci. 2012;68(10):1323–1331. https://doi.org/10.1002/ps.3374

AUTHORS:
Marie E.A. Bissong[1,2]
Collins N. Ateba[1,3]

AFFILIATIONS:
[1]Department of Microbiology, School of Biological Sciences, North-West University, Mmabatho, South Africa
[2]Department of Biomedical Science, University of Bamenda, Bambili, Cameroon
[3]Food Security and Safety Niche Area, Faculty of Agriculture, Science and Technology, North-West University, Mmabatho, South Africa

CORRESPONDENCE TO:
Marie Bissong

EMAIL:
mabissong@yahoo.com

DATES:
Received: 19 Mar. 2018
Revised: 26 Oct. 2018
Accepted: 02 July 2019
Published: 26 Sep. 2019

HOW TO CITE:
Bissong MEA, Ateba CN. Detection of virulent thermophilic *Campylobacter* species in communal chickens. S Afr J Sci. 2019;115(9/10), Art. #4771, 5 pages. https://doi.org/10.17159/sajs.2019/4771

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
John Butler-Adam

KEYWORDS:
Campylobacter, virulence, genotype, prevalence, chicken

FUNDING:
North-West University

Detection of virulent thermophilic *Campylobacter* species in communal chickens

Campylobacter is a foodborne pathogen found in the gut of poultry and other animals, which thereby act as reservoirs for human infection. *Campylobacter* produces various virulence factors which contribute to the establishment of disease. The aim of this study was to detect virulence genes in *Campylobacter* isolated from chicken faeces. A total of 408 poultry cloacal samples were analysed through culture, PCR and DNA sequencing. *Campylobacter* were detected in 24 (5.9 %) of the samples, of which 21 were *C. jejuni* while 3 were *C. coli.* The *cdtB* gene was detected in all *Campylobacter* isolates while *cdtA* and *cdtC* genes were detected in 20 (83.3%) and 18 (75.0%) of the isolates, respectively. The *cadF* gene was detected in 18 (85.7%) *C. jejuni* but not in *C. coli*. Eight genotypes designated G1 to G8 were detected among the isolates. Genotypes G1 and G7 possessed all three *cdt* genes, a potential for producing functionally active toxins. A large proportion (70.8%) of the isolates possessed both genes for toxin production and adhesion. The high prevalence of *cdt* and *cadF* genes in *Campylobacter* species in this study indicates the pathogenic potential of the isolates. There is a need to implement control measures to curb the spread of these virulent strains.

Significance:

- Virulent thermophilic campylobacters cause severe health complications in humans.
- Although chicken meat is the main source of human infections, there is limited data on the occurrence of *Campylobacter* in chicken faeces in southern Africa.
- This study adds to literature on the occurrence of virulent strains of *Campylobacter* species, notably through the deposition of the DNA sequences of our local strains into GenBank®.
- Communal chickens may serve as a source of virulent thermophilic campylobacters to humans, resulting in serious public health implications.

Introduction

Campylobacter species are Gram-negative curved bacteria which normally occur in the gastrointestinal tract of many domestic animals including poultry.[1] These animals serve as potential reservoirs for human infection. If proper farm management techniques are not enforced, *Campylobacter* species can contaminate foods of animal origin and eventually gain access into the food chain.[2-4] Undercooked meat, especially poultry, has been reported to serve as the main vehicle for the transmission of these organisms.[5,6] Despite the fact that many species have been detected, it is the thermophilic species, most notably *C. jejuni* and *C. coli,* that are responsible for human infection.[1] These species cause diseases in humans ranging from mild self-limiting diarrhoea to severe dysentery and, in some cases, the disease presents with severe complications such as Guillain-Barre Syndrome.[7-9]

Campylobacter virulence mechanisms include motility, mucus colonisation, toxin production, attachment, internalisation and translocation.[10] These virulence factors and toxins are known to contribute to the survival and establishment of diseases in susceptible hosts.[10] The flagella are used for motility of the bacterium and are known to enhance attachment of the bacterium to epithelial cells.[7,11] The involvement of the bacterial flagella contributes to their pathogenicity in the host's gut.[11] Invasiveness is also an important mechanism that enhances virulence in *Campylobacter* species.[10] These bacteria become internalised into the sub-mucosal cells by means of the flagella and certain proteins.[10] Invasion causes tissue damage and inflammation leading to gastroenteritis.[10,12]

Campylobacter adhesion to fibronectin (CadF) is an outer membrane protein which aids in the binding of *C. jejuni* and *C. coli* to intestinal epithelial cells,[13] thus enhancing the colonisation of the gut. Bacterial toxins also play a key role in disease pathogenesis. *Campylobacter* species produce several cytotoxins, notably the cytolethal distending toxin (CDT) which causes direct damage to the DNA of the host cell and eventually leads to cell death.[14] The CDT gene cluster consists of three protein subunits: *cdtA, cdtB* and *cdtC*. Although very little is known about the functions of the *cdtA* and *cdtC* gene fragments, *cdtB* has recently been identified to have nuclease activities.[14] However, considering the fact that there are only a few *Campylobacter* genes that have been reported to be directly involved in virulence,[15] the importance of the current study cannot be overemphasised.

Campylobacter is considered to be the most common cause of bacterial gastroenteritis worldwide.[16] The rate of *Campylobacter* infections worldwide has been increasing and it is estimated that incidence of campylobacteriosis exceeds those of salmonellosis and shigellosis.[17,18] This increase, as well as the expanding spectrum of diseases caused by *Campylobacter* species, necessitates a clearer understanding of the epidemiology and control of campylobacteriosis, especially in developing countries. Campylobacteriosis is common in Africa, especially among children under 5 years, with a prevalence of 7.7–21%.[19-21]

It is worth noting that the agropastoral industry is fast developing in South Africa in particular and in sub-Saharan Africa in general. Unfortunately, animals intended for human consumption are often natural reservoirs for pathogenic bacteria in humans. Infections with thermophilic *Campylobacter* present with severe complications in humans. Despite these health impacts, there are limited data on the prevalence of *Campylobacter* in clinical, environmental and food products in southern Africa.[22-25] There is a need to exploit other sources, particularly animal species that are natural hosts of *Campylobacter* species, for the presence of potentially virulent strains.

The present study was designed in order to isolate campylobacters from chicken cloacal samples and to determine the virulence profile of the isolates. Although other genes are responsible for the expression of pathogenicity in *Campylobacter*, the *cdt* gene subtypes (A, B and C) responsible for the expression of the cytolethal distending toxin and the *cadF* gene expressing products of adhesion were selected as targets for the present study.

Materials and methods

Study location

Samples were collected at two poultry farms in Mafikeng, North West Province of South Africa. Laboratory analyses were conducted in the Molecular Microbiology Laboratory of the Department of Microbiology, North-West University, Mafikeng Campus. Ethical clearance for the study was obtained from the Mafikeng Animal Research Ethics Committee (MAREC) of the North-West University, South Africa.

Sample collection

A total of 408 chicken cloacal swabs were collected using the culture swab transport system (COPAN, Italy) and transported on ice to the laboratory for analysis. The samples were analysed for the presence of *Campylobacter* species within 24 h.

Isolation of Campylobacter *species*

Each swab was rinsed in 5 mL of 2% (w/v) sterile peptone water (Biolab, Modderfontein, South Africa) and 100-μL aliquots spread-plated on *Campylobacter* selective agar supplemented with 5% (w/v) sheep blood and *Campylobacter* selective supplement (Merck, Darmstadt, Germany). The plates were incubated under microaerophilic (10% CO_2) conditions at 42 °C for 24–72 h. Presumptive isolates were sub-cultured and stored in brain heart infusion broth with 20% (v/v) glycerol at -80 °C for further analysis.

Biochemical identification of Campylobacter *species*

All isolates were identified using standard biochemical tests.[26] Isolates were Gram stained and observed for Gram-negative spiral rods. The oxidase test was performed using the oxidase test reagent obtained from Pro-Lab Diagnostics (Merseyside, UK) as indicated in the manufacturer's protocol. The catalase test was performed as previously reported.[26] The *Campylobacter* test kit (Oxoid, Basingstoke, Hampshire, UK) was used to identify thermophilic campylobacters based on the manufacturer's description.

Molecular characterisation of Campylobacter *species*

Extraction of DNA

Pure isolates were inoculated into 10 mL of nutrient broth (Merck, Darmstadt, Germany) and incubated aerobically at 37 °C for 24 h in a shaking incubator to enhance bacterial growth. Genomic DNA was extracted using the Zymo Research Genomic DNA™ Tissue MiniPrep Kit (Zymo Research Corp, Irvine, CA, USA) based on the manufacturer's instructions and the samples were stored at -20 °C.

Detection of *C. jejuni* and *C. coli*

All isolates were screened for the presence of bacterial 16S rRNA gene fragments and identification of *C. jejuni* and *C. coli* was achieved using the *hipO* and *asp* gene primers, respectively (Table 1). Polymerase chain reaction (PCR) amplifications were performed using the model Bio-Rad C1000 Touch™ Thermal Cycler (Bio-Rad, Johannesburg, South Africa). The reactions were prepared in 25-μL volumes made up of 12.5 μL of One Taq^R Quickload 2X Master Mix with standard buffer (Inqaba Biotec Ltd, Pretoria, South Africa), 1 μL of template DNA, 11 μL nuclease free sterile water and a final primer concentration of 0.2 μM. All PCR reagents were obtained from Inqaba Biotec Ltd (Pretoria, South Africa). PCR conditions comprised an initial denaturation at 95 °C for 6 min, followed by 35 cycles of denaturation at 94 °C for 50 s, annealing at 55 °C for 40 s and extension at 72 °C for 50 s. A final extension step was performed at 72 °C for 3 min. PCR products were stored at 4 °C until electrophoresis.

Detection of virulence genes

Virulence gene determinants associated with pathogenesis in *Campylobacter* species were determined by specific PCR analysis using the primers listed in Table 1. Cycling was performed using a DNA thermal cycler. The conditions for amplification were the same as those for the identification genes, except for *cdtC* and *cadF* genes, which were annealed at 50 °C.

Table 1: List of primers used for PCR amplification

Primer	Target gene	Sequence 5'–3'	Size (bp)	Reference
Identification genes				
hipO-F	*hipO*	GACTTCGTGCAGATATGGATGCTT	344	27
hipO-R		GCTATAACTATCCGAAGAAGCCATCA		
CC18F	*asp*	GGTATGATTTCTACAAAGCGAG	500	28
CC519R		ATAAAAGACTATCGTCGCGTG		
27F	16S rRNA	AGAGTTTGATCCTGGCTCAG	1400	29
1492R		TACCTTG TTACGACTT		
Virulence genes				
CDTAF	*cdtA*	CCTTGTGATGCAAGCAATC	370	30
CDTAR		ACACTCCATTTGCTTTCTG		
CDTBF	*cdtB*	CAGAAAGCAAATGGAGTGTT	620	30
CDTBR		AGCTAAAAGCGGTGGAGTAT		
CDTCF	*cdtC*	CGATGAGTTAAAACAAAAAGATA	182	30
CDTCR		TTGGCATTATAGAAAATACAGTT		
CADF-F	*cadF*	TTGAAGGTAATTTAGATATG	400	31
CADF-R		CTAATACCTAAAGTTGAAAC		

Agarose gel electrophoresis

PCR products were resolved by electrophoresis on a 1% (w/v) agarose gel at 80 V for 3 h, using 1X TAE buffer (40 mM Tris, 1 mM EDTA and 20 mM glacial acetic acid, pH 8.0). A 100-bp DNA molecular weight marker (Fermentas, Glen Burnie, MD, USA) was included in each gel. The gels were stained in 0.001 μg/mL of ethidium bromide for 15 min and the amplicons visualised using a ChemiDoc Imaging System (Bio-RAD, Hercules, CA, USA).

DNA Sequencing

16S rRNA PCR amplicons were cleaned using a GeneJet PCR Purification kit (Thermo Fisher Scientific, Vilnius, Lithuania) and sequenced at Inqaba Biotec (Pretoria, South Africa). The identities of the isolates were confirmed through a BLAST search of the US National Center for Biotechnology Information (NCBI) database (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Results and discussion

Bacterial virulence is an important issue that needs scientific investigation. Several virulent determinants are associated with the survival and pathogenesis of *Campylobacter* species. This study was designed to determine the prevalence of *Campylobacter* species in chicken faeces and to identify virulence genes in the isolates. The *cdt* gene subtypes (A, B and C) responsible for the expression of the cytolethal distending toxin and the *cadF* gene expressing products of adhesion were selected for the study.

The occurrence of *Campylobacter* in both clinical and environmental samples from southern Africa has been reported.[22-25] Diergaardt and colleagues reported the presence of *Campylobacter* spp. in both drinking and environmental water sources in Pretoria.[22] In a study conducted in the Limpopo Province, *Campylobacter* species were found to be common causes of gastroenteritis in primary school children.[23] Meanwhile, Uaboi-Egbenni and colleagues detected the presence of *Campylobacter* in

goat faeces.[32] In the Mafikeng area, *Campylobacter* has been detected in commercial milk as well as chicken meat.[24,25] Despite these reports, information on the occurrence of virulent *Campylobacter* in chicken faeces in sub-Saharan Africa is scarce. In the present study, samples were collected from two commercial farms in Mafikeng. These farms sell live chickens to the local population and the chickens are slaughtered at the customers' homes. The handling of live chickens might enhance the spread of enteric pathogens including *Campylobacter* from chickens to humans, especially as standard hygiene practices might not be observed. In such scenarios, the presence of virulent campylobacters in chicken might present a serious public health hazard, more especially as these thermophilic bacteria can survive higher temperatures.

Isolation and molecular identification of Campylobacter *species*

In this study, *Campylobacter* was isolated from 24 of the 408 (5.9%) chicken cloacal samples. However, of the samples inoculated, growth was observed in only 198 (48.5%), and out of this number, 126 (63.6%) were presumptively identified as *Campylobacter* based on growth on selective media and colony morphology. Isolates that presented with the characteristic mucoid cream to pink colonies (Figure 1) were tested using species-specific primers. The isolated DNA was tested for the presence of bacterial 16S rRNA gene fragments and all 126 (100%) presumptive isolates were positive – an indication that the genetic material was all from bacterial sources. Results of the PCR using specific primers showed that only 24 of 126 (19.0%) presumptive isolates were *Campylobacter* species; with 21 (16.7%) and 3 (2.4%) possessing the *hipO* and *asp* genes, respectively, specific for *Campylobacter jejuni* and *Campylobacter coli,* respectively (Figure 2).

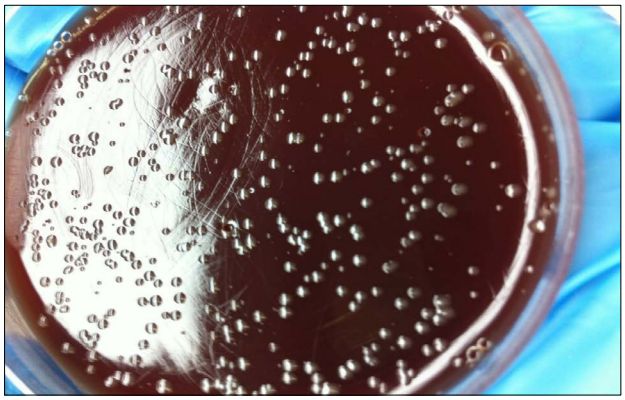
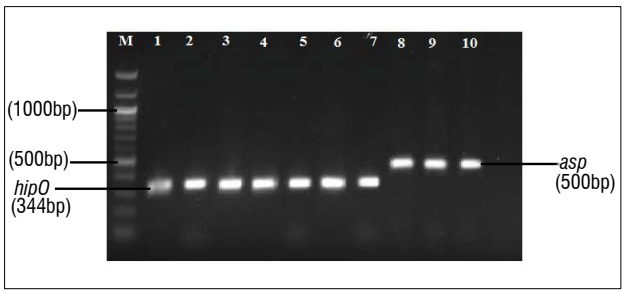


Figure 1: Colonies of *Campylobacter* on blood agar.



Lane M, DNA marker (100 base pairs); Lanes 1–7, hipO *gene of* C. jejuni*; Lanes 8–10,* asp *gene of* C. coli

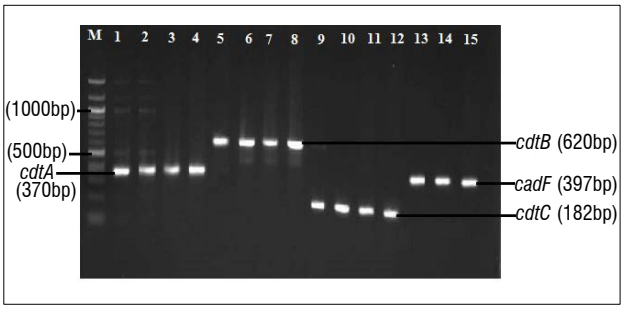
Figure 2: Agarose gel of the *hipO* and *asp* genes.

DNA Sequencing

Four isolates that were positive for the *hipO* gene were randomly selected and sent for DNA sequencing to confirm their identities. The identities of the isolates were confirmed through a BLAST search of the NCBI database and all the isolates were confirmed to be *Campylobacter jejuni.* In addition, one of the isolates that grew on the selective media but did not amplify with the campylobacter-specific primers was also sent for DNA sequencing using the universal 16S rRNA primer. This isolate was found to be *Lactobacillus reuteri*. Similar observations have been recorded in which isolates supposedly identified as *Campylobacter*, were later confirmed by PCR and or DNA sequencing to belong to other genera. In a related study, Diergaardt et al.[22] confirmed by sequencing that only 3 out of 22 presumptive isolates were actually *Campylobacter*. In another study, out of 75 samples examined, only 1 sample (1.33%) was positive by cultural method while 5 samples (6.66%) showed the presence of *Campylobacter* species by PCR.[33] Consequently, these reports confirm the superior efficacy of PCR over culture-based techniques for the rapid screening of both clinical and environmental samples. The sequences of the isolates were deposited in GenBank®, the sequence database of the NCBI, with the following accession numbers: MF872608 (*Campylobacter jejuni*, strain MOLcc14, 99% identity), MF872609 (*Campylobacter jejuni*, strain MOLcc20, 99% identity), MF872610 (*Campylobacter jejuni*, strain MOLcc29, 99% identity), MF872611 (*Campylobacter jejuni*, strain MOLcc30, 99% identity) and MF872612 (*Lactobacillus reuteri*, strain MOLcl43, 99% identity).

PCR detection of virulence gene determinants of Campylobacter *species*

To identify the virulence determinants of *Campylobacter* species isolated from chicken cloacal samples, four genes (*cdtA, cdtB, cdtC* and *cadF*) were targeted. The *cdt* gene is responsible for the expression of the cytolethal distension toxin which induces cell cycle arrest and promotes DNA damage.[34,35] Consequently, its presence is associated with the severity of *Campylobacter* disease in humans. Figure 3 shows the amplification of the virulence genes of *C. jejuni*. In this study, the *cdtB* gene was detected in all (24, 100%) *Campylobacter* isolates while *cdtA* and *cdtC* genes were detected in 83.3% and 75% of the isolates, respectively. These results concur with the findings of studies by Gonzalez-Hein et al.[36], Carvalho et al.[37] and Casabonne et al.[38] who detected *cdtB* genes in all *Campylobacter* species isolated from various sources. In contrast, Montwedi and Ateba[25] reported relatively lower detection rates of *cdtA* (18/50, 36%) and *cdtC* (9/50, 18%) genes from chicken meat samples. The *cadF* gene encodes a protein that interacts with the host's fibronectin, consequently playing an important role in the adhesion and colonisation of the bacterium in the host's gut. In this study, the *cadF* gene was detected in the majority (18, 85.7%) of *C. jejuni* isolates, in line with previous studies which revealed close to 100% detection of the *cadF* gene.[36,38,39] However, none of the *C. coli* isolates from the present study showed the presence of the *cadF* gene, possibly because only three *C. coli* isolates were tested. Despite the importance of the *cadF* gene, other genes such as the *flaA* and *ciaB* genes (not investigated in this study) are equally important in the adhesion, colonisation and, consequently, survival of bacteria in the gut.[38]



Lane M, DNA marker (100 base pairs); Lanes 1–4, cdtA *gene; Lanes 5–8,* cdtB *gene; Lanes 9–12,* cdtC *gene; Lanes 13–15,* cadF *gene*

Figure 3: Agarose gel of the *cdtA, cdtB, cdtC* and *cadF* gene fragments.

Genotyping of Campylobacter *species*

Genotyping was based on the combination of various identification and virulence genes. A total of eight genotypes designated G1 to G8 were detected among *Campylobacter* isolates (Table 2). G1 was the most prevalent (17, 71.4%) genotype among all *Campylobacter* species. Two genotypes (G1 and G7) possessed all three genes (*cdtA, cdtB* and *cdtC*) for toxin expression. These genotypes make up the majority of the

Table 2: Proportion of different genotypes of *Campylobacter* isolates. Genotyping was based on the various combinations of identification (*) and virulence (#) genes.

Genotype	Species	Gene combinations for *Campylobacter* isolates							Number (%)
		*16rRNA	**hipO*	**asp*	#*cdtA*	#*cdtB*	#*cdtC*	#*cadF*	
G1	*C. jejuni*	+	+	-	+	+	+	+	15 (71.4)
G2	*C. jejuni*	+	+	-	+	+	+	-	1 (4.8)
G3	*C. jejuni*	+	+	-	-	+	+	-	2 (9.5)
G4	*C. jejuni*	+	+	-	-	+	-	-	1 (4.8)
G5	*C. jejuni*	+	+	-	+	+	-	+	1 (4.8)
G6	*C. jejuni*	+	+	-	+	+	-	-	1 (4.8)
G7	*C. coli*	+	-	+	+	+	+	-	2 (66.7)
G8	*C. coli*	+	-	+	-	+	-	-	1 (33.3)

+, present; -, absent

Campylobacter isolated in this study. The CDT is a tripartite toxin with CDTB being the enzymatically active subunit; however, all three gene products are required for the toxin to be functionally active.[34] Consequently, the majority of *Campylobacter* isolated in this study have the potential for producing functionally active toxins. As a result, we envisage a study to investigate the actual expression of toxins by *Campylobacter* isolates from environmental samples. In addition, most of the *Campylobacter* isolates as represented by G1, G3 and G5 possessed both genes for toxin production and adhesion. Interestingly, all the four *C. jejuni* strains that were sequenced belong to G1. The high prevalence of the *cdt* and *cadF* genes in *Campylobacter* species in this study is an indication of the pathogenic potential of the isolates.

Conclusion

Thermophilic campylobacters were successfully isolated from chicken cloacal swabs and the gene sequences of isolates deposited in GenBank. Highly virulent *Campylobacter* species were detected in chicken faeces in the study area, which may present a serious health hazard. Consequently, there is a need to implement control measures to curb the spread of these virulent strains.

Acknowledgements

We are grateful to the North-West University Postdoctoral Fellowship Programme for providing funds for this study and to the Department of Microbiology, North-West University (Mafikeng Campus) for providing facilities to conduct the study. We also acknowledge the assistance of the workers at Molelwane Farm in sample collection.

Authors' contributions

M.E.A.B.: Design of methodology, sample collection and laboratory analysis, data validation, writing and revision of the manuscript. C.N.A.: Conception and design of the study, data validation, revision and final approval of the manuscript.

References

1. Ryan KJ, Ray CG. Sherris medical microbiology: An introduction to infectious diseases. 4th ed. New York: McGraw Hill; 2004. p. 378–380.
2. Linscott AJ. Food-borne illnesses. Clin Microbiol Newsl. 2011;33(6):41–45.
3. Azevedo I, Albano H, Silva J, Teixeira P. Food safety in the domestic environment. Food Control. 2014;37:272–276. https://doi.org/10.1016/j.foodcont.2013.09.058
4. Meunier M, Guyard-Nicodeme M, Dory D, Chemaly M. Control strategies against *Campylobacter* at the poultry production level: Biosecurity measures, feed additives and vaccination. J Appl Microbiol. 2015;120:1139–1173. https://doi.org/10.1111/jam.12986
5. Humphrey T, O'Brien S, Madsen M. Campylobacters as zoonotic pathogens: A food production perspective. Int J Food Microbiol. 2007;117(3):237–257. https://doi.org/10.1016/j.ijfoodmicro.2007.01.006
6. Silva J, Leite D, Fernandes M, Mena C, Gibbs PA, Teixeira P. *Campylobacter* spp as a foodborne pathogen: A review. Front Microbiol. 2011;2:200. https://doi.org/10.3389/fmicb.2011.00200
7. Konkel ME, Monteville MR, Rivera-Amill V, Joens LA. The pathogenesis of *Campylobacter jejuni*-mediated enteritis. Curr Issues Intest Microbiol. 2001;2(2):55–71.
8. Rao MR, Naficy AB, Savarino SJ, Abu-Elyazeed R, Wierzba TF, Peruski LF, et al. Pathogenicity and convalescent excretion of *Campylobacter* in rural Egyptian children. Am J Epidemiol. 2001;154:166–173. https://doi.org/10.1093/aje/154.2.166
9. Nyati KK, Nyati R. Role of *Campylobacter jejuni* infection in the pathogenesis of Guillain-Barre Syndrome: An update. BioMed Res Int. 2013;2013, Art. ID 852195, 13 pages. https://doi.org/10.1155/2013/852195
10. Quinn PJ, Markey BK, Carter ME, Demnelly WJ, Leonard FC. Veterinary microbiology and microbial disease. 8th ed. Oxford: Blackwell Publishing; 2001.
11. Guerry P. *Campylobacter* flagella: Not just for motility. Trends Microbiol. 2007;15:456–461. https://doi.org/10.1016/j.tim.2007.09.006
12. Grant CCR, Michael EK, Witold C Jr, Lucy ST. Role of flagella in adherence, internalization and translocation of *Campylobacter jejuni* in non-polarised and polarised epithelial cell cultures. Infect Immun. 1993;61:1764–1771.
13. Konkel ME, Garvis SG, Tipton SL, Anderson Jr DE, Cieplak Jr W. Identification and molecular cloning of a gene encoding a fibronectin-binding protein (CadF) from *Campylobacter jejuni*. Mol Microbiol. 1997;24:953–963. https://doi.org/10.1046/j.1365-2958.1997.4031771.x
14. Lee RB, Hassane DC, Cottle DL, Pickett CL. Interactions of *Campylobacter jejuni* cytolethal distending toxin subunits CdtA and CdtC with HeLa cells. Infect Immun. 2003;71:4883–4890. https://doi.org/10.1128/iai.71.9.4883-4890.2003
15. Lara-Tejero M, Galán JE. *CdtA, CdtB*, and *CdtC* form a tripartite complex that is required for cytolethal distending toxin activity. Infect Immun. 2001;69(7):4358–4365. https://doi.org/10.1128/iai.69.7.4358-4365.2001
16. US Centers for Disease Control and Prevention (CDC). Surveillance for foodborne disease outbreaks, United States, 2014: Annual report. Atlanta, GA: US Department of Health and Human Services, CDC; 2016.
17. Altekruse SF, Stern NJ, Fields PI, Swerdlow DL. *Campylobacter jejuni* – an emerging foodborne pathogen. Emerg Infect Dis. 1999;5:28–35. https://doi.org/10.3201/eid0501.990104
18. Platts-Mills JA, Kosek M. Update on the burden of *Campylobacter* in developing countries. Curr Opin Infect Dis. 2014;27(5):444–450. https://doi.org/10.1097/qco.0000000000000091

19. Mason J, Iturriza-Gomara M, O'Brien SJ, Ngwira BM, Dove W, Maiden MC, et al. *Campylobacter* infection in children in Malawi is common and is frequently associated with enteric virus co-infections. PLoS ONE. 2013;8, e59663, 5 pages. https://doi.org/10.1371/journal.pone.0059663

20. Koulla-Shiro S, Loe C, Ekoe T. Prevalence of *Campylobacter* enteritis in children from Yaounde (Cameroon). Cent Afr J Med. 1995;41:91–94.

21. Lindblom GB, Ahren C, Changalucha J, Gabone R, Kaijser B, Nilsson LA, et al. *Campylobacter jejuni/coli* and enterotoxigenic *Escherichia coli* (ETEC) in faeces from children and adults in Tanzania. Scand J Infect Dis. 1995;27:589–593. https://doi.org/10.3109/00365549509047073

22. Diergaardt SM, Venter SN, Spreeth A, Theron J, Brozel VS. The occurrence of campylobacters in water sources in South Africa. Water Res. 2004;38:2589–2595. https://doi.org/10.1016/j.watres.2004.03.004

23. Samie A, Obi CL, Barret LJ, Powell SM, Guerrant RL. Prevalence of *Campylobacter* species, *Helicobacter pylori* and *Arcobacter* species in stool sample from the Venda region, Limpopo, South Africa: Studies using molecular diagnostic methods. J Infect. 2007;54(6):558–566. https://doi.org/10.1016/j.jinf.2006.10.047

24. Mabote KI, Mbewe M, Ateba CN. Prevalence of *Campylobacter* contamination in fresh chicken meat and milk obtained from markets in the North-West Province, South Africa. J Hum Ecol. 2011;36(1):23–28. https://doi.org/10.1080/09709274.2011.11906413

25. Montwedi MG, Ateba CN. Use of the *cdt* gene specific PCR in determining virulence properties of *Campylobacter jejuni* isolated from chicken meat samples obtained in some supermarkets in Mafikeng, NWP, South Africa. Life Sci J. 2012;9(3):2696–2701.

26. Prescott H. Laboratory exercise in microbiology. 5th ed. New York: McGraw Hill; 2002.

27. Persson S, Olsen KE. Multiplex PCR for identification of *Campylobacter coli* and *Campylobacter jejuni* from pure cultures and directly on stool samples. J Med Microbiol. 2005;54(11):1043–1047. https://doi.org/10.1099/jmm.0.46203-0

28. Linton D, Lawson AJ, Owen RJ, Stanley J. PCR detection, identification to species level, and fingerprinting of *Campylobacter jejuni* and *Campylobacter coli* direct from diarrheic samples. J Clin Microbiol. 1997;35:2568–2572.

29. Frank JA, Reich CI, Wilson BA, Sharma S, Olsen GJ. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol. 2008;74(8):2461–2470. https://doi.org/10.1128/aem.02272-07

30. Datta S, Niwa H, Itoh K. Prevalence of 11 pathogenic genes of *Campylobacter jejuni* by PCR in strains isolated from humans, poultry meat and broiler and bovine faeces. J Med Microbiol. 2003;52:345–348. https://doi.org/10.1099/jmm.0.05056-0

31. Konkel ME, Gray SA, Kim BJ, Garvis SG, Yoon J. Identification of the enteropathogens *Campylobacter jejuni* and *Campylobacter coli* based on the *cadF* virulence gene and its product. J Clin Microbiol. 1999;38:510–517.

32. Uaboi-Egbenni PO, Bessong PO, Samie A, Obi CL. Prevalence and antimicrobial susceptibility profiles of *Campylobacter jejuni* and *coli* isolated from diarrheic and non-diarrheic goat faeces in Venda region, South Africa. Afr J Biotechnol. 2011;10(64):14116–14124. https://doi.org/10.5897/ajb10.1662

33. Kumar PA, Kumar S. Comparative analysis of cultural and PCR based assays for detection of *Campylobacter* spp. in human stool samples. Proc Natl Acad Sci India. 2015;85(3):839–844. https://doi.org/10.1007/s40011-015-0565-2

34. Asakura M, Samosornsuk W, Hinenoya A, Misawa N, Nishimura K, Matsuhisa A, et al. Development of a cytolethal distending toxin (*cdt*) gene-based species-specific multiplex PCR assay for the detection and identification of *Campylobacter jejuni*, *Campylobacter coli* and *Campylobacter fetus*. FEMS Immunol Med Microbiol. 2008;52:260–266. https://doi.org/10.1111/j.1574-695x.2007.00369.x

35. Bolton DJ. *Campylobacter* virulence and survival factors. Food Microbiol. 2015;48:99–108.

36. González-Hein G, Huaracán B, Garcia P, Figueroa G. Prevalence of virulence genes in strains of *Campylobacter jejuni* isolated from human, bovine and broiler. Braz J Microbiol. 2013;44(4):1223–1229. https://doi.org/10.1590/s1517-83822013000400028

37. Carvalho AF, Silva DM, Azevedo SS, Piatti RM, Genovez ME, Scarcelli E. Detection of CDT toxin genes in *Campylobacter* spp. strains isolated from broiler carcasses and vegetables in São Paulo, Brazil. Braz J Microbiol. 2013;44(3):693–699. https://doi.org/10.1590/s1517-83822013000300005

38. Casabonne C, Gonzalez A, Aquili V, Subils T, Balague C. Prevalence of seven virulence genes of *Campylobacter jejuni* isolated from patients with diarrhea in Rosario, Argentina. Int J Infect. 2016;3(4), e37727, 6 pages. https://doi.org/10.17795/iji-37727

39. Zhang T, Luo Q, Chen Y, Li T, Wen G, Zhang R, et al. Molecular epidemiology, virulence determinants and antimicrobial resistance of *Campylobacter* spreading in retail chicken meat in Central China. Gut Pathog. 2016;8, Art. #48, 9 pages. https://doi.org/10.1186/s13099-016-0132-2

Agricultural practices and their potential role in mycotoxin contamination of maize and groundnut subsistence farming

AUTHORS:
Sylvia Phokane[1,2]
Bradley C. Flett[1,3]
Edson Ncube[1]
John P. Rheeder[4]
Lindy J. Rose[2]

AFFILIATIONS:
[1]Grain Crops Institute, Agricultural Research Council, Potchefstroom, South Africa
[2]Department of Plant Pathology, Stellenbosch University, Stellenbosch, South Africa
[3]Unit for Environmental Sciences and Management, North-West University, Potchefstroom, South Africa
[4]Institute of Biomedical and Microbial Biotechnology, Cape Peninsula University of Technology, Cape Town, South Africa

CORRESPONDENCE TO:
Edson Ncube

EMAIL:
NcubeE@arc.agric.za

DATES:
Received: 04 Apr. 2019
Revised: 13 June 2019
Accepted: 15 July 2019
Published: 26 Sep. 2019

HOW TO CITE:
Phokane S, Flett BC, Ncube E, Rheeder JP, Rose LJ. Agricultural practices and their potential role in mycotoxin contamination of maize and groundnut subsistence farming. S Afr J Sci. 2019;115(9/10), Art. #6221, 6 pages. https://doi.org/10.17159/sajs.2019/6221

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:
mycotoxigenic fungi, storage, mycotoxins, survey, South Africa

FUNDING:
National Research Foundation (South Africa); The Maize Trust (South Africa); Agricultural Research Council (South Africa)

Mycotoxigenic fungi are common pathogens of maize and groundnuts; they produce mycotoxins which reduce the yield and quality of these grain crops. Numerous agricultural practices including crop rotation and storage methods have been shown to impact mycotoxin accumulation. Therefore, the farming and storage practices in maize and groundnut subsistence farming systems in Pongola, Vryheid, Jozini, Manguzi and Mbazwana Districts of northern KwaZulu-Natal (South Africa) were surveyed to determine their potential role in promoting or mitigating mycotoxin contamination. A questionnaire about agricultural farming practices and storage facilities was presented to 65 subsistence maize and/or groundnut farmers. At least 90% of the farmers surveyed were not aware of mycotoxins and their consequences to animal and human health. The majority of the farmers did not practise crop rotation. However, they practised intercropping and sorted damaged and mouldy grain (maize and groundnuts) before storage. The damaged or mouldy grain was largely used as animal feed, thereby exposing animals to an increased risk of mycotoxicoses. Metal tanks and *inqolobane* (a type of wooden structure) were identified as the most common storage structures. Harvested homegrown maize was mostly used for the farmers' own consumption but also sometimes sold to the local community. The implementation of mycotoxin awareness campaigns is necessary, particularly in these districts. The storage facilities used by the subsistence farmers allowed increased moisture and insect invasion. The need for the surveillance of mycotoxins in subsistence-farmed food crops is vital.

Significance:

- The main finding of this study is the extent of post-harvest losses and mycotoxin contamination of maize produced by smallholder farmers in South Africa.
- We further identify methods to manage the risk of mycotoxin exposure to smallholder farmers and their communities as well as reduce post-harvest losses.

Introduction

Maize (*Zea mays* L.) and groundnuts (*Arachis hypogaea* L.) are produced by subsistence farmers, particularly in the northern KwaZulu-Natal Province of South Africa.[1,2] Maize is an important staple food and groundnuts serve as a protein and fat supplement for subsistence farmers.[3,4] Both maize and groundnut may be contaminated with mycotoxins, produced by fungi, prior to and after harvesting.[5-7] Mycotoxin contamination follows infection by mycotoxigenic fungi, of which the most common are *Fusarium* and *Aspergillus* species[1,2,8] that can contaminate maize and groundnut with fumonisins and aflatoxins, respectively. Ingestion of mycotoxin-contaminated food and feed can cause mycotoxicoses in humans and animals.[9,10] Fumonisins have been associated with a high incidence of oesophageal cancer in rural areas in South Africa due to the preference for mouldy kernels to produce traditional *umqombothi* beer.[11] Mycotoxicoses may also develop in cattle that consume contaminated feed.[12,13] During 2011, an estimated 100 dogs died in South Africa's Gauteng Province due to the ingestion of aflatoxin-contaminated feed.[14]

Agricultural practices such as crop rotation, irrigation, early planting and use of transgenic hybrids are employed by commercial farmers to reduce mycotoxin contamination of crops.[15] Moreover, some subsistence farmers in Tanzania and Zimbabwe recently applied these agricultural practices and a reduction in mycotoxin contamination was reported.[16,17] Hand sorting of maize before storage was also reported as a good measure to reduce fungal infection and subsequent mycotoxin contamination at storage.[18,19] Unlike in a commercial setting, many subsistence farmers do not apply these agronomic practices, concentrating only on sorting their grain after harvest into visually healthy and mouldy grain.[20,21] In areas in the Eastern Cape and Limpopo Provinces, mouldy grain is not discarded but used for traditional beer, thereby posing a risk of mycotoxin contamination.[3,22]

Limited information exists on storage of these crops by subsistence farmers and the associated mycotoxin risks. Storage of improperly dried grain, accompanied by high temperatures, causes rapid proliferation of mycotoxigenic fungi which results in reduced quality, nutrition and dry matter and higher mycotoxin levels.[23-25] Contamination at storage by fungi can occur when the moisture content is above 13% and temperatures are between 10 °C and 40 °C.[26] Therefore, the use of ventilated storage systems to reduce mycotoxin contamination is recommended, together with appropriate post-harvest control technologies to minimise mycotoxin contamination in the food chain.[27-29] For example, storage in moisture-free, dry wooden pallets, ventilated drying on polythene sheets and hand sorting led to a decrease in aflatoxin contamination of kernels at storage.[30] Storage facilities are some of the control points that have to be re-evaluated in the value chain; good storage facilities will lead to good marketable agricultural products.

Subsistence farmers incur economic losses due to pre-harvest and post-harvest contamination of grain crops caused by fungal species and insect pests. The present work continues earlier studies[1,2] that identified hotspots for fumonisin and aflatoxin contamination. Good farming practices and proper pre-harvest handling of maize and groundnuts, together with good storage practices, have been demonstrated to minimise the risk of fungal contamination. Hence, we aimed to identify pre- and post-harvest practices that could potentially contribute to mycotoxin contamination of maize and groundnuts produced in KwaZulu-Natal.

Materials and methods

Geographical areas surveyed

Agricultural extension officers from the South African Department of Agriculture and Rural Development assisted with the selection of five districts in northern KwaZulu-Natal and identification of households within districts where maize and groundnuts were planted. Global Positioning System (GPS) was used to detect and mark different localities within the districts. Subsistence farmers growing maize or groundnuts were interviewed in all five districts: Jozini ($n=7$), Manguzi ($n=17$), Mbazwana ($n=13$), Pongola ($n=17$) and Vryheid ($n=11$). All farmers in all five districts planted maize and all farmers in Jozini, Manguzi and Mbazwana planted groundnuts as well. No farmer in Vryheid planted groundnuts and there was only one identified groundnut farmer in Pongola.

Questionnaires

The agricultural farming practices, storage facilities and grain consumption for each farmer were determined through a survey. Questionnaires were drafted in English and translated into isiZulu, the predominant local language. These questionnaires were approved by the South African Medical Research Council. Both closed- and open-ended questions were asked randomly to ensure adequacy of the questionnaire. The questions were on the awareness of mycotoxins, crop rotation, intercropping, residue removal, sorting of damaged and mouldy grain, end result of the sorted grain, types of storage facilities, consumption and trading of homegrown maize and groundnut. An awareness of mycotoxins, nematodes and fungal pathogens was determined as well as whether participants were aware of negative health implications caused by fungal pathogens. Additional explanations of questions were provided when needed and included non-scientific descriptions such as mould growth for fungal infection and mycotoxin contamination. All the farmers were also informed that fungal infection may be associated with mycotoxin contamination.

Interviews

Before the interviews, the farmers were informed about the significance of the survey. The first author interviewed each farmer according to the questions stated on the questionnaire. Gathering of information was done in collaboration with local extension officers. An opportunity was granted for questions after the interviews and appropriate management strategies were discussed with the farmers and local extension officers.

Statistical analyses

The data obtained from the questionnaires were analysed using a chi-square test for independence. One-way analysis of variance (ANOVA) was used to test only the numerical entries. The significance level for both tests was set at a 95% confidence level with $p<0.05$ indicating a significant difference. The tested null hypothesis (H_0) for the chi-square test was that the factor evaluated is independent of the different districts surveyed. Conversely, the alternative hypothesis (H_a) was that the factor is dependent on the different districts surveyed. The null hypothesis was accepted if $p>0.05$ and rejected if $p<0.05$.[31]

Results

Mycotoxin awareness

None of the farmers in Jozini, Manguzi and Mbazwana were aware of mycotoxins (Figure 1).

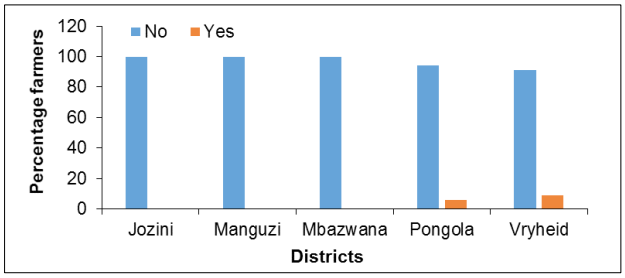


Figure 1: Mycotoxin awareness of subsistence farmers in five districts of northern KwaZulu-Natal (chi-square: d.f.=4; $p=0.17766$).

Only 6% and 9% of farmers in Pongola and Vryheid, respectively, had an idea of what mycotoxins could be, but did not know the cause of these mycotoxins and their implications on animal and human health (Figure 1). Mycotoxin awareness and maize districts were therefore independent ($p=0.1766$).

Residue removal

Of the farmers in Jozini, 43% removed residues from the soil before planting their groundnuts, 54% of the farmers in Manguzi did so, while all of the farmers in Pongola but none of the farmers in Mbazwana removed crop residues before planting groundnuts (Figure 2). Residue removal and groundnut districts were therefore independent ($p=1.769$)

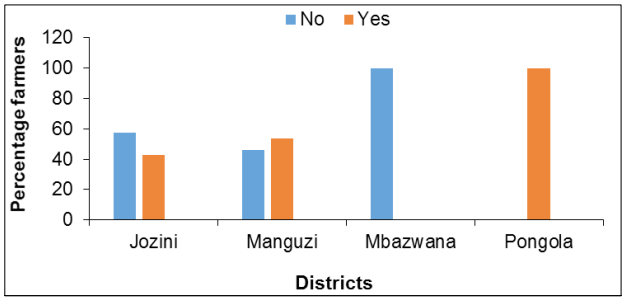


Figure 2: Residue removal before planting groundnuts by subsistence farmers in four districts of northern KwaZulu-Natal (chi-square: d.f.=3; $p=1.769$).

Crop rotation and intercropping

The majority of groundnut farmers did not practise crop rotation: 71%, 92% and 100% in Jozini, Manguzi and Mbazwana, respectively, did not rotate their groundnuts with any other crop (Figure 3). Only farmers in Pongola (100%) practised crop rotation (Figure 3) and therefore crop rotation and groundnut-farming districts were dependent on each other ($p=0.0122$). Farmers in all districts did not rotate maize with other crops (data not shown), but a variety of crops including beans, groundnuts and pumpkins were intercropped with maize. Maize was widely intercropped with groundnut in the Manguzi and Mbazwana Districts by 53% and 92% of farmers, respectively. Some farmers in all surveyed maize districts only planted maize (data not shown). Intercropping and the districts in which maize farmers were surveyed were, therefore, dependent on each other ($p<0.001$) (data not shown). Only farmers in the Pongola District did not intercrop groundnuts with other crops, whereas farmers in the other districts intercropped with crops such as spinach (*Spinacia oleracea* L.) and cowpeas (*Vigna unguiculata* L.) (data not shown). Therefore, intercropping was dependent on the groundnut-farming districts surveyed ($p=0.0071$) (data not shown).

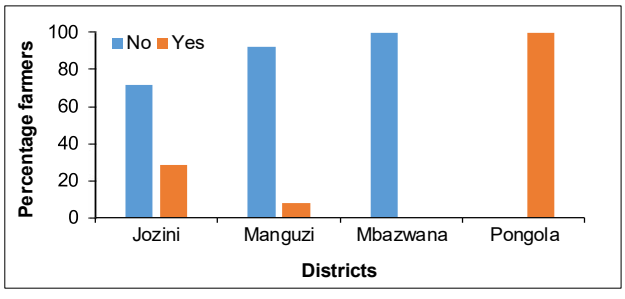


Figure 3: Rotation of groundnuts with other crops by subsistence farmers in four districts of northern KwaZulu-Natal (chi-square: d.f.=4; $p=0.0122$).

Grain sorting before storage

All the maize farmers in all districts surveyed sorted their maize into apparently healthy, mouldy and damaged maize before storage (results not shown). All the groundnut farmers in Jozini and Manguzi and 10% in Mbazwana also sorted their groundnuts into apparently healthy, mouldy

and damaged groundnuts before storage (Figure 4). Sorting and groundnut districts surveyed are therefore independent variables ($p=0.610$).

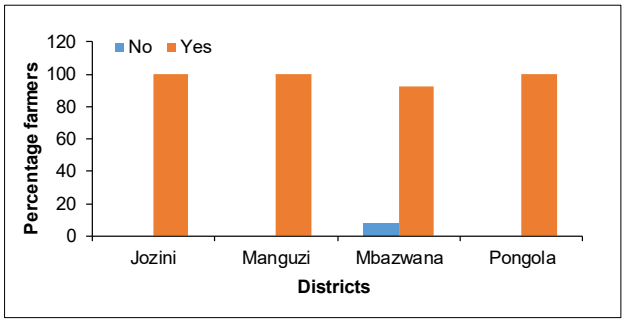


Figure 4: Sorting of damaged and mouldy groundnuts by subsistence farmers in four districts of northern KwaZulu-Natal (chi-square: d.f.=3; $p=0.610$).

End result of mouldy and damaged grain

All the farmers in Jozini fed the mouldy and damaged maize kernels to chickens (*Gallus gallus domesticus*) only. Some farmers in the other four districts used the mouldy and damaged maize as chicken feed, but also discarded the grain. Additionally, 59% of farmers in Pongola and 55% of farmers in Vryheid fed the mouldy and damaged grain to other domestic animals such as pigs (*Sus domesticus*), cattle (*Bos taurus*) and goats (*Capra aegagrus hircus*). Furthermore, 18%, 8% and 9% of the farmers in Manguzi, Mbazwana and Vryheid, respectively, consumed the mouldy and damaged maize (Figure 5). The end-users of mouldy and damaged maize kernels and the maize districts surveyed were, therefore, dependent ($p=0.0009$). For groundnuts, all the farmers in Pongola and some farmers in other districts fed the mouldy and damaged groundnuts to chickens only. Less than 30% of farmers in Manguzi and Mbazwana discarded the mouldy and damaged groundnuts. In contrast with maize farmers, more groundnut farmers in Manguzi (50%) consumed the mouldy and damaged groundnuts. Also, 60% of groundnut farmers in Jozini consumed mouldy and damaged groundnuts (data not shown). The end-users of mouldy and damaged groundnuts and groundnut districts surveyed were also dependent variables ($p=0.0396$) (data not shown).

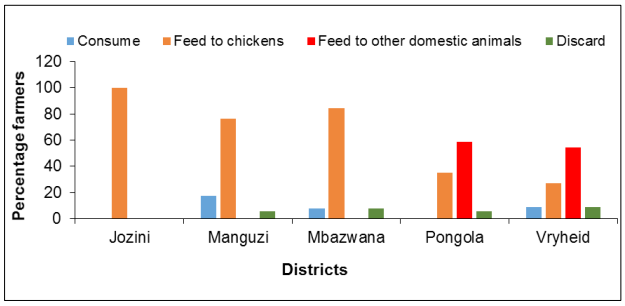


Figure 5: End result of damaged and mouldy maize produced by subsistence farmers in five districts of northern KwaZulu-Natal (chi-square: d.f.=12; $p=0.0009$).

Storage facilities

A type of storage facility widely used in all the northern KwaZulu-Natal districts surveyed was an *inqolobane*, which is the isiZulu name for a widely ventilated wooden storage facility (Figure 6a). Metal drums were used by maize farmers only; some metal drums were ventilated and others unventilated (Figure 6b). Groundnuts were planted in small quantities in comparison to maize, hence groundnuts were easily and most commonly placed in bags which were stored in the farmers' homes (Figure 6c). In fact, in all the groundnut-farming districts (Jozini, Manguzi and Mbazwana), farmers stored groundnuts in their homes only (data not shown). Metal tanks were used to store maize by some subsistence farmers in Jozini, Pongola and Vryheid (Figure 7). The storage facilities and maize districts surveyed were dependent variables ($p=0.0014$).

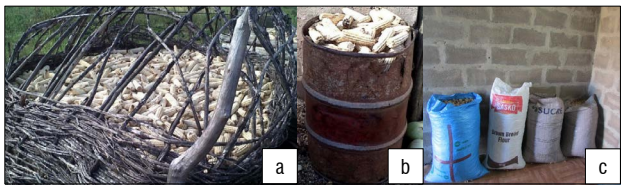


Figure 6: Common storage facilities for maize utilised by subsistence farmers in districts of northern KwaZulu-Natal: (a) *inqolobane*, (b) metal drum and (c) groundnut bags.

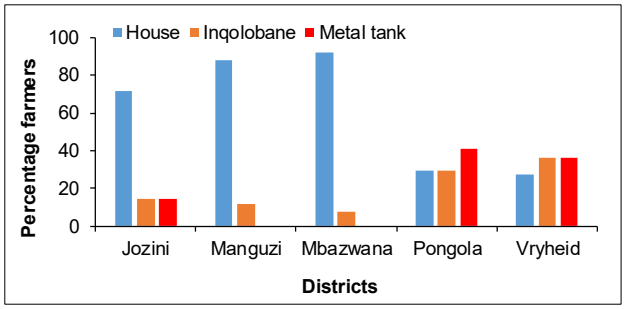


Figure 7: Storage facilities utilised by subsistence farmers in five districts of northern KwaZulu-Natal (chi-square: d.f.=8; $p=0.0014$).

Consumption and trading of grain

Farmers from all districts either only consumed or both consumed and sold their homegrown maize (Figure 8). Consumption with trading of homegrown maize and maize-farming districts were independent ($p=0.1766$). Half the farmers in Mbazwana only consumed their homegrown groundnuts and the other half both sold and consumed their homegrown groundnuts. Over 60% of farmers in Jozini and Manguzi only consumed their homegrown groundnuts (data not shown). Consumption with trading of homegrown groundnuts and groundnut-farming districts were also independent ($p=0.635$) (data not shown). All the farmers in Jozini and Manguzi only sold their homegrown maize to the local community; farmers in Mbazwana, Pongola and Vryheid also sold their homegrown maize to the nearest markets (data not shown). Maize trading areas and maize districts were dependent on each other ($p=0.0046$) (data not shown).

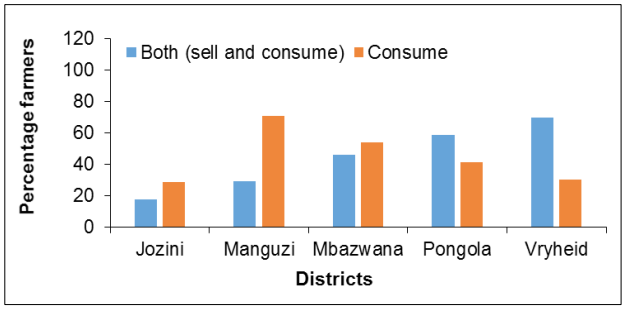


Figure 8: Consumption and trading of harvested homegrown maize by subsistence farmers in five districts of northern KwaZulu-Natal (chi-square: d.f.=4; $p=0.1766$).

Discussion

Numerous crop production and post-harvest practices have been found to influence mycotoxin accumulation in grain crops. In this study, the majority of groundnut farmers and all the maize farmers surveyed did not practise crop rotation. Furthermore, nearly half of the groundnut farmers did not remove plant residues before planting. Crop rotation can help

reduce available inoculum for subsequent infection when non-host crops are employed.[32] In a recent study, conservation agriculture – commonly described as practices that maintain permanent soil cover (no removal of plant residues) and minimum soil disturbance – did not increase the risk of maize ear rots and mycotoxin production.[33] The storage facilities used by both maize and groundnut farmers favour fungal entry which increases the risk for mycotoxin contamination. Improving maize and groundnut subsistence farming and grain storage is crucial in mitigating the risk of mycotoxin contamination within the particular communities surveyed. Good-quality maize-based and groundnut-based products are not only necessary for consumption but also for trade. Hence, it was important to conduct a survey on the current farming practices in order to determine which potentially contribute to increased risk of mycotoxin contamination. This information will help to determine possible intervention strategies that could cause a reduction in the risk of mycotoxin contamination.

The lack of mycotoxin awareness in these districts indicates that humans and livestock may be consuming mycotoxin-contaminated maize and groundnuts daily which places them at a high health risk. Incidentally, some agricultural practices used by subsistence farmers, such as sorting of damaged and mouldy grain from storage, may have assisted in limiting mycotoxin exposure. Crop residues also harbour mycotoxigenic fungi,[34] hence it is vital to remove crop residues before planting so that they do not serve as an inoculum source. Destruction or removal of infected crop residues from the field has been found to reduce fungal inoculum.[35]

The practice of rotating maize and groundnuts with other crops may be associated with the variation in soil types of the districts surveyed as this directly determines the crops that can be successfully cultivated. For example, the Manguzi and Mbazwana Districts had sandy soil types which mostly favour the cultivation of groundnuts over maize. Light-textured soils which include deep, well-drained sandy and loamy sand soils at a pH between 5.3 and 7.3 favour significant groundnut yields.[36] The majority of farmers do not employ crop rotation, possibly because of a lack of knowledge of the advantages. Pest and disease cycles are broken by crop rotation, thereby reducing fungal infestation and subsequent mycotoxin contamination in the field.[37] Farmers prefer to grow the same crop throughout, especially when it can be sustainably produced under prevailing conditions. However, rotating crops potentially increases crop yield and the root system health is maintained by the reduced inoculum potential of soil-borne pathogens.[38] Intercropping has also been shown to reduce contamination of maize with mycotoxins.[39]

The manner in which farmers sorted groundnuts was determined by the quantity of groundnuts harvested and/or whether this would be kept for household consumption or sold for additional income. Mycotoxin contamination was reduced in the former Transkei region by sorting damaged/mouldy grain from apparently healthy grain.[40] This study reported that fumonisin concentration decreased by 71% after removing highly infected maize kernels. Also, washing and sorting of maize kernels was found to reduce fumonisin contamination by 84%.[8] In the Rombo District of Tanzania the sorting of maize also led to a reduction in fumonisin contamination.[18] Therefore, it is good practice that the majority of the farmers in the northern KwaZulu-Natal sort their maize and groundnut to decrease contamination at storage. Mouldy and damaged maize was used to feed domestic livestock while most farmers across all districts fed the mouldy and damaged maize to chickens. Mycotoxin-contaminated feed generally affects the growth of chickens.[41]

Farmer preference dictated the use of specific storage facilities in the different districts. The choice of a storage facility may be due to problems experienced at storage relating to the different districts; for example, the use of tanks and drums to prevent mice damage specifically. Storage facilities used by farmers in the surveyed districts in northern KwaZulu-Natal are the same as those used by other farmers in sub-Saharan African countries[30] and some of these storage facilities do not promote proper drying of maize and thus enhance interaction with insects, thereby promoting fungal infection and mycotoxin production[30]. The application of a pesticide to control stored-maize insect pests was proved to be an ineffective method compared to other post-harvest methods.[42]

Most farmers use wooden granaries for storage; these structures are widely used, possibly because of the ease of construction and for drying maize ears. However, this structure allows invasion by insect pests and rodents as it is not covered on top. Insects damage maize ears during feeding, thereby facilitating fungal invasion and infection.[43] Therefore, maize cannot be stored for prolonged periods under such conditions. Farmers could be advised to use metal silos[44,45] and hermetic storage containers[46]; these storage structures are airtight and, therefore, prevent any pathogen or pest from invading the stored maize[42]. Subsistence farmers prefer the traditional storage systems as they are cheaper to construct and maintain, although their use can cause high post-harvest losses.[45] The specific storage practices employed were dictated by the quantity of maize produced. For instance, in high maize production areas such as Vryheid and Pongola, maize was predominantly stored in tanks.

Subsistence farmers consume high quantities of homegrown maize, as much as 300 g per person per day,[47] and also sell the homegrown maize and groundnuts to the local community. Hence their exposure to mycotoxins is potentially higher than that of consumers in cities and towns. Furthermore, subsistence farmers have to contend with supermarkets present in local communities and small towns, which sell their good-quality products, especially maize meal and bread, at reduced costs.[48] Also, pressure is placed on subsistence farmers to produce safe and healthy food due to new regulations for deoxynivalenol and fumonisin B_1 and B_2 limits in maize. The South African government implemented new regulations, setting maximum levels of 2000 μg/kg for deoxynivalenol and 4000 μg/kg for fumonisin B_1 and B_2.[49] Subsistence farmers were not aware of these regulations. The monitoring of these regulations in an informal environment is unclear and possibly impractical; however, the supply chain will need to be regulated for quality and safety[10] considering the potential for trade between subsistence farmers. Therefore, there is a need to determine the extent of mycotoxin contamination of these crops. Additionally, limited information is available on control methods to reduce the risk of mycotoxin contamination of food crops.

Conclusion

Mycotoxin contamination of maize and groundnuts produced through subsistence farming systems can be reduced by following good agricultural farming and storage practices such as crop rotation and sorting before storage, respectively, thus, improving the health and economic status of subsistence farmers and the communities involved. The implementation of good farming practices can be effortless; however, access to adequate storage facilities may not be feasible. Therefore, support in this regard is of utmost importance in subsistence farming. Furthermore, it is vital, to minimise mycotoxin contamination, that knowledge of good agricultural practices be transferred to subsistence farmers as well as agricultural extension officers. This knowledge transfer can form part of mycotoxin awareness campaigns to inform farmers of the threats and effects of mycotoxins on humans and animals. Additional surveillance is required to continuously monitor and advise on mycotoxin contamination and potential exposure in subsistence farming.

Acknowledgements

We acknowledge the financial support received from The Maize Trust, National Research Foundation of South Africa (Thuthuka grant no. 84162) and the Agricultural Research Council of South Africa. We thank the Department of Agriculture and Environmental Affairs in the KwaZulu-Natal Province for their collaboration and support provided; Ms Gugu Khali and Ms Yvonne Maila for assisting with sample collection in all the districts; and Ms Nicolene Thiebaut for statistical analysis of the results.

Authors' contributions

S.P.: Research design; fieldwork and sample collection; laboratory analysis of samples; data analysis; and writing article drafts. B.C.F.: Research design; research supervision; reviewing article drafts. E.N.: Obtaining funding; research design; research supervision; reviewing article drafts. J.P.R.: Obtaining funding; research design; reviewing article drafts. L.J.R.: Research design; research supervision; reviewing article drafts.

References

1. Ncube E, Flett BC, Waalwijk C, Viljoen A. *Fusarium* spp. and levels of fumonisins in maize produced by subsistence farmers in South Africa. S Afr J Sci. 2011;107:33–39. http://dx.doi.org/10.4102/sajs.v107i1/2.367
2. Ncube E, Flett BC, Waalwijk C, Viljoen A. Occurrence of aflatoxins and aflatoxin-producing *Aspergillus* spp. associated with groundnut production in subsistence farming systems in South Africa. S Afr J Plant Soil. 2010;27:195–198.
3. Shephard GS, Burger H-M, Gambacorta L, Krska R, Powers SP, Rheeder JP, et al. Mycological analysis and multimycotoxins in maize from rural subsistence farmers in the former Transkei, South Africa. J Agric Food Chem. 2013;61:8232–8240. http://dx.doi.org/10.1021/jf4021762
4. Sarvamangala C, Gowda MVC, Varshney RK. Identification of quantitative trait loci for protein content, oil content and oil quality for groundnut (*Arachis hypogaea* L.). Field Crop Res. 2011;122:49–59. http://dx.doi.org/10.1016/j.fcr.2011.02.010
5. Beukes I, Rose LJ, Shephard GS, Flett BC, Viljoen A. Mycotoxigenic *Fusarium* species associated with grain crops in South Africa – A review. S Afr J Sci. 2017;113, Art. #2016-0121, 12 pages. http://dx.doi.org/10.17159/sajs.2017/20160121
6. Njoroge SMC, Matumba L, Kanenga K, Siambi M, Waliyar F, Maruwo J, et al. Aflatoxin B_1 levels in groundnut products from local markets in Zambia. Mycotoxin Res. 2017;33(2):113–119. http://dx.doi.org/10.1007/s12550-017-0270-5
7. Mupunga I, Lebelo SL, Mngqawa P, Rheeder JP, Katerere DR. Natural occurrence of aflatoxins in peanuts and peanut butter from Bulawayo, Zimbabwe. J Food Prot. 2014;77:1814–1818. http://dx.doi.org/10.4315/0362-028X.JFP-14-129
8. Van der Westhuizen L, Shephard GS, Rheeder JP, Burger H-M, Gelderblom WCA, Wild CP, et al. Simple intervention method to reduce fumonisin exposure in a subsistence maize-farming community in South Africa. Food Add Contam. 2010;27:1582–1588. http://dx.doi.org/10.1080/19440049.2010.508050
9. Gelderblom WCA, Jaskiewicz R, Marasas WFO, Thiel PG, Horak RM, Vleggaar R, et al. Fumonisins-novel mycotoxins with cancer-promoting activity produced by *Fusarium moniliforme*. Appl Environ Microbiol. 1988;54:1806–1811.
10. Stoev SD. Food safety and increasing hazard of mycotoxin occurrence in foods and feeds. Crit Rev Food Sci Nutr. 2013;53:887–901. http://dx.doi.org/10.1080/10408398.2011.571800
11. Shephard GS, Van der Westhuizen L, Gatveni PM, Somdyala, NM, Vismer HF, Marasas WFO. Fumonisin mycotoxins in traditional Xhosa maize beer in South Africa. J Agric Food Chem. 2005;53:9634–9637. http://dx.doi.org/10.1021/jf0516080
12. Njobeh PB, Dutton MF, Åberg AT, Haggblom P. Estimation of multi-mycotoxin contamination in South African compound feeds. Toxins. 2012;4:836–848. http://dx.doi.org/10.3390/toxins4100836
13. Marasas W, Gelderblom W, Shephard G, Vismer H. Mycotoxicological research in South Africa 1910-2011. World Mycotoxin J. 2011;5:89–102. http://dx.doi.org/10.3920/WMJ2011.1322
14. Arnot LF, Duncan NM, Coetzer H, Botha CJ. An outbreak of canine aflatoxicosis in Gauteng province, South Africa. J S Afr Vet Assoc. 2012;83:2–4. http://dx.doi.org/10.4102/jsava.v83i1.2
15. Bruns HA. Controlling aflatoxin and fumonisin in maize by crop management. J Toxicol. 2003;22:153–173. http://dx.doi.org/10.1081/TXR-120024090
16. Nyangi C, Beed F, Mugula JK, Boni S, Koyano E, Mahuku G, et al. Assessment of pre-harvest aflatoxin and fumonisin contamination of maize in Babati District, Tanzania. Afr J Food Agric Nutr Dev. 2016;16:11039–11053. http://dx.doi.org/10.18697/ajfand.75.ILRI06
17. Ndemera M, Landschoot S, De Broeve M, Nyanga LK, De Saeger S. Effect of agronomic practices and weather conditions on mycotoxins in maize: A case study of subsistence farming households in Zimbabwe. World Mycotoxin J. 2018;11:421–436. http://dx.doi.org/10.3920/WMJ2017.2227
18. Kimanya ME, De Meulenaer B, Tiisekwa B, Ugullum C, Devlieghere F, Van Camp J, et al. Fumonisins exposure from freshly harvested and stored maize and its relationship with traditional agronomic practices in Rombo district, Tanzania. Food Add Contam A. 2009;26:1199–1208. http://dx.doi.org/10.1080/02652030902922784
19. Odongo N. Sorting is an affordable technology that can reduce mycotoxin contamination to safe levels. Afr J Food Agric Nutr Dev. 2016;16:1–2.
20. Van der Westhuizen L, Shephard GS, Rheeder JP, Burger H-M, Gelderblom WCA, Wild CP, et al. Optimising sorting and washing of home-grown maize to reduce fumonisin contamination under laboratory-controlled conditions. Food Control. 2011;22:396–400. http://dx.doi.org/10.1016/j.foodcont.2010.09.009
21. Matumba L, Van Poucke C, Ediage EN, Jacobs B, De Saeger S. Effectiveness of hand sorting, flotation/washing, dehulling and combinations thereof on the decontamination of mycotoxin contaminated white maize. Food Add Contam. 2015;32:960–969. http://dx.doi.org/10.1080/19440049.2015.1029535
22. Phoku JZ, Dutton MF, Njobeh PB, Mwanza M, Egbuta MA, Chilaka CA. *Fusarium* infection of maize and maize-based products and exposure of a rural population to fumonisin B1 in Limpopo Province, South Africa. Food Add Contam A. 2012;29:1743–1751. http://dx.doi.org/10.1080/19440049.2012.708671
23. Mylona K, Sulyok M, Magan N. Relationship between environmental factors, dry matter loss and mycotoxin levels in stored wheat and maize infected with *Fusarium* species. Food Add Contam. 2012;29:1118–1128. http://dx.doi.org/10.1080/19440049.2012.672340
24. Janse van Rensburg B, McLaren NW, Flett BC. Grain colonization by fumonisin-producing *Fusarium* spp. and fumonisin synthesis in South African commercial maize in relation to prevailing weather conditions. Crop Prot. 2017;102:129–136. http://dx.doi.org/10.1016/j.cropro.2017.08.019
25. Di Domenico AS, Christ D, Hashimoto EH, Busso C, Coelho SRM. Evaluation of quality attributes and the incidence of *Fusarium* sp. and *Aspergillus* sp. in different types of maize storage. J Stored Prod Res. 2015;61:59–64. http://dx.doi.org/10.1016/j.jspr.2014.12.001
26. Atanda SA, Pessu PO, Agoda S, Isong IU, Adekalu OA, Echendu MA, et al. Fungi and mycotoxins in stored foods. Afr J Microbiol Res. 2011;5:4373–4382.
27. Mutegi C, Wagacha M, Kimani J, Otieno G, Wanyama R, Hell K, et al. Incidence of aflatoxin in peanuts (*Arachis Hypogaea* Linnaeus) from markets in Western, Nyanza and Nairobi Provinces of Kenya and related market traits. J Stored Prod Res. 2013;52:118–127. http://dx.doi.org/10.1016/j.jspr.2012.10.002
28. Atukwase A, Kaaya AN, Muyanja C. Dynamics of *Fusarium* and fumonisins in maize during storage – A case study of the traditional storage structures commonly used in Uganda. Food Control. 2012;26:200–205. http://dx.doi.org/10.1016/j.foodcont.2012.01.016
29. Waliyar F, Osiru M, Ntare BR, Kumar KVK, Sudini H, Traore A, et al. Post-harvest management of aflatoxin contamination in groundnut. World Mycotoxin J. 2014;8:245–252. http://dx.doi.org/10.3920/WMJ2014.1766
30. Seetha A, Munthali W, Msere HW, Swai E, Muzanila Y, Sichone E, et al. Occurrence of aflatoxins and its management in diverse cropping systems of central Tanzania. Mycotoxin Res. 2017;33:323–331. http://dx.doi.org/10.1007/s12550-017-0286-x
31. Lombaard C, Van der Merwe L, Kele T, Mouton S. Elementary statistics for business and economics. Cape Town: Pearson Education South Africa; 2011. p. 57–361.
32. Marocco A, Gavazzi C, Pietri A, Tabaglio V. On fumonisin incidence in monoculture maize under no-till, conventional tillage and two nitrogen fertilisation levels. J Sci Food Agric. 2008;88:1217–1221. http://dx.doi.org/10.1002/jsfa.3205
33. Mabuza LM, Janse van Rensburg B, Flett BC, Rose LJ. Accumulation of toxigenic *Fusarium* species and *Stenocarpella maydis* in maize grain grown under different cropping systems. Eur J Plant Pathol. 2018;152:297–308. http://dx.doi.org/10.1007/s10658-018-1475-y
34. Munkvold GP. Cultural and genetic approaches to managing mycotoxins in maize. Annu Rev Phytopathol. 2003;41:99–116. http://dx.doi.org/10.1146/annurev.phyto.41.052002.095510
35. Wambacq E, Vanhoutte I, Audenaert K, De Gelder L, Haesaert G. Occurrence, prevention and remediation of toxigenic fungi and mycotoxins in silage: A review. J Sci Food Agric. 2016;96:2284–2302. http://dx.doi.org/10.1002/jsfa.7565
36. Okello DK, Kaaya AN, Bisikwa J, Were M, Oloka HK. Management of aflatoxins in groundnuts: A manual for farmers, processors, traders and consumers in Uganda. Entebbe: National Agricultural Research Organisation; 2010. p.1–38.

37. TerAvest D, Carpenter-Boggs L, Thierfelder C, Reganold JP. Crop production and soil water management in conservation agriculture, no-till, and conventional tillage systems in Malawi. Agric Ecosyst Environ. 2015;212:285–296. http://dx.doi.org/10.1016/j.agee.2015.07.011

38. Nel AA, Lamprecht SC. Crop rotational effects on irrigated winter and summer grain crops at Vaalharts. S Afr J Plant Soil. 2011;28:127–133. http://dx.doi.org/10.1080/02571862.2011.10640023

39. Van Asselt ED, Azambuja W, Moretti A, Kastelein P, De Rijk TC, Stratakou I, et al. A Dutch field survey on fungal infection and mycotoxin concentrations in maize. Food Add Contam A. 2012;29:1556–1565. http://dx.doi.org/10.1080/19440049.2012.689997

40. Mogensen JM, Sørensen SM, Sulyok M, Van der Westhuizen L, Shepherd GS, Frisvad JC, et al. Single-kernel analysis of fumonisins and other fungal metabolites in maize from South African subsistence farmers. Food Add Contam A. 2011;28:1724–1734. http://dx.doi.org/10.1080/19440049.2011.611823

41. Smith EE, Kubena LF, Braithwaite RB, Harvey RB, Phillips TD, Reine AH. Toxicological evaluation of aflatoxin and cyclopiazonic acid in broiler chickens. Poult Sci. 1992;71:1136–1144. http://dx.doi.org/10.3382/ps.0711136

42. Chigoverah AA, Mvumi BM. Efficacy of metal silos and hermetic bags against stored-maize insect pests under simulated smallholder farmer conditions. J Stored Prod Res. 2016;69:179–189. http://dx.doi.org/10.1016/j.jspr.2016.08.004

43. Dafoe NJ, Thomas JD, Shirk PD, Legaspi ME, Vaughan MM, Huffaker A, et al. European corn borer (*Ostrinia nubilalis*) induced responses enhance susceptibility in maize. PLoS ONE. 2013;8(9), e73394, 18 pages. https://doi.org/10.1371/journal.pone.0073394

44. Tefera T, Kanampiu F, De Groote H, Hellin J, Mugo S, Kimenju S, et al. The metal silo: An effective grain storage technology for reducing post-harvest insect and pathogen losses in maize while improving smallholder farmers' food security in developing countries. Crop Prot. 2011;30:240–245. http://dx.doi.org/10.1016/j.cropro.2010.11.015

45. Gitonga ZM, De Groete H, Kassie M, Tefera T. Impact of metal silos on households' maize storage, storage losses and food security: An application of a propensity score matching. Food Policy. 2013;43:44–55. http://dx.doi.org/10.1016/j.foodpol.2013.08.005

46. Murashiki TC, Chidewe C, Benhura MA, Manema LR, Mvumi BM, Nyanga LK. Effectiveness of hermetic storage technologies in limiting aflatoxin B_1 and fumonisin B_1 contamination of stored maize grain under smallholder conditions in Zimbabwe. World Mycotoxin J. 2018;11:459–469. http://dx.doi.org/10.3920/WMJ2017.2288

47. Shephard GS, Marasas WFO, Burger H-M, Somdyala NIM, Rheeder JP, Van der Westhuizen L, et al. Exposure assessment for fumonisins in the former Transkei region of South Africa. Food Add Contam. 2007;24:62–1629. http://dx.doi.org/10.1080/02652030601101136

48. D'Haese M, Van Huylenbroeck G. The rise of supermarkets and changing expenditure patterns of poor rural households case study in the Transkei area, South Africa. Food Policy. 2005;30:97–113. https://dx.doi.org/10.1016/j.foodpol.2005.01.001

49. South African Department of Health. Foodstuffs, Cosmetics and Disinfectants Act, 1972 (Act 54 of 1972) – Regulations Governing Tolerances for Fungus-produced Toxins in Foodstuffs: Amendment. Government Gazette no. 40250 [document on the Internet]. 05 September 2016 [cited 2017 Feb 08]. Available from: https://www.gov.za/sites/default/files/gcis_document/201609/40250gon987.pdf

AUTHORS:
Bradley Robinson[1]
Lee Pote[1]
Candice Christie[1]

AFFILIATION:
[1]Department of Human Kinetics and Ergonomics, Rhodes University, Makhanda, South Africa

CORRESPONDENCE TO:
Lee Pote

EMAIL:
l.pote86@gmail.com

DATES:
Received: 28 Nov. 2018
Revised: 25 Apr. 2019
Accepted: 06 May 2019
Published: 26 Sep. 2019

HOW TO CITE:
Robinson B, Pote L, Christie C. Strength and conditioning practices of high school rugby coaches: A South African context. S Afr J Sci. 2019;115(9/10), Art. #5837, 6 pages. https://doi.org/10.17159/sajs.2019/5837

ARTICLE INCLUDES:
☒ Peer review
☒ Supplementary material

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

EDITOR:
John Butler-Adam

KEYWORDS:
training, sport, survey

FUNDING:
None

Strength and conditioning practices of high school rugby coaches: A South African context

Although rugby union as a sport is well established, the strength and conditioning practices of high school level players are not well known. Therefore, the main purpose of this study was to examine the current strength and conditioning practices that coaches implement at South African high school level rugby. A secondary purpose was to compare practices between high schools of different socio-economic status in South Africa. An online survey or in person interview (depending on the school) was conducted and 43 responses were received: from 28 coaches at schools among the top 100 rugby schools in South Africa for 2016 and from 15 coaches at no-fee public schools in the Eastern Cape Province of South Africa. Most coaches (72%) performed some form of physical testing, with the most common parameter tested being 'speed'. The most common strength and conditioning practices utilised included flexibility (stretching), speed (free sprinting), agility (cone drills), plyometric (box drills) and resistance (weight lifting) training. Unlike the no-fee schools, the top 100 rugby schools implemented conditioning practices similar to best-known international practices. Furthermore, no-fee school coaches did not have the qualifications necessary to administer the correct training techniques. Education and upskilling on the best strength and conditioning practices for school level coaches need to be improved, particularly in less privileged schools. Such improvement is crucial to the transformation goals set out by the South African Rugby Union, which would benefit from player development in lower socio-economic schools.

Significance:

- The top 100 rugby playing schools in the country implement similar strength and conditioning practices to the best known international practices.
- Education of coaches is key to the implementation of correct strength and conditioning practices, specifically at underprivileged schools.

Introduction

Rugby union (hereafter referred to as rugby) is a popular sport with over 8.5 million players around the globe. Because of the increasing popularity and high profile of elite rugby, research on the sport has increased with two key objectives: optimising performance[1-5] and reducing the risk of injury[6-8]. As a consequence, there is increasing importance on developing the physical characteristics of modern players through the principles of strength and conditioning.[9,10] Physical characteristics such as improved strength, power, speed, cardiovascular fitness and body composition can contribute towards the success of a player and a team.[11] Improving these attributes has thus become a crucial part of training and development; understanding the techniques for an effective physical conditioning programme can give players and teams a competitive edge.[12] Furthermore, rugby has a high incidence of injury due to the physical nature of the game[8] and in South Africa it presents an above-average overall injury risk (69 injuries in 1000 hours exposure) compared to other popular sports such as cricket (2 injuries in 1000 hours exposure)[13]. The physiological and morphological adaptations of strength and conditioning do not only benefit rugby performance, but also reduce the risk of injury and are thus important to consider within the training paradigm.[6,8] South African high school rugby is fiercely competitive and reflects the modern professionalism of the sport at an elite level.[7] Sponsorships, televised matches and bursaries have all increased rugby competition in high schools and schools have made major investments to improve their results.[14] Such investments include hiring professional coaches, with strength and conditioning specialists becoming more frequent in top rugby-playing schools. The expertise provided by such specialists is paramount in developing young players into competitive athletes.

Despite the increasing popularity of rugby in South Africa, the sport is plagued with political interference and racial division.[15] Interventions such as racial quotas have been introduced to reverse the apartheid policies of racial division and systemic disadvantage.[16] However, despite such interventions, the previously disadvantaged people of colour, who represent 90% of the South African population, remain the minority in the elite level rugby teams.[15] There is a distinct lack of black players who fulfil their potential and reach an elite level of rugby. Development of these players needs to start from grassroots level.[17] However, the South African schooling system is steeped in inequalities that favour the minority, and these inequalities are reflected in rugby structures.[10,18] The majority of school learners, predominantly black learners, face unfavourable socio-economic circumstances at home and have to strive to overcome similar circumstances in school as well.[10,18] Schools that cater for poverty-struck communities also struggle for educators, resources and facilities.[19] As a result, these schools lack the structures and knowledge to develop their learners into competitive athletes who could pursue a sport as a career. Combining the principles of strength and conditioning with those of nutrition could assist these players to achieve the physical development required to reach an elite level of rugby performance.[20] However, the extent to which strength and conditioning exercises have been adopted in a South African high school context is yet to be investigated. If effective transformation is to be achieved, the nation cannot rely on the few top rugby playing schools in the country to provide the players, especially as these teams are represented pre-dominantly by white players.[10] There is much to be learned from the success of the top rugby playing schools. Their expertise can lead adolescent development across the country and the practices of these schools can form a fundamental starting point from which research can expand.

The purpose of this study was therefore twofold: the primary aim was to describe the current strength and conditioning practices utilised by coaches at a high school level and a secondary aim was to compare these practices to those of a smaller sub-sample of rural school coaches.

Methods

The study was a cross-sectional descriptive study which required participation in an online questionnaire. The online questionnaire was created using www.kwiksurvey.com and is available as Appendix 1 in the Supplementary material. The questionnaire was designed following a narrative literature review of previous strength and conditioning surveys specific to intermittent and contact sports similar to rugby. Questionnaires examining similar issues were given specific attention, particularly those regarding high school strength and conditioning practices and rugby-related research.[18,21-24] However, due to the limited literature available, particularly that in regard to rugby-specific questions, newly constructed questions were verified for construct validity through pilot testing with local strength and conditioning specialists as well as sports scientists.

The survey itself consisted of four sections:

1. General: This section surveyed the background information of the coaches, their degrees and qualifications obtained and their general knowledge of strength and conditioning principles.
2. Sports conditioning: This section focused on the coaches' use of physical testing, flexibility development, speed development, agility development, plyometric training, and strength and resistance training.
3. Rugby specific: This section dealt with rugby-specific strength and conditioning related questions that gave insight into position-specific practices.
4. General comments: This section allowed coaches an opportunity to include any information they might deem useful to the study.

The survey consisted of both quantitative and qualitative questions. Quantitative questions provided specific data that were transformed into useable statistics. These questions included a combination of nominal and interval levels of measurements that reflected the response frequencies. The qualitative questions revealed specific trends or opinions and ensured the participants were not limited in their answers.

Participants were the head strength and conditioning specialists or, in the case of no strength and conditioning specialist, coaches who were responsible for the strength and conditioning of their high school boy rugby teams. Participants were recruited from 12 private schools, 16 public schools and 15 no-fee schools. The private and public high school teams chosen were from the top 100 rugby teams in the country in 2016. Because the no-fee schools are of major interest in the growth and transformation of South African rugby, it was decided that these schools would be approached in the Eastern Cape region surrounding Makhanda (previously Grahamstown) where the researchers were situated and thus comprised a sub-sample of convenience. Most rural schools have no access to computers or the Internet so schools in close proximity were approached for pragmatic reasons. Ethical clearance was granted by the Department of Human Kinetics and Ergonomics Ethical Standards Committee for research involving human participants (Rhodes University). All participants were informed of the benefits and risks of the study before consenting to participate.

The survey was emailed to the strength and conditioning specialist or coach in charge of conditioning for the Under-19 First team of the selected school. Email addresses were obtained via the schools' websites. The email sent contained information on participation and a link to the online survey. It was explained in the email that recipients were under no obligation to complete the questionnaire and could decline or withdraw at any point. If a coach was unavailable via email, a face-to-face interview was arranged where possible. Because of the multilingual nature of the South African education system, it was expected that there would be some language barriers. Thus, face-to-face interviews were also deemed necessary in assisting with the understanding of certain concepts and questions. Participants were not influenced in any way during the interviews, but explanations were provided for terms where necessary.

The survey consisted of both fixed response (quantitative) questions and open-ended (qualitative) questions. The responses were exported by the survey host (KwikSurveys) to a Microsoft Excel spreadsheet. Quantitative data were analysed using the data analysis tool provided by KwikSurveys. Thematic analysis was used for qualitative data to identify patterns of meaning across the data sets that provide insight into the relevant question being addressed. This analysis was performed using the process of data familiarisation, data coding, theme development and revision. Patterns or themes were identified in accordance with the methods used in previous studies on strength and conditioning principles.[24] Braun and Clarke[25] identified a six-phase process for thematic analysis that was closely followed in the current study. Answers to the qualitative questions were collated and content analysed for specific major and minor themes. Categorical and ordinal data were reported as percentages of the responses. Univariate analysis was used to describe the basic features of the data and was stratified by the types of schools in the study.

Results

From the list of top 100 South African rugby schools, contact details were gathered for 66 of the schools (15 private and 51 public). Contact details were not available online for 44 of the schools. Of the 66 schools approached, 28 coaches (from 12 private schools and 16 public schools) responded to the survey. This response rate of 42% was deemed acceptable as it was higher than the rate obtained in a previous survey on conditioning practices.[21,26] Additionally, 15 no-fee schools were included in the study; these surveys were completed via face-to-face and telephonic interviews. The total number of responses was therefore 43 (private = 12, public = 16, no-fee = 15). The basic demographic data of the respondents are shown in Table 1.

Table 1: Basic demographic data of the participants

	Private	Public	No-fee	Total
	n=12	*n*=16	*n*=15	*n*=43 (%)
Gender				
Men	12	16	15	43 (100)
Women	0	0	0	0 (0)
Age (years)				
<30	6	8	2	16 (37)
30–39	6	7	6	19 (44)
>40	0	1	7	8 (19)
Experience (years)				
<5	4	3	2	9 (20)
5–10	5	8	4	17 (40)
>10	3	5	9	17 (40)
Highest level of education				
Matric	0	0	7	7 (16)
Undergraduate	2	0	1	3 (7)
Postgraduate	9	14	6	29 (68)
Other†	1	2	1	4 (9)
S&C Certifications				
NSCA	2	1	0	3 (7)
Other†	1	2	0	3 (7)

S&C, strength and conditioning; NSCA, National Strength and Conditioning Association
†Other refers to additional qualifications, such as certificates or diplomas in health and wellness.

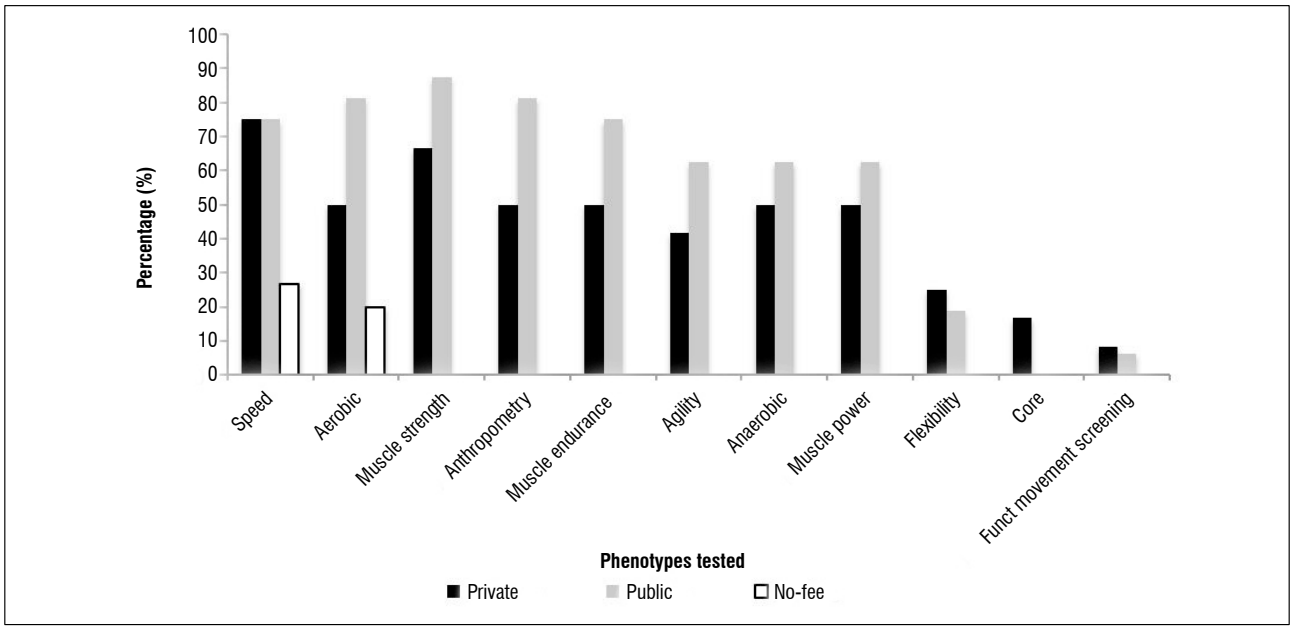


Figure 1: Different physical parameters tested (%) across the schools (n=31).

According to the coaches' responses, the most commonly tested measure was speed (58%), which remained the most popular for private schools (75%) and no-fee schools (27%) alike (Figure 1). The public schools indicated a preference for testing muscular strength (88%), with the only other parameter tested by no-fee schools being aerobic capacity (Figure 1). Overall, the most common tests were for speed (40-m and 100-m sprint tests), aerobic capacity (multi-stage fitness test), anaerobic capacity (repeat sprint ability test), muscular strength (one-to-five repetition maximum test), muscle endurance (maximum push-up, pull-up and sit-up tests), agility (Illinois agility test), muscle power (standing broad and vertical jump tests) and flexibility (sit and reach test).

Flexibility, speed, agility, plyometrics and resistance training were the most common practices. The most common types of equipment used for these training exercises are shown in Table 2.

Table 2: Most common (%) equipment used for each type of training

Type of training	Private (%)	Public (%)	No-fee (%)
Flexibility	Resistance bands (100%)	Resistance bands and foam roller (69%)	None
Speed	Resistance bands (75%)	Sled (63%)	Cones (40%)
Agility	Agility poles (88%)	Agility poles (88%)	Cones (33%)
Plyometrics	Boxes (75%)	Boxes (88%)	Cones (33%)
Resistance training	Free weights (100%)	Free weights (100%)	Body weight (53%)

Flexibility training

All 43 coaches reported that their teams did some type of flexibility training. Dynamic exercises were most common overall (n=38; private = 100%, public = 100%, no-fee = 67%). Static exercises were the most common form of flexibility training for no-fee schools (80%) and the second most common overall (n=36; private = 83%, public = 88%).

Speed training

All 43 coaches reported that their teams did some type of speed development training. Most (33) coaches reported holding speed development sessions on a weekly basis (private = 58%, public = 94%, no-fee = 73%) and almost all (39) reported free sprinting as the most common method for speed development (private = 83%, public = 88%, no-fee = 100%).

Agility training

Of the 43 respondents, 37 reported prescribing agility exercises (private = 92%, public = 100%, no-fee = 67%) and 24 prescribed agility exercises on a weekly basis (private = 58%, public = 56%, no-fee = 53%). Only 10 coaches reported using agility training more frequently than once per week (private = 25%, public = 38%, no-fee = 7%), and 3 coaches prescribed agility exercises on a monthly basis (private = 8%, public = 6%, no-fee = 7%). The most common type of agility training exercises were those involving sudden changes in direction (for example, Illinois agility drill, X-agility drill, L-agility drill).

Plyometric training

A total of 34 coaches prescribed plyometric exercises (private = 83%, public = 100%, no-fee = 53%). Various forms of jumping exercises were the most common method of plyometric training for private (58%) and no-fee schools (33%) while box drills were most popular with public school coaches (63%). Overall box drills were most common with 50% of private school coaches and 27% of no-fee school coaches using them. Of the 34 coaches prescribing plyometric training, 26 did so on a weekly basis (private = 67%, public = 75%, no-fee = 40%).

Resistance training

Resistance training was prescribed by 36 coaches (private = 100%, public = 100%, no-fee = 53%). The most common frequency of resistance training sessions was three times per week (n=16; private = 42%, public = 63%, no-fee = 7%). Of the 36 coaches, 35 required their players to take part in resistance training during the pre-season (private = 100%, public = 100%, no-fee = 47%); 32 continued resistance training in-season (private = 92%, public = 94%, no-fee= 40%); and 23 during the off-season (private = 92%, public = 69% and no-fee = 7%).

About half the coaches (n=20; private=58%, public=69%, no-fee=13%) reported using different conditioning practices for different positions (Table 3).

Table 3: Position-specific conditioning results

	Private *n*=7	Public *n*=11	No-fee *n*=2	Total (%) *n*=20
Front row				
Strength	4	8	2	14 (70)
Core	1	5	0	6 (30)
Scrummaging exercises	2	2	0	4 (20)
Aerobic training	0	1	0	1 (5)
Collision conditioning	1	0	0	1 (5)
Second row				
Strength	3	4	2	9 (45)
Explosive power	3	4	0	7 (35)
Core	0	2	0	2 (10)
Aerobic training	0	1	0	1 (5)
Collision conditioning	1	0	0	1 (5)
Flexibility	0	1	0	1 (5)
Back row				
Explosive power	2	6	0	8 (40)
Strength	2	2	2	6 (30)
Ruck specific exercises	2	3	0	5 (25)
Aerobic training	1	3	0	4 (20)
Anaerobic training	2	1	0	3 (15)
Speed	0	2	0	2 (10)
Half backs				
Passing drills (weighted)	2	2	1	5 (25)
Speed	1	3	0	4 (20)
Aerobic training	1	2	0	3 (15)
Explosive power	0	1	2	3 (15)
Decision-making / reactive agility	1	2	0	3 (15)
Strength	0	2	1	3 (15)
Wrestling	2	0	0	2 (10)
Agility	0	1	0	1 (5)
Flexibility	0	1	0	1 (5)
Centres				
Explosive power	3	6	2	11 (55)
Speed	1	6	1	8 (40)
Strength	1	5	0	6 (30)
Agility	0	1	0	1 (5)
Collision conditioning	1	0	0	1 (5)
Wrestling	1	0	0	1 (5)
Outside backs				
Speed	4	7	1	12 (60)
Power	1	5	2	8 (40)
Agility / reactive agility	1	3	0	4 (20)
Strength	0	1	0	1 (5)

Discussion

We sought to determine the strength and conditioning practices of South African school boy rugby coaches, specifically of those schools ranked highly in rugby performance and those from less privileged, no-fee schools. Overall, the no-fee school coaches were less likely to implement agility, plyometrics and resistance training, which could be explained by a lack of knowledge and limited or no access to equipment and/or facilities. As these exercises are key for injury prevention[6,8] and performance[5,21] in rugby, this limitation can have a negative impact on these players. The fact that no-fee schools lack the resources, facilities and skilled coaches to implement specific conditioning practices is somewhat surprising as there has been a big push over the past decade, by the then Department of Sport and Recreation South Africa, to eradicate these problems.[17] However, this scenario needs to be investigated further as these reasons were not explored in this study.

The difference in principles and techniques prescribed by coaches is related to the coaches' education, the schools' level of rugby play, and the socioeconomic standpoint of the school and the community which it serves. The no-fee school coaches, who have no formal education in strength and conditioning, arguably have less understanding of the advantages of physical development. Additionally, they may not have adequate knowledge or the appropriate facilities to implement effective strength and conditioning programmes for their players. Furthermore, the players from these schools are often plagued with socioeconomic disparities which can hinder general health and well-being.[27] A coach cannot prescribe the same physical training for those who are malnourished as for healthy adolescents who are well nourished.[28] Furthermore, children from no-fee schools are already expending more energy by virtue of the fact that they walk to school and do physically demanding chores at home such as collecting water and firewood.[29] These factors need to be taken into consideration when determining training programmes.

The coaches surveyed in the top 100 rugby schools for 2016 were also from the nation's best rugby playing schools historically. These schools employ specialised strength and conditioning coaches who tend to align their practices with the best-known international trends in rugby.[18] The no-fee schools cannot be expected to achieve the same level of strength and conditioning expertise with fewer resources and coaches who have no background in strength and conditioning. This highlights the complexities of sport and athletic development in a country such as South Africa where inequality and poverty prevails.

The majority (18 or 64%) of the top 100 rugby school coaches implemented different conditioning practices for different positions, compared with only two (13%) of the no-fee school coaches. The purpose of position-specific conditioning practices is to train the player according to the physical demands of their position, which is crucial in rugby, which has varying positional requirements.[30] The principles of 'specificity' and 'individualisation' are key aspects of developing an effective strength and conditioning programme to suit the needs of each player.[18] In no-fee schools, the lack of importance in developing each individual could be due to time constraints and/or an absence of knowledge. To achieve a high level of success in rugby, every individual in the team needs to reach their peak physical capabilities to meet the demands of their position and benefit the team as a whole.

In the top rugby schools, strength training was indicated as the largest focus for the front row (67%) and second row (39%) which makes sense as these positions require large amounts of force production at low velocities during scrums – a crucial aspect of their gameplay.[31] Explosive power was reported as the most important aspect of training for the back row (39%) and centres (50%) as these positions require a great capacity for power in tackling, scrumming, rucking, mauling and breaking through tackles.[32] The main focus for half backs was weighted passing drills (22%), despite the fact that a previous study in youth rugby players indicated no significant differences in passing performance variables between weighted ball and non-weighted ball training interventions.[33] Despite this lack of evidence, coaches seem to believe in an added benefit of performing weighted passes for half back positions. There might be a

case for progressive implementation of weighted ball passes to reduce shoulder injuries in half backs, for whom frequent forceful passing actions are required. This possibility is addressed in the survey in the subsequent questions on injury prevention; however, weighted ball exercises are not specifically mentioned. The biggest focus for outside backs by coaches was on speed development (61%), which is crucial as they often are required to beat the opposition by means of speed.[34]

Conclusions

This study provides a rich source of ideas to improve current training practices and expand on current knowledge on high school boys' rugby. It is clear from the lack of adequate conditioning in no-fee schools that methods need to be created to provide opportunities for learning for the coaches less resourced than those at wealthier schools. The South African Rugby Union should employ qualified individuals to promote player development among the no-fee schools and educate the coaches. Failing that, structures should be implemented to allow qualified strength and conditioning specialists in the top rugby playing schools to assist coaches at underprivileged schools. But first it should be determined whether no-fee schools are not administering adequate strength and conditioning practices due to a lack of knowledge or a lack of access to facilities, equipment and resources.

Authors' contributions

B.R. was responsible for conceptualising research goals and aims, developing the methodology, administering the survey and applying statistical techniques to analyse the data. L.P. was responsible for conceptualising research goals and aims, developing the methodology, preparing and creating the published work (initial draft) as well as reviewing and revising the writing. C.C. was responsible for conceptualising research goals and aims, developing the methodology, validating the experimental process and results, student supervision, writing revisons as well as overall project leadership and management.

References

1. Hendricks S, Sarembock M, Jones B, Till K, Lambert M. The tackle in South African Youth Rugby Union – Gap between coaches' knowledge and training behaviour. Int J Sports Sci Coach. 2017;12:708–715. https://doi.org/10.1177/1747954117738880
2. Sewry N, Verhagen E, Lambert M, Van Mechelen W, Brown J. Players' and coaches' knowledge and awareness of the BokSmart Safe Six injury prevention program: An ecological cross-sectional questionnaire study. BMJ Open. 2017;7, e018575, 7 pages. http://dx.doi.org/10.1136/bmjopen-2017-018575
3. Tee JC, Lebatie F, Till K, Jones B. Injury incidence and characteristics in South African school first team rugby: A case study. S Afr J Sports Med. 2017;29, Art.#1532, 7 pages. http://dx.doi.org/10.17159/2078-516x/2017/v29i0a1532
4. Duthie GM. A framework for the physical development of elite rugby union players. Int J Sports Physiol Perform. 2006;1(1):2–13. http://dx.doi.org/10.1123/ijspp.1.1.2
5. Sedeaud A, Marc A, Schipman J, Tafflet M, Hager J, Toussaint J. How they won Rugby World Cup through height, mass and collective experience. Br J Sports Med. 2012;46:580–584. http://dx.doi.org/10.1136/bjsports-2011-090506
6. Faigenbaum AD, Schram J. Can resistance training reduce injuries in youth sports? Strength Con J. 2004;26:16–21. http://dx.doi.org/10.1136/bjsm.2009.068098
7. Brown JC. Safer rugby through BokSmart? Evaluation of a nationwide injury prevention programme for rugby union in South Africa [doctoral dissertation]. Cape Town: University of Cape Town; 2014.
8. Lloyd RS, Faigenbaum AD, Stone MH, Oliver J, Jeffrey I, Brewer C, et al. Position statement on youth resistance training: The 2014 International Consensus. Br J Sports Med. 2014;48:498–505. http://dx.doi.org/10.1136/bjsports-2013-092952
9. Lambert MI, Durandt J. Long-term player development in rugby – how are we doing in South Africa? S Afr J Sports Med. 2010;22:67–68. http://dx.doi.org/10.17159/2078-516X/2010/v22i3a312
10. Arkell R. Body size, socioeconomic status, and training background of a selected group of u16 South African rugby union players (2010-2013): The impact on national selection [master's thesis]. Cape Town: University of Cape Town; 2015.
11. Lambert MI. BokSmart: Aspects of physical conditioning for rugby [document on the Internet]. c2009 [cited 2019 Apr 25]. Available from: http://images.supersport.com/content/BokSmart%20-%20Aspects%20of%20physical%20conditioning%20for%20rugby.pdf
12. Garraway W, Lee A, Hutton S, Russel E, Brewer C. Impact of professionalism on injuries in rugby union. Br J Sports Med. 2000;34:348–351. http://dx.doi.org/10.1136bjsm.34.5.348
13. Bathgate A, Best J, Craig G, Jamieson M. A prospective study of injuries to elite Australian rugby union players. Br J Sports Med. 2002;36:265–269. http://dx.doi.org/10.1136/bjsm.36.4.265
14. Hunter M. The race for education: Class, white tone, and desegregated schooling in South Africa. J Hist Sociol. 2016;29:319–358. https://doi.org/10.1111/johs.12097
15. Du Toit J. Playing time of professional senior rugby players across all levels of South African rugby, 2007-2012: Implications for transformation [master's thesis]. Cape Town: University of Cape Town; 2014.
16. Du Toit J, Durandt J, Joshua J, Masimla H, Lambert M. Playing time between senior rugby players of different ethnic groups across all levels of South African rugby. S Afr J Sports Med. 2012;24:81–84. https://doi.org/10.17159/SAJSM.327
17. Department of Sports and Recreation South Africa (DSRSA). Transformation Charter for South African Sport. Pretoria: DSRSA; 2011. Available from: https://www.srsa.gov.za/sites/default/files/Transformation%20Charter%20-%20FINAL%20Aug%202012.pdf
18. Jones TW, Smith A, Macnaughton LS, French DN. Strength and conditioning and concurrent training practices in elite rugby union. J Strength Cond Res. 2016;30:3354–3366. https://doi.org/10.1519/JSC.0000000000001445
19. Klein M, Fröhlich M, Pieter A, Emrich E. Socio-economic status and motor performance of children and adolescents. Eur J Sport Sci. 2016;16:229–236. https://doi.org/10.1080/17461391.2014.1001876
20. Weakley JJ, Till K, Darrall-Jones J, Phibbs P, Read D, Jones B. Strength and conditioning practices in adolescent rugby players: Relationship with changes in physical qualities. J Strength Cond Res. In press 2017. https://doi.org/10.1519/JSC.0000000000001828
21. Duehring MD, Feldmann CR, Ebben WP. Strength and conditioning practices of United States high school strength and conditioning coaches. J Strength Cond Res. 2009;23:2188–2203. https://doi.org/10.1519/JSC.0b013e3181bac62d
22. Pote L, Christie C. Strength and conditioning practices of university and high school level cricket coaches: A South African context. J Strength Cond Res. 2016;30:3464–3470. https://doi.org/10.1519/JSC.0000000000001432
23. Jones TW, Smith A, Macnaughton LS, French DN. Variances in strength and conditioning practices in elite rugby union between the northern and southern hemispheres. J Strength Cond Res. 2017;31:3358–3371. https://doi.org/10.1519/JSC.0000000000001773
24. Ebben WP, Carroll RM, Simenz CJ. Strength and conditioning practices of national hockey league strength and conditioning coaches. J Strength Cond Res. 2004;18:889–897. https://doi.org/10.1519/14133.1
25. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3:77–101. https://doi.org/10.1191/1478088706qp063oa
26. Williams S, Trewartha G, Kemp S, Stokes K. A meta-analysis of injuries in senior men's professional rugby union. Sports Med. 2013;43:1043–1055. https://doi.org/10.1007/s40279-013-0078-1
27. Williams DR, Priest N, Anderson NB. Understanding associations among race, socioeconomic status, and health: Patterns and prospects. Health Psychol. 2016;35:407. https://doi.org/10.1037/hea0000242
28. Ryan B. The efficacy of a community based eccentric hamstring strengthening program in peri-urban black South African soccer players [master's thesis]. Grahamstown: Rhodes University; 2015.

29. Muthuri SK, Wachira L, Leblanc AG, Francis C, Sampson M, Onywera V, et al. Temporal trends and correlates of physical activity, sedentary behaviour, and physical fitness among school-aged children in sub-Saharan Africa: A systematic review. Int J Environ Res Public Health. 2014;11:3327–3359. https://doi.org/10.3390/ijerph110303327

30. Van Wyk J. The relationship between training/match load and injuries in academy players during a provincial under 19 rugby union season [master's thesis]. Cape Town: University of Cape Town, 2015.

31. Bevan HR, Bunce PJ, Owen NJ, Bennet M, Cook C, Cunningham D, et al. Optimal loading for the development of peak power output in professional rugby players. J Strength Cond Res. 2010;24:43–47. https://doi.org/10.1519/JSC.0b013e3181c63c64

32. Nicholas CW. Anthropometric and physiological characteristics of rugby union football players. Sports Med. 1997;23:375–396. https://doi.org/10.2165/00007256-199723060-00004

33. Hooper JJ, James SD, Jones DC, Lee DM, Gal JM. The influence of training with heavy rugby balls on selected spin pass variables in youth rugby union players. J Sci Med Sport. 2008;11:209–213. https://doi.org/10.1016/j.jsams.2006.09.005

34. Morehen JC, Routledge HE, Twist C, Morton JP, Close GL. Position specific differences in the anthropometric characteristics of elite European Super League rugby players. Eur J Sport Sci. 2015;15:523–529. https://doi.org/10.1080/17461391.2014.997802

AUTHORS:
Muhammad Aziz[1,2]
Muhammad Anwar[2]
Shazia Iqbal[3]

AFFILIATIONS:
[1]School of science, University of Management and Technology Lahore, Lahore, Pakistan
[2]Institute of Biochemistry, University of Balochistan, Quetta, Pakistan
[3]Department of Chemistry, Balochistan University of Information Technology, Engineering and Management Sciences, Quetta, Pakistan

CORRESPONDENCE TO:
Muhammad Aziz

EMAIL:
Aziz1sh@Hotmail.com

DATES:
Received: 13 Oct. 2018
Revised: 08 Apr. 2019
Accepted: 08 Apr. 2019
Published: 30 Sep 2019

HOW TO CITE:
Aziz M, Anwar M, Iqbal S. Theoretical evaluation of valeraldehyde. S Afr J Sci. 2019;115(9/10), Art. #5666, 4 pages. https://doi.org/10.17159/sajs.2019/5666

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Priscilla Baker

KEYWORDS:
DFT spectroscopy calculation, pentanal, aldehyde

FUNDING:
None

Theoretical evaluation of valeraldehyde

The aim of this study was to compute the theoretical (software-based) spectroscopic properties of valeraldehyde as a member of the aldehyde family. The structural, thermochemical, electrical and spectroscopic properties of valeraldehyde were investigated using a quantum calculation approach. The infrared, ultraviolet-visible, Raman, and vibrational self-consistent field calculations and thermochemistry were calculated using the computational software GAMESS (General Atomic and Molecular Electronic Structure System) and the nuclear magnetic resonance predictions were calculated. The calculated energy gap between the lowest unoccupied molecular orbital (LUMO) and the highest occupied molecular orbital (HOMO) was 164.892.eV, which means valeraldehyde is a poor electrical conductor. The band gap is a non-neglected parameter for optical material. Results of the current computational analysis are useful to predict even a complex aldehyde precursor.

Significance:

- The computational strategy used successfully here could also be useful for investigation of other aldehyde polymers.

Introduction

Valeraldehyde is a colourless member of the aldehyde family. Applications involving aldehydes are vast. Valeraldehyde holds prime importance in the food industry as a flavour enhancer and in the polymer industry.[1] Vegetable glycerin produces aldehydes after heat contact in electronic cigarettes.[2] Darker beers produce disproportionately high levels of Strecker aldehydes.[3] However, the European Food Agency has prohibited the addition of flavouring agents[4], including valeraldehyde. Although aldehydes occur naturally in cells and are generated during intermediary metabolism of natural compounds, metabolism of drugs and xenobiotics may result in toxic interactions, which is why the pharmaceutical industry has a keen interest in aldehyde studies.[5]

The aldehyde under consideration here is valeraldehyde. Valeraldehyde is neither a simple nor a complex member of the aldehyde family. With the introduction of each ethyl (C_2H_5), the molecule can alter its intrinsic properties to some extent, so more data are needed to provide guidance to scientists working with this molecule. Kinetic modelling studies have been devoted to understanding the oxidation chemistry of aldehydes, because of their importance as intermediate and product species in alkane and biofuel oxidation.[6]

Through the use of computer-aided material design, the end product and the choice of material from which the product is made are interdependent activities. In this way, complex products can be created easily and directly from simulating the base result, with minimal requirements for the design or manufacturing tool.[7,8] Computational spectroscopic study is an analytical tool for identifying organic compounds. It is hoped that experimental methods in combination with computational methods will be a powerful tool for determining the absolute configuration of molecules. Through the use of computational software, vibrational spectroscopic simulations have long assisted in solving chemical problems and the industrial use of vibrational spectroscopy is highlighted by many.[9] Vibrational spectroscopy is based on the methods of theoretical chemistry consolidated into effective computer programs to determine the structural properties of a compound. Advancements in computational analytical chemistry base software has paved the way to predicting results before undertaking experimental analysis, which creates greater confidence in the results of the analysis.[10] Different types of computational methods and models are available. Selecting the appropriate approach for a particular molecule is the task of the computational chemist. Density functional theory is used to determine electronic properties with reference to β-functionalisation of aldehydes, as reported by Liu et al.[11] To our knowledge, this is the first report of infrared (IR), ultraviolet-visible (UV-Vis), Raman and vibrational self-consistent field (VSCF) spectrograms of valeraldehyde using density functional theory. The International Academy of Quantum Molecular Science accepts the results of theoretical calculations, which are very near to those of real or experimental calculations.

Spatial arrangement of valeraldehyde

The molecule contains 5 carbon atoms, 1 simple oxygen atom which shows the presence of a formyl group, 10 hydrogen atoms and 1 double bond – 16 atoms in total (Figure 1). The molecular formula is $C_5H_{10}O$ and SMILES (simplified molecular input line entry system) notation is CCCCC=O. The structure has no stereoisomers. Its molecular weight is 86 g/mol and estimated dipole moment is 2.4 D.

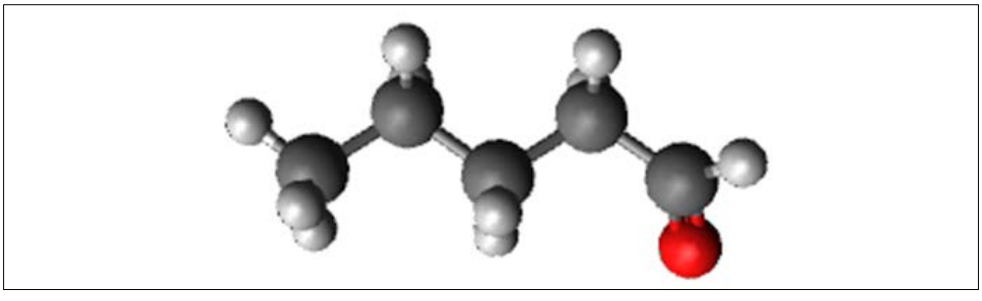


Figure 1: The arrangement of valeraldehyde.

Computational method

The molecular structure of valeraldehyde was drawn with the help of Avogadro 1.2.[12] Then geometry of the molecule was optimised by applying the universal force field. Verification of theoretical results was done through KnowItAll Informatics Systems, an online system that compares results with previous deposits. We selected GAMESS (General Atomic and Molecular Electronic Structure System) for computational calculation. The simulation program MestreNova was used for prediction of carbon and hydrogen nuclear magnetic resonance (NMR) and Gabedit 2.50 was used to visualise the results. All the software programs used are free for academic use.

Parameter use

For determination of electronic spectra, calculation frequency in basic set STO-3G with Becke three-parameter Lee Yang-Parr (B3LYP) hybrid function code was used. This method increases the time for computation but produces better results. Parameters for theoretical Raman and UV-VIS calculations were the same.

Computational results

Most of the quantum chemical results for molecular vibrations of molecules with more than a few atoms were described by the Born–Oppenheimer approximation. According to the Born–Oppenheimer assumption, the electronic and nuclear motion of molecules can be treated separately.[13] Computational scientists apply different computational capabilities to modelling and simulation, data analysis, and visualisation of results in their exploration of molecules. The successful outcome of the theoretical calculation is used as a reference for future work.

IR theoretical value

Figure 2 shows the IR spectrogram of valeraldehyde. The IR spectrogram is widely used for measurement, quality control, dynamic measurement and identification of the functional group. In the group frequency region of the spectrogram, peaks can be seen which correspond to functional groups.

From the IR spectrogram of valeraldehyde, for identification of the aldehyde group, formyl group stretching and C-H stretching, two sharp peaks of moderate intensity are observed near 2941 cm^{-1} and 2720 cm^{-1}. But from the theoretical calculation, the peak is observed at 2007 cm^{-1}.

C=O stretching vibration: 1725–715 cm^{-1} (strong)

C-H deformation vibration: 900–700 cm^{-1} (limited diagnostic value).

UV-Vis theoretical value

As can be seen in Figure 3, the UV-Vis spectrogram of valeraldehyde shows a parabolic curve. There are many applications of UV-Vis spectrophotometric studies. In this study, we used the UV-Vis spectrogram to examine electronic energy transition and excitation, that is, any molecule raised to a higher electron level means that an electron has been shifted from an orbital of lower energy to an orbital of higher energy.

Raman theoretical value

Figure 4 shows the Raman spectrogram of valeraldehyde. Presence of the aldehyde group is confirmed by medium C-H stretching in the functional group region at 3735 cm^{-1}.

Raman spectroscopy also reveals the molecular structure. Rotational and vibrational motions are unique attributes of every molecule, analogous to the fingerprints of humans. Each atom of the molecule is connected by a bond and each bond requires a different frequency, which creates different types of peaks – this region of the IR spectrum is called the fingerprint region.

As mentioned previously, aldehydes play a vital role in beverage production and determine the aqueous solution. Raman spectrophotometry is a superior tool for aqueous solutions because water can absorb the IR rays and thus distort the results of IR spectroscopic investigations. The Raman spectrogram is therefore widely used to complement the IR results.

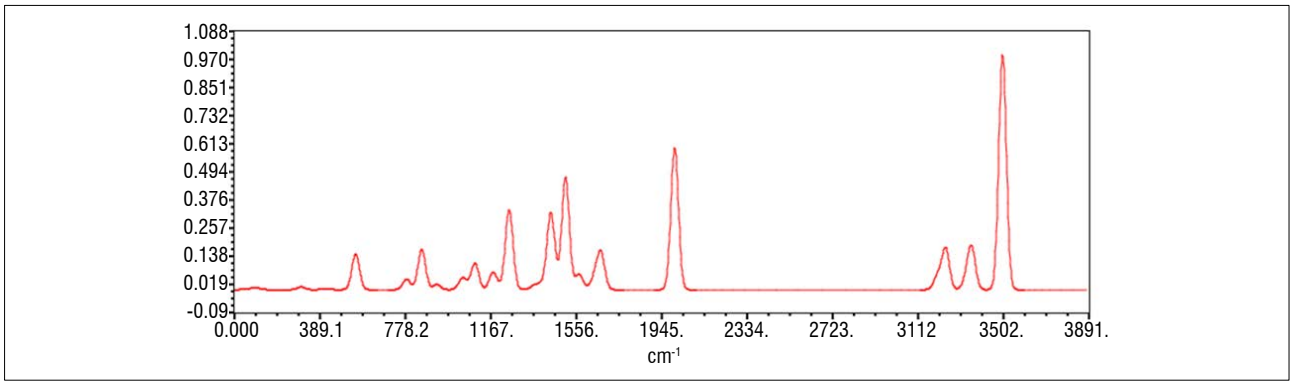


Figure 2: The infrared spectrogram of valeraldehyde lies between 0 and 3600 cm^{-1}.

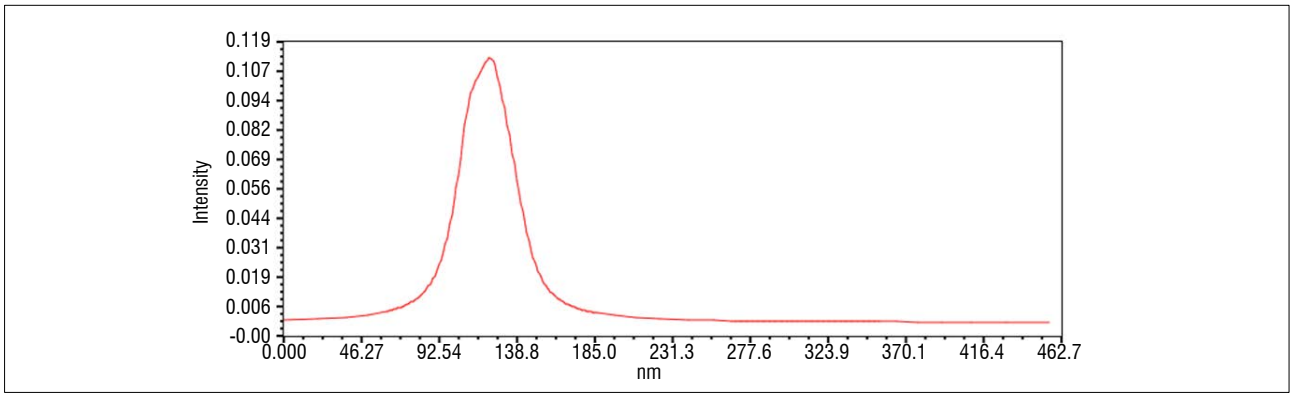


Figure 3: Ultraviolet-visible spectrogram of valeraldehyde.

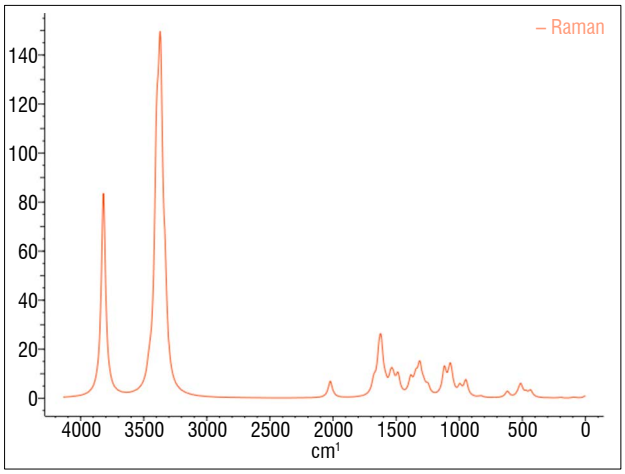


Figure 4: Raman spectrogram of valeraldehyde.

NMR prediction

NMR provides information on the structure, dynamics and reaction state. NMR also confirms the three-dimensional structure at a molecular level and thus limits the need for X-ray crystallography. The overriding goal of NMR is to produce magnets with higher field strengths in an effort to gain sensitivity and chemical shift dispersion for measurement. Three NMR analyses were needed for confirmation of the structure of the valeraldehyde molecule: one for the presence of carbon (Figure 5a), one for the presence of hydrogen (Figure 5b) and one for the presence of oxygen (Figure 5c), as the molecular formula is $C_5H_{10}O$.

NMR is a tool used in the food industry for quality control and research. As mentioned above, aldehydes are a main food ingredient and foodomics requires assessment of food. IR and Raman spectroscopy provide information on structure only, whereas NMR provides information on the chemical composition of food, allowing sources of variation to be identified.

Electronic properties

Table 1 shows the band gap of valeraldehyde. The band gap or energy gap indicates the difference between energy at the ground state and energy at the first excited state and is calculated by subtracting the energy in electron volts (eV) at the lowest unoccupied molecular orbital (LUMO) from that at the highest occupied molecular orbital (HOMO).

Table 1: Band gap of valeraldehyde

Calculation	Orbital	Energy (eV)
1	LUMO+4	287.210
2	LUMO+3	275.511
3	LUMO+2	267.366
4	LUMO+1	237.897
5	LUMO	57.457
6	HOMO	-107.435
7	HOMO-1	-199.247
8	HOMO-2	-209.243
9	HOMO-3	-211.020
10	HOMO-4	-215.611

HOMO, highest occupied molecular orbital; LUMO, lowest unoccupied molecular orbital

For valeraldehyde, the minimum calculated energy gap between LUMO and HOMO is 164.892 eV, which means that it is a poor electrical conductor and acts more like an insulator.

Aldehyde C-H bond dissociation energy is much smaller. The direct influence of electronegativity due to the hydrogen CO bond stabilises any negative intermediate. Here the essential gap is the distinction of electron affinity and upright ionisation potential. A narrow energy gap indicates a tranquil electron transition and a wide energy gap confers high thermodynamic solidness of the compound. As indicated by Koopmans' hypothesis, with regard to Hartree–Fock estimations, ionisation potential and electron affinity can be considered (short) energies of HOMO and LUMO, individually. At a presumption, the relation between band gap, which is classically associated with Fermi level to conduction band gap, and the HOMO-LUMO gap may be significant. An adverse feature of obtaining results from a molecular assessment (HOMO-LUMO) and a continuous model assessment of the solid state band gap is that it neglects unlike

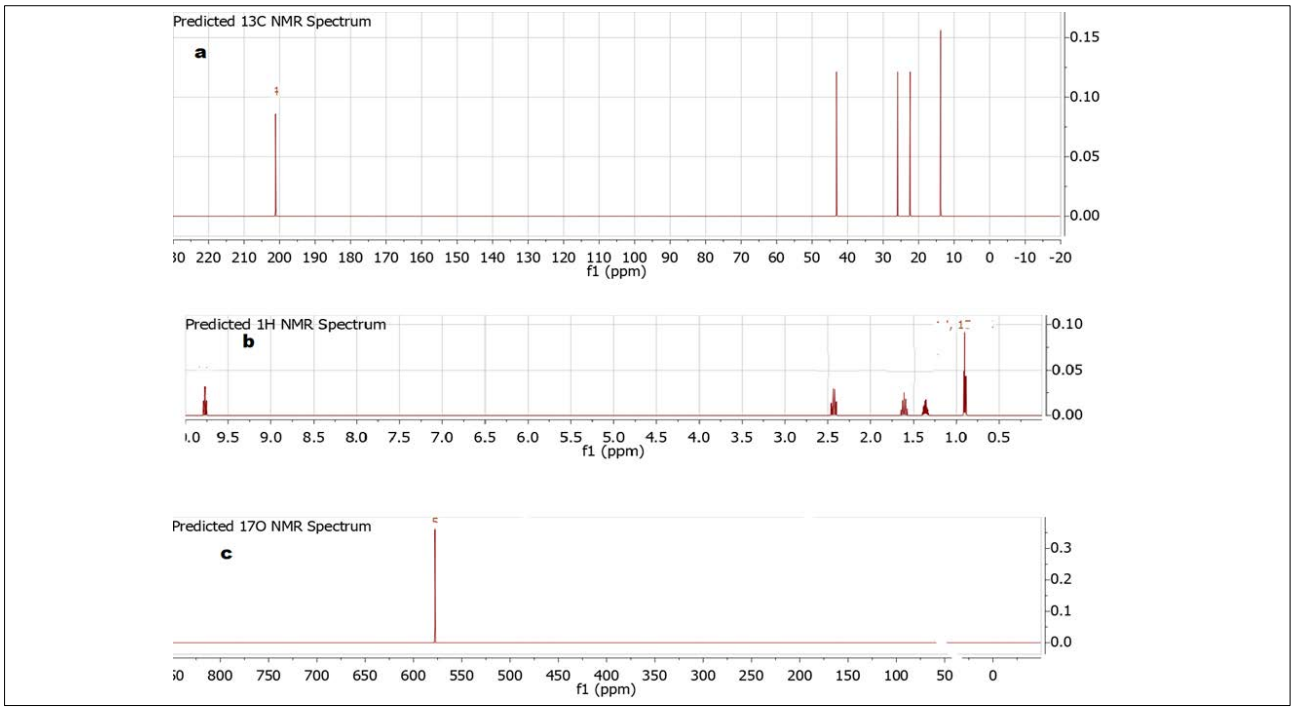


Figure 5: Nuclear magnetic resonance (NMR) spectra of valeraldehyde: (a) carbon, (b) hydrogen and (c) oxygen.

results. Unfortunately, these two models meet up somewhat badly in the nanodomain, because neither interpretation is very good in this crossover domain, giving rise to a possible new frontier area for research.

Thermochemistry

Thermochemical data provide information on the stability and reactivity of molecules. The enthalpy of the valeraldehyde molecule is -704086.66 kJ/mol, at a temperature of 25 °C and pressure of 1 atm. This result indicates that the molecule is exothermic in nature.

Conclusion

The current computational analysis is useful to predict even a complex aldehyde precursor and this approach may be useful for the assessment of other aldehyde molecule polymers. The resultant IR and Raman spectra of the compound were compared with the standard compound. Extra peaks in the spectrogram are indicative of impurities present in the compound. The field of computational analysis is still not fully developed enough to determine the vibrational property of organic molecules. The success of this theoretical approach opens a pathway to applying a correct algorithm force field for the assignment to aldehyde family. The key benefit of computational vibration analysis is an edge in the ease with which the electron shift, donation and withdrawing effects can be observed. The choice of open-source software allows for easier reproducibility of the methods. More recently, aldehyde nanomaterials are under development.[14] Further, it is necessary to calculate the polarisation density, the magnitude of the displacement vector, the dipole moment, the dielectric constant, the electric susceptibility, and the refractive index for a complete description of electrical and optical properties of alkanes.

Acknowledgement

We thank Muhammad Yunus and Sohail Nadeem for a layout of the manuscript.

Authors' contributions

M.Az.: Sample analysis, writing the initial draft. M.An.: Supervision. S.I.: data validation.

References

1. Kubisa P, Neeld K, Starr J, Vogl O. Polymerization of higher aldehydes. Polymer. 1980;21(12):1433–1447. https://doi.org/10.1016/0032-3861(80)90145-7
2. El Mubarak MA, Danika C, Vlachos NS, Farsalinos K, Poulas K, Sivolapenko G. Development and validation of analytical methodology for the quantification of aldehydes in e-cigarette aerosols using UHPLC-UV. Food Chem Toxicol. 2018;116:147–151. https://doi.org/10.1016/j.fct.2018.04.021
3. Gibson B, Aumala V, Heiniö R-L, Mikkelson A, Honkapää K. Differential evolution of Strecker and non-Strecker aldehydes during aging of pale and dark beers. J Cereal Sci. 2018;83:130–138. https://doi.org/10.1016/j.jcs.2018.08.009
4. Amending and correcting Annex II to Regulation (EC) nº 1333/2008 of the European Parliament and of the Council as regards the use of certain food additives. Off J Eur Union. 2013;L129(438):28–33.
5. Kuykendall JR, Kuykendall NS. Aldehydes. In: McQueen CA, editor. Comprehensive toxicology. 3rd ed. Oxford: Elsevier; 2018. p. 352–388.
6. Pelucchi M, Namysl S, Ranzi E, Frassoldati A, Herbinet O, Battin-Leclerc F, et al. An experimental and kinetic modelling study of n-C4C6 aldehydes oxidation in a jet-stirred reactor. Proc Combust Inst. 2019;37(1):389–397. https://doi.org/10.1016/j.proci.2018.07.087
7. Gutteridge PA, Turner J. Computer aided materials selection and design. Mater Design. 1982;3(4):504–510. https://doi.org/10.1016/0261-3069(82)90161-3
8. Smith P, Rennie A. Computer aided material selection for additive manufacturing materials. Virtual Phys Prototyp. 2010;5(4):209–213.
9. Siesler HW. Vibrational spectroscopy. In: Reference module in materials science and materials engineering. Amsterdam: Elsevier; 2016.
10. Barone V. Computational spectroscopy. In: Reference module in chemistry, molecular sciences and chemical engineering. Amsterdam: Elsevier; 2013.
11. Liu Y, Feng J, Liu X, Wang J. Theoretical study with DFT on the mechanism of visible light-driven β-functionalization of aldehydes. Comput Theor Chem. 2018;1123:154–160. https://doi.org/10.1016/j.comptc.2017.11.021
12. Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. J Cheminform. 2012;4(1), 17 pages. https://doi.org/10.1186/1758-2946-4-17
13. Oppenheimer MBJR. Zur Quantentheorie der Molekeln [On the quantum theory of molecules]. Annalen der Physik. 1927;389(20):457–484. German.
14. Chandiramouli R. Structural and electronic properties of germanane nanosheet upon molecular adsorption of alcohol and aldehyde molecules: DFT comparative analysis. J Mol Liq. 2017;242:571–579.

AUTHORS:
Munyaradzi. T. Matose[1]
Mladen Poluta[1]*
Tania S. Douglas[1]

AFFILIATIONS:
[1]Division of Biomedical Engineering, University of Cape Town, Cape Town, South Africa
*Current: Department of Health, Western Cape Government, Cape Town, South Africa

CORRESPONDENCE TO:
Munyaradzi Matose

EMAIL:
munya.matose@gmail.com

DATES:
Received: 02 Nov. 2018
Revised: 04 Jan. 2019
Accepted: 11 Aug. 2019
Published: 26 Sep. 2019

HOW TO CITE:
Matose MT, Poluta M, Douglas TS. Natural ventilation as a means of airborne tuberculosis infection control in minibus taxis. S Afr J Sci. 2019;115(9/10), Art. #5737, 4 pages. https://doi.org/10.17159/sajs.2019/5737

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:
ventilation rate, airborne infection control, tracer-gas concentration, window configuration, transmission risk

FUNDING:
None

Natural ventilation as a means of airborne tuberculosis infection control in minibus taxis

Airborne infection control measures are used extensively in health-care settings to curtail the spread of airborne infectious diseases. Few such measures are applied in public congregate spaces outside health facilities, such as those associated with public transport. In minibus taxis – a popular form of public transport in South Africa – poor ventilation creates conditions that allow for transmission of airborne diseases, particularly tuberculosis. In this study, we focused on developing quantitative ventilation profiles for the 16-seater Toyota Quantum Ses'fikile model commonly used in the Cape Town metropole. We studied the ventilation rates achievable in an occupied taxi under varying operational conditions, such as driving speed and open window configurations, which were based on observations made during preliminary taxi journeys. Two open-window configurations were found to provide ventilation rates close to or exceeding WHO recommended per-person requirements for high-risk clinical areas and are therefore likely to be effective in reducing the risk of tuberculosis transmission.

Significance:

- The results obtained augment the limited data available on the role that natural ventilation can play in reducing TB transmission in minibus taxis.
- Ventilation rates were shown to depend on both the taxi speed and specific open window configuration, countering the notion that simply opening a random selection of windows provides adequate reduction in the transmission risk.

Introduction

With some of the highest recorded infection and prevalence rates globally,[1] tuberculosis (TB) is a leading cause of both mortality and morbidity in South Africa, and particularly so in the Western Cape. From an infection control perspective, individuals most likely to transmit TB are those who are infected but undiagnosed and not receiving treatment and therefore highly contagious.[2] By the time many of these individuals present themselves at health-care facilities, they may inadvertently have transmitted TB to several people.

In addressing the TB burden in South Africa,[3] it is essential to consider the congregate settings outside health-care facilities as hotspots of transmission. A study conducted in a high incidence township in Cape Town estimated that transmission of TB outside the households of infected individuals contributed significantly to the spread of the disease.[4] Public transport studies conducted in Cape Town, Peru and Tanzania have shown that the risk of transmission in modes such as minibus taxis is high.[5-7] In South Africa, minibus taxis are an important component of the public transport system, accounting for 68% of public transport commutes to work nationally.[8] In the Western Cape, minibus taxis were found in a 2013 survey[8] to account for 39% of public transport commutes to work, followed by trains and buses at 34% and 18%, respectively.

In this study, ventilation refers to bringing and distributing outdoor air into or out of the interior space using natural forces such as wind pressures, thereby diluting the number of infectious particles within the indoor space.[1] In minibus taxis, ventilation is typically achieved through window opening during transit and door opening when passengers enter and exit the vehicle. Poor ventilation increases risk to passengers as conditions within the minibus taxi allow for transmission of airborne diseases. Studies exploring the achievable ventilation in taxis are limited[5-7,9] and the ventilation resulting from the combined changes in taxi speeds and window opening configurations has not been examined.

Both experimental and mathematical methods have been used in the study of airborne disease transmission risk in congregate settings (health care, public transport, prisons and public buildings)[10-14], and include airflow measurement and visualisation techniques[15], tracer gas techniques[9], the Wells–Riley deterministic model[16] and modifications to that model[17,18], dose-response numerical models[19,20], and computational fluid dynamics modelling[21]. A 2012 study by Johnstone-Robertson[5] investigated the risk of TB transmission in minibus taxis in South Africa by monitoring the CO_2 concentration as an indicator for ventilation (or lack thereof). Ventilation characteristics were explored in a stationary taxi for four ventilation scenarios using repeated CO_2 tracer gas tests and the Wells–Riley numerical model to determine the transmission probabilities for these scenarios. The study by Johnstone-Robertson[5] showed that ventilation rates from 1.31 L/s to 25.09 L/s could be achieved in a stationary taxi with three-open-window configurations.

In the current study, we examined the ventilation profiles associated with varying open-window configurations in operational transit, and aimed to determine whether ventilation rates above those recommended by the World Health Organization (WHO) could be achieved. This study differed from previous studies[5-7,9] in that the number of occupants in the taxi was controlled, the taxi speed was controlled and monitored, and the distinct open window configurations used were based on observed configurations in operational taxis.

Methods

Prior to establishing minibus taxi operating conditions for testing, 23 preliminary journeys were taken in taxis operating within the Cape Town metropole during which operating conditions that would potentially affect ventilation

were recorded. The average taxi speeds on highways and main roads, passenger occupancy (the number and location of passengers within the taxi) between peak morning and late afternoon rush hours, open window patterns (configurations, opening and closing frequency, duration of opening and the degree of window opening), as well as the routes taken, were recorded on journeys to and from the Cape Town central business district. From the preliminary data, the experimental taxi speeds, window configurations and passenger occupancy were determined.

CO_2 tracer gas decay tests were used to estimate ventilation rates. The experiments were carried out in a fully occupied minibus taxi (emulating high passenger occupancy rates observed in the preliminary journeys), with seating arrangement shown in Figure 1. The aim was to determine the range of ventilation rates achievable by changing the open window configurations at operating speeds of 40 km/h, 80 km/h and 100 km/h.

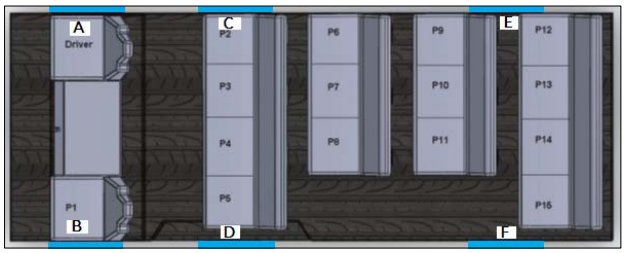


Figure 1: Top view of taxi interior with window and seating arrangements shown; A–F are the windows and P1–15 are the passenger seats.

Six window configurations (three double and three multiple open window configurations), shown in Table 1, were tested to provide a range of ventilation rates; the windows were always opened fully during the experiments. Taxi occupants comprised the driver, a researcher and 14 volunteer participants. Each configuration was tested twice at each speed.

Table 1: Window configuration for experiments

Window configuration	Open windows
1	A, B
2	C, D
3	E, F
4	A, B, E, F
5	A, B, C, D
6	C, D, E, F

CO_2 decay was measured with a data-logging CO_2/humidity/temperature sensor (Extech®SD800) with a resolution of 1 ppm, accuracy of ±40 ppm (<1000 ppm), ±5% reading (>1000 ppm) and range of 0–4000 ppm. The data logger was placed centrally, where the CO_2 concentration is highest. The use of a single CO_2 sensor was a limitation in the study, as more sensors would have yielded localised ventilation estimates that could reveal the higher risk seat positions with lowest ventilation rates per window configuration. The taxi speed and position were tracked using a front-mounted GPS unit linked to a separate data logger. The sampling frequency for the CO_2 sensor was set at 0.2 Hz and that of the GPS unit was set at 1 Hz, the highest that each would allow.

The CO_2 concentration was allowed to rise to above 2000 ppm when the windows were closed, while the participants fanned the air to ensure uniform CO_2 distribution; they then opened the windows to create each configuration while the taxi maintained the test speed. The gas was allowed to decay for 5 min or until the concentration decreased to within 200 ppm of baseline concentrations. The process was repeated at the same speed until all the test window configurations were completed within a 1-h experimental time frame. The taxi travelled on routes that would allow movement at the predetermined speeds: 40 km/h off the highway (non-residential roads) and both 80 km/h and 100 km/h on the highway. Ventilation was expected to increase as either the taxi speed increased or the total number of open windows increased.

The straight-line slope of the natural log (ln) of the CO_2 concentration decay curve was used to calculate ventilation rate for each window configuration.

Results

Ventilation data were recorded for the six window configurations in a fully occupied taxi over 7 days. A sample of the CO_2 concentration decay data for a journey with a target speed of 100 km/h is shown in Figure 2 with the accompanying recorded taxi speed. Each peak on the CO_2 concentration decay curve represents the moment before a window configuration was changed (shown in sequence from Configurations 1 to 6 in Table 1).

In Figure 2, the average speed achieved when the tests were conducted was 89 km/h, with fluctuations reflecting the conditions encountered on the highway. The corresponding average speeds attained for the target speeds of 40, 80 and 100 km/h were within 2.5%, 5% and 11%, respectively.

Figure 3 shows the average ventilation rate at each target speed compared against WHO recommended rates for three settings. The ventilation rates varied with changes in open window configurations, with Configurations 1, 4 and 5 achieving the highest rates at all speeds. Among the two-open-window configurations (Configurations 1, 2 and 3), Configuration 1 alone

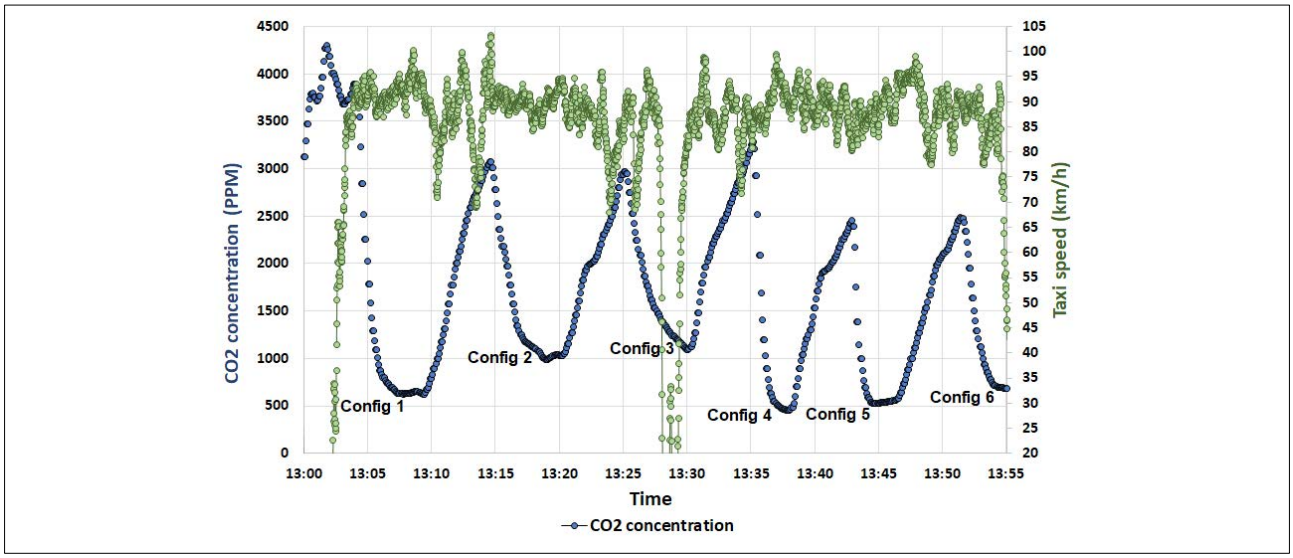


Figure 2: Taxi speed and CO_2 concentration decay curves for the 100 km/h experiment.

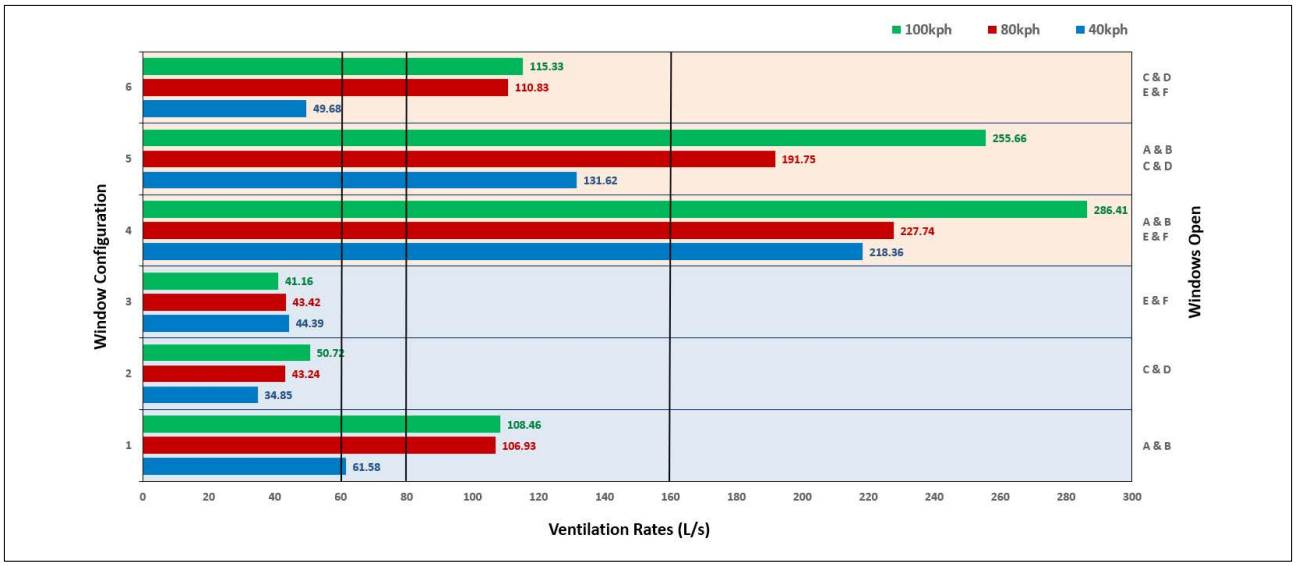


Figure 3: Effect of taxi speed on average ventilation rates at different window configurations (with open windows indicated on the right). World Health Organization recommended ventilation rates for general wards and outpatient departments (60 L/s), new health-care facilities and major renovations (80 L/s), and airborne disease precaution rooms (160 L/s) are indicated with vertical lines.

resulted in nearly double the ventilation rates seen in Configurations 2 and 3 at all test speeds. For the four-open-window configurations (Configurations 4, 5 and 6), Configuration 4 resulted in nearly double the rates produced by Configuration 6 at all speeds. For each individual configuration, apart from Configuration 3 (in which the rear windows were open), as the speed increased, the average ventilation rate increased. Configuration 1, with two open windows, resulted in ventilation rates similar to those of Configuration 6 (with four open windows) implying that the ventilation flow rates could depend as much on interior flow resulting from window configuration as on the number of open windows. Configurations 4 and 5 provided ventilation rates ranging from 132 L/s to 286 L/s.

Discussion

In Table 2, WHO recommended per-person ventilation rates for different kinds of health-care settings are provided[12]; these recommended rates are compared with the ventilation rates achieved in the interior of the taxi during the test periods.

The six configurations showed a wide ventilation range (34.85–286.41 L/s) when combined with variations in taxi speed. At all three test speeds, Configuration 1 produced ventilation rates that matched and exceeded that suggested for general wards and outpatient departments (60 L/s), whereas Configuration 6 only produced similar results at 80 km/h and 100 km/h. Configuration 4 produced the best ventilation rate, exceeding that suggested for airborne infection prevention rooms (160 L/s) at all three test speeds with a minimum ventilation rate of 218 L/s measured at 40 km/h. Configuration 5 similarly produced ventilation rates exceeding the highest recommended ventilation rates with 190 L/s at 80 km/h and 255 L/s at 100 km/h, whilst also exceeding the new health-care facility minimum ventilation rate (80 L/s) at 40 km/h with a ventilation rate over 130 L/s. Thus ventilation is a viable means of airborne infection control[22] in the confined environment studied. Configurations 4 and 5 are recommended for minibus taxis.

The high CO_2 concentrations, exceeding 4000 ppm (Figure 2), agree with the findings of Andrews et al.[9], confirming the high transmission risk environment in taxis. We have shown the impact of open window configurations on ventilation during operational transit, augmenting the ventilation profiles explored by Johnstone-Robertson[5] for a stationary taxi.

Despite the ventilation capabilities achieved in the study, passengers might not tolerate the fully open window configurations, for example due to discomfort caused by temperature changes, safety concerns, or inclement weather. Overloading would also alter the ventilation rates achieved with the window configurations considered in this study; taxis are often congested or overloaded during peak-traffic commuting journeys.[8] The potential role of airborne infection control devices (similar to those employed in health-care settings) to address the suboptimal ventilation scenarios must be considered; these devices include ultraviolet germicidal irradiation devices, filtration devices or mechanical ventilation.

Table 2: Performance of window configurations at different speeds against World Health Organization (WHO) guidelines

		40 km/h				80 km/h				100 km/h			
		Configuration				Configuration				Configuration			
		1 AB	4 ABEF	5 ABCD	6 CDEF	1 AB	4 ABEF	5 ABCD	6 CDEF	1 AB	4 ABEF	5 ABCD	6 CDEF
WHO recommended rate	General wards and outpatient departments (60 L/s)	=	>	>	<	>	>	>	>	>	>	>	>
	New health-care facilities and major renovations (80 L/s)	<	>	>	<	>	>	>	>	>	>	>	>
	Airborne disease precaution rooms (160 L/s)	<	>	<	<	<	>	>	<	<	>	>	<

<, = and > represent values below, matching and exceeding WHO recommendations, respectively

Conclusion

Our results show that in a fully occupied taxi, ventilation rates that match and surpass the WHO recommended ventilation guidelines for airborne infection settings can be achieved at a range of speeds. Thus ventilation is a viable means of airborne infection control[22] in such confined environments. However, driver and passenger education on the TB transmission risks in taxis and the importance of maintaining ventilation is necessary. Despite the ventilation rates achieved in the study, scenarios in which natural ventilation cannot be maintained, are still of concern, as passenger behaviour in operating taxis might not support the fully open window configurations. Taxi loads during peak-traffic commuting journeys are also of concern. The potential use of additional measures to address suboptimal ventilation scenarios must be considered.

Authors' contributions

M.T.M.: Study and manuscript conceptualisation, study design, data collection and analysis, manuscript preparation. M.P.: Study conceptualisation, funding acquisition, initial study supervision, manuscript editing. T.S.D.: Manuscript conceptualisation, project management, manuscript editing.

References

1. World Health Organization (WHO). Global tuberculosis report. Geneva: WHO; 2017.
2. Calligaro G, Esmail A, Mnguni T, Mottay I, Dheda K. The diagnostic gap: Characterising the profile of undiagnosed infectious tuberculosis patients in the community. S Afr Respir J. 2016;22(4):93–98. https://doi.org/10.7196/SARJ.2016.v22i4.100
3. Pooran A, Pieterson A, Davids M, Theron G, Dheda K. What is the cost of diagnosis and management of drug resistant tuberculosis in South Africa? PLoS ONE. 2013;8(1), e54587, 11 pages. https://doi.org/10.1371/journal.pone.0054587
4. Wood R, Racow K, Bekker LG, Morrow C, Middelkoop K, Mark D, et al. Indoor social networks in a South African township: Potential contribution of location to tuberculosis transmission. PLoS ONE. 2012;7(6), e39246, 5 pages. https://doi.org/10.1371/journal.pone.0039246
5. Johnstone-Robertson S. Calculating the risk of infection of *Mycobacterium tuberculosis* in endemic settings [master's thesis]. Stellenbosch: Stellenbosch University; 2012.
6. Zamudio C, Krapp F, Choi HW, Shah L, Ciampi A, Gotuzzo E, et al. Public transportation and tuberculosis transmission in a high incidence setting. PLoS ONE. 2015;10(2), e0115230, 8 pages. https://doi.org/10.1371/journal.pone.0115230
7. Hella J, Morrow C, Mhimbira F, Ginsberg S, Chitnis N, Gagneux S, et al. Tuberculosis transmission in public locations in Tanzania: A novel approach to studying airborne disease transmission. J Infect. 2017;75(3):191–197. https://doi.org/10.1016/j.jinf.2017.06.009
8. Statistics South Africa. National Household Travel Survey 2013. Pretoria: Department of Transport; 2014.
9. Andrews JR, Morrow C, Wood R. Modeling the role of public transportation in sustaining tuberculosis transmission in South Africa. Am J Epidemiol. 2013;177(6):556–561. https://doi.org/10.1093/aje/kws331
10. Gammaitoni L, Nucci MC. Using a mathematical model to evaluate the efficacy of TB control measures. Emerg Infect Dis. 1997;3(3):335–342. https://doi.org/10.3201/eid0303.970310
11. Kale SR, Veeravalli SV, Punekar HD, Yelmule MM. Air flow through a non-airconditioned bus with open windows. Sadhana. 2007;32(4):347–363. https://doi.org/10.1007/s12046-007-0029-3
12. World Health Organization (WHO). Natural ventilation for infection control in health-care settings. Geneva: WHO; 2009.
13. Zhang Z, Chen X, Mazumdar S, Zhang T, Chen Q. Experimental and numerical investigation of airflow and contaminant transport in an airliner cabin mockup. Build Environ. 2009;44(1):85–94. https://doi.org/10.1016/j.buildenv.2008.01.012
14. Johnstone-Robertson S, Lawn SD, Welte A, Bekker LG, Wood R. Tuberculosis in a South African prison – a transmission modelling analysis. S Afr Med J. 2011;101(11):809–813.
15. Cao X, Liu J, Jiang N, Chen Q. Particle image velocimetry measurement of indoor airflow field: A review of the technologies and applications. Energy Build. 2014;69:367–380. https://doi.org/10.1016/j.enbuild.2013.11.012
16. Noakes CJ, Sleigh PA. Mathematical models for assessing the role of airflow on the risk of airborne infection in hospital wards. J R Soc Interface. 2009;6(suppl.6):S791–S800. https://doi.org/10.1098/rsif.2009.0305.focus
17. Rudnick SN, Milton DK. Risk of indoor airborne infection transmision estimated from carbon dioxide concentration. Indoor Air. 2003;13(3):237–245. https://doi.org/10.1034/j.1600-0668.2003.00189.x
18. Yan Y, Li X, Shang Y, Tu J. Evaluation of airborne disease infection risks in an airliner cabin using the Lagrangian-based Wells–Riley approach. Build Environ. 2017;121:79–92. https://doi.org/10.1016/j.buildenv.2017.05.013
19. Issarow CM, Robin W, Nicola M. Modelling the risk of airborne infectious disease using exhaled air. J Theor Biol. 2015:100–106. https://doi.org/10.1016/j.jtbi.2015.02.010
20. Sze To GN, Chao CYH. Review and comparison between the Wells–Riley and dose-response approaches to risk assessment of infectious respiratory diseases. Indoor Air. 2009;20(1):2–16. https://doi.org/10.1111/j.1600-0668.2009.00621.x
21. Hathway EA, Noakes CJ, Sleigh PA, Fletcher LA. CFD Simulation of airborne pathogen transport due to human activities. Build Environ. 2011;46:2500–2511. https://doi.org/10.1016/j.buildenv.2011.06.001
22. Meadow JF, Altrichter AE, Kembel SW, Kline J, Mhuireach G, Moriyama M, et al. Indoor airborne bacterial communities are influenced by ventilation occupancy and outdoor air source. Indoor Air. 2014;24(1):41–48. https://doi.org/10.1111/ina.12047

AUTHORS:
Bontle Mataboge[1]
Amélie Beaudet[2,3]
Jason L. Heaton[1,4]
Travis R. Pickering[5]
Dominic Stratford[2]

AFFILIATIONS:
[1]Evolutionary Studies Institute, University of the Witwatersrand, Johannesburg, South Africa
[2]School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand, Johannesburg, South Africa
[3]Department of Anatomy, University of Pretoria, Pretoria, South Africa
[4]Department of Biology, Birmingham-Southern College, Birmingham, Alabama, USA
[5]Department of Anthropology, University of Wisconsin, Madison, Wisconsin, USA

CORRESPONDENCE TO:
Amélie Beaudet

EMAIL:
beaudet.amelie@gmail.com

DATES:
Received: 27 May 2019
Revised: 26 July 2019
Accepted: 12 Aug. 2019
Published: 26 Sep. 2019

HOW TO CITE:
Mataboge B, Beaudet A, Heaton JL, Pickering TR, Stratford D. Endostructural assessment of a hominin maxillary molar (StW 669) from Milner Hall, Sterkfontein, South Africa. S Afr J Sci. 2019;115(9/10), Art. #6404, 4 pages. https://doi.org/10.17159/sajs.2019/6404

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☐ All data included
☒ On request*
☐ Not available
☐ Not applicable

*Data are available upon application from Dr Bernhard Zipfel, Curator of Fossil and Rock Collections at the University of the Witwatersrand (Bernhard.Zipfel@wits.ac.za)

EDITOR:
Maryna Steyn

KEYWORDS:
Homo, late Pliocene-early Pleistocene, maxillary first molar, enamel-dentine junction, EDJ, enamel thickness

FUNDING:
Palaeontological Scientific Trust; DST-NRF Centre of Excellence in Palaeosciences

Endostructural assessment of a hominin maxillary molar (StW 669) from Milner Hall, Sterkfontein, South Africa

The site of the Sterkfontein Caves, South Africa, is one of the richest early hominin fossil-bearing sites in Africa. Recent excavations in the Milner Hall locality have contributed to the discovery of new hominin specimens, including StW 669, a right permanent maxillary first molar (M^1). StW 669 was excavated from the T1 deposits, which consist of a mixture of sediments from Members 2 and 5 of the Sterkfontein Formation. Accordingly, the deposits have the potential to contain remains of *Australopithecus*, *Paranthropus* and *Homo*. In this study, we employed micro-focus X-ray tomography in order to assess dental tissue proportions, enamel thickness distribution and enamel-dentine junction morphology as approaches to investigate the taxonomy of StW 669. We compare our results to those generated on the teeth of *Australopithecus africanus*, *Paranthropus robustus*, *Homo erectus*, *Homo antecessor, Homo neanderthalensis* and *Homo sapiens*. Our results suggest that StW 669 shares quantitative and qualitative affinities with M^1s of *Homo* in terms of tissue proportions (i.e. two- and three-dimensional average and relative enamel thickness of 1.2–1.3 mm and 18.4, respectively) and enamel thickness distribution (i.e. thickest enamel on the lingual aspect of the protocone). However, data on the enamel-dentine junction morphology of StW 669 are inconclusive as to the tooth's taxonomic affinities. Pending additional morphometric analyses, our studies of inner morphology of the crown of StW 669 support its attribution to *Homo*.

Significance:

- The Sterkfontein Caves have contributed significantly to our understanding of early human evolution and continue to do so. This study highlights the specific value of the Milner Hall locality as a valuable store of hominin fossils. Moreover, we tentatively clarify the enigmatic taxonomic status of StW 669, a right permanent maxillary first molar, excavated from the T1 deposit of Milner Hall. Pending additional morphometric evidence, our preliminary data on tissue proportions, enamel thickness distribution and enamel-dentine junction morphology, suggest an attribution of StW 669 to early *Homo*. This result is significant given the historical contention concerning the presence of early *Homo* at Sterkfontein.

Introduction

The Milner Hall is a deep underground chamber within the Sterkfontein Caves (South Africa) that extends about 100 m in an east-west direction.[1] The complex stratigraphic context of the Milner Hall fossiliferous depositional sequence – in which an early distal accumulation of the 3.67-Ma-old Member 2 (T3) and 2.18-Ma Oldowan artefact-bearing sediments from Member 5 (T2) contribute to the formation of the uppermost fossil- and artefact-bearing deposit, T1[1-3] – affords the potential for fossils of *Australopithecus*, *Paranthropus* and early *Homo* to be represented in the unit. Four hominin fossils – a manual proximal phalanx (StW 668), a right permanent maxillary first molar (M^1) (StW 669) and two cranial vault fragments (StW 671 and StW 672) from the T1 depositional unit of the Central Underground Deposits excavation site of Milner Hall (STK-MH1) – have been described previously.[2,4]

The description and metrical analyses of the hominin remains excavated from T1 revealed an enigmatic mix of unique, primitive and derived morphological traits, some of which indicated affinities with the genus *Homo*.[2,4] More specifically, the dimensions and shape of the crown of StW 669 were found to be more consistent with those of *Homo* than of *Australopithecus* and *Paranthropus*, and most similar to the Olduvai *Homo habilis* specimen OH 6 (Tanzania) and to the *Homo naledi* specimen UW 101-1688 from Rising Star (South Africa).[2] The attribution of StW 669 to *Homo* sp. is of particular interest for assessing the presence of early *Homo* at Sterkfontein and improving our understanding of the early hominin taxonomic diversity in South Africa.[5,6]

In association with significant improvements in analytical tools, the development of advanced imaging techniques, more specifically micro-focus X-ray tomography, offers the opportunity to assess taxonomically relevant endostructural information non-invasively.[7] Variation in tissue proportions, enamel thickness distribution and morphology of the enamel-dentine junction (EDJ) recorded in fossil hominins has been particularly useful for discriminating hominin taxa[8-10] and for clarifying long-standing uncertainties on the assignment of isolated fossil teeth[11].

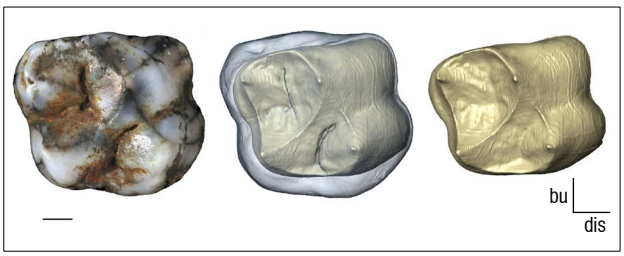
In this context, we provide here additional evidence from tooth crown endostructural analysis in order to clarify the taxonomic status of StW 669 and to tentatively contribute to ongoing debates about hominin morphological and taxonomic diversity at Sterkfontein during the late Pliocene and early Pleistocene.

Material and methods

StW 669 is a rootless right M^1 described by Stratford et al.[2] The occlusal surface of StW 669, which has a square outline (i.e. mesiodistal diameter:buccolingual diameter ratio of 1.02), is relatively unworn but the enamel formation has been disrupted at the tip of the paracone, at the buccal occlusal margin of the paracone, and on the mesial face of the tooth.[2] The occlusal area is small and relatively symmetric, and the areas of the paracone, hypocone and metacone are smaller than the areas of the protocone, hypocone and metacone.[2] The paracone is particularly small and the metacone area:paracone area ratio is high (i.e. 1.26).[2]

To discuss the taxonomic attribution of StW 669, qualitative and quantitative results derived from the endostructural three-dimensional (3D) analysis of StW 669 are compared to published data on tissue proportions, enamel distribution and enamel-dentine junction morphology of the permanent M^1 crowns of *Australopithecus africanus*[9,12], *Paranthropus robustus*[9], *Homo erectus*[13], *Homo antecessor*[14], *Homo neanderthalensis*[11,14,15], European Middle Pleistocene *Homo*[16] and modern *Homo sapiens*[11,14,17-19]. Unfortunately, to the best of our knowledge, there is no information available on the endostructure of the first molar of *Homo naledi*. As only a few specimens could be directly compared to StW 669 in terms of 3D tissue proportions (see Table 1), we also performed additional two-dimensional (2D) measurements on 2D sections that are analogous to previous studies (Table 2).

StW 669 was scanned with a Nikon XTH 225/320 LC dual-source industrial CT scanner at the microfocus X-ray tomography facility of the Evolutionary Studies Institute at the University of the Witwatersrand (Johannesburg, South Africa) using the following parameters: 70 kV, 100 μA, 3142 projections, an acquisition time of two frames per second and an isotropic voxel size of 26.7 μm (Figure 1). The dental tissues were obtained through a semi-automatic threshold-based segmentation and 3D mesh reconstructions via the software Avizo v.9.0 (Visualization Sciences Group Inc., Hillsboro, OR, USA) and verified by two users (B.M. and A.B.; Figure 1). The dental crown was digitally closed following the protocol detailed in Beaudet et al.[20] This coronal volume embeds both the dentine material and pulp cavity. In addition to the 3D volume, we extracted a virtual section of StW 669 passing through the mesial dentine horns to assess the tissue proportions in 2D following previously published protocols.[9,14,21]



Scale bar: 2 mm

Figure 1: Original picture[2] (left) and virtual rendering of the enamel cap (light blue, middle) and of the enamel-dentine junction (light yellow, right) of StW 669.

Three variables were measured in 3D and 2D[8,9,21]: volume (Ve, mm^3) and area (c, mm^2) of the enamel cap; volume (Vcdp, mm^3) and area (b, mm^2) of the coronal dentine that includes the coronal aspect of the pulp chamber; and surface (SEDJ, mm^2) and length (e, mm) of the enamel-dentine junction. We then calculated two indices of enamel thickness: 3D (3D AET, mm) and 2D (2D AET, mm) average enamel thickness, calculated

Table 1: 3D crown tissue proportions (including average and relative enamel thickness) of StW 669 and comparative fossil and extant hominins. For samples with more than one assigned specimen, mean as well as range (in parentheses) are given.

Taxa/specimen	*n*	Ve (mm^3)	Vcdp (mm^3)	SEDJ (mm^2)	3D AET (mm)	3D RET
StW 669		352.3	364.3	268.1	1.3	18.4
Australopithecus africanus[9]	1	382.3	491.1	240.6	1.6	20.1
Paranthropus robustus[9]	2	617.2 (592.0–642.4)	521.5 (498.2–544.8)	315.0 (248.2–381.8)	2.1 (2.4–1.7)	25.4 (20.6–30.1)
Homo erectus[13]	1	275.0	331.6	243.8	1.1	16.3
Homo antecessor[14]	2	301.8 (279.2–324.3)	345.7 (288.4–402.8)	233.5 (204.7–262.2)	1.3 (1.2–1.4)	18.7 (16.8–20.6)
Homo neanderthalensis[11,14]	3	283.6 (259.5–341.9)	362.9 (272.3–460.4)	255.9 (217.3–317.8)	1.1 (1.1–1.2)	15.9 (13.9–18.7)
Homo sapiens[11]	5	206.9 (185.2–229.9)	294.6 (285.8–304.2)	195.1 (145.4–255.0)	1.1 (0.8–1.6)	16.6 (12.6–23.5)

Sources: Published data from Olejniczak et al.[9], Zanolli et al.[11,13], Martín-Francés et al.[14]
Ve, volume of the enamel cap; Vcdp, volume of the coronal dentine that includes the coronal aspect of the pulp chamber; SEDJ, surface of the enamel-dentine junction; 3D AET, 3D average enamel thickness; 3D RET, 3D relative enamel thickness

Table 2: 2D crown tissue proportions (including average and relative enamel thickness) of comparative fossil and extant hominins. For samples with more than one assigned specimen, mean as well as range (in parentheses) are given.

Taxa/specimen	*n*	c (mm^2)	a (mm^3)	e (mm^2)	2D AET (mm)	2D RET
StW 669		24.6	68.7	20.1	1.2	18.4
Australopithecus africanus[9]	1	33.9	80.5	22.1	1.5	22.5
Paranthropus robustus[9]	2	50.3 (49.8–50.7)	106.0 (105.5–106.6)	23.4 (23.1–23.6)	2.2 –	28.8 (28.6–29.0)
Homo erectus[13]	1	28.9	75.5	22.2	1.3	19.1
Homo antecessor[14]	4	23.0 (20.5–23.0)	65.8 (60.1–72.2)	20.6 (19.2–22.1)	1.1 (1.1–1.2)	17.1 (16.1–18.4)
Homo neanderthalensis[8]	5	23.0 (21.0–28.0)	66.9 (58.2–77.8)	22.4 (21.1–23.5)	1.0 (0.9–1.2)	15.5 (13.8–16.9)
European Middle Pleistocene *Homo*[16]	1	23.3	64.2	21.5	1.1	16.9
Homo sapiens[14,19]	37	42.9 (32.5–59.4)	68.1 –	20.6 (17.7–24.1)	1.2 (1.0–1.5)	18.8 (14.0–23.9)

Sources: Published data from Olejniczak et al.[8,9], Zanolli et al.[13], Martín-Francés et al.[14], Smith et al.[16,19]
C, area of the enamel; a, area of the coronal dentine that includes the coronal aspect of the pulp chamber; e, length of the enamel-dentine junction; 2D AET, 2D average enamel thickness; 2D RET, 2D relative enamel thickness

as the ratio between Ve/c and SEDJ/e; 3D (3D RET) and 2D (2D RET) relative enamel thickness, obtained through the ratios 3D AET/(Vcdp)$^{1/3}$*100 and 2D AET/b$^{1/2}$*100 and allowing direct, scale-free comparisons.[8-9,21] Intra- and inter-observer tests for measurement accuracy run by two observers in previous published studies using a similar analytical protocol revealed differences less than 5%.[22]

We assessed the 3D distribution of enamel across the crown by computing the distances between the occlusal and the enamel-dentine junction surfaces through the 'Surface Distance' module on Avizo v.9.0.[20] The distances recorded were visualised at the outer enamel surface using a colour scale ranging from dark blue ('thinner') to red ('thicker').

Results

Tissue proportions

3D and 2D tissue proportions of StW 669 and comparative specimens are shown in Table 1 and Table 2. The volume of enamel (Ve) in StW 669 is higher than in *H. erectus*, *H. antecessor*, *H. neanderthalensis* and modern humans, but lower than in *P. robustus,* and fits more closely the value for *A. africanus*. Conversely, coronal dentine and pulp volume (Vcdp) in StW 669 is higher than in modern humans but lower than in *A. africanus* and *P. robustus* and approximates more closely the figures of *H. erectus*, *H. antecessor* and *H. neanderthalensis*. The surface of the enamel-dentine junction is intermediate between *A. africanus*, the *Homo* species and *P. robustus*. In general, the 2D enamel area (c), coronal dentine and pulp area (a) as well as the EDJ length (e) in StW 669 are lower than in *A. africanus*, *P. robustus* and *H. erectus,* but fits the range of, or at least are close to, the values reported for *H. antecessor*, *H. neanderthalensis*, European Middle Pleistocene *Homo* and *H. sapiens*.

In terms of average (3D AET) and relative (3D RET) enamel thickness, values recorded for StW 669 are lower than in the single representative of *A. africanus* and the two specimens of *P. robustus* and higher than in *H. erectus* and *H. neanderthalensis,* but are close to the values of *H. antecessor* (Table 1). Both 3D AET and 3D RET fall within the range of modern humans. The 2D AET fits the range of fossil and modern humans but is lower than in *A. africanus* and *P. robustus*. The 2D RET is consistent with the *H. antecessor* and *H. sapiens* ranges of variation, but is higher than in *H. neanderthalensis* and lower than in *A. africanus*, *P. robustus* and *H. erectus*.

Enamel distribution

The cartography that maps topographic enamel thickness variations at the outer crown surface of StW 669 is shown in Figure 2. Isolated spots correspond to enamel removal following wear (even if moderate here, i.e. there is no exposed dentine)[2] or disruptions in the enamel formation (i.e. deep, irregularly shaped pits)[2]. The thickest enamel in StW 669 is found along the oblique ridge connecting the metacone and the protocone, on the lingual aspect of the protocone and on the distal aspect of the hypocone. In these respects, the enamel distribution of StW 669 fits the fossil and extant human condition by sharing thick enamel on the lingual aspect of the crown[13,14,17] but differs from *Australopithecus* and *Paranthropus* that have thicker enamel at the cusp tips[9].

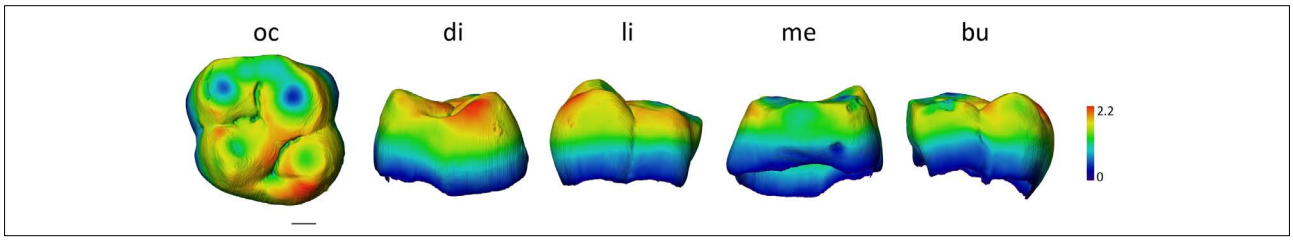
Enamel-dentine junction shape

The StW 669 EDJ shape is shown in Figure 3. The EDJ reflects the uncomplicated topography described for the occlusal surface.[2] The EDJ shows an oval-rectangular outline. An uninterrupted oblique ridge links the metacone to the protocone and separates central and distal fossae, the former being larger than the latter. The mesial marginal ridge is sharp and particularly high compared to the lateral and distal ones. The dentine horns are equally elevated. The protocone is more distally set compared to the paracone while the hypocone is more lingually set compared to the protocone. Carabelli's trait expression corresponds to grade 1 (i.e. a groove is present).[23] The StW 669 EDJ morphology is thus comparable to previous descriptions of *Australopithecus*[12] and *Homo*[13-15] M^{1}s.

Discussion

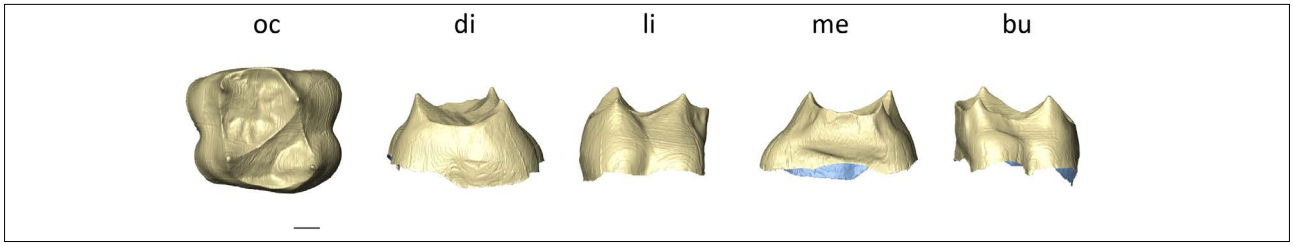
With hominin remains dated to either 3.67 Ma or 2.18 Ma, the Milner Hall fossil assemblage has the potential to provide further evidence useful for assessing early hominin palaeobiodiversity and the taxonomic context of the Sterkfontein hominin-bearing deposits. Consistent with previous descriptions and measurements of the outer morphology,[2] our micro-tomographic-assisted revision of StW 669 and characterisation of the endostructure reveal additional *Homo* affinities of the tooth. The identification of fossil *Homo* remains is critical in the context of Sterkfontein (and of South Africa as a whole) given the taxonomic ambiguity of specimens previously assigned to early *Homo* (e.g. StW 53)[24] and the fragmentary nature of fossil specimens identified as *Homo* (e.g. SK 847)[25]. Moreover, the diagnosis of the hominin remains found in the stone tools bearing deposits of T1 of Milner Hall may have implications for discussing the identity of the toolmakers at Sterkfontein.

Unfortunately, our data are not sufficient to determine with certainty to which species of *Homo* StW 669 should be assigned. The lack of consensus on the diversity of early humans, particularly in South Africa[6], represents another limiting factor. However, given the success of previous studies in identifying human fossil remains in South Africa by examining



Scale bar: 2 mm

Figure 2: Enamel thickness cartography of StW 669 in occlusal (oc), distal (di), lingual (li), mesial (me) and buccal (bu) views. Maximum thickness in mm.



Scale bar: 2 mm

Figure 3: Virtual rendering of the enamel-dentine junction of StW 669 in occlusal (oc), distal (di), lingual (li), mesial (me) and buccal (bu) views.

the molar crown inner structural organisation[26], future quantitative analysis of the enamel-dentine junction shape using morphometric methods[10-13] will be useful for refining the taxonomic status of StW 669.

Acknowledgements

We are indebted to B. Zipfel (Johannesburg) for granting access to the fossil specimen. We thank K. Jakata (Johannesburg) for microtomographic acquisitions. For technical and/or scientific discussion/collaboration we are grateful to R. Clarke (Johannesburg). We also thank the two anonymous reviewers for their comments that significantly improved this contribution and Associate Editor Maryna Steyn for managing the publication of this article. We thank the DST-NRF for sponsoring the Micro-XCT facility at Necsa, and the DST-NRF and Wits University for funding the microfocus X-ray CT facility in the Evolutionary Studies Institute (www.wits.ac.za/microct). The support of the DST-NRF Centre of Excellence in Palaeosciences (CoE-Pal) and the Palaeontological Scientific Trust (PAST) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the Centre of Excellence in Palaeosciences.

Authors' contributions

B.M.: Data collection; sample and data analysis; validation; writing. A.B.: Conceptualisation; methodology; sample and data analysis; validation; writing; student supervision; project leadership. J.L.H.: Writing. T.R.P.: Writing. D.S.: Conceptualisation; writing; student supervision; project management; funding acquisition.

References

1. Stratford DJ, Grab S, Pickering TR. The stratigraphy and formation history of fossil-and artefact-bearing sediments in the Milner Hall, Sterkfontein Cave, South Africa: New interpretations and implications for palaeoanthropology and archaeology. J Afr Earth Sci. 2014;96:155–167. https://doi.org/10.1016/j.jafrearsci.2014.04.002
2. Stratford D, Heaton JL, Pickering TR, Caruana MV, Shadrach K. First hominin fossils from Milner Hall, Sterkfontein, South Africa. J Hum Evol. 2016;91:167–173. https://doi.org/10.1016/j.jhevol.2015.12.005
3. Granger DE, Gibbon RJ, Kuman K, Clarke RJ, Bruxelles L, Caffee MW. New cosmogenic burial ages for Sterkfontein Member 2 *Australopithecus* and Member 5 Oldowan. Nature. 2015;522:85–88. https://doi.org/10.1038/nature14268
4. Beaudet A, Heaton JL, L'Abbé EN, Pickering TR, Stratford D. Hominin cranial fragments from Milner Hall, Sterkfontein, South Africa. S Afr J Sci. 2018;114(11/12), Art. #5262, 6 pages. https://doi.org/10.17159/sajs.2018/5262
5. Clarke RJ. *Homo habilis*: The inside story. In: Sahnouni M, Semaw S, Garaizar JR, editors. Proceedings of the II Meeting of African Prehistory. Burgos: Consorcio CENIEH; 2017. p. 25–51.
6. Grine FE, Smith HF, Heesey CP, Smith EJ. Phenetic affinities of Plio-Pleistocene *Homo* fossils from South Africa: Molar cusp proportions. In: Grine FE, Fleagle JG, Leakey RE, editors. The first humans: Origin and early evolution of the genus *Homo*. Dordrecht: Springer; 2009. p. 49–62. https://doi.org/10.1007/978-1-4020-9980-9
7. Macchiarelli R, Bayle P, Bondioli L, Mazurier A, Zanolli C. From outer to inner structural morphology in dental anthropology. The integration of the third dimension in the visualization and quantitative analysis of fossil remains. In: Scott GR, Irish JD, editors. Anthropological perspectives on tooth morphology, genetics, evolution, variation. Cambridge: Cambridge University Press; 2013. p. 250–277. https://doi.org/10.1017/CBO9780511984464.011
8. Olejniczak AJ, Smith TM, Feeney RNM, Macchiarelli R, Mazurier A, Bondioli L, et al. Dental tissue proportions and enamel thickness in Neandertal and modern human molars. J Hum Evol. 2008;55:12–23. https://doi.org/10.1016/j.jhevol.2007.11.004
9. Olejniczak AJ, Smith TM, Skinner MM, Grine FE, Feeney RNM, Thackeray JF, et al. Three-dimensional molar enamel distribution and thickness in *Australopithecus* and *Paranthropus*. Biol Lett. 2008;4:406–410. https://doi.org/10.1098/rsbl.2008.0223
10. Skinner MM, Gunz P, Wood BA, Hublin J-J. Enamel-dentine junction (EDJ) morphology distinguishes the lower molars of *Australopithecus africanus* and *Paranthropus robustus*. J Hum Evol. 2008;55:979–988.
11. Zanolli C, Kullmer O, Kelley J, Bacon AM, Demeter F, Dumoncel J, et al. Evidence for increased hominid diversity in the Early to Middle Pleistocene of Indonesia. Nat Ecol Evol. 2019(3):755–764. https://doi.org/10.1038/s41559-019-0860-z
12. Braga J, Thackeray JF, Subsol G, Kahn JL, Maret D, Treil J, et al. The enamel-dentine junction in the postcanine dentition of *Australopithecus africanus*: Intra-individual metameric and antimeric variation. J Anat. 2010;216:62–79. https://doi.org/10.1111/j.1469-7580.2009.01154
13. Zanolli C. Molar crown inner structural organization in Javanese *Homo erectus*. Am J Phys Anthropol. 2015;156:148–157. https://doi.org/10.1002/ajpa.22611
14. Martín-Francés L, Martinón-Torres M, Martínez de Pinillos M, García-Campos C, Modesto-Mata M, Zanolli C, et al. Tooth crown tissue proportions and enamel thickness in Early Pleistocene *Homo antecessor* molars (Atapuerca, Spain). 2018;13(10), e0203334, 29 pages. https://doi.org/10.1371/journal.pone.0203334
15. Martin RMG, Hublin JJ, Gunz P, Skinner MM. The morphology of the enamel-dentine junction in Neanderthal molars: Gross morphology, non-metric traits, and temporal trends. J Hum Evol. 2017;103:20–44. https://doi.org/10.1016/j.jhevol.2016.12.004
16. Smith TM, Olejniczak AJ, Zermeno JP, Tafforeau P, Skinner MM, Hoffmann A, et al. Variation in enamel thickness within the genus *Homo*. J Hum Evol. 2012;62:395–411. https://doi.org/10.1016/j.jhevol.2011.12.004
17. Kono RT, Suwa G, Tanijiri T. A three-dimensional analysis of enamel distribution patterns in human permanent first molars. Arch Oral Biol. 2002;47:867–875.
18. Kono RT. Molar enamel thickness and distribution patterns in extant great apes and humans: new insights based on a 3-dimensional whole crown perspective. Anthropol Sci. 2004;112:121–146.
19. Smith TM, Olejniczak AJ, Reid DJ, Ferrell RJ, Hublin J-J. Modern human molar enamel thickness and enamel-dentine junction shape. Arch Oral Biol 2006;51:974–95. https://doi.org/10.1016/j.archoralbio.2006.04.012
20. Beaudet A, Dumoncel J, Thackeray JF, Bruxelles L, Duployer B, Tenailleau C, et al. Upper third molar internal structural organization and semicircular canal morphology in Plio-Pleistocene South African cercopithecoids. J Hum Evol. 2016;95:104–120. https://doi.org/10.1016/j.jhevol.2016.04.004
21. Martin L. Significance of enamel thickness in hominoid evolution. Nature. 1985;314(6008):260–263. https://doi.org/10.1038/314260a0
22. Bondioli L, Bayle P, Dean C, Mazurier A, Puymerail L, Ruff C, et al. Morphometric maps of long bone shafts and dental roots for imaging topographic thickness variation. Am J Phys Anthropol. 2010;142:328–334.
23. Ortiz A, Skinner MM, Bailey SE, Hublin JJ. Carabelli's trait revisited: An examination of mesiolingual features at the enamel-dentine junction and enamel surface of *Pan* and *Homo sapiens* upper molars. J Hum Evol. 2012;63:586–596. https://doi.org/10.1016/j.jhevol.2012.06.003
24. Grine FE. The alpha taxonomy of *Australopithecus africanus*. In: Reed KE, Fleagle JG, Leakey RE, editors. The paleobiology of *Australopithecus*. Dordrecht: Springer; 2013. p. 74–104. https://doi.org/10.1007/978-94-007-5919-0
25. Clarke RJ. The cranium of the Swartkrans hominid, SK 847 and its relevance to human origins [PhD dissertation]. Johannesburg: University of the Witwatersrand; 1977.
26. Zanolli C, Pan L, Skinner MM, Dumoncel J, Beaudet A, De Beer F, et al. What is South African early *Homo*? New insights from the molar endostructural signature. Paper presented at: Meeting of the European Society for the Study of Human Evolution; 2018 September 13–15; Faro, Portugal.