South African Journal *of* Science

volume 116
number 7/8

Aerobiology in South Africa and the launch of SAPNET

Identification of solar periodicities in southern African baobab $\delta^{13}C$ record

Potential global warming impact of wood-based residential building in South Africa

COVID-19: Perspectives from the humanities and social sciences

ASSAf
ACADEMY OF SCIENCE OF SOUTH AFRICA

Volume 116
Number 7/8
July/August 2020

South African Journal *of* Science

eISSN: 1996-7489

Guest Leader

More eyes on COVID-19: Invited Commentary

Commentary

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
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

Statement

Review Article

Research Article

Research Letter

Cover caption
Baobab trees. In an article on page 65, Kotzé identifies characteristic solar periodicities from carbon isotope data of 1000-year-old baobab trees.

‘More eyes on the problem’: What the social sciences and humanities allow us to see and do in response to COVID-19

The COVID-19 pandemic has bared the fragile and unequal social and economic structures underlying South Africa, and in this issue of the Journal we foreground what contributions the humanities and social sciences can make to charting a path into an improved future.

We are fortunate that Jonathan Jansen, President of the Academy of Science of South Africa, gathered a group of multidisciplinary essays each written by an expert. Jansen, Distinguished Professor in the Faculty of Education at Stellenbosch University, is a leading intellectual, fearless author, and acclaimed researcher. At his invitation, we present a diverse series of Invited Commentaries that will have lasting value for our country in the difficult period that lies ahead. There are challenges but opportunities too.

When the world’s most well-known infectious diseases expert was interviewed about the alarming rate of COVID-19 infections in his country (the USA), Anthony Fauci appealed for caution ‘until we have more eyes on the problem’. It is an apt metaphor for what this collection of Invited Commentaries seeks to do – to provide ‘more eyes’ on the pandemic by drawing in perspectives from the social sciences and the humanities. Until now, the views of scientists such as epidemiologists, virologists and immunologists have enjoyed prominence in advising government on its responses to the pandemic.

The dominance of medical scientists in the initial response to the pandemic is of course understandable given the imminence of the viral threat to human health and human lives. But the limits of medical science evidence alone soon became evident.

Social distancing as a mitigating factor did not take account of crowded human settlements. Restriction of the number of mourners at funerals was made without attending to rituals of mourning and meaning-making in communities. Return of children to schools was gazetted without insights into how schools function as organisations and what that implies for re-opening under strict conditions. Enforcing of lockdown regulations was often done with little regard for the values of the Constitution and the rights of citizens. These were clearly not issues that could be resolved using epidemiological data alone.

Unsurprisingly, the wheels of scientific decorum came off within the science community even as debates raged in the broader public about the easing of lockdown restrictions.

We asked some of the leading social science and humanities scholars in South Africa to throw new eyes on the problem of COVID-19 from the vantage point of their particular disciplines by answering two questions:

1. What does your discipline allow us to better understand with respect to the social and human(e) aspects of COVID-19?
2. What do the insights from your discipline allow us to do differently in the context of the pandemic?

We bring together perspectives from only 11 fields or disciplines given limitations of time and space; these are economics, anthropology, law, political science, ethics, education, history, sociology, religion studies, language and philosophy.

Here – for the first time – is a collection of the best ideas for understanding the social nature of the COVID-19 pandemic and its consequences for humanity. The contributions are at once intellectual and practical as they speak to the many other facets of COVID-19 which we now know is as much a scientific puzzle as it is a social problem in these uncertain times.

HOW TO CITE:
Jansen J. ‘More eyes on the problem’: What the social sciences and humanities allow us to see and do in response to COVID-19. S Afr J Sci. 2020;116(7/8), Art. #8501, 1 page. https://doi.org/10.17159/sajs.2020/8501

More eyes on COVID-19: Perspectives from History

The need for history in a time of plague

AUTHOR:
Howard Phillips[1]

AFFILIATION:
[1]Department of Historical Studies, University of Cape Town, Cape Town, South Africa

CORRESPONDENCE TO:
Howard Phillips

EMAIL:
howard.phillips@uct.ac.za

HOW TO CITE:
Phillips H. More eyes on COVID-19: Perspectives from History: The need for history in a time of plague. S Afr J Sci. 2020;116(7/8), Art. #8495, 1 page. https://doi.org/10.17159/sajs.2020/8495

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
COVID-19, South Africa, social sciences, humanities, history

PUBLISHED:
29 July 2020

Despite the backward-looking, frozen-in-past-time connotation which the term ‘history’ invariably conjures up, it is necessary to recognise that history or, more precisely, our knowledge (such as it is) of the past, is closely tied to the present too. Not only does it explain how the present has been reached – every current problem has its source in the past – but it is also able to offer comparative historical examples, helping us to keep the present in perspective. To use a motoring metaphor, history provides both a rear-view mirror that shows us where we come from and by what route, and side-view mirrors that reveal where we are on the road in relation to others. To embark on a trip without either would be as foolhardy as drafting policies in a vacuum, without the benefit of memory or an awareness of the wider context.

The critical value of having such historically informed perspectives is well demonstrated when this lens is applied to the COVID-19 pandemic enveloping our country, for South Africa is no stranger to pandemics and so ought to be able to draw on these historical encounters to good effect today. These predecessors of COVID-19 include repeated epidemics of smallpox in the 18th and 19th centuries, bubonic plague (1901–1907), so-called ‘Spanish’ flu (1918–1919), polio (1944–1963) and ongoing HIV/Aids (1982–).

Studying these makes clear that severe pandemics invariably produce at least five broad reactions or results because of the direct, frightening threat which they pose to life:

1. they highlight the basic features of any society and its modus operandi, especially its shortcomings and fault-lines, in ways which are difficult to ignore;
2. they reveal underlying social and cultural attitudes which are not normally on public display;
3. they accelerate trends and tendencies already in train but not yet at full pace;
4. they introduce new, unanticipated developments into society; and
5. they trigger a zeal for medical research and for reform of the deficiencies exposed by the epidemic experience, although usually the latter is short-lived.

Knowledge about these earlier pandemics and the responses they evoked also give us a yardstick by which to judge what is novel and distinctive about a new pandemic and what is not.

Even though COVID-19 is still raging about us, all of these predictable reactions and results are already manifest in South Africa in some way – for instance, the manner in which dire poverty and overcrowding have been highlighted as a standard part of the daily experience of many South Africans, both at home and aboard public transport; the attitude of finger-pointing beginning to emerge towards ‘others’ who have been identified as COVID-19 positive, in the course of which the mantra of social cohesion has been superseded by that of social distancing; the way in which lockdown measures have accelerated the failure of already faltering companies like SAA and Edcon, tipping them over the edge, while at the same time calling forth an unprecedented roll-out of social relief by the state to the unemployed; how apparently sound small businesses and cultural institutions and activities have been very hard hit financially out of the blue, some to the point of collapse; how modes of social gathering have been altered, possibly forever; and how politicians have been vocal about their commitment, in President Ramaphosa’s words, ‘to forge a new economy and not merely return the economy to where it was before COVID-19 struck’. How lasting such commitments are, time will tell.

All of the above – and more – chime in exactly with the kind of responses which a historical study of previous epidemics would have led us to anticipate and prepare for. Moreover, such a study of the Spanish flu and the way in which that catastrophe (which killed 6% of the South African population in 6 weeks) accelerated the slow-moving process of creating a national department of health for the first time in 1919, should alert us to the likelihood that the COVID-19 disaster will be followed by a push to implement a national health insurance system in South Africa speedily.

In other words, history can not only explain the backstory to the present and put a current situation into perspective, but can also be a good guide to the future. The best futurologist is an historically informed one. To put it more felicitously (and with a revisionist nod to George Orwell’s *1984*): ‘Who understands the present, shapes the future; who understands the past, shapes the present.’

AUTHOR:
Tania Ajam[1]

AFFILIATION:
[1]School of Public Leadership, Stellenbosch University, Stellenbosch, South Africa

CORRESPONDENCE TO:
Tania Ajam

EMAIL:
tajam@spl.sun.ac.za

HOW TO CITE:
Ajam T. More eyes on COVID-19: Perspectives from Economics: The economic costs of the pandemic – and its response. S Afr J Sci. 2020;116(7/8), Art. #8490, 2 pages. https://doi.org/10.17159/sajs.2020/8490

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
COVID-19, South Africa, social sciences, economics, economic impact analysis

PUBLISHED:
29 July 2020

More eyes on COVID-19: Perspectives from Economics

The economic costs of the pandemic – and its response

From an economics perspective, the COVID-19 shock is unprecedented and very different from other global financial shocks. For the first time since the 1918 Spanish flu epidemic, the South African economy has been hit by real supply and demand shocks that have struck both domestic production and global supply chains, and simultaneously depressed demand in both the domestic and global economies. Lockdown halted consumption (e.g. in retail) to promote social distancing. Productive activities in most sectors ceased with consequent loss of jobs, workers furloughed or salaries cut – all of which induced a massive demand shock and loss of business and consumer confidence. There is a very real danger that these shocks to the real economy will morph into a financial crisis at a time when the South African economy has already been bedevilled by a secular decline in output growth, high unemployment (especially among the youth), precarious informal sector livelihoods, abject poverty and obscene inequality. The scenario is one of pure uncertainty rather than one of estimable probability.

It is crucial to distinguish the economic consequences of the health impacts from the pandemic itself, and the economic consequences of COVID-19 lockdown policy responses. There is no doubt that the public health imperatives must take centre stage during such an emergency. But it is ironic that, from an economic perspective, the economic costs of the COVID-19 response are likely to exceed the economic costs of the pandemic itself by several orders of magnitude. The quantum of these economic costs would be determined by the trajectory of infections (as contagion surges in 'hotspot' locations and subsequently ebbs in waves), the effectiveness of government's response strategy, the duration and coverage of a lockdown, and the phased exit strategy employed. Some of these costs would be felt immediately, others would manifest more in the medium term and beyond. It has been argued that the pandemic, left unchecked, would affect working adults, reducing their productivity for a month or two, probably keeping many away from work while killing a small number of persons, mainly the aged who are out of the workforce. This rationale has driven the more laissez-faire approaches based on 'herd immunity' once 60–70% of the population had already been infected.

Direct costs of the pandemic would include funding the public health response (borne by both the public sector and private sector businesses) and loss of productivity due to illness and death of economically active workers, which would undermine production and reduce consumption and exports. Lockdown responses acknowledge that it is the movement of people which spreads the virus. In order to slow the infection rate as much as possible early on in order to prevent the public health system from being overwhelmed, the lockdown responses aim to promote social distancing through the complete cessation of economic activity, bar a few essential services, such as health care and financial services. The costs of infection control are the incomes lost, both now and in the foreseeable future, as a result of control measures

The gross domestic product (GDP) – the aggregate production of a country – encompasses the value of final outputs of all its businesses, households, individuals, and its public sector. As the virus infects individuals and families, it impacts on the household sector, the business sector, labour markets, the public sector, the balance of payments, foreign investment, prices and money supply. All these impacts could converge to cause a decline in economic growth (a contraction in GDP), the magnitude of which is difficult to forecast at present, with forecasts ranging from 6% to 16%.

As households, firms and government in lockdown reduce their production, they also reduce exports (especially in sectors such as mining and agriculture) and their imports. The effect of COVID-19 on the balance of trade is ambiguous: both imports and exports are likely to decline, but which will fall more is still uncertain. With loss of investor confidence and sovereign credit downgrade, foreign investment might come to a sudden stop, or substantial risk premiums would be required to induce investment, mainly from volatile speculative investors ('hot money'). Trade partners which would have imported South African products and services are also hit by the pandemic, and exports decline. If trade partners develop relationships with other suppliers when South African export supply chains are disrupted, the decline in export losses may become permanent. These interacting supply and demand shocks will impact on the prices of final goods and services. They will also affect input prices and costs of intermediate inputs, inflation rates, wages, exchange rates, and financial, residential and other asset prices as firms and individuals are forced to divest assets in order to remain liquid or to survive.

The above exposition of some of the economic transmissions and contagion mechanism of the COVID-19 pandemic is highly simplified – but it does illustrate the magnitude and complexity of the shocks which virtually simultaneously hit South Africa and the world, and how economics as a discipline can shed light on how the pandemic shock was transmitted throughout the economy, and its differentiated impact on workers, consumers, firms, the fiscus and more.

Although there may have been public health benefits from an early pre-emptive lockdown strategy, it is clear that the longer the duration of a lockdown, the less effective it is likely to be from a public health perspective. Social distancing and self-isolation during lockdown is only possible in middle-class suburbia. It is simply not practical in the overcrowded informal settlements and townships where access to water and sanitation has shamefully been lacking for decades. On the other hand, the economic costs of a lockdown increase the longer lockdown continues. Extending the lockdown beyond a certain point runs the risk of still being able to contain the pandemic *and* paying the full economic costs of the lockdown.

Government has designated a system of lockdown levels ranging from Level 5 (most restrictions on social and economic activities) to Level 1 (normal, unconstrained operations), with certain services designated as essential or protected. By 28 May 2020, government had already announced a risk-based lockdown exit strategy, with sectors like mining soon becoming operational. Sector-specific health protocols will have to be negotiated – and more importantly enforced – to ensure that workers' lives are not put at preventable risk (e.g. provision of personal protective equipment and appropriate training).[1]

While many of the lockdown regulations were sensible from both a public health and economic perspective, in other cases they appeared to be irrational. Examples include the ban on e-commerce, alcohol sales (including wine exports) and tobacco sales, attempts to microregulate winter clothing sales and attempts to restrict NGOs rolling out feeding schemes when government itself was unable to reach all the needy and excluded groups like foreign nationals, unlike the NGOs. After popular backlash, these regulations were reversed. On the other hand, the Competition Commission was quite proactive in combatting price gouging. For managing the risks to the livelihoods and laying the foundations for inclusive, environmentally sustainable and digitally equitable growth and innovation, effective regulation which enables civil society and the private sector to co-create a post-COVID future would be a prerequisite.

Public discourse on immediate responses to the pandemic's economic impact centre on borrowing, drawing down cash reserves, reprioritising spending, credit guarantee schemes and judicious application of the South African Reserve Bank's balance sheet.[2] These responses can be only short term at best. In the medium term, the tax system can be the only sustainable bedrock for financing post COVID-19 reconstruction. Inclusive growth is the only way South Africa can dig itself out of this economic hole. Such a recovery cannot simply be focused on increasing the rate of growth but must also shift the direction and nature of growth to a more employment-intensive, equitable, climate friendly and digitally smart trajectory.

Charting the way forward calls for a fundamental reconceptualisation of the roles and relationships between the public sector, private sector and civil society to forge new social compacts. This will require evidence-based analysis from a range of sub-disciplines of economics, such as fiscal policy and public economics, intergovernmental fiscal relations, municipal finance, monetary policy, exchange rate policy, trade policy, industrial policy, labour market policy, social policy and social security, regional development, innovation systems and R&D, political economy, institutional analysis, energy policy, infrastructure financing, health and education financing, financial economics (banking and credit markets), development finance, and network industries regulation (e.g. information and communication technology and game theory). The pandemic will spawn a thousand PhDs in Economics. This modest contribution has attempted to sketch the broad outlines of that research agenda.

References

1. Department of Cooperative Government. Disaster Management Act: Regulations: Alert level 3 during Coronavirus COVID-19 lockdown [webpage on the Internet]. 2020 May 28 [cited 2020 Jun 08]. Available from: https://www.gov.za/documents/disaster-management-act-regulations-alert-level-3-during-coronavirus-covid-19-lockdown-28

2. Sachs M. Macro-fiscal considerations in response to the COVID-19 crisis [webpage on the Internet]. 2020 April 04 [cited 2020 Jun 08]. Available from: https://covid19economicideas.org/2020/04/07/macro-fiscal-considerations-in-response-to-the-covid-19-crisis/

Check for updates

AUTHOR:
Thaddeus Metz[1]

AFFILIATION:
[1]Department of Philosophy, University of Johannesburg, Johannesburg, South Africa

CORRESPONDENCE TO:
Thaddeus Metz

EMAIL:
tmetz@uj.ac.za

HOW TO CITE:
Metz T. More eyes on COVID-19: Perspectives from Philosophy: How philosophy bears on COVID-19. S Afr J Sci. 2020;116(7/8), Art. #8494, 1 page. https://doi.org/10.17159/sajs.2020/8494

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
COVID-19, humanities, philosophy, ethics, epistemology

PUBLISHED:
29 July 2020

More eyes on COVID-19: Perspectives from Philosophy

How philosophy bears on COVID-19

Philosophy is rational enquiry that addresses fundamental matters of human life and that transcends science in some way. For example, biologists and chemists appeal to physical facts when explaining what transpires in the world, but philosophers (and specifically ontologists) consider whether there is only a physical realm and whether there is evidence of anything spiritual such as God. Cosmologists claim to know some facts about the universe, whereas philosophers (epistemologists) try to ascertain how it is they know and why astrologers do not know. Sociologists describe how people behave using value-free language, while philosophers (ethicists) prescribe how they morally ought to behave and distinguish between good and bad ways of living.

What follow are three examples of ways in which philosophers are particularly well qualified to address difficult questions pertaining to COVID-19. My aim is not to provide answers here, but rather to demonstrate that compelling answers are not obvious and would require sustained and careful philosophical enquiry.

How should we allocate scarce resources during a pandemic? Ethicists, and specifically philosophers of justice, argue about how to allocate benefits and burdens in ways that are fair. Such issues abound in the context of COVID-19. Consider, here, debates concerning how to balance the interests of the elderly against those of the young. Most who die from COVID-19 are older than 60, while Africa has a relatively large population of young people. How should trade-offs be made between them?

Specifically, if both an elderly person and a youthful person need a ventilator to survive, but resources are scarce such that only one of them can receive a ventilator, who should it be? Should one flip a coin in such cases, because all lives have an equal dignity? Or should we favour the young, because the old have already had lives to lead? Or should we favour the old, because they are entitled to greater respect in virtue of their personhood (wisdom, accomplishment) and because they have paid much more tax into the health-care system?

Beyond this dilemma involving just two persons, there are broader, generational conflicts. For instance, is saving the lives of thousands of elderly people worth impairing the livelihoods of a much greater number of youth through a lockdown? If you are tempted to say that life always trumps livelihood, does it follow that driving cars should be forbidden due to the tens of thousands of lives that are lost in accidents each year?

Must we obey all the government's rules about COVID-19? Ethicists, and specifically political and legal philosophers, argue about when and why we are obligated to obey the government. Sometimes its laws and policies appear unjust, and you, the reader, are invited to pick your favourite example pertaining to COVID-19. Are we obligated to obey a government's rules even when we reasonably disagree with them?

One might be tempted to say that we are obligated to obey only the just laws and policies, and not any of the unjust ones. However, imagine what would happen if people disobeyed the government any time they thought its rules are unjust. Think about what would happen to tax collection, for instance. Values such as the rule of law, order, and peace would be gravely threatened.

In contrast, then, one might suggest that one has a duty to obey all of a government's decisions, whether just or unjust. Or one might think, more specifically, that one has a duty to obey any law or policy that has been ratified by an elected majority. Maybe rebellion against dictatorships is justified, whereas it is not against democracies. However, imagine you are in a group of three people, where, after some debate at a park, two of them vote to take away your shoes and you vote to keep them. The decision was made democratically, but are you obligated to abide by it? Similarly, remember that slavery and Jim Crow laws in the USA were democratically adopted; white people simply outnumbered black people. Did slaves really have a moral obligation to obey their masters? What difference might there be between these cases and your favourite COVID-19 example above?

Whom should we believe about COVID-19? Epistemologists argue about when and why it is appropriate to believe claims about the nature of ourselves and the world. There have been a variety of sources offering competing perspectives on how the coronavirus is spread, how it has affected people's lives and health, and how best to stem the pandemic. The government says one thing in policy briefings, a majority of scientists might have formed a view, a minority of scientists invariably dissent from the majority, international NGOs have their view, religious leaders have theirs, citizens on YouTube who appear to have done a lot of research have theirs, and you of course have yours. When these various viewpoints are incompatible, how should you proceed?

You might be drawn to hold one of two extremes. On the one hand, you might think, 'It's my life, and so I'm going to believe whatever I want or what makes me feel good'. However, it's not just your life that is affected by how you live; the choices you make can radically influence the course of other people's lives, even end them.

On the other hand, you might think, 'I'm not qualified to form an opinion; I'll just leave it to others to judge'. However, you do have a life to live, and so the question becomes: which others should you believe when making choices, including, say, about whether to send your children back to school?

In between these extremes is the approach that we should form beliefs in the light of what the experts tell us. Who counts as an expert? Presumably someone who knows a lot about a topic. OK. But how can we tell who that is? You might be inclined to say that we can know someone is an expert because other experts say so. But how can you know that *those* people are experts? Because they say so, or because still other self-proclaimed experts say so? Presumably not, but then how is one to identify those with expertise?

To conclude, most readers will have real difficulty providing what they deem to be firm answers to these questions backed up by reasoning that others would find compelling. Philosophers are in the business of searching for such answers. No – we do not always find them. However, we do spend 45+ years of our lives trying to. Might we therefore count as experts, or at least deserve a hearing?

AUTHOR:
Solomon Benatar[1,2]

AFFILIATIONS:
[1]Emeritus Professor of Medicine and Senior Scholar, University of Cape Town, Cape Town, South Africa
[2]Adjunct Professor, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada

CORRESPONDENCE TO:
Solomon Benatar

EMAIL:
solomon.benatar@uct.ac.za

HOW TO CITE:
Benatar S. More eyes on COVID-19: Perspectives from Ethics: The most powerful health-promoting forces in COVID-19 are social. S Afr J Sci. 2020;116(7/8), Art. #8492, 2 pages. https://doi.org/10.17159/sajs.2020/8492

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
COVID-19, South Africa, social sciences, humanities, ethics, social determinants of health

PUBLISHED:
29 July 2020

More eyes on COVID-19: Perspectives from Ethics

The most powerful health-promoting forces in COVID-19 are social

As the COVID-19 pandemic rages through the world, all aspects of life globally are being disrupted by mounting death rates[1] and governments' responses. The first ethical lesson has been the realisation that the increasing instability of the world, characterised by diverging trajectories[2] of health and well-being, with a minority (25%) benefitting from spectacular human development and progress, and a majority (75%) suffering from inadequate human and social development, is amplified in South Africa as a failing state, with its even wider disparities and continuing, pervasive poverty, hunger, unemployment and heavy burden of disease.

The second lesson relates to the complexity of the challenge for the government and people of a middle-income country seeking a balance between efforts to: (1) mitigate and control the pandemic for long enough to prepare already inadequate overall health facilities to save as many lives as possible, and (2) prevent severe damage to our fragile and crumbling economy in order to avoid deaths from starvation and other neglected health needs.

Our politicians, who are abandoning their moral legacy (like those in another retrogressing country[3]), are not well equipped to take difficult evolving decisions as the pandemic unfolds, without support from a range of available scholars in science, the humanities and medicine. Knowledge of the science[4] and dynamics of socio-political-economic influences on health and disease are crucial to the wise use of knowledge to improve the lives and health of people at both individual and population levels.

Despite these shortcomings, some admirable attempts are being made to utilise both our well-funded private and poorly funded public health-care sectors to face the immediate challenges. The spirit in which the best and most committed of our health professionals are working, embraces high standards of evidence-based medical practice. Although less adequately taught in our medical schools, the ethics of clinical duties of care and the art of medicine are also manifest, having been nurtured during many decades of clinical experience in caring for the world's largest proportions of patients with both HIV/Aids and tuberculosis.[5] Inspiring confidence, trust and measured hope are important in everyday health-care practices, and of special importance during public health emergencies. These are best achieved through the application of knowledge with clear, unambiguous communication across diverse barriers by coordinated health-care teams, with empathic understanding of the contextual nature of personal suffering and appreciation of the uniqueness of each person with respect for patient autonomy.

A significant ethical challenge highlighted by the pandemic, is the failure to openly acknowledge a weakness of the popular notion of a 'right'. Conceptually a 'right' can only be considered as one side of a coin, the other being a co-relative 'responsibility'. Rights cannot be met without identifiable and accountable bearers of responsibilities with the ability to do so. The relevance of this at the level of the easing of lockdown restrictions is that national public cooperation is needed to ensure that all are aware, for example, that 'your right not to be infected by me requires me to wear a mask, sanitise and respect social distancing, and my right not to be infected by you requires that you do the same'. By imposing some rigid and poorly conceived rules that provoke frustration and anger, our government is regrettably losing a crucial opportunity to enable all its people to embrace an ethics of good character and responsibility that could contribute to solidarity and social capital. At a higher level, achieving the 'right to health care for all' implies a societal responsibility. Despite success in ensuring equitable access to treatment for HIV/Aids, universal access to broader health-care needs and rights remains an important unfulfilled societal responsibility[6] with implications that extend to considerations of the global political economy.

In addition to all the above, it should be noted that a public health lens enables us to 'see' that the most powerful health-promoting forces are social. Their effect on health is exemplified by the improved living conditions and use of sanitoria during the 19th century resulting in a ten-fold reduction in mortality from tuberculosis long before effective drug treatment was developed to complete the cycle towards the potential of curing almost all patients with this disease.[7] It has been estimated that social forces, even in a wealthy country like Canada, account for 50% of the causal factors impeding good health. This proportion is much greater in Africa, disadvantaged by a legacy of previous exploitation[8] that continues through internal and external processes[9].

Extension of the interpersonal health ethics discourse since the 2003 SARS epidemic, to include public health ethics, has enabled careful examination of tensions between individual rights and the common good (e.g. quarantine), as well as the evaluation of arguments about how best to balance these conflicting, but mutually valued, ethical perspectives.[10] Greater attention to social justice involves transparent and accountable processes for the allocation of limited health-care resources.[11]

A study of public health lessons from the SARS and Ebola epidemics revealed the ill-preparedness of the World Health Organization and the global community for large/sustained disease outbreaks.[12] Seven themes that were identified as *ethical lessons stemming from such moral failures* and requiring rectification, are also of crucial importance in South Africa. These include recognition that: health systems are fragile and need strengthening to prevent and mitigate future epidemics and pandemics; there is a need for improved surveillance/response capacities and improved communication and community engagement to build trust; effective and rapid response requires leadership at international, national and local levels; and market-based systems do not cater adequately for neglected diseases.[13]

The impact of both the pandemic and of governments' responses, that most profoundly affect the poor majority in our country and globally, are amplified by a multifaceted complex global/planetary crisis[14] within an ecological system stretched to the limits where multiple tipping points[15] into chaos threaten the future of us all. These insights also help to clarify what striving ethically for health means in the world in which the COVID-19 pandemic has emerged and spread so dramatically. This context[16] comprises a multitude of upstream crises that generate considerations of the ethics of the global political economy, international trade, development aid and the creation of crippling debts, and of cruel industrial animal farming[17] and wet markets with their implications for our humanity and our ecosystem. Such problems are aggravated within an energy-intensive market civilisation, driven by belief in endless economic growth, consumerism, the profit motive and free-riding on the environment, with damaging effects on health, especially in low- and middle-income countries.[18]

All the above should be viewed through framings[19] that reveal the ethical dilemmas and power relations[20] relevant to population health and well-being (the socio-political underpinnings[21] of which have previously been identified), and the need for a paradigm shift[22] from a competitive anthropocentric focus towards a cooperative ecological perspective on all aspects of life and health. This agenda for a 'new normal' in a post COVID-19 world could be advanced globally and locally through education and public discourse to foster widespread construction of a more collaborative concept of global health ethics as the rationale for mutual caring[23], in a country and a planet in the throes of entropy[24].

Innovative social action to facilitate sustainable survival[25] is potentially feasible, given human imagination, ingenuity, determination and global political will by those with a vision for the future.

References

1. Johns Hopkins University. Maps: Tracking global cases [webpage on the Internet]. c2020 [updated daily; cited 2020 Jun 08]. Available from: https://coronavirus.jhu.edu/data
2. Labonté R, Schrecker T. The state of global health in a radically unequal world: Patterns and prospects. In: Benatar S, Brock G, editors. Global health and global health ethics. Cambridge: Cambridge University Press; 2011. p. 24–36. https://doi.org/10.1017/CBO9780511984792.003
3. Packer G. We are living in a failed state. The Atlantic. 2020 June; Ideas [cited 2020 Jun 08]. Available from: https://www.theatlantic.com/magazine/archive/2020/06/underlying-conditions/610261/
4. Abdool Karim SS. The South African response to the pandemic. N Engl J Med. 2020;382, e95. https://10.1056/NEJMc2014960
5. Mayosi BM, Benatar SR. Health and health care in South Africa – 20 Years after Mandela. N Engl J Med. 2014;371:1344–1353. https://10.1056/NEJMsr1405012
6. Benatar S, Gill S. Universal access to healthcare: The case of South Africa in the comparative global context of the Late Anthropocene era. Int J Health Policy Manag. 2020 In press. https://10.34172/ijhpm.2020.28
7. Benatar SR, Upshur R. Tuberculosis and poverty: What could (and should) be done? Int J Tuberc Lung Dis. 2010;14(10):1215–1221.
8. Caplan G. The betrayal of Africa. Toronto: Anansi Press; 2008.
9. Stop the bleeding campaign – Africa [webpage on the Internet]. c2015 [cited 2020 Jun 08]. Available from: https://www.globaltaxjustice.org/en/latest/stop-bleeding-campaign-africa-2015
10. Singer PA, Benatar SR, Bernstein M, Daar AS, Dickens BM, MacRae SK, et al. Ethics and SARS: Lessons from Toronto. BMJ. 2003;327:1342. https://doi.org/10.1136/bmj.327.7427.1342
11. Benatar SR, Ashcroft R. International perspectives on resource allocation. In: Quah SR, editor. International encyclopedia of public health. 2nd ed. Waltham, MA: Academic Press; 2017. p. 316–321. https://doi.org/10.1016/B978-0-12-803678-5.00380-5
12. Smith MJ, Upshur REG. Ebola and Learning Lessons from Moral Failures: Who Cares about Ethics? Public Health Ethics. 2015;8(3):305–318. https://doi.org/10.1093/phe/phv028
13. Haffajee RL, Mello MM. Thinking globally, acting locally — The U.S. response to Covid-19. N Engl J Med. 2020;382, e75. https://10.1056/NEJMp2006740
14. Gill SR, Benatar SR. Reflections on the political economy of planetary health. Rev Int Polit Econ. 2020;27(1):167–190. https://doi.org/10.1080/09692290.2019.1607769
15. Pearce F. As climate change worsens, a cascade of tipping points looms. Yale Environment 360. 2019 December 05 [cited 2020 Jun 08]. Available from: https://e360.yale.edu/features/as-climate-changes-worsens-a-cascade-of-tipping-points-looms
16. Benatar S, Brock G, editors. Global health and global health ethics. Cambridge: Cambridge University Press; 2011. https://doi.org/10.1017/CBO9780511984792
17. Benatar D. Our cruel treatment of animals led to the coronavirus. The New York Times. 2020 April 13; Opinion [cited 2020 Jun 08]. Available from: https://www.nytimes.com/2020/04/13/opinion/animal-cruelty-coronavirus.html?referringSource=articleShare
18. Gill S. Globalization, market civilization and disciplinary neoliberalism. In: Hovden E, Keene E, editors. The globalization of liberalism. Millennium Series. London: Palgrave Macmillan; 2002. https://doi.org/10.1057/9780230519381_7
19. Pogge TW. World poverty and human rights. 2nd ed. Cambridge: Polity Press; 2008.
20. Benatar S, Upshur R, Gill S. Understanding the relationship between ethics, neoliberalism and power as a step towards improving the health of people and our planet. Anthropocene Rev. 2018;5(2):155–176. https://doi.org/10.1177/2053019618760934
21. Baudot J, editor. Building a world community: Globalisation and the common good. Copenhagen / Seattle, WA: Royal Danish Ministry of Foreign Affairs / University of Washington Press; 2001.
22. Benatar SR, Gill S, Bakker I. Making progress in global health: The need for new paradigms. Int Affairs. 2009;85:347–371. https://doi.org/10.1111/j.1468-2346.2009.00797.x
23. Benatar SR, Daar AS, Singer PA. Global health ethics: The rationale for mutual caring. Int Affairs. 2003;79:107–138. http://doi.org/10.1111/1468-2346.00298
24. Sandy MA. The Sixth Extinction: An unnatural history. The AAG Review of Books. 2017;5(3):159–161. https://10.1080/2325548X.2017.1315239
25. Bensimon CM, Benatar SR. Developing sustainability: A new metaphor for progress. Theor Med Bioeth. 2006;27(1):59–79. https://10.1007/s11017-005-5754-1

Check for updates

AUTHOR:
Jess Auerbach[1]*

AFFILIATIONS:
[1]Faculty of Education, Stellenbosch University, Stellenbosch, South Africa
*Current: Senior Lecturer, Anthropology, North West University, Potchefstroom, South Africa

CORRESPONDENCE TO:
Jess Auerbach

EMAIL:
jess.m.auerbach@gmail.com

HOW TO CITE:
Auerbach J. More eyes on COVID-19: Perspectives from Anthropology: *What* people believe is a lot less important than *that* they believe it. S Afr J Sci. 2020;116(7/8), Art. #8491, 1 page. https://doi.org/10.17159/sajs.2020/8491

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
COVID-19, South Africa, social sciences, humanities, anthropology

PUBLISHED:
29 July 2020

More eyes on COVID-19: Perspectives from Anthropology

What people believe is a lot less important than *that* they believe it

Anthropology compels us all to see the world from many different perspectives at once. On the rare occasions, such as a pandemic, where we all need to adopt very specific habits from the intimate to the public, these different perspectives must be taken seriously and must inform policy at every level.

Health is not only biophysical, but also emotional, spiritual, environmental and social. Dignity, security, and purpose cannot be achieved alone, but through collective everyday experiences that are currently being radically altered.

In South Africa, the virus has both brought us together and shown up lines of separation that were entrenched throughout our history and have altered very little since 1994. The things that will keep us safe – physical distancing, handwashing, strong immune systems – are simply not possible for many citizens.

Nonetheless, in facing COVID-19, South Africa has a unique opportunity to prove that it will use this historic moment to do well by its population. To respond with respect, compassion, and a recognition of the basic intelligence of all of us: nobody wants their loved ones to get sick and possibly die. What, therefore, can anthropology teach us about an appropriate response to the pandemic?

Firstly, the term ‘social distancing’ should ideally be replaced with the more accurate ‘physical distancing’. This reminds us that human beings are inherently social, and cannot thrive without community. In reality almost nobody is ‘socially distancing’ in South Africa right now. Rather, a vocal ‘some of us’ have the privilege of shifting our human, learning and income-generating connection online.

If we acknowledge that hygiene and physical distancing are the fundamental requirements that currently underpin our constitutionally assured ‘right to life’ we must act strongly. Water, nourishment and *data* must then be free resources for the duration of the crisis. If everyone can eat enough, wash their hands and maintain connections to their loved ones, we are far more likely to emerge from this unbroken.

Belief must be taken seriously. *What* people believe is a lot less important than *that* they believe it. Belief informs actions from the micro to the macro and back again. Religion, news (fake or otherwise), science, statistics and government all demand belief and if people buy into something – regardless of whether or not it is also factual – they will act on it.

Be it online or in person, most people turn to sources that they trust. These may be religious leaders, community level activists, social media stars, educators, or the favourite uncle in the corner store with an interesting opinion. The people we *choose* to listen to direct our behaviour in powerful ways. Partnerships must be established at every level of society, to ensure necessary information is shared in a way that is not antagonistic to existing belief systems. The information must be consistent in translation and modelled in the actions of leadership.

The violent structure of South African society should be acknowledged. COVID-19 is not happening in an historical vacuum, and not all diseases and experiences have been responded to equally to date. The country was not shut down for HIV or tuberculosis or gun warfare or inequality, but the virus will nonetheless run its course on trails worn smooth by these conditions. Acknowledging that this may not be perceived as ‘fair’ and explaining why this time the response is different as the government tries to do better, will help people to make the sacrifices being asked of them with less resentment.

It is important to validate rites of passage. So far, we have only focused on the end of life, and guidelines have been given for funerals. Much more is needed. We must reconsider and develop new rituals for births, transitions into adulthood, marriages and separations, graduations, promotions, and even birthdays. These events help everyday life to continue to hold significance. They make us feel connected, enriched, supported – as a part of something bigger than ourselves, rather than apart from it.

Many of us need help to find meaning: ‘meaning’ makes people happier and more secure, and therefore much more likely to support safe policy and action. ‘Beating the pandemic’ is theoretically a strong motivation, but until someone we know personally dies, COVID-19 will remain abstract for most of us not on its frontlines. Meaning must be given to this rupture.

A national campaign calling for the reimagination of a just and equitable South Africa that is different from the country we left behind could be unifying and productive. It could provide a rare opportunity for *everyone* to reimagine the social contract of citizenship, the individual and the collective and to ensure that when this passes we do not return to where we were before. Anthropology can draw on an extraordinary record of life affirming processes and practices in the face of challenge to support this process going forward.

We live within environments. Our environments should also now be part of our plans. Lockdown has forced many to pay much more attention to spaces around us as we have been more present where we usually come just to sleep. Again, there is a rare opportunity to mobilise towards protection and sustainability – the ecosystems, the animal, bird and insect life, the safety not only of humanity, but of everything else in our world as well.

Finally, our understanding of value needs to be expanded. This is something that anthropology has quietly documented for more than a century. The cultural repertoires that exist within South Africa extend far beyond ubuntu. As we try to mop up the economy, it will be critical to expand the definition of ‘value’ beyond the financial towards the *human, the relational, the social*.

Check for updates

AUTHOR:
Rajend Mesthrie[1]

AFFILIATION:
[1]Department of Linguistics, School of African and Gender Studies, Anthropology and Linguistics, University of Cape Town, Cape Town, South Africa

CORRESPONDENCE TO:
Rajend Mesthrie

EMAIL:
rajend.mesthrie@uct.ac.za

HOW TO CITE:
Mesthrie R. More eyes on COVID-19: Perspectives from Linguistics: Pay attention to how people are talking about the pandemic in different languages. S Afr J Sci. 2020;116(7/8), Art. #8497, 1 page. https://doi.org/10.17159/sajs.2020/8497

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
COVID-19, South Africa, social sciences, humanities, linguistics, speech, communication

PUBLISHED:
29 July 2020

More eyes on COVID-19: Perspectives from Linguistics

Pay attention to how people are talking about the pandemic in different languages

As linguists we claim that language underlies all human activities. Yet we have entered a cold new world with the freezing of the very interactions which must have given rise to language and which are in turn enhanced by it. Placing masks on mouths and noses is a masking of communication too. No linguistic or applied linguistic textbook quite prepares us for the shutdown of speech in an age of viruses. Here I highlight matters of language and communication that have come to the fore in these times.

We need to keep the lights and the technologies of communication on. Older humanities specialists have always been slightly suspicious of the instruments of technology, of their speed – which seems to overtake the thought processes of their users, coupled somehow with their built-in obsolescence. But COVID-19 has forced us to reconsider. *London Calling* to the faraway towns during the war of a century ago was an important, albeit one-way, means of radio communication for survival. The ubiquitous cellphone, computer and television networks of today have made possible a degree of transmission, discussion and action that would have been unthinkable in the age of radio. They have greatly facilitated modes of survival.

But in using them we miss the nuances of face-to-face interaction that they mask or only approximate in a mechanical fashion. It is the use of eye contact, gesture, facial cues, bodily posture and immediate feedback that makes our communication human. Even for more public and relatively one-way communication, we still require the insights of humanities specialists – as interpreters of official and unofficial stories. We have to be alert to poor arguments, as when politicians appear to put their parties and business interests above those of public safety, or when they disobey the fundamental rules they themselves put in place. Our own President has proven an able communicator and leader, showing the wisdom of humility during a global crisis. Critical language awareness helps us sort out the possible from the impossible, and the plausible from the patently false. At the same time, it is important to listen to the voices of those affected most: the weak and poor, the ones with least access to the technologies of independence, information and power.

In this regard, we need to pay attention to how ordinary people are talking about the pandemic in different languages. Communication of meaning does not rest with the scientists, health specialists or presidents alone. 'Hydroxychloroquine' is simply not usable as an everyday word: even a president tongue-tripped over this polysyllabic poser, and apparently not because he is an avid user of the drug. Applied linguists in southern Africa showed a decade ago how people responded in everyday discourse to the ravages of HIV/Aids, moving from silence, taboo and euphemism to circumlocution and paraphrase. This was a new lexicon of coping.

I will cite just a few examples from the present crisis. Unsurprisingly, words like *ikhorona* and *ikhovidi* (and close pronunciation variants) have entered the isiXhosa language at high speed. Zukile Jama (personal communication 2020 May 25) notes that people make links with past epidemics in some of the new terminology (e.g. *ubhubhane*, which links it to the word for the plague and/or Spanish flu of almost exactly a century ago). The word seems to resemble the English word *bubonic flu*, but this might be coincidence, as its root form is *-bhubha* 'to perish, die'. The entry in Doke and Vilakazi's *Zulu-English Dictionary* of 1972 for this verb now takes on a poignant reading: 'to perish, die, be destroyed, become no more, suffer annihilation'. From the virus's side (virus as subject) the causative form *-bhubhisa* is even more chilling: 'to destroy, kill, wipe out, annihilate'. Add to this the noun suffix *-ane*, which marks prolonged activity and the full original force of the word *ubhubhane* can be felt. Thabo Ditsele explains a new term for 'quarantine' – *diagelo* – that he came across in Setswana social media in relation to the present crisis (personal communication 2020 May 22). The word has undergone an extension of meaning from its original use within the traditional African church for 'spiritual cleansing by confinement in the home of a church elder'.

COVID-19 has shown us how fragile we humans really are and how presumptuous our use of language can be. Our once clever metaphors – like our airplanes – have been grounded. Can we ever speak lightly of having the travel bug again? Will we think twice before describing someone as a pest or when using the word 'pestilence'? Or speaking of someone as having an infectious sense of humour? May we ever again describe some facile joke or observation on social media as 'going viral'? The virus is silent and invisible, but 'COVID' and 'Corona' will enter the all-time human lexicon of suffering, ahead of 'tsunami', 'AIDS', 'Holocaust', 'WWII' and 'the plague'. Language innovation serves to record our human hopes, but even more so our fragility in the face of bigger forces in the universe.

AUTHOR:
Barney Pityana[1]

AFFILIATION:
[1]Professor Emeritus of Law, University of South Africa, Pretoria, South Africa

CORRESPONDENCE TO:
Barney Pityana

EMAIL:
bpityana@gmail.com

HOW TO CITE:
Pityana B. More eyes on COVID-19: Perspectives from Religion Studies: How Christian theology helps us make sense of the pandemic. S Afr J Sci. 2020;116(7/8), Art. #8498, 1 page. https://doi.org/10.17159/sajs.2020/8498

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
COVID-19, South Africa, social sciences, humanities, religion studies

PUBLISHED:
29 July 2020

More eyes on COVID-19: Perspectives from Religion Studies

How Christian theology helps us make sense of the pandemic

Religion Studies starts with the assumption that a study of religion does not entail a personal religious conviction but is justifiable as an intellectual pursuit of knowledge about the human condition. It also recognises that there are a variety of ways in which humanity understands and responds to the impetus of religion from within (e.g. faith) or from without (e.g. the environment).

All religion is about seeking – a search for understanding and for answers to questions or a world view that is persuasive and becomes a way of life. The sudden arrival of COVID-19 once again places such searching questions at the centre of human lives. What does the study of religion allow us to understand in relation to the pandemic?

Without taking anything away from the broader study of religions, I limit my focus to my area of specialisation which is Christian theology. Christian theology works with tools in history, philosophy, culture and language studies, and psychology to express this idea of the human, especially in its relationship with and understanding of God, as well as in its moral life. Christian theology, it has been said, is first and foremost about the human and the totality of human relationships. It is through seeking to be fully human that one seeks to understand what God wills for one's life.

For that reason, Christian theology is about the imagination. It is by the power of the imagination that so much of Christian life makes any sense. It is accordingly in myth and symbols, story and meaning, worship and human expression in music, dance and movement, poetry and art, that the story of faith can be truly told and faithfully understood. Mouton describes this phenomenon in apt terms as 'the ability of the human imagination to *redescribe* reality, to *rename* experiences, to *retell* their stories from new angles'[1]. Theology thereby helps one not to be confined to the obvious, natural and material, but to explore the metaphysical world beyond with as much confidence.

To be a Christian is an acknowledgement of God as a 'presence' in your life and in the world, or as one's response to the moment of encounter with a 'living God'. It is a life of confession of Christ in word and deed to be the Saviour of the world, *salvator mundi.* Contrary to what some may believe, the Christian faith is not a transactional religion or about bargaining with God. It is an acknowledgement of God in one's life and in the world. It is rather a compelling reflection and outworking of a life in the midst of a messy world. It makes God 'visible' to the inhabited world and evangelism becomes an offer and an invitation to share the life of Christ.

What this means is that a belief in God and of God's Son Jesus Christ is to be alive to The Other, their significance in making one fully human, the energy derived from relationships, and the moral effect of living together in a community of sharing and loving. This idea of The Other is not only about the recognition or an acknowledgement of God. It is also about the knowledge that The Other is one's neighbour, who has needs and wants.

Christian theology further helps one to understand that not everything that we experience today is known or knowable. It is always being open to new knowledge and understandings that we become human. It is in the surprises of life that the knowledge of God gives substance and meaning.

And yet, Christian theology teaches that human living is never an act in futility but rather one infused with purpose. That purpose is to fulfil God's intentions in Creation. As a moral substance, being human is about taking responsibility and understanding human agency. As agents of God, humanity is continuously a part of the making and unmaking of Creation. If that is so, then choices are made in such a way as to fulfil the act of God.

That explains why Christian thought is able to denounce evil and injustice, and campaign for a fair and a just world. It is out of that ethic of love and peace with justice that a fair and just distribution of wealth is demanded and lasting and viable solutions to the human predicament are sought. Religion and theology therefore is itself a part of science. It lives and breathes science.

Finally, COVID-19 is ultimately never just about the person affected or dying in isolation at a point in time. It is about community and family. The work of caring, of assurance and hope is an ongoing task for those who understand the dynamism of community that is formed and shaped by a selfless community.

After COVID-19 systems must be put in place, families must continue to live wholesome lives, loved ones must be memorialised and humanised and community that has been broken must be formed and re-formed. Religious communities formed as churches, mosques, synagogues, temples, or in ceremonies in African traditional belief systems, for example, become important in sustaining the value of human life beyond death.

Reference

1. Mouton E. Christian theology in the university: On the threshold or in the margin? Herv Teol Stud. 2008;64(1):431–445.

AUTHOR:
Adam Habib[1]

AFFILIATION:
[1]Vice-Chancellor and Principal, University of the Witwatersrand, Johannesburg, South Africa

CORRESPONDENCE TO:
Adam Habib

EMAIL:
Adam.Habib@wits.ac.za

HOW TO CITE:
Habib A. More eyes on COVID-19: Perspectives from Political Science: Insights from the political management of COVID-19. S Afr J Sci. 2020;116(7/8), Art. #8499, 2 pages. https://doi.org/10.17159/sajs.2020/8499

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
COVID-19, South Africa, social sciences, humanities, political science

PUBLISHED:
29 July 2020

More eyes on COVID-19: Perspectives from Political Science

Insights from the political management of COVID-19

It is often said that crises reveal who we truly are. This is as true for societies as it is for individuals. COVID-19 has revealed the bipolarity of South Africa; its deep inequalities and the schizophrenic character that this engenders. The society aspires to and revels in being recognised as world class, politically sophisticated, and socially caring. When this is rendered asunder by the structural realities of our social and economic context, the society quickly degenerates, manifesting in polarised political divides and reflecting embittered social actors embroiled in accusatory and ideologically laden contestation.

When COVID-19 arrived on our shores, our political authorities quickly assembled the medical fraternity's best minds to advise them on how to respond. The President very quickly engaged in consultations with opposition parties and with social actors including business, labour and civic players like religious leaders. A private–public partnership in the form of a Solidarity Fund was announced, and a number of billionaire families including the Oppenheimers, Ruperts and Motsepes pledged a billion rand each. This was rapidly followed by government's decision to impose one of the most rigid and extreme lockdowns announced anywhere in the world. We collectively beamed at this decisive, world-class response. We bathed in the praise of the World Health Organization, of the recognitions of our public–private partnerships, and in the realisation that the political responses of our President and government were based on evidence, data and world-class science.

Then COVID-19 exposed the crude underbelly of South Africa. The lockdown exposed our deep economic divides which manifested in starvation, food riots and increasing incidents of malnourishment in the hospitals, while the affluent retreated into secluded suburbs or fled to their holiday homes. The President's leadership and the government's decisive action were quickly unravelled by the skills-compromised civil service and its acute inability to execute decisions like, among others, the payment of social support grants and the distribution of food parcels. It was also undermined when overzealous police and soldiers abused residents. The social pact itself began to erode as businesses closed, workers lost their jobs, and a variety of stakeholders, including academics and scientists, began to advocate for the opening up of the economy. As all of this occurred, the political consensus within the state and among the ruling and opposition parties quickly disintegrated.

Much has been written about these successes and failures. Yet little has been written about the thinking around the political management of these challenges. To be fair, there has been some public puzzlement about, and maybe even derision of, some of the actions of the state and its regulations on what can and cannot be bought, or when one is allowed to exercise. But beyond criticism at the silliness of some of these regulations, there has been very little attempt to understand what is going on and why it is happening. In a sense we seemed to have assumed that this is simply the result of the idiocy of individual ministers and government officials.

But is a deeper understanding not warranted? Is this not the result of the challenge of politically managing a pandemic in a deeply unequal society? The problem is that pandemics have differential effects in unequal societies. Managing it therefore will also have unequal consequences, thereby making government vulnerable to charges that its actions are more prejudicial to the poor than it is to the middle classes and the rich. This is a serious political conundrum in a politically polarised society, especially for a party that derives much of its support from the poor and marginalised.

One solution to this challenge is to be seen to be imposing as many penalties on the rich as is inflicted on the poor. This was particularly explained to me by Garth Stevens, academic psychologist and Dean of the Faculty of Humanities at the University of the Witwatersrand (Wits). As Stevens puts it, you cannot manage the pandemic in local townships without closing the shebeens. But to do so, while allowing middle-class persons to sip on their chardonnays in their suburban homes, is politically unpalatable. The political response is therefore to impose a comprehensive purchase ban of alcohol on all citizens. Similarly, enforcing social distancing in townships without acting against suburbanites is again politically unpalatable. The result was a comprehensive ban on exercise and then its limitation to between 6:00 and 9:00. Regulation after regulation, however irrational, begins to take on a more significant meaning if it is understood from the lens of mitigating the political consequences of managing the pandemic in an unequal society.

But the challenge of managing the unequal effects of the pandemic is not limited to the politicians of the ruling party. It also infects the understandings of a variety of other stakeholders. This was brought home to me recently as we tried to shift to emergency remote learning at Wits so as not to lose the academic year for our students. This was quickly opposed by student leaders at Wits and nationally and even by some academics and politicians. In their view, social justice requires that if everyone cannot learn, no one should. This, in a sense, was the advancement of a vindictive populist politics of impoverishing all, very much akin to that advanced by the Khmer Rouge in Cambodia, or that entertained in China during its Cultural Revolution.

The problem with this strategy of comprehensive imposition of penalties is that it provokes resistance across the board. Government is then forced to make concessions which are then perceived by some political stakeholders as capitulating to entrenched privileged interests. It therefore would be far better if government and or other stakeholders adopted a more pragmatic, yet progressive, strategy to manage the pandemic. This would hold that social justice does not require a reversion to the lowest common denominator. Instead, it would hold that social justice requires an awareness of inequality, and a conscious attempt to mitigate its consequences through addressing the inequities. One example of this is how Wits and other universities established a computer loan facility, dispatched computers to thousands of students who did not have a device, and arranged 30G of data for students at no cost to themselves. Another is the ZAR500 billion stimulus package – which is in fact a relief measure – that the President announced with the lockdown. This, in a sense, is a political agenda of constructive social justice that is directed to building the capacities of and lending a helping hand to disadvantaged communities, rather than advancing a vindictive populist politics of impoverishing all.

Another element of such a pragmatic yet progressive strategy is harnessing the managerial capacity of the private sector to effect decisions. One of the central challenges in this pandemic is the almost non-existent capacity of the state to deliver and execute decisions. Yet the one feature of highly unequal societies like South Africa is that capacity exists, but that it is not in the public but rather in the private sector. Would it not make sense to harness this managerial and logistics capacity of the private sector for the execution of decisions on the pandemic? Could the enormous logistics capacity in the private sector not be used to distribute food parcels, or should the technological and managerial capacities of the financial sector not be deployed for the payment of social grants and implementing other elements of the relief package? Separately, could the procurement and production capacities of the domestic and international pharmaceutical companies not be deployed for the purchase of testing kits and for the production of drug therapies and vaccines as and when these become available? Could our private hospitals not be enlisted to treat public patients on a cost basis? What is being suggested here is for the state not to be obsessed with centralising the administration and management of the pandemic in a context where its execution capacity is so limited. Rather it would be more prudent for it to harness the distributed capacities in the public, private and civic sector to enable an efficient execution of programmes to manage the pandemic.

The additional benefit of harnessing private sector managerial, logistics and technological capacities is that it could become the kernel of the social pact that the President has spoken of for so long. Social pacts are not always realised in grand negotiating forums. Rather they sometimes evolve in the actual practise of small day-to-day multisectoral partnerships in the heat of a crisis. This is what would in effect occur in this case. More importantly, such a social pact is going to be absolutely necessary if we are to manage the after-effects of the pandemic. Its economic, fiscal and ultimately social consequences are likely to live with us for some time, and it is important that we be able to bring together the collective capacities and energies of South Africa to undertake the trade-offs that are going to be required and to manage the consequences thereof.

Finally, if this pragmatic but progressive agenda of change is to be successful, it cannot be undertaken by stealth. We cannot articulate in public a ‘populist politics of impoverishing all’, and then implement by stealth a pragmatic constructive social justice agenda. This kind of duplicitous political engagement of saying one thing in public and another in private has to come to an end. The President and those around him have to lead with courage, packaging an agenda of constructive change in a language that the broader citizenry would understand. They would have to also openly challenge more populist interpretations of social justice that seem to prevail in both the opposition and ruling parties. The executive leadership and stewardship of this pandemic must thus involve a politically educative and consciousness raising element as much as it would entail a management and execution of decisions that have been decided by the appropriate authorities.

South Africa and the world are in a challenging moment where a virus has not only brought life as we know it to a standstill, but has also exposed the dark underbelly of what we have become as a nation and a world. If we are to come through this with our collective innate humanity intact, then we have to manage this pandemic with less rigid ideology, and a greater pragmatism. This does not mean that we need to abandon our desire for social justice. Rather it requires a recognition that we operate in a world that does exist, rather than one we wish existed. It requires a constructive social agenda of lending a helping hand, rather than one of impoverishing all. Most of all, it requires managing the pandemic in a manner that enables it to serve as a bridge of praxis that would help us to reimagine and rebuild our country and our world in a more inclusive direction. Are these after all not the central lessons we need to learn from both the heroic social justice struggles of the last century and the more recent enriching, yet simultaneously marginalising, experiences of our contemporary era of globalisation?

More eyes on COVID-19: A legal perspective

The unforeseen social impacts of regulatory interventions

AUTHOR:
Thuli Madonsela[1]

AFFILIATION:
[1]Law Trust Chair in Social Justice, Stellenbosch University, Stellenbosch, South Africa

CORRESPONDENCE TO:
Thuli Madonsela

EMAIL:
tmadonsela@sun.ac.za

HOW TO CITE:
Madonsela T. More eyes on COVID-19: A legal perspective: The unforeseen social impacts of regulatory interventions. S Afr J Sci. 2020;116(7/8), Art. #8527, 2 pages. https://doi.org/10.17159/sajs.2020/8527

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
COVID-19, South Africa, social sciences, humanities, law

PUBLISHED:
29 July 2020

It is important to consider the social injustice impact of the COVID-19 responses since the first of them were promulgated on 18 March 2019. It is clear that some of the *Disaster Management Act* regulations constituted unfair discrimination by unnecessarily and unfairly imposing restrictions that inflicted a disproportionately higher burden on poor and historically disadvantaged groups and communities. Examples in this regard were restrictions on public transport, warm cooked food, limiting most commerce to e-commerce, and restricting childcare to parents, and even then subject to a court order if parents live apart. Subordinate instruments such as directions, guidelines and municipal bylaws have elicited similar and sometimes more serious concerns regarding impacts on social life, in particular family functionality.

For me, reality struck when the father of one of our students died. None of his medical colleagues could attend the funeral. This was because the COVID-19 regulations do not only restrict the permissible number of funeral attendees, it also specifies who can obtain a permit to attend a funeral across municipal districts. Even the pastor had to conduct the funeral service via Zoom because he was in a different municipal district from the one in which the funeral took place. Regulation 35 of the *Disaster Management Act* only permits nuclear family members and parents to travel across districts for a funeral. Even a mother- and father-in-law are not included among those who may obtain a funeral travel permit.

But nothing had prepared me for the moment the COVID-19 regulations struck in my own family. This happened when someone I have always regarded as my brother-in-law suddenly passed on. He was the husband of a cousin I grew up regarding as my sister, as my father raised her after her own father died. I realised that I could not legally obtain a permit to travel across provinces. Firstly, flights are only allowed for business while road travel would require a permit, which cannot be obtained for funerals of cousins or their spouses.

In advising government, it is imperative therefore to include a social science lens on the process of conceptualising and drafting the COVID-19 regulations and subordinate instruments. Quite frankly a social science lens is imperative for all policy designs in responding to a pandemic. It is my considered view that the experts consulted on the COVID-19 pandemic should transcend doctors and scientists. This is in recognition of the fact that society is a system, and as in all systems, things are interconnected. By government's own admission, COVID-19 is not only a public health threat, but also a threat to society and the economy. The paradigm that therefore limits the threat of COVID-19, and the responding regulations, to concerns about public health and the economy is therefore myopic.

Children's rights have been particularly disrupted by the restrictions on social intercourse. The implications are particularly dire for single parents and others in the lower social classes who cannot afford paid childcare. The challenge is compounded by the closure of early child development centres. The impact on education has hardly been mentioned by the President, yet it is one of the most devastating consequences of the COVID-19 curbing regulations and related instruments.

It is worth noting that the government's own impact assessment tool, the Social and Economic Impact Assessment Systems (SEIAS), is informed by the same policy design paradigm. SEIAS prescribes the assessment of the likely social and economic impact of any policy or law before Cabinet approves it. It does not appear that SEIAS is consistently applied before regulations and subordinate instruments are approved.

There is also the broader equality duty and related social justice commitment that all state action must comply with in the light of section 9 of the Constitution read with sections 1, 7(2), 195, 237 and the Preamble. These collectively enjoin government to advance equality and other human rights in the pursuit of a society based on democratic governance, social justice and human rights. To do so, government needs expertise and instruments such as the 9-Dimensional Social Justice Impact Assessment Matrix, which requires the leveraging of data analytics to assess the likely disparate impact of any planned policy or law. Experts should help the crafting of an alternative achievement of the purpose without the predicted unfair discrimination or human rights violation or to design a compensation strategy to be implemented concurrently with the disruptive regulation or related instrument.

It seems to me that the drafters of the *Disaster Management Act*, the anchor legislation for the COVID-19 regulations and subordinate instruments, also regarded a social science lens as essential, regardless of the nature of the disaster at hand. Section 5(1)(e) of the *Disaster Management Act* includes academics among the civil society groupings, including experts who should form part of the Disaster Management Advisory Forum.

While the *Disaster Management Act* does not specifically mention lawyers, by default, lawyers are involved because legislative drafting is principally undertaken by State Law Advisors. But this is a problem because State Law Advisors are too close to the executive to see things laterally. There is also the reality that state employees tend to do as they are told and advise as expected. There is not sufficient independence, in other words, to speak truth to power. For this reason, it is important that lawyers from outside be also involved in advising. There may also be inadequate expertise in areas such as impact assessment, including social justice and human rights impact prediction.

Anecdotal information suggests that the opinions of outside experts, including lawyers, are sought now and then by individual Ministers and the presidency. Herein lies one of the challenges of the COVID-19 regulations and subordinate instruments – the problem of incoherence and impact unconsciousness. For example, when the regulations allowed shopping for essential items from supermarkets, public transport was only restricted to essential services.

Also, the impact of closing early child development centres and schools on child nutrition was never detected until the problem of hunger for children became real. The same applies to the implications of poverty and inequality on digital inclusion as education was migrated to online platforms. The impact will be felt in historically disadvantaged communities long after COVID-19.

Even that restriction to essential services was flawed because the operation of taxis and lift hailing services was restricted to the early hours of the morning and evening. Such unintended consequences could have been limited if the multisectoral experts operated under one roof. In this arrangement, the likely systemic impact of any planned regulatory intervention could be assessed from multiple angles before adopted.

There is no gainsaying the fact that the COVID-19 regulatory approach taken to date has saved lives. Yet the paucity of an impact consciousness has likely exacerbated poverty, inequality, mental health challenges, family dysfunctionality and societal vulnerabilities.

It is my considered view that a Multidisciplinary COVID-19 Advisory Forum could help and needs to be established urgently. That structure should include lawyers, educators, sociologists, psychologists, social workers, statisticians, economists, development experts and others. Having pandemic experts and doctors has clearly not been enough.

More eyes on COVID-19: Perspectives from Sociology

The social life of a virus

AUTHOR:
Vasu Reddy[1]

AFFILIATION:
[1]Dean's Office, Faculty of Humanities, University of Pretoria, Pretoria, South Africa

CORRESPONDENCE TO:
Vasu Reddy

EMAIL:
vasu.reddy@up.ac.za

HOW TO CITE:
Reddy V. More eyes on COVID-19: Perspectives from Sociology: The social life of a virus. S Afr J Sci. 2020;116(7/8), Art. #8496, 1 page. https://doi.org/10.17159/sajs.2020/8496

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
COVID-19, South Africa, social problems, social life, social epidemics, sociology

PUBLISHED:
29 July 2020

This brief provides a sociological perspective which is the study of how individuals behave in groups and how human behaviour is shaped by groups. The fact that our societies are complex and are made up of a collection of individual people, means that there is a certain connection between the structures and social problems of society and the behaviour of the people who shape it. Human behaviour is complex, far from robotic, but indeed malleable in the way we can respond to social realities and engage with the structures and social problems facing us, such as this COVID-19 pandemic.

The government-imposed 'self-isolation' has certainly activated a range of perspectives that go beyond the biomedical. There is no single story that fully captures the diverse meanings this virus represents. However, the social dimensions of the virus open up an archive of insights related to the physical sense of our experience as social beings. At a personal and collective level, self-isolation – a dominant theme in current language – is not merely about social distancing as a public health measure. A shutdown, with lockdowns and shelter-at-home orders, closed borders and blocked airspace implies a form of self-imposed exile. These tactics highlight the social, and, by forcing us to physically distance from each other, will have implications for our location and our relationships in the world. COVID-19 has created much dis-ease not simply about *what* we know, but also about *how* we come to know our diverse spaces as workers, students, parents, activists, professionals, etc. If, as sociology teaches us, our social world depends on our relationships with individuals and groups, which includes the freedom to move and interact with one another, then we are undoubtedly also living in precarious times. The regulations in place direct us to disengage (and by extension to be disembodied) from our social environments brought about by a disease that will be long and lasting.

The last century has not seen a disease spread (locally and globally) at this rate or on this scale. It is as hasty and impatient as humans are. Self-isolation and self-quarantine limit people in volume and area while time also characterises the incubation period of infections. Livelihoods, communities, small business enterprises, and families have been disrupted in ways that have changed the fabric and texture of our social lives, with social ills such as crime also a factor. As a first pandemic of zoonotic origin since 2013, COVID-19 marshals a history and social context of similar diseases such as H5N1 avian influenza in 1997; Severe Acute Respiratory Syndrome (SARS) in 2002; a H1N1 variant in 2009/2010; Middle East Respiratory Syndrome (MERS) in 2012; and the H7N9 bird flu in 2013. While biomedical science remains at the coalface of the management, treatment, and public health efforts to curb infections, there are indeed other views that matter. I am reminded of the Croatian-born historian of life sciences Mirko Draž en Grmek (1924–2000) who does not deny the reality of individual illnesses but instead emphasises the 'cultural fabrication' of disease as a concept. The point here is that we cannot ignore the social and cultural implications of *how* a disease comes to be, *how* it is shaped and *how* it impacts people as individuals and in groups.

For these reasons, COVID-19 is not just purely a medical pandemic – it is also a social phenomenon whose uncertainty continues to disrupt our social order and risks shaping our social and public imagination. The virus has given governments licence to organise their security apparatus to monitor, shape and modify human mobility and behaviour, including subjecting people to surveillance. The virus has forced international bodies to pursue 'disease diplomacy' that impacts the global health security regime. And it compels us to challenge how we integrate the dis-ease into our social, economic, political and creative 'culture' and histories. The irony is that 'the social life of a virus' has more to tell us beyond the science of the virus, viz. its anatomy, taxonomy and architecture, but also about the virus' impact on human and social costs, including life and its linked meaning, death. The latter connects to an argument made by a prominent Johannesburg-based philosopher, Achille Mbembe, in his philosophical argument *Necropolitics* where he motivates that the political order also increasingly reconstructs itself as a form of organisation for death that results from war and other forms of violence, including in this argument, disease.

This virus and its meanings therefore move us to bigger questions beyond epidemiology. It directs us to query our existence and our very being during this time. It prompts us to ask deeper questions about ways of knowing and our sense of the world we are in. Because we are social beings, our current mental state and our divided emotions trigger our core. More than that, the virus' socio-economic impact is a big unknown as it is surely to impact the local fiscus and the global economy in unprecedented ways. For instance, how is the outbreak felt differently across lines of race, class, gender, age, disability and geography? How has it changed our social lives and relations with one another beyond families, communities and the wider world? How can the decisions South Africans make today inform our responses to the next pandemic? There are no easy answers to these questions as they warrant deeper investigation. Beyond compromising health systems, economic processes, and the glaring challenges of testing and treatment capacity, this pandemic has also manifested stigma and racist ideologies directed, for example, toward China (in relation to the outbreak's supposed origin in a Wuhan wet market), inflaming tension between China and the USA.

Given that there is no single story, we are only scratching the surface of what is a larger project to delve deeper into more sustained arguments about the virus. It is clear that the social life of COVID-19 is arresting inasmuch as it is assaulting. We are learning to renegotiate our world and appreciate parallels with other interconnected social epidemics that are prevalent – poverty, inequality, violence, climate change, the burden of care to name a few. In its social dimensions the epidemic reveals forms of violence, resistance, resilience and new ways of realigning to a new normal.

In spite of the challenges, COVID-19 has also emboldened opportunities to stave off the crisis in a recognition of our intricate human and emotive connections to the social in how we heal, show solidarity, and how we mourn and grieve. As a social species we are locked in the midst of a struggle to regain our social health, an innate human desire to reconnect with life, and with the social groups that give meaning to our identification and relationships in the world. The social life of a virus tells us, sociologically speaking, as of every science, that there are hidden meanings to be revealed and studied.

More eyes on COVID-19: Perspectives from Education Studies

Schools as organisations and the science of re-opening

AUTHOR:
Jonathan Jansen[1]

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

CORRESPONDENCE TO:
Jonathan Jansen

EMAIL:
jonathanjansen@sun.ac.za

HOW TO CITE:
Jansen J. More eyes on COVID-19: Perspectives from Education Studies: Schools as organisations and the science of re-opening. S Afr J Sci. 2020;116(7/8), Art. #8502, 1 page. https://doi.org/10.17159/sajs.2020/8502

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
COVID-19, South Africa, social sciences, humanities, education studies

PUBLISHED:
29 July 2020

One of the most public of controversies has been about when and how to re-open schools after the pandemic lockdown. The fierce debate on the subject is understandable. In mid-May 2020, about 1725 million learners in more than 156 countries worldwide were affected by school closures. Everywhere in the world, parents send their children to school to learn on the assumption that their loved ones will be safe and secure. COVID-19 has, at the very least, caused us to question that assumption.

How therefore can education as a field of study help answer questions about the re-opening of schools? One of the most powerful insights available comes from studies of schools as organisations.

Schools are complex places. They have been described as ecologies, in that many different kinds of human actors exist in this placed called school. They interact with each other and often depend on each other. These humans occupying the organisation govern and manage, teach and learn, serve and administer, lead and follow. What this means is that to simply focus on children is to miss the presence and interactions among teachers, secretaries, kitchen staff, cleaners, caretaking staff, delivery personnel, principals, deputies and parents who pass through the organisation throughout the day.

This observation has serious consequences in a pandemic. It means that an invisible virus can enter, live and thrive in this ecosystem, whether or not children are vectors of the disease.

Schools are compact places. There is a solid wall and fence (most times) cordoning off the school from the outside world, a secure and guarded entrance, locked doors into the schools, and occupied classrooms. In most South African schools, those classrooms are packed with learners, even more so when departmental budgets were sliced. Schools determine class sizes depending on the number of teachers they can pay. This observation has direct consequences for how to think about social distancing, especially at that point when all children are back at school.

Schools are contrived places. Children do not roam around freely. They are confined within and move between classrooms. There are breaks that bring students out of classrooms into larger congregations and call them back into confined spaces. The curriculum distributes teachers to some classes as specialists and to others as supervisors when a teacher is absent. Between classes, teachers live in staffrooms that become more or less occupied depending on the time of day or the calling of special events such as staff meetings. In classrooms, students share pencils; in the school library, books; and on the playground, balls and bats.

This observation means that schools are highly mobile places in which streams of human beings move past each other, touch each other, hug and tackle each other.

Schools are sometimes chaotic places. A landmark study describes many South African schools as '(dis) organisations'. Timetables are unpredictable. Teacher absenteeism is high. Basic resources are in short supply. Students come late and leave early. At any point in the school day there are children outside and strangers hanging around the plant. The fact that school principals and teachers are on the school grounds does not mean that active teaching is being done.

This observation implies that the organisational discipline required for managing people, executing plans and organising resources for mitigation purposes would be severely compromised in such dysfunctional contexts.

Why do these aspects of schools as organisations even matter in a pandemic? Because simply talking about opening schools without accounting for how more or less than 1000 humans live and learn in this bounded organisation would have serious implications for the spread of a virus at close quarters. There are no regulations that can fully or consistently manage these many living, moving and interacting elements of an organisation for 6 h a day and for 5 days a week over the months that the pandemic rages in the broader society.

The organisation of a school has a direct bearing on the social and educational lives of those inside of them. School climate studies, for example, have shown that a healthy, positive school environment invariably leads to feelings of well-being and improved academic attainment. Now place children coming out of an extended lockdown in these complex, compact and contrived environments during a pandemic and immediately there could expected to be fears, anxieties and other kinds of distresses that are likely to affect living and learning inside these organisations.

One policy option is to invest in the re-organisation of schools to receive children during a pandemic – as in the case of a phased return of children by grade. Another is to close the schools until such time that safe, secure and well-managed organisations can be put in place given the high risk of infection under the conditions described.

What the complexity of schools as organisations demonstrate is that re-opening decisions based on epidemiological judgements alone can place thousands of lives at risk.

COVID-19 and the academe in South Africa: Not business as usual

AUTHORS:
David W. Hedding[1]
Michelle Greve[2]
Gregory D. Breetzke[3]
Werner Nel[4]
Bettine Jansen van Vuuren[5]

AFFILIATIONS:
[1]Department of Geography, University of South Africa, Pretoria, South Africa
[2]Department of Plant and Soil Sciences, University of Pretoria, Pretoria, South Africa
[3]Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa
[4]Department of Geography and Environmental Science, University of Fort Hare, Alice, South Africa
[5]Centre for Ecological Genomics and Wildlife Conservation, Department of Zoology, University of Johannesburg, Johannesburg, South Africa

CORRESPONDENCE TO:
David Hedding

EMAIL:
heddidw@unisa.ac.za

HOW TO CITE:
Hedding DW, Greve M, Breetzke GD, Nel W, Jansen van Vuuren B. COVID-19 and the academe in South Africa: Not business as usual. S Afr J Sci. 2020;116(7/8), Art. #8298, 3 pages. https://doi.org/10.17159/sajs.2020/8298

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
research, National Research Foundation, higher education institutions, South Africa

PUBLISHED:
29 July 2020

The famous R.E.M. song laments 'It's the end of the world as we know it, I had some time alone, I feel fine…'. Many South Africans would agree that COVID-19 signals the end of the world (or business) as we know it, and through the lockdown we have certainly had some time alone. But contrary to the lyrics, all may not be fine, especially for South Africa's scientific community.

The novel coronavirus SARS-CoV-2 has impacted every economic and social sector[1] across the globe, including higher education in South Africa. Every student and staff member at a higher education institution will have been affected in some way and to varying degrees; not one person will emerge from this unscathed. It is impossible to predict every short- and long-term impact of the COVID-19 pandemic, but we will experience the aftershocks for a long time to come. Here we discuss some of these impacts, ranging from undergraduate level to large research projects, and we offer suggestions on how to mitigate some of the damage.

At undergraduate and Honours levels, several higher education institutions have had to scramble to place study material online for students. Out of necessity, contact universities have had to develop innovative and flexible ways to offer both theory and practical components to students, and find alternative forms of formative (and most likely summative) assessment. However, academic staff at contact universities typically have little, if any, experience or training in the pedagogy or delivery of online learning. Thus, academics with teaching responsibilities will have had to upskill and familiarise themselves quickly with online learning platforms and all that they entail, including increased administration. The #FeesMustFall protests of 2015 may have prepared some faculties or universities hardest hit by earlier student protests for this transformation, but the total shutdown of almost all sectors of society has presented unprecedented challenges. In addition, COVID-19 has compounded the difficulties experienced at some universities in South Africa which were negatively impacted by staff or student strikes at the start of 2020. An encouraging aspect of this exercise is, however, that the shift to online teaching forces academics to truly interrogate and re-evaluate their curricula. It encourages a move away from 'rote learning' to focus more on problem solving, critical thinking and applied understanding by using a holistic and integrated approach, because traditional assessment techniques are no longer viable. But in a country where a large percentage of students depend on financial assistance to make ends meet, where data costs are high and even a mobile connection may not be readily available to all, and where devices such as laptop computers are seen as a luxury, it is not surprising that contact universities have faced push-back from students who have argued that universities cannot expect them to continue with online learning without providing the necessary resources.[2] To address the resource issue, universities have negotiated with several cellular networks to make data available to students (at a cost to the university, thereby forcing universities to reshuffle their financial budgets and/or asking the general public to donate to discretionary funds), and various universities are already providing devices to disadvantaged students. Moreover, a limited number of educational websites have been made data-free to students, although the largest online platform (Blackboard and its various platforms) is not hosted in South Africa and, therefore, cannot currently be accessed for free. (There are plans to move it to South Africa, but this will not be in place for several months.)

Notwithstanding all these remedial efforts, they do little to help students living in remote areas where electricity supply is inconsistent and network coverage is poor. In addition, not all university staff own a laptop or a personal home computer; neither do all have proper access to the Internet from home. Moreover, the devices that they do have may need to be shared with a spouse or with children who are being home-schooled. In particular, historically disadvantaged universities in rural areas have fewer resources to support their students and teaching staff. Even better resourced universities, such as the University of South Africa, which is an open distance learning institution and arguably best suited to address the challenges that this pandemic presents, has had to grapple with making laptops available for staff and devising mechanisms to run its internal operations and administration remotely, as well as to find alternatives to traditional sit-down examinations for hundreds of thousands of students. A notable number of courses have switched from traditional sit-down examinations to continuous assessment; however, large classes (sometimes in excess of 400 or 500 students) in many faculties renders this extremely difficult and its administration near impossible. In addition, many disciplines require compulsory experimental training to complete professional courses or retain accreditation with international bodies; these activities remain prohibited under lockdown. Here, the answers remain elusive.

While consideration and attention has been focused on moving teaching content online – and rightly so – the neglected 'elephant in the room' concerns the broader impact of COVID-19 on research in South Africa. These impacts include those on supervising postgraduate students, meeting research output targets, submitting new grants to secure the next cycle of research, and meeting research funding and project deliverables. What the impact will be on the country's overall research output remains to be seen. Some research[3] has argued that academics have more time for research during lockdown, with data sheets being taken out of proverbial bottom drawers, dusted off, and turned into publications. However, the research outcome will be unique to every person. For example, working under heightened anxiety may limit research productivity.[4] Academics with young children will feel this especially as they juggle childcare and work in the same household. Other research[5] suggests that women's productivity is likely to suffer more than men's during the pandemic because, even in many higher earning households, women remain the primary caregivers and, as such, childcare and home-schooling fall predominantly on their shoulders. On the other hand, the psychological effects of enforced solitude may be severe for those having to isolate without a partner or family. And equally, early-career researchers may be affected more by the lockdown than senior researchers with established laboratories.[5] These examples serve to illustrate that we live in a complex world with a plethora of real-world problems, and that the experiences of students and staff during the COVID-19 pandemic cannot be generalised. It is, therefore, inappropriate to highlight any one particular group as being more or less vulnerable than another, but rather to see the collective in this situation together.

From a practical perspective, many research projects will be compromised by lockdown and social distancing regulations. The limitations will affect, in some way or another, Honours students who now have no access to a laboratory on campus (or the field) to run their experiment, to many A-rated researchers who will likewise struggle to achieve the objectives of funded research and may fail to meet international obligations. While academics from the natural and physical sciences will mainly be affected by the lack of access to field study sites and laboratory facilities during the various phases of lockdown, many social scientists may be affected far longer as social distancing strategies persist. Social science research often relies on interviews, focus groups, and survey questionnaires; thus, these researchers face the risk of exposing themselves, or the communities in which they work, to the virus. Most geospatial modelling predictions indicate that many poorer communities will be hardest hit by the disease. Therefore, post-COVID-19, they may not wish to participate in research, even after the pandemic and life has returned to the 'new' normal in the following months and years.

On the other hand, COVID-19 may also present opportunities for research, particularly for social scientists. The proliferation of COVID-19-related research being disseminated through preprints attests to this opportunity.[6] Wherever possible, postgraduate students and researchers from all disciplines are required to be innovative in terms of running experiments, collecting data and redesigning postgraduate projects. Examples might include using remote-sensing methods for long-term monitoring studies, mining older data sets, extracting information from large online data sources such as the Global Biodiversity Facility or the Scientific Committee on Antarctic Research's various data portals (to name but a few), or conducting reviews or meta-analyses of existing studies.

The overarching point is that academia, including all researchers and administrators, and associated bodies such as the South African National Research Foundation (NRF), should be conscious of the impacts that this pandemic will have on every academic and student in terms of research, and will have to devise strategies to facilitate the research of all who are impacted. Academia, as a whole, needs to be cognisant of the plight of all researchers and be aware of claiming that specific groups are disproportionally affected. How individuals experience this period will be highly individualistic. In different ways, this pandemic will touch everyone, and it is the responsibility of academia to show compassion during this difficult time. Field and laboratory experiments may fail, collection of data in long-term studies will suffer, the submission of journal articles for review may be delayed, and the submission of grant applications may be deferred (although several institutions have indicated that they will be submitting more grant applications than in other years), and this is to be expected. The question is what can (and should) be done to benefit the entire research community during this difficult time? What mechanisms and strategies can higher education institutions and the research sector in general put in place to facilitate the continuation of research and save the aspirations of students and staff alike?

Strict time frames to completion are imposed on postgraduate students, both by the NRF by limiting the number of years of student financial support as well as by universities that require motivations from students unable to complete their degrees within the allocated time. Students whose research can be conducted entirely online or is conceptual in nature, or whose data have already been gathered, may be less affected by the lockdown. However, students who still need to collect data in the field, or to perform laboratory experiments, will experience significant challenges in completing their research on time. Where fieldwork is season dependent, students may lose an entire year. A case in point concerns students funded through the South African National Antarctic Programme. The annual Marion Island Relief Voyage in 2020 was cancelled, and unless a second voyage can be scheduled later in the year, students will lose an entire year's data. Long-term data sets that have continuous sampling over several decades will suffer. The NRF and various universities should carefully assess requests for extending student support. Despite cost implications, it is certain that the long-term benefits would outweigh those costs. These 'costed' extensions are urgently needed, particularly for postgraduate students in South Africa.[7] Simply deferring student registrations may not be the most sensible option as numerous students will disappear from the system (the leak in the pipeline will increase notably) as background and socio-economic circumstances[8] may not allow students to spend a year being idle and simply re-enter the pipeline next year.

The majority of NRF grants are cyclical; unspent money has to be returned at the end of each year. Where projects are in the middle of a cycle, researchers can motivate for a roll-over of unspent funds only under specific financial rules and these should be carefully re-examined. In addition, where projects were due to end in 2020, consideration might be given to allow projects to extend for another year so that they can be completed (the so-called 'no-cost extensions' being considered by the Royal Society of London and the Wellcome Trust, the German Research Foundation and the Swedish Research Council). This is especially required for projects that require field-based work, or projects involving foreign or international partners who will not be able to travel for months. Travel bans will also prevent international conference attendance and thus the presentation of the latest research, during which international collaborations are set up. Collectively, working from home during lockdown, heightened anxiety, and travel bans – coupled with countries coming out of lockdowns at different times – may also affect the number of local and international (as part of larger multinational collaborations) grant applications that are submitted.

Another possibility may be to extend deadlines for grant proposals or stipulate no deadlines. In South Africa, as elsewhere[9], the NRF extended the closing date for the 'One Call' from 30 April to 15 May 2020, but this was only done after the internal closing dates for most of the institutions had passed. Thus, the benefit of this extension sits with the designated university authorities, some simply using the extension to extend the time available for screening and providing feedback to their own researchers. The German Research Foundation provides a workable alternative because most of its funding calls have no deadline. We suggest that the NRF consider a second funding call in September or October 2020, once there has been time to scrutinise the number of researchers who submitted funding applications during the 2020 One Call. In addition, a reallocation of some existing subsidies to universities should be considered. For example, higher education institutions that pay out publication subsidies from the Department of Higher Education and Training directly to academics' salaries should consider retaining these funds to help researchers support their own research and fund postgraduate students.[10] In short, we need flexibility in a time of crisis.

The South African government has pledged billions of rands to kickstart the economy during and after COVID-19 and this itself may have severe implications for how much funding will be channelled into academia and research. If COVID-19 has any lessons, it is the importance of research across various spheres (natural and social sciences, economics, medicine, etc). We earnestly hope that Treasury and those government departments that drive research will continue to support these endeavours and even increase funding to expand research. Governments across the world are basing their decisions on advice from scientists and research panels, and it seems that how well countries fare in curbing the spread of COVID-19 depends on how well they implement research advice. Cutting funds from research departments may therefore be fundamentally flawed, not only now but certainly in any future crisis.

In conclusion, the pandemic has taught us that academics must be innovative in the way we do our science and facilitate learning. We also need to be compassionate to fellow academics and students: we are all in this together. Several research groups have set up platforms of communication (e.g. Zoom, WhatsApp, Google Group) to remain connected, to plan research and to maintain group morale and a research culture while working remotely.[11] We suggest that now is the time to forge strong supportive collaborations enabling South African researchers to stand together and support one another, particularly in light of possible future austerity measures.[12] Collectively, we need to ensure the well-being of our colleagues, of our postgraduate students who fear that they may not complete their degrees on time, and of our undergraduate students, many of whom feel rudderless. And foremost, we need to safeguard our own physical and mental health.

References

1. Carruthers J. Sustainability in an era of emerging infectious diseases. S Afr J Sci. 2020;116(3/4), Art. #8043. https://doi.org/10.17159/sajs.2020/8043
2. Molosankwe B. UJ students reject planned online teaching to start next week. The Star. 2020 April 17, News.
3. Fleming N. Shut-in scientists are spending more time on research papers. Nature Index. 2020 April 9, News.
4. Seedat S. COVID-19: Let's not forget about the pervasive mental health impact and let us act. NRF Science Matters, 2020 May, p. 8–13.
5. Minello A. The pandemic and the female academic. Nature. 2020 April 17, Worldview. http://dx.doi.org/10.1038/d41586-020-01135-9
6. Anon. Scientific research on the coronavirus is being released in a torrent. The Economist. 2020 May 7, Science & Technology
7. Husby A, Modinos G. COVID-19: Research after the pandemic. Nature. 2020;580(7803):185. http://dx.doi.org/10.1038/d41586-020-01031-2
8. Mangoma A, Wilson-Prangley A. Black tax: Understanding the financial transfers of the emerging black middle class. Dev S Afr. 2019;36(4):443–460. http://dx.doi.org/10.1080/0376835X.2018.1516545
9. Matthews D. Funders extend deadlines as coronavirus disrupts research. Times Higher Education. 2020 March 21, News.
10. Hedding DW. Payouts push professors towards predatory journals. Nature. 2019;565(7739):267. http://dx.doi.org/10.1038/d41586-019-00120-1
11. Powell K. Science-ing from home in the coronavirus era. Nature. 2020;580(7803):419–421. http://dx.doi.org/10.1038/d41586-020-00935-3
12. Czerniewicz L. The struggle to save and re-make public higher education. University World News. 2020 April 30, South Africa.

AUTHOR:
David R. Walwyn[1]

AFFILIATION:
[1]Graduate School of Technology Management, Faculty of Engineering, University of Pretoria, Pretoria, South Africa

CORRESPONDENCE TO:
David Walwyn

EMAIL:
david.walwyn@up.ac.za

HOW TO CITE:
Walwyn DR. Teaching on the edge of chaos: Report on 'The future of universities in a post-COVID-19 world'. S Afr J Sci. 2020;116(7/8), Art. #8404, 2 pages. https://doi.org/10.17159/sajs.2020/8404

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
higher education, technology, SARS-CoV-2

PUBLISHED:
29 July 2020

Teaching on the edge of chaos: Report on 'The future of universities in a post-COVID-19 world'

You might well ask, what more could or should be said about COVID-19? Don't we all have COVID-fatigue? Is it not clear that everything, not just universities, will change post-pandemic?

Yes, to all of these questions, but in this latest edition (number 5) of the ASSAf Presidential Roundtables on 'Science, Scholarship and Society', Professor Jonathan Jansen, ASSAf President and Distinguished Professor in Education (Stellenbosch University), chaired a lively, engaging and important discussion on 'The future of universities in a post-COVID-19 world'. The panel consisted of Professor Ruksana Osman (Deputy Vice Chancellor: Academic, University of the Witwatersrand), Professor Laura Czerniewiecz (Director of the Centre for Innovation in Learning and Teaching, University of Cape Town) and Mr Stafford Masie (General Manager of WeWork South Africa): two academics with backgrounds in pedagogy and teaching methods, and a technopreneur with a background in private education and entrepreneurship.

Universities are both barometers and leaders; they make visible the dynamics and tensions within society, and at the same time they influence the society of the future. It is broadly acknowledged that knowledge and power are inextricably linked, and if not knowledge and power, then certainly knowledge and agency.[1,2] A university education strengthens individual capability, allowing graduates to participate more actively in public life, and to reap the benefits of social and economic freedoms that accompany such agency. The future of universities is, therefore, an important question. As articulated by Walker[3], 'the choices we make about higher education are also choices about what kind of society we wish to build'.

Universities perform a diverse set of functions, including teaching, research and community outreach. The panel discussions, however, focused almost exclusively on the teaching function, with specific attention to *how* modules are taught, rather than *what* is taught. In particular, panel members presented their views on the longer-term implications of digital technologies, the use of which has been accelerated by the pandemic, on pedagogy or teaching method.

This scope was outlined by Jansen, whose introduction noted that the campus closure due to COVID-19 was not a unique event. For instance, the #FeesMustFall period had already catalysed the adoption of online learning techniques and the development of infrastructure necessary for students to complete their courses without having to attend lectures or tutorials. However, the disruptions due to COVID-19, accompanied by the growing viability of alternatives to conventional teaching methods, were unprecedented in his view, incubating three important questions which he posed to the panel: 'Is the brick and mortar university, where students have a social and educational experience, a thing of the past? Does this mean the end of the lower tier universities? Is this the end of the residential university?'

All three panellists agreed that this was a pivotal moment presenting significant challenges and requiring innovative responses and solutions, developed collaboratively across disciplines and sectors. Face-to-face tuition, delivered through lectures (this format is called synchronous learning), is no longer the baseline; new approaches are required. But exactly what solutions, and what dangers to avoid, were widely dissimilar and provided the real interest of the discussions.

Learning moments and game changers for teaching practice

In her presentation, Professor Osman listed her top three learning points from the pandemic: it has highlighted the failure of South Africa's development project (broadly conceptualised as extending access to higher education across race, class and gender); it has aggravated the high pressure of the academic environment which is eroding the possibility of a work–life balance; and it has shown that both faster communication and shorter decision-making is possible (silos can be crumbled!).

In her view, the pivotal moment of COVID-19 is the need to rethink teaching practice (pedagogy), and hence revise the way in which teachers are educated. With her professional background in education, it is not surprising that pedagogy was the focus of her talk. She observed that we need to 'rethink teaching practice, develop approaches which support learning without the need to attend lectures, deepen professional practice and change the way in which we train teachers'. Osman also raised the question as to whether online resources could include material from top tier universities, supporting the 'equalisation of knowledge practices'.

In summarising, she noted that COVID-19 offers the opportunity to explore alternatives. It is not one or the other, but blended approaches that will be important. Moreover, she stressed that we need to continue the traditions of collegiality and collaboration as the core of our approaches to the development of those alternatives.

Gramsci, social justice and open education

Professor Czerniewiecz raised an entirely different set of issues, informed mainly by her concern over the marketisation of higher education. Quoting Gramsci's renowned saying 'I'm a pessimist because of intelligence, but an optimist because of will', she described the COVID-19 moment as 'teaching on the edge of chaos' and facing a future that 'no-one knows what it will look like'. Although the transition might appear as a phased move to online education, in fact it is an emergency response to the collapse of campus-based teaching.

Admitting that she might be accused of adopting a dystopian, apocalyptic perspective, she declared that big capital's initiatives in the sector are undermining the public good characteristic of higher education. The combination of marketisation – the trend whereby universities are encouraged to act like private businesses rather than public institutions – and budget cuts had threatened the value and existence of public education, particularly in the USA (see recession_reality.com) – a trend which would further entrench inequality and weaken the values of social justice.

Czerniewiecz noted that the pandemic had been a gift to companies that offered online higher education and teaching tools, generating new billionaires and multibillion-dollar corporations, including the likes of Zoom and Turnitin. The advent of surveillance capitalism had invaded personal lives and rights to privacy, allowing technology to digitise human behaviour and enable unethical profit-seeking. In her view, the problem was not technology per se, but the business models within which it was being applied.

On the optimistic side of the Gramsci quote, Czerniewiecz observed that digital technologies will support blended approaches to learning and enrich such experiences. She considered that the tension between technology, education, equality and social justice has become a wicked problem which will require a multidisciplinary effort to understand and resolve. She stressed that South Africa has an almost unique perspective in the sense that our education policies had retained imperatives for the attainment of social justice. Many countries had already abandoned such values in favour of a strictly instrumental view of higher education (its value lies only in the extent to which the educated are able to contribute economically).

Finally, Czerniewiecz considered that we now have the opportunity to build high-quality online and blended learning materials, and especially to adopt open education as the dominant practice, resonating with the earlier call by Osman to improve access. However, such perspectives on open education and the danger of surveillance capitalism were not shared by the final speaker on the panel, Mr Stafford Masie, the General Manager of WeWork South Africa, who stridently challenged the public good sentiments presented by the previous speakers.

Big tech will deliver a 'record label moment' to higher education

Masie began with the statement 'I love listening to academics talking about their universities', and then proceeded to ignore or demolish their arguments. Labelling academics and their institutions (public universities) as short sighted, resistant to change, and unaware of broader trends, his criticisms were explicit and unambiguous. Referring to the impact of COVID-19 on education, he noted that 'this is going to be massive … this is going to be a reboot of education'. The pandemic is accelerating everything, and he saw the universities as being in denial.

Informed by his entrepreneurial and technology-rich background, Masie declared that the pandemic had 'laid the university model bare' and that this was their 'record label moment' (the latter refers to the failure of record companies to acknowledge and hence respond to the threat of digital streaming on their core business model). The COVID-19 shock had forced parents and scholars to evaluate alternatives to conventional practices including working from home, online teaching, online retail and other services.

More profoundly, these experiences, he believed, had led many parents and students to realise that the cost of formal education was 'not worth it'. Even if their initial experiences of such platforms had been disappointing, Masie stressed that the universities should not underestimate the power of entrepreneurs to fill the gaps. He warned complacent academics that big tech is 'on the doorstep' and that big tech companies such as Google and Apple were targeting education and health care as new business areas. By injecting higher education with cognition and thinking in a skeuomorphic way, technology-based companies could replace public higher education with online learning platforms.

Masie is clearly a proponent of disruptive change – hoping perhaps that his strong opinions on the present model for higher education will be performative and lead to the opportunities that he outlined. In his final remarks, he expressed a more collaborative proposal, in which he outlined the importance of co-creation, and a new architecture for developing solutions, allowing non-education sector participants in addition to the existing incumbents. 'Only the whole of humanity can solve the problem.'

This report attempts to convey the intensity of the discussions, with substantial differences in opinion on how higher education can, or will be, forced to respond to the COVID-19 disruption of its teaching activities. All presenters agreed, not surprisingly, that this is a pivotal moment – which presents both a threat and an opportunity. The opportunity, in their view, lies in transforming the baseline for teaching from the face-to-face synchronous approaches of the past, to asynchronous, blended and active learning, which were not specifically detailed but indirectly noted as necessary additions to present teaching practices and designs.

The panel's differences in opinion, however, were profound and reminiscent of the conceptual frameworks widely used in the field of transition studies, and particularly the insights of the multi-level perspective.[4,5] This perspective states that opportunities arise when an exogenous event (in this case, COVID-19) destabilises the landscape, which, in turn, weakens the dominance of the existing socio-technical regime (in this case, the public universities). At such moments, niche actors (in this case, the educational entrepreneurs) can successfully challenge the regime and become dominant.

Given the strong arguments in favour of higher education as a public good[6], and as a means of rebuilding social justice, the demise of public universities would be an undesirable outcome for South Africa. The market would be unlikely to respect the plea for a pedagogy of engagement which goes beyond both physical and epistemological borders, as made by Osman in her closing remarks. Neither would it be concerned about the historical goal for education as a means of raising public good overall, rather than to increase the access of one group at the expense of another. It may be that we, as academics, are indeed 'teaching on the edge of chaos' in Czerniewiecz's words, but it is critical that this chaos does not degenerate into Masie's 'record label moment'. At this point, we should be reminded of another of Gramsci's quotes: 'The old world is dying, and the new world struggles to be born; now is the time of monsters.'

References

1. Maton K. Knowledge and knowers: Towards a realist sociology of education. Milton Park, Abingdon: Routledge; 2014.
2. Foucault M. Power/knowledge: Selected interviews and other writings, 1972-1977. New York: Pantheon; 1980.
3. Walker M. No curriculum transformation without epistemic justice capabilities. Pretoria: University of Pretoria; 2019.
4. Geels FW. Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study. Res Policy. 2002;31(8–9):1257–1274.
5. Geels FW. Socio-technical transitions to sustainability. In: Shugart H, editor. Oxford research encyclopedia of environmental science. Oxford: Oxford University Press; 2018. p. 1–37.
6. Walker M. Dimensions of higher education and the public good in South Africa. High Educ. 2018;76(3):555–569. https://doi.org/10.1007/s10734-017-0225-y

The emerging public health risk of extended electronic device use during the COVID-19 pandemic

AUTHORS:
Verusia Chetty[1]
Alvin Munsamy[2]
Saul Cobbing[1]
Diane van Staden[2]
Rowena Naidoo[3]

AFFILIATIONS:
[1]Discipline of Physiotherapy, University of KwaZulu-Natal, Durban, South Africa
[2]Discipline of Optometry, University of KwaZulu-Natal, Durban, South Africa
[3]Discipline of Biokinetics, Exercise and Leisure Sciences, University of KwaZulu-Natal, Durban, South Africa

CORRESPONDENCE TO:
Verusia Chetty

EMAIL:
Chettyve@ukzn.ac.za

HOW TO CITE:
Chetty V, Munsamy A, Cobbing S, Van Staden D, Naidoo R. The emerging public health risk of extended electronic device use during the COVID-19 pandemic. S Afr J Sci. 2020;116(7/8), Art. #8530, 2 pages. https://doi.org/10.17159/sajs.2020/8530

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
musculoskeletal health, visual health, sleep impairment, digital eye syndrome, health awareness

PUBLISHED:
29 July 2020

The experience of being in lockdown, isolated, and socially distant, even from family and friends, has become the ‘new normal’ during the COVID-19 pandemic. People across the world are increasingly using electronic devices (e-devices) to connect socially, for education, and for work. Anecdotally, voices have emerged that predict that COVID-19 will revolutionise how we interact beyond the pandemic, enabling people to connect over large geographical areas. However, the increased use of e-devices may have far-reaching consequences on body systems, including the musculoskeletal and visual systems. These complications are often overlooked due to the focus on the medical management of COVID-19. Attention should be paid to these complications as they have the potential to become a serious public health issue.

Musculoskeletal impairments such as back, neck, shoulder and wrist pain are associated with sustained postures during e-device use. This association has been highlighted in children, young adults and higher education students, and has the potential to become more severe as people age.[1-5] Varied postures, specifically increased head and neck flexion[6], as well as the incorrect placement of the e-device on surfaces, are of concern. The size and weight of e-devices can further cause fatigue, and negatively affect the biomechanics of users.[7] Furthermore, e-device use is associated with other negative behaviours, such as reduced time exercising and playing. Overuse of e-devices in children has been shown to be associated with obesity[8], impaired physical and cognitive development, as well as sleep problems[9]. These challenges have the potential to add a significant burden to already over-stretched primary health-care systems.[10]

Additionally, users’ environments also influence musculoskeletal impairments. Many people in low- and middle-income countries live in resource-scarce home environments. They lack a suitable ergonomic set-up for the correct use of e-devices, or information on managing screen time and exercise. These users are also more likely to use a smartphone for online engagement, as opposed to a laptop, and may sit on low beds, floors or outside the home to access data coverage for online platforms. In addition, the restrictions on daily free movement are likely to result in people living more sedentary lifestyles, which can, in turn, increase the prevalence of other health challenges such as diabetes and obesity, and impair child brain development.

Prolonged screen time further leads to associated eye health concerns for people of all ages. Even before COVID-19, the rising prevalence of myopia, or short-sightedness, was acknowledged as a global public health problem, predicted to affect five billion people by 2050.[11] Central to this increasing incidence are environmental factors such as increased time indoors, increased educational demands, and increased use of e-devices. Lockdown measures, precipitating a move towards working remotely or learning online, have forced people to spend significant time indoors and in front of e-devices. This prolonged use of e-devices, particularly at close working distances, results in a condition known as digital eye syndrome (DES).[12,13] DES produces clinical symptoms such as headaches, caused by stress on the accommodative and binocular vision systems, as well as reduced blink rate and poor blink quality.[12-14] This is compounded by the blue light emission from light-emitting diode (LED) devices. Blue light also affects sleep latency and duration by reducing melatonin production, which disrupts circadian rhythms.[12] The impact of excessive e-device use may, therefore, extend beyond DES and musculoskeletal problems, to increasingly disrupted sleep patterns, which are associated with emotional distress and cognitive deficits.[12-14]

Usage of e-devices is likely to increase during the COVID-19 pandemic. This, in turn, increases the likelihood that associated visual, musculoskeletal and developmental impairments will become more prevalent and severe, thus adding a significant burden to already over-stretched primary health-care systems. We, therefore, need to develop strategies to provide better information on how to adopt preventative measures that address both eye health and the musculoskeletal impairments associated with the increased use of e-devices. These strategies should include innovative ways to use e-devices and programmes that educate communities on appropriate measures to counteract the adverse effects of screen time. Public health initiatives should involve collaboration with various sectors, including community partners, to establish and integrate contextually tailored health awareness programmes into countries across the globe. An immediate public health response for health-care advocates, employees and other COVID-19 response structures should include advice on contemporary e-device use and ensure this critical information is integrated into the package for essential protective principles used to combat the spread of the COVID-19 infection.

References

1. Gustafsson E, Thomee S, Grimby-Ekman A, Hagberg M. Texting on mobile phones and musculoskeletal disorders in young adults: A five-year cohort study. Appl Ergon. 2017;58:208–214. https://doi.org/10.1016/j.apergo.2016.06.012

2. Straker L, Harris C, Joosten J, Howie EK. Mobile technology dominates school children’s IT use in an advantaged school community and is associated with musculoskeletal and visual symptoms. Ergonomics. 2018;61(5):658–669. https://doi.org/10.1080/00140139.2017.1401671

3. Jun D, Zoe M, Johnston V, O’Leary S. Physical risk factors for developing non-specific neck pain in office workers: A systematic review and meta-analysis. Int Arch Occup Environ Health. 2017;90(5):373–410. https://doi.org/10.1007/s00420-017-1205-3

4. Woo EHC, White P, Lai CWK. Musculoskeletal impact of the use of various types of electronic devices on university students in Hong Kong: An evaluation by means of self-reported questionnaire. Man Ther. 2016;26:47–53. https://doi.org/10.1016/j.math.2016.07.004

5. Arshad MA, Shamsudin MZ, Jamalil M, Mustafa A. Laptop use and upper extremities musculoskeletal disorders among higher learning students. 2020;01(1):1–4.

6. Lee S, Kang H, Shin G. Head flexion angle while using a smartphone. Ergonomics. 2015;58(2):220–226. https://doi.org/10.1080/00140139.2014.967311

7. Pereira A, Miller T, Huang Y-M, Odell D, Rempel D. Holding a tablet computer with one hand: Effect of tablet design features on biomechanics and subjective usability among users with small hands. Ergonomics. 2013;56(9):1363–1375. https://doi.org/10.1080/00140139.2013.820844

8. Marsh S, Mhurchu CN, Maddison R. The non-advertising effects of screen-based sedentary activities on acute eating behaviours in children, adolescents, and young adults. A systematic review. Appetite. 2013;71:259–273. https://doi.org/10.1016/j.appet.2013.08.017

9. Domingues-Montanari S. Clinical and psychological effects of excessive screen time on children. J Paediatr Child Health. 2017;53(4):333–338. https://doi.org/10.1111/jpc.13462

10. Henschke N, Harrison C, McKay D, Broderick C, Latimer J, Britt H, et al. Musculoskeletal conditions in children and adolescents managed in Australian primary care. BMC Musculoskelet Disord. 2014;15(1):164. https://doi.org/10.1186/1471-2474-15-164

11. Holden BA, Fricke TR, Wilson DA, Jong M, Naidoo KS, Sankaridurg P, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. Ophthalmology. 2016;123(5):1036–1042. https://doi.org/10.1016/j.ophtha.2016.01.006

12. Sheppard AL, Wolffsohn JS. Digital eye strain: Prevalence, measurement and amelioration. BMJ Open Ophthalmol. 2018;3(1), e000146. https://doi.org/10.1136/bmjophth-2018-000146

13. Coles-Brennan C, Sulley A, Young G. Management of digital eye strain. Clin Exp Optom. 2019;102(1):18–29. https://doi.org/10.1111/cxo.12798

14. Jaiswal S, Asper L, Long J, Lee A, Harrison K, Golebiowski B. Ocular and visual discomfort associated with smartphones, tablets and computers: What we do and do not know. Clin Exp Optom. 2019;102(5):463–477. https://doi.org/10.1111/cxo.12851

Check for updates

AUTHOR:
Marina Joubert[1]

AFFILIATION:
[1]Centre for Research on Evaluation, Science and Technology (CREST), Stellenbosch University, Stellenbosch, South Africa

CORRESPONDENCE TO:
Marina Joubert

EMAIL:
marinajoubert@sun.ac.za

HOW TO CITE:
Joubert M. From top scientist to science media star during COVID-19 – South Africa's Salim Abdool Karim. S Afr J Sci. 2020;116(7/8), Art. #8450, 4 pages. https://doi.org/10.17159/sajs.2020/8450

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
science communication, science and mass media, public visibility in science, public credibility in science, scientific celebrities

PUBLISHED:
29 July 2020

From top scientist to science media star during COVID-19 – South Africa's Salim Abdool Karim

Around the world, the COVID-19 pandemic has turned a handful of leading scientists into highly visible public figures. Anthony Fauci is the media star in the USA; Roberto Burioni in Italy; and in Sweden, it is Anders Tegnell. In Germany, Christian Drosten has become a household name, while Hugo Lopez-Gatell is highly visible in Mexico.

In South Africa, we witnessed a remarkable surge in the public prominence of Professor Salim Abdool Karim following his appointment, in mid-April 2020, to lead a Ministerial Advisory Committee advising government on combating COVID-19. Data from Pear Africa, a South African media monitoring company, show that Abdool Karim featured in 545 print, broadcast and online media items during April 2020, compared with 20 in April 2019.

Tracing Abdool Karim's journey towards becoming a trusted public voice of science on COVID-19 reveals meaningful insights into the relationships between scientists, policymakers and the South African public. It also illustrates how Abdool Karim follows an international cultural trend whereby charismatic scientists approach celebrity status under certain conditions, giving them unique power in shaping public trust in science.[1]

For this case study, I explored the characteristics of Abdool Karim's engagement with the South African public during the COVID-19 pandemic in the light of two key dimensions of quality in science communication, namely 'visibility' (including accessibility) and 'credibility' (encompassing expertise, trust and relevance). Both are needed for people to make decisions about costs, risks, benefits and ethics.[2] Furthermore, I considered the effect of his communication skill (efficacy) and his willingness to engage (attitude).

Public visibility in science

To become visible in academic circles, researchers must publish highly cited work in scholarly journals. In contrast, visibility in the public sphere depends on a high media profile.[3] Media prominence amplifies scientists' views and stimulates public interest, to the point that some become influential thought leaders and even iconic celebrities.[4,5]

A signal of public visibility that approaches celebrity status is when the media starts to take an interest in a scientist's personal life.[3] Furthermore, public visibility in the media is sustained by controversy.[6,7] Both these factors are illustrated in the case of Abdool Karim.

Public visibility in science: The case of South Africa's Salim Abdool Karim

Abdool Karim has long been recognised as one of the most visible scientists in South Africa.[8] He knows that his public visibility results from his media profile and he pays attention to his media appearances:

> *I am not surprised that I'm seen as publicly visible. In 2016, I appeared in about 455 articles and news items and I was on eTV about 20 times and SABC 16 times. Our communication department keeps track of all my media appearances and, given the amount of time it takes, it is extremely useful to have those statistics.*[9]

For him, engaging with the media (and the public) is an integral part of his role as a scientist:

> *Our job is to do good science and to communicate it. It is part and parcel of the job.*[9]

With a long track record of science advocacy, he is acutely aware of the power of the media to drive policy change. Around 2000, Abdool Karim was one of the scientists who spoke out when then South African President, Thabo Mbeki, questioned the science of HIV/Aids. He describes this period as a turning point:

> *I could not stay quiet. I knew that we had to convey the scientific viewpoint in a way that people could understand – clearly and succinctly. If we did not challenge Mbeki's views, how would people be able to make sound judgements? We had to fight back with the help of the media.*[9]

Those who publicly disagreed with the government's Aids denialism were branded as enemies of the state. Abdool Karim was chastised for his views during the 2000 International Aids Conference in Durban when Mbeki's health minister, Manto Tshabalala-Msimang, accused him and colleague Hoosen Coovadia of being 'disloyal' and 'traitors'.[10]

Twenty years later, facing the COVID-19 pandemic, the South African government turned to scientists for advice on how to handle the crisis, with Abdool Karim taking the lead. The contrast with his previous experience is clear:

> *The minister has been contacting us, he wants to involve us, he is seeking the opposite of what Mbeki and Tshabalala-Msimang wanted.*[10]

The South African media generally welcomed the establishment of a committee of leading scientists as a signal that government 'has chosen to respond to the Coronavirus armed with the best available information and evidence'[11]. On 13 April 2020, Abdool Karim participated in a two-hour (live) televised public briefing. Resulting media headlines referred to him as a 'world-class scientist'[12] and 'a great mind at work'[13]. Journalists praised the way he explained the scientific basis for the government's response to COVID-19:

> *He had millions of South Africans glued to their screens, collectively eating out of the palm of his hand. After the presentation … there was a palpable shift in the public's understanding of what informed some of governments' decisions around its response.*[14]

> *In a calm and reasoned two-hour-long knowledge explosion, Prof Salim Abdool Karim became our Dr Fauci by helping us, at last, to make sense of our corona fears.*[15]

Some South Africans took to social media to thank him and #ProfKarim trended briefly on Twitter:

> *#ProfKarim South Africa's COVID-19 Hero!*[16]

> *The amount of info that he gave is priceless. Everyone in South Africa should watch this. Thank you #ProfKarim.*[17]

> *#ProfKarim is our very own Dr Anthony Fauci. This man gives SA peace in his explanations. What a great resource to SA at this time!*[18]

Over the following weeks, journalists lined up to interview him and Abdool Karim participated in many media briefings and webinars.

The face of COVID-19 science in South Africa, Professor Salim Abdool Karim (photo: Madelene Cronjé, Bhekisisa Centre for Health Journalism) .

Indicative of his new level of fame, journalists wanted to know more about the man behind the scientist. Several articles focused on his childhood, his life story and his family, with some journalists highlighting his friendly demeanour and charisma.[15,19]

Public credibility in science

Public credibility in science is complex and nuanced. Definitions of what makes scientists credible tend to focus on their productivity, leadership, and the requirement to limit their public communication to peer-reviewed evidence within their own field of expertise.[20] Peters[21] defines public credibility in science as consisting of 'expertness' (the ability of the source to provide accurate information) and 'trustworthiness' (the readiness of a source to convey information clearly and completely). Expertise (i.e. the knowledge and ability to be accurate) is associated with the communicators' credentials, such as institutional affiliations, leadership positions and academic recognition. Importantly, expertise on its own may earn audiences' respect, but not necessarily their trust. Audiences decide whom to trust based on the perceived intent (truthfulness) of the communicator and whether the communicator is deemed to be sincere and friendly, thereby demonstrating concern for the audience (warmth). Therefore, people will decide whom to trust based on perceived competence and warmth, suggesting that 'scientific communication can be more effective by drawing on both dimensions of communicator credibility'.[22]

During a time of crisis, there are several 'warmth'-related factors that will contribute to or detract from the credibility of scientists as public communicators. These factors include ongoing interaction with the public, and communication characterised as inclusive, transparent, frank, truthful, succinct, consistent and compassionate, as well as the acknowledgment of uncertainty and ambiguity that are often inherent to a crisis.[23-25] Credibility is undermined when communicators avoid expressing their own feelings of concern, because they may be perceived by the public as cold and uncaring.

The following examples show how Abdool Karim achieves public credibility in science, based on the criteria outlined above.

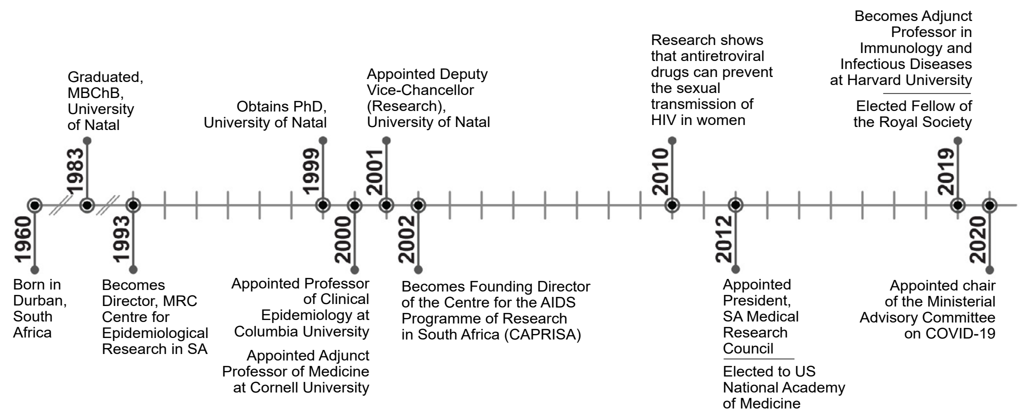
He has the academic credentials

When journalists introduce Abdool Karim, they often refer to his imposing academic track record that positions him as a credible expert:

> *Karim's CV spans four universities and several institutions. He is a Fellow of the Royal Society. His most recent posting is that of director of the Centre for the Aids Programme of Research in South Africa (Caprisa), based in Durban. His other titles include*

Salim Abdool Karim was born on 29 July 1960 in Durban, South Africa.
Now, close to his 60th birthday, his 62-page CV consists of impressive lists of academic achievements, appointments at leading research organisations and scholarly prizes and awards.

A few highlights are reflected in the timeline below:



Adjunct Professor of Immunology and Infectious Diseases at Harvard University and Adjunct Professor of Medicine at Cornell University.[12]

Some of his awards include The World Academy of Sciences Prize in Medical Sciences, the African Academy of Science's Olusegun Obasanjo Prize for Scientific Discovery and Technological Innovation and the Kwame Nkrumah Continental Scientific Award, the most prestigious scientific award in Africa.[13]

He stays in his lane and sticks to the evidence

Abdool Karim is adamant about sticking to his role as advisor to the government, and formulating advice based on scientific evidence:

I think there is confusion; I am not leading anything. I'm just the person giving the advice. I don't work for the government. I make no decisions. I simply provide information, that's all, and I'm very happy to do so.[14]

Is BCG helping us? I would love that to be the case but I'm really sceptical that it has any influence. I will wait for the data to make up my mind firmly one way or the other.[26]

He speaks clearly

Jargon is a barrier when it comes to public understanding of science.[27] Abdool Karim is clearly aware of the need to package his science messages simply and clearly:

As scientists, we have our own language and terminology, which make it very hard to explain some things to the public. We must learn to get away from our scientific terminology and talk clearly and simply – that is the most important thing.[9]

We find that people often just need that little clarity. That little clarity changes the amount of stress they are under or the way in which they are looking at the problem. With a clearer understanding they are able to function better and relax. In some small way, we are happy to make that contribution.[28]

He uses metaphors and analogies effectively to help people grasp key messages:

When the virus enters a community, it spreads like wildfire. … You need people on the ground that are looking for fires. If we see one, we can prevent it. If we get there too late, then we have to put out a raging fire.[29]

We have not escaped this bullet, we have only postponed its impact.[30]

He stands up for science

As during the earlier period of Aids denialism, Abdool Karim believes that scientists can combat fake news by speaking out publicly:

So, we do quite a lot of Q&As because we believe that if you empower people with facts, then the fake news will just fade into the background. And ignorance and over reaction become less of an issue.[28]

He delivers unpopular messages with compassion

While scientific evidence is not always reassuring or what people want to hear, Abdool Karim delivers 'hard' messages with empathy:

Our job sometimes is to look at all the bad options and give the least bad one, saying: Of all the bad options we have, this one is probably the one we should go with. Here is the scientific evidence for it.[14]

What we hope for is that the number of new cases will steadily decline and the new cases will decrease. But I am sorry to tell you that that is very unlikely.[31]

He emphasises uncertainty

He acknowledges the challenges when scientists are called on to give policy input and public advice during this pandemic:

Whether we're right or wrong, time will tell. History will judge us. I don't claim that we are not making mistakes. If you don't make mistakes in tackling a disease of this nature that means you're not being sufficiently proactive.[30]

As South Africa braces for an expected surge of cases over the coming weeks, I describe the national COVID-19 response as 'sailing a ship while building it'.[32]

He is competent, but not cold

When addressing the public, Abdool Karim mostly speaks in the first person (showing personal concern) and connects with people on an emotional level:

Once we end the lockdown, and we must remember that none of us have immunity against this virus, we are all at risk.[29]

I am sorry to say that life is not going to be what it was like before. … We are going to have to learn to live in a way where we will lose that soft touch that comes from being close to those we love, those we care about, because in order to protect them we are going to have to keep some distance.[33]

Going public is not risk free

A high public profile goes hand-in-hand with some level of reputational risk that could result from media scrutiny, disdain from peers and controversies.[4,34] Notably, scientists with a high standing in academic circles are better able to weather the storms that could result from publicity.[4,9,35]

In the case of Abdool Karim, *Business Day* columnist Steven Friedman[36] took issue with his insistence that South Africa could not avoid a severe COVID-19 epidemic, accusing him of political double-speak. Also, Abdool Karim's research funding came under scrutiny with claims on social media that he received ZAR944 million in funding from the Bill and Melinda Gates Foundation. He dismissed these claims as 'fake news' and 'nonsense'.[37]

Conclusion

These examples of how Abdool Karim has engaged with the South African public about COVID-19, mostly via the mass media, illustrate the criteria that have been suggested as measures of effective public communication of science. They show that journalists and the public respond positively to simple and compelling messages that are delivered clearly and with empathy. In order to achieve this result, scientists need to have credibility and advanced communication skills, but they must also be willing to invest time in working with the media and engaging with the public.

References

1. Caulfield T, Fahy D. Science, celebrities and public engagement. Issues in Science and Technology. 2016;23(4) [cited 2020 Jun 03]. Available from: https://issues.org/perspective-science-celebrities-and-public-engagement/

2. Nisbet MC, Scheufele DA. What's next for science communication? Promising directions and lingering distractions. Am J Bot. 2009;96(10):1767–1778. https://doi.org/10.3732/ajb.0900041

3. Fahy D, Lewenstein B. Scientists in popular culture: The making of celebrities. In: Bucchi M, Trench B, editors. Routledge handbook of public communication of science and technology. New York: Routledge; 2014. p. 83–96.
4. Goodell R. The visible scientists. Boston, MA: Little, Brown and Company; 1977.
5. Fahy D. The new celebrity scientists. New York: Rowman & Littlefield; 2014.
6. Horst MA. Field of expertise, the organization, or science itself? Scientists' perception of representing research in public communication. Sci Commun. 2013;35(6):758–779. http://dx.doi.org/10.1177/1075547013487513
7. Molinatti G, Simonneau L. A socioenvironmental shale gas controversy: Scientists' public communications, social responsibility and collective versus individual positions. Sci Commun. 2015;37(2):190–216. https://doi.org/10.1177/1075547014560827
8. Joubert M, Guenther L. In the footsteps of Einstein, Sagan and Barnard: Identifying South Africa's most visible scientists. S Afr J Sci. 2017;113(11/12), Art. #2017-0033. http://dx.doi.org/10.17159/sajs.2017/20170033
9. Joubert M. Factors influencing the public communication behaviour of publicly visible scientists in South Africa [dissertation]. Stellenbosch: Stellenbosch University; 2018. https://scholar.sun.ac.za/handle/10019.1/103268
10. Malan M. The facts beat the quacks: Our #Covid19SA vs. our #HIV response. Bhekisisa. 2020 March 25 [cited 2020 Jun 03]. Available from: https://bhekisisa.org/article/2020-03-25-coronavirus-reporting-journalism-hiv-denialism-evidence-based-response-south-africa-cyril-ramaphosa-thabo-mbeki/
11. Chetty K. What you need to know about the team of experts helping to fight coronavirus in SA. IOL. 2020 April 02 [cited 2020 Jun 03]. Available from: https://www.iol.co.za/news/politics/what-you-need-to-know-about-the-team-of-experts-helping-to-fight-coronavirus-in-sa-46000952
12. Evans S. Leading the charge: World-class scientist heads SA's Covid-19 response. News24. 2020 April 14 [cited 2020 Jun 03]. Available from: https://www.news24.com/news24/southafrica/news/leading-the-charge-world-class-scientist-heads-sas-covid-19-response-20200414
13. Nkgadima R. South Africans bow to a great mind at work. Daily News. 2020 April 15;3/Metro
14. Mkize V. Government decisions. City Press. 2020 April 19;5/News.
15. Retief H. Prof. 'Slim' bring kennis en hoop [Prof Slim brings knowledge and hope]. Rapport. 2020 April 19;10/News.
16. @keithf1960. #ProfKarim South Africa's [tweet]. 2020 April 13 [cited 2020 Jun 03]. Available from: https://twitter.com/keithf1960
17. @nengelbrecht036. Dear #eNCA … [tweet]. 2020 April 13 [cited 2020 Jun 03]. Available from: https://twitter.com/nengelbrecht036
18. @TonyVenturini. #ProfKarim is our … [tweet]. 2020 April 13 [cited 2020 Jun 03]; Available from: https://twitter.com/TonyVenturini
19. Sibisi G. The good professor. Drum. 2020 April 30;14–15.
20. Dunwoody S, Ryan M. The credible scientific source. Journal Mass Commun Q. 1987;64:21–27.
21. Peters HP. The credibility of information sources in West Germany after the Chernobyl disaster. Public Underst Sci. 1992;1(3):325–343. https://doi.org/10.1088/0963-6625/1/3/006
22. Fiske ST, Dupree C. Gaining trust as well as respect in communicating to motivated audiences about science topics. Proc Natl Acad Sci USA. 2014;111(Suppl. 4):13593–13597. https://doi.org/10.1073/pnas.1317505111
23. Seeger MW. Best practices in crisis communication: An expert panel process. J Appl Comm Res. 2006;34(3):232–244. https://doi.org/10.1080/00909880600769944
24. Canfield KN, Menezes S, Matsuda SB, Moore A, Mosley Austin AN, Dewsbury BM, et al. Science communication demands a critical approach that centers inclusion, equity, and intersectionality. Front Commun. 2020;5, Art. #2. https://doi.org/10.3389/fcomm.2020.00002
25. Neeley L. How to talk about the coronavirus. The Atlantic. 2020 March 31 [cited 2020 Jun 03]. Available from: https://www.theatlantic.com/ideas/archive/2020/03/how-talk-about-coronavirus/609118/
26. Pijoos I. Expert 'sceptical' about BCG vaccine for Covid-19: 'It's meant to prevent severe forms of TB'. Herald Live. 2020 April 14 [cited 2020 Jun 03]. Available from: https://www.heraldlive.co.za/news/2020-04-14-expert-sceptical-about-bcg-vaccine-for-covid-19-its-meant-to-prevent-severe-forms-of-tb/
27. Bullock OM, Colón Amill D, Shulman HC, Dixon GN. Jargon as a barrier to effective science communication: Evidence from metacognition. Public Underst Sci. 2019;28(7):845–853. https://doi.org/10.1177/0963662519865687
28. Botha N. #OnTheFrontLine with the Abdool Karims. Bizcommunity. 2020 May 04 [cited 2020 Jun 03]. Available from: https://www.bizcommunity.com/Article/196/858/203501.html
29. Ellis E. Lockdown has bought us time, Professor Karim says, as South Africa bucks the trend. Daily Maverick. 2020 April 14 [cited 2020 Jun 03]. Available from: https://www.dailymaverick.co.za/article/2020-04-14-lockdown-has-bought-us-time-expert-says-as-sa-bucks-the-trend/
30. Karrim A. Coronavirus: Don't underestimate the enemy, says Prof Karim. News24. 2020 April 17 [cited 2020 Jun 03]. Available from: https://www.news24.com/news24/southafrica/news/coronavirus-dont-underestimate-the-enemy-says-prof-karim-20200417
31. Mahlati Z. SA cannot escape flood of infections, warns Prof Karim. IOL. 2020 April 15 [cited 2020 Jun 03]. Available from: https://www.iol.co.za/news/politics/sa-cannot-escape-flood-of-infections-warns-prof-karim-46719138
32. Abdool Karim SS. The South African response to the pandemic. N Engl J Med. 2020;382, e95. https://doi.org/10.1056/NEJMc2014960
33. Karrim A. Covid-19 in SA: Life is not going to be what it once was, says Prof Karim. News24. 2020 April 15 [cited 2020 Jun 03]. Available from: https://www.news24.com/news24/southafrica/news/covid-19-in-sa-life-is-not-going-to-be-what-it-once-was-says-prof-karim-20200415
34. Rödder S. The ambivalence of visible scientists. In: Rödder S, Franzen M, Weingart P, editors. The sciences' media connection–public communication and its repercussions. Dordrecht: Springer; 2012. p. 155–178.
35. Dunwoody S, Brossard D, Dudo A. Socialization or rewards? Predicting US scientist-media interactions. Journal Mass Commun Q. 2009;86(2):299–314. https://doi.org/10.1177/107769900908600203
36. Friedman S. Dangerous when scientists doublespeak like politicians. Business Day. 2020 May 13;7/Opinion.
37. Comins L. Professor Karim dismisses R944m Bill and Melinda Gates Foundation funding claims as fake news. IOL. 2020 May 11 [cited 2020 Jun 03]. Available from: https://www.iol.co.za/mercury/news/professor-karim-dismisses-r944m-bill-and-melinda-gates-foundation-funding-claims-as-fake-news-47798953

'Preventing the next pandemic' – A 2020 UNEP Frontiers Series Report on zoonotic diseases with reflections for South Africa

AUTHORS:
Bianca Wernecke[1,2]
Danielle A. Millar[3]
Michele Walters[4]
Andre Ganswindt[5]
Luthando Dziba[6,7]
Caradee Y. Wright[3,8]

AFFILIATIONS:
[1]Environment and Health Research Unit, South African Medical Research Council, Johannesburg, South Africa
[2]Department of Environmental Health, University of Johannesburg, Johannesburg, South Africa
[3]Environment and Health Research Unit, South African Medical Research Council, Pretoria, South Africa
[4]Biodiversity and Ecosystem Services, Council for Scientific and Industrial Research, Pretoria, South Africa
[5]Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Pretoria, South Africa
[6]Conservation Services Division, South African National Parks, Pretoria, South Africa
[7]Department of Zoology, Nelson Mandela University, Port Elizabeth, South Africa
[8]Department of Geography, Geoinformatics and Informatics, University of Pretoria, South Africa

CORRESPONDENCE TO:
Caradee Wright

EMAIL:
cwright@mrc.ac.za

HOW TO CITE:
Wernecke B, Millar DA, Walters M, Ganswindt A, Dziba L, Wright CY. 'Preventing the next pandemic' – A 2020 UNEP Frontiers Series Report on zoonotic diseases with reflections for South Africa. S Afr J Sci. 2020;116(7/8), Art. #8531, 4 pages. https://doi.org/10.17159/sajs.2020/8531

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
human health, zoonoses diseases, environmental health, One Health, Planetary Health

PUBLISHED:
29 July 2020

Zoonoses account for about 25% of the infectious disease burden in low-income countries.[1] Poverty might increase the risk for zoonotic disease where the active human–livestock and human–wildlife interfaces can increase the likelihood of disease transmission.[1] A combined disease burden exists for people in areas such as tropical and subtropical Africa, where there is likelihood of co-infection with zoonotic diseases and other pathogenic or infectious diseases, such as malaria, tuberculosis and HIV.[1] Many endemic zoonoses remain widely neglected in such settings, undetected and underreported, because their impacts are borne largely by impoverished and marginalised communities.[2] Due to these unique contexts, the prevention and management of emerging and endemic zoonotic diseases in many African countries is a complex undertaking needing evidence-based guidance.[1,3]

In early 2020, the United Nations Environment Programme (UNEP) and the International Livestock Research Institute (ILRI) took on the urgent task to provide an up-to-date, rapid scientific assessment on zoonotic diseases as part of the UNEP's *Frontiers* Report Series.[4-6] The goal of the report is to provide relevant information for policymakers on how to 'prevent the next pandemic' by interrogating what is known about zoonotic diseases and how best one can break the chain of transmission. As the world presently faces the SARS-CoV-2 pandemic, this timely report helps decision-makers with evidence-based actions, not only to flatten the curve of COVID-19 incidence, but to answer questions about zoonoses in general and plan for the future. In this Commentary, we give a brief overview of UNEP's latest report[7] and then relate some of the key messages and recommendations for policymakers to a South African context.

An overview of zoonotic diseases

Many microorganisms are shared between people and other animals, but only some of these microorganisms, known as pathogens, cause disease.[7] Most current zoonotic infections in people are transmitted from domesticated wildlife, livestock, and pets – true wildlife zoonoses are rarer but can be significant and catastrophic. Coronavirus disease 2019 (COVID-19) is ongoing; however, in recent decades there have also been zoonotic influenza (Bird Flu), pandemic human influenza (H1N1), Middle East respiratory syndrome (MERS) and severe acute respiratory syndrome (SARS) – most of which have a suspected or proven domestic or peri-domestic wildlife reservoir. The three types of zoonotic diseases – i.e. **emerging** (e.g. Zika virus 2015–2016 in the Americas), **outbreak or epidemic** (e.g. Rift Valley fever 2015–2016 in Mauritania), and **neglected zoonotic diseases** (e.g. pig tapeworm) – can (and have) all contribute(d) to significant public health and related societal impacts. While seven major anthropogenic drivers for zoonotic disease emergence have been identified (Figure 1), it is not necessarily a simple task to determine which driver(s) lead to a zoonotic occurrence.

Between January 2017 and July 2018, South Africa experienced the largest listeriosis outbreak in the world to date.[8,9] The National Institute for Communicable Diseases (NICD) detected a total of more than 1000 laboratory-confirmed cases and 200 fatalities,[10,11] all of which were linked to eating contaminated polony (a type of bologna sausage). Polony was confirmed to contain *Listeria monocytogenes* after undergoing genome sequencing.[10,11] In response to the outbreak, a national multisectoral task force coordinated investigations and response activities. Consequently, national authorities took measures to limit further infections, including issuance of safety recall notices and compliance notices, and communication about risks with vulnerable groups. In December 2017, listeriosis was made a notifiable medical condition in South Africa.[12] Listeriosis, like many foodborne diseases, is a zoonosis.

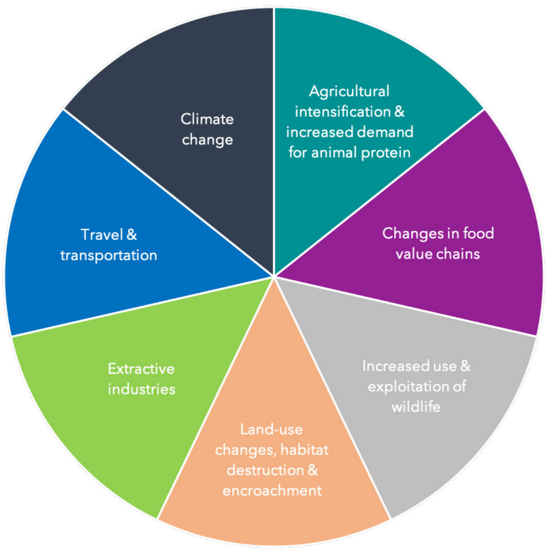


Figure 1: Seven major anthropogenic drivers of zoonotic disease emergence.

Coronaviruses, common elements and origins of coronavirus pandemics

'Corona' refers to the crown-like arrangement of spike-shaped proteins on the surface of membranes of coronaviruses. In an infographic, the UNEP report gives an insightful timeline of various coronaviruses and other zoonotic disease occurrences from 1931 to the present day.[7] Among some of these coronaviruses, such as MERS and porcine epidemic diarrhoea (PED) virus in 1971, there are common elements and origins: epidemics are driven by agricultural intensification/increased demand for animal protein; they have high economic costs, particularly due to social and economic disruptions; it seems that bats are natural reservoir hosts; and transmission is often associated with traditional markets or value chains. These elements are often interlinked, as described next.

Linkages between habitat loss and zoonoses

Human population increase has led to greater encroachment of humans into natural habitats, thereby proportionately increasing the risk of animal-to-human disease transmission.[7] Changes in biodiversity and landscape conversion lead to changes in natural food cycles, often with an increase in vectors and pests and a decrease in predators. This causes a change in the species composition found in certain areas. Animals such as rodents are often present and considered opportunistic animals which frequently can represent vectors for disease or pathogen reservoirs. The 'dilution effect' suggests that the greater the species diversity, the fewer the opportunistic animals and pathogen reservoirs, translating into decreased disease transmission.[7] However, this theory is dependent on animal and human population density and the transmission mode of the pathogen. The coevolution theory suggests that the change in landscape causes an adaptive change of migratory and resident animals, including pathogens, which increases the risk of disease spill-over and transmission. While there is a large variety of zoonotic pathogens, the greatest risk factor for transmission spill-over remains increased exposure to animals.

Linkages between trade and use of wildlife and zoonoses

Wildlife harvesting, trade and consumption are age-old activities, with wild animals being hunted and captured for human subsistence, for recreation, and for sale of animal body parts and/or their derivatives.[7] Today, many households in Africa, Asia and Latin America harvest and consume wild meat, such as herbivores, rodents and reptiles. Although less documented, the harvesting and consumption of aquatic mammals (e.g. manatees and whales) and turtles is increasing too. The consumption of wild meat is driven by, among other factors, an increasing human population and an increased demand for protein-rich food and income. Not everyone may have access to, or be able to satisfy these needs with traditional resources, such as land, labour and livestock.

Much wild meat and many live wild animals are sold in traditional markets, which lack adequate biosafety measures, thus increasing the risk for zoonotic disease emergence and transmission. While legal ventures of wild meat production are increasing, there is also an increase in illegal wild meat production – an industry that employs many people and contributes to food and job security.[7] Wild animals that are subject to trade often live in captivity and close quarters. As such, viruses spread between the diverse species and have higher pandemic potential as the viruses are considered to have greater 'host plasticity', which increases the possibility that they can spread to humans. Wildlife use, trade and consumption brings zoonotic risks. For example, zoonoses can be introduced at any point during the transfer of animal meat between hunting and handling. There is also the risk of transmission when meat is opportunistically harvested, such as in the case of the 2013–2016 Ebola outbreak in Guinea, Liberia and Sierra Leone.

Reflections for South Africa

Here, we reflect upon the Ten Key Messages put forward by the UNEP 'Preventing the next pandemic' report (Table 1) and how they may be discussed and applied in South Africa. In terms of knowledge to understand current systems and particularly those dealing with host populations, we need continued research in South Africa. As we put greater pressure on ecological systems, the way these systems respond changes, meaning that what was true yesterday can be different today.[13] If we are to keep a 'finger on the pulse' of future disease outbreaks, we need research (especially baseline data) and monitoring of zoonotic diseases which takes these changes into account. Monitoring systems form the basis for tools to help respond to future pandemics. Data, including spatial data on drivers of disease, need to be made freely available across research fields such as wildlife ecology, human health, agricultural health and environmental health. Such information would prepare us better to respond with urgency to regional outbreaks.

In South Africa, while a scientific assessment as mentioned in the UNEP/ILRI report is necessary to explore the role of wildlife and domesticated animals in disease transmission, it is also necessary to look into respective 'animal groups' from a regional economic (and maybe spiritual) value perspective, as well as from a species diversity and conservation point of view. A case in point of the latter are countries where most of a species resides, like the southern white rhino in South Africa. Perhaps it is timeous to discuss global financial support mechanisms for those regions (or countries) in the world that by default must look after certain ecological hotspots or flagstone species.

Table 1: One-liners and Ten Key Messages emerging from the 'Preventing the next pandemic' report[7]

Message title	One-liner message
De-risking food systems	More evidence-based scientific assessments, such as this one, are needed to examine the environmental and zoonotic context of the current pandemic, as well as the risk of future zoonotic disease outbreaks.
Urgency	Urgent action is needed to develop knowledge and tools to prevent future pandemics.
Report audience	This scientific, policy-relevant assessment is designed for decision-makers, businesses and civil society and explores the role of wild and domesticated animals in emerging zoonoses.
Scope of the problem	About 60% of human infections have an animal origin; most zoonoses happen indirectly (via food).
Outbreak frequency and predictability	The frequency of pathogens jumping from animals to humans is increasing, and pandemic outbreaks are a predictable outcome based on unsustainable human activities.
Connectivity and complexity	The links among the wider environment, biodiversity and emerging infectious diseases are interconnected and complex.
Disease drivers	There are seven human-mediated factors driving the emergence of zoonotic disease and unsustainable use of natural resources and animals: travel, change in food supply and accelerated climate change are some of these factors.
Impact and cost	Emerging zoonotic diseases threaten human and animal health, economic development, and the environment, and this disproportionately in lower-income countries.
Policy options	Recommended policy response actions to 'build back better' after COVID-19 include awareness raising, disease monitoring and surveillance, multi-stakeholder approaches, financial accounting and transformed agricultural practices as well as the introduction of the 'One Health' approach.
One Health	Uniting medical, veterinary, and environmental expertise will help achieve sustainable health for people, animals, and the environment alike.

To determine a holistic scope of the problem of zoonotic diseases and how to prevent the next outbreak, South Africa may need to discuss and prioritise next steps. We need a clear understanding of the 'food system': what people are eating, the sources of food, and how food is processed, stored and distributed, with a clear distinction between in-country consumption and/or export. In South Africa's present situation, the impacts of lockdown (and poverty in general) may force a reliance on wild meat, particularly for those who were dependent on the tourism industry to sustain a livelihood, which has largely closed. Increasing unemployment drives people towards harvesting of wildlife. Recent action by government to pass new laws on what constitutes domesticated animals[14] and what is being considered under the *Meat Safety Act*[15] means we will be brought closer to or be in greater contact with more species of animals. If we are to intensify pressures on wildlife in this way and start incorporating more wildlife into our diets, stringent testing for infection in these species should occur and potential disease and species-specific protocols should be put in place. Culture is also an important driver in South Africa. Countries reliant on traditional medicines and foods (also wild meat harvests) and those with an expanding demand for animal protein, including insects, for human and animal consumption may be at risk.[16]

Ecological disruption caused by an ever-growing human population, and its demand for and pressure on resources, may allow us to predict that future disease outbreaks will occur. In ecological systems in general, however, we are becoming less able to predict patterns and processes particularly in 'food webs' or 'assembly rules' and future projections (in this case around communities and food webs) are becoming increasingly messy.[13] This has obvious consequences for our ability to monitor or understand zoonotic disease flows. In South Africa, these flows are complicated by issues of poverty, health inequity, landscape alterations, waste management and pollution, and the quadruple burden of disease. In terms of connectivity and complexity, the problem is intricate indeed but also constantly changing. As the predictability of biological and ecological processes decline, so too does our understanding of the connectivity and complexity at any point in time, particularly into the future. Hence, the need to mainstream and implement a 'One Health' approach.

An approach for the future

The latest UNEP report adopts 'One Health' as a collaborative, multisectoral and transdisciplinary approach – working at local, regional, national and global levels – to achieve optimal health and well-being outcomes, recognising the interconnections between people, animals, plants and their shared environment (Figure 2).[17] Other approaches, such as 'Planetary Health'[18] and 'One Welfare'[19] also exist; however, UNEP considered 'One Health' as an easily understood umbrella term.

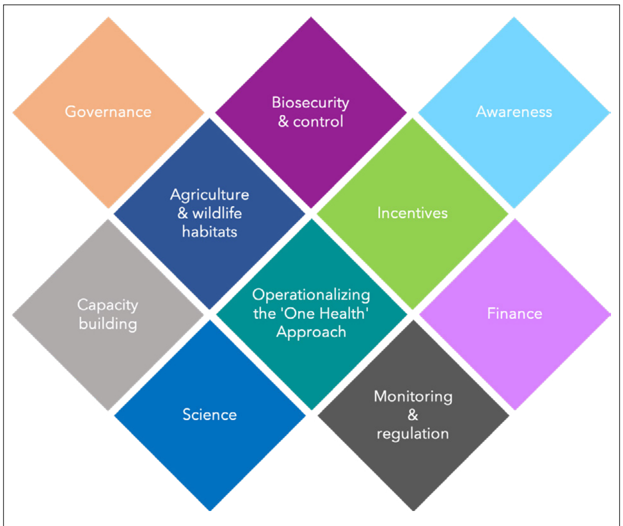


Figure 2: UNEP's Ten science-based policy recommendations that focus on a One Health approach.

South Africa's 'One Health' Programme, led by the US Centers for Disease Control and Prevention with partners, as well as other initiatives such as AHEAD (Animal & Human Health for the Environment and Development)[20] have done work to apply the approach in South Africa. UNEP's support of the 'One Health' approach now gives new impetus to finding ways in which South Africa can holistically embrace the philosophy. For example, a 'One Health' approach in policy development is critical for the future and has been sorely lacking. South Africa needs to adopt greater interdepartmental cooperation and joint decision-making when it comes to policy development. It is also essential that 'One Health' provides a suitable inclusive approach to assess the environmental context due to our country's unique, interconnected multi-layer diversity (e.g. ethnic, cultural, religious, social, economic and ecological diversity).

Promisingly, several actions and impacts of 'One Health' implementation in Africa do exist. Collaborative projects implementing joint livestock and child vaccination campaigns in pastoralist communities in Chad resulted in economic savings for the Chadian public health and animal health ministries and improved vaccination coverage of children and women who would otherwise have no access to health care.[21] Shared human and veterinary laboratories to diagnose brucellosis in febrile patients in Mali resulted in brucellosis being considered as a differential diagnosis for febrile illness (along with malaria and typhoid fever) in an area where raw milk consumption is still prevalent.[21] Lastly, Uganda's environmental health practitioners embody 'One Health' for healthy people, animals and environment.[22] They investigate zoonotic disease outbreaks, ensure disease vector control, and provide communities with health education on vaccination of pets and food safety. These are encouraging success stories, from which South Africa should learn.

It is imperative that South Africa uses the rich interconnected multi-layer diversity in a call to all environmental health practitioners, public health professionals, doctors, veterinarians, conservationists, people living in communities, industries and small businesses selling animal products, and others to discuss, debate, and embrace 'One Health' as promoted by the UNEP report to prevent future epidemics and pandemics, locally and internationally. A solid evidence-based policy foundation is needed, as are human and financial capital, technological advances and advocacy, to facilitate creation of the tools and platforms for effective response and crisis mitigation.

Acknowledgements

Two SAMRC environmental health researchers (B.W. and C.Y.W.) gave direct input to the UNEP/ILRI 2020 report and thank UNEP for this opportunity.

References

1. Asante J, Noreddin A, El Zowalaty ME. Systematic review of important bacterial zoonoses in Africa in the last decade in light of the 'One Health' concept. Pathogens. 2019;8(50):1–29. https://doi.org/10.3390/pathogens8020050

2. Halliday JEB, Allan KJ, Ekwem D, Cleaveland S, Kazwala RR, Crump JA. Endemic zoonoses in the tropics: A public health problem hiding in plain sight. Vet Rec. 2015;176(9):220–225. http://dx.doi.org/10.1136/vr.h798

3. Kemunto N, Mogoa E, Osoro E, Bitek A, Kariuki Njenga M, Thumbi SM. Zoonotic disease research in East Africa. BMC Infect Dis. 2018;18(1):545–554. https://doi.org/10.1186/s12879-018-3443-8

4. United Nations Environment Programme (UNEP). UNEP Frontiers 2018/19 report: Emerging issues of environmental concern. Nairobi: UNEP; 2019. Available from: https://wedocs.unep.org/bitstream/handle/20.500.11822/27538/Frontiers1819.pdf?sequence=1&isAllowed=y

5. United Nations Environment Programme (UNEP). UNEP Frontiers 2017 report: Emerging issues of environmental concern. Nairobi: UNEP; 2017. Available from: https://wedocs.unep.org/bitstream/handle/20.500.11822/22255/Frontiers_2017_EN.pdf?sequence=1&isAllowed=y

6. United Nations Environment Programme (UNEP). UNEP Frontiers 2016 report: Emerging issues of environmental concern. Nairobi: UNEP; 2017. https://doi.org/10.18356/4392feb8-en

7. United Nations Environment Programme (UNEP) and International Livestock Research Institute (ILRI). Preventing the next pandemic: Zoonotic diseases and how to break the chain of transmission. Nairobi: UNEP and ILRI; 2020. Available from: https://wedocs.unep.org/bitstream/handle/20.500.11822/32316/ZP.pdf?sequence=1&isAllowed=y

8. Kaptchouang Tchatchouang C-D, Fri J, De Santi M, Brandi G, Fiorella Schiavano G, Amagliani G, et al. Listeriosis outbreak in South Africa: A comparative analysis with previously reported cases worldwide. Microorganisms. 2020;8(1):135. https://doi.org/10.3390/microorganisms8010135

9. Thomas J, Govender N, McCarthy KM, Erasmus LK. Outbreak of listeriosis in South Africa associated with processed meat. N Engl J Med. 2020;382(7):632–643. https://doi.org/10.1056/NEJMoa1907462

10. South African National Institute for Communicable Diseases (NICD). Situation report on listeriosis outbreak, South Africa, 2017 [document on the Internet]. c2017 [cited 2020 Jun 20]. Available from: https://www.nicd.ac.za/wp-content/uploads/2017/12/NICD-Situation-report-on-listeriosis-outbreak-South-Africa_18-December-2017.pdf

11. South African National Listeria Incident Management Team. Listeriosis outbreak situation report [document on the Internet]. c2018 [cited 2020 Jun 20]. Available from: https://www.nicd.ac.za/wp-content/uploads/2018/07/Listeriosis-outbreak-situation-report-_26July2018_fordistribution.pdf

12. World Health Organization (WHO). Listeriosis – South Africa. Disease Outbreak News [document on the Internet]. c2018 [cited 2020 Jun 20]. Available from: http://www.who.int/csr/don/28-march-2018-listeriosis-south-africa/en/

13. Berger J, Wangchuk T, Briceno C, Vila A, Lambert, JE. Disassembled food webs and messy projections: Modern ungulate communities in the face of unabating human population growth. Front Ecol Evol. 2020;8(1):128. https://doi.org/10.3389/fevo.2020.00128

14. Somers MJ, Walters M, Measey J, Strauss WM, Turner AA, Venter JA, et al. The implications of the reclassification of South African wildlife species as farm animals. S Afr J Sci. 2020;116(1/2), Art. #7724. https://doi.org/10.17159/sajs.2020/7724

15. Alden C, Harvey R. South African proposal to breed wildlife for slaughter courts disaster. The Conversation. 2020 Jun 14. Available from: http://theconversation.com/south-african-proposal-to-breed-wildlife-for-slaughter-courts-disaster-140399

16. Mishyna M, Chen J, Benjamin O. Sensory attributes of edible insects and insect-based foods – Future outlooks for enhancing consumer appeal. Trends Food Sci Technol. 2020;95:141–148. https://doi.org/10.1016/j.tifs.2019.11.016

17. One Health Commission. Why One Health? What is One Health? [webpage on the Internet]. c2020 [cited 2020 Jun 20]. Available from: https://www.onehealthcommission.org/en/why_one_health/what_is_one_health/

18. Whitmee S, Haines A, Beyrer C, Boltz F, Capon AG, de Souza Dias BF, et al. Safeguarding human health in the Anthropocene epoch: Report of the Rockefeller Foundation–Lancet Commission on planetary health. The Lancet. 2015;386:1973–2028. https://www.doi.org/10.1016/S0140-6736(15)60901-1

19. Bourque T. One Welfare. Can Vet J. 2017;58(3):217–218.

20. Osofsky S. Times evoke: Preventing the next pandemic — We can make this a ‘never again’ moment. India News. The Times of India. 2020 April 07. Available from: https://timesofindia.indiatimes.com/preventing-the-next-pandemic-we-can-make-this-a-never-again-moment/articleshow/74974789.cms

21. Kamani TM, Kazwala R, Mfinanga S, Haydon D, Keyyu J, Lankester F, et al. Viewpoint: One Health: A concept led by Africa, with global benefits. Vet Rec. 2015;176(19):496–497. https://doi.org/10.1136/vr.h2461

22. Musoke D, Ndejjo R, Atusingwize E, Halage AA. The role of environmental health in One Health: A Uganda perspective. One Health. 2016;2(1):157–160. https://doi.org/10.1016/j.onehlt.2016.10.003

Public Statement on COVID-19

The Academy of Science of South Africa (ASSAf) is the only statutory academy in the country established to provide government and the general public with evidence-based advice on issues of pressing national concern. Through its membership of outstanding scientists from across the disciplines, Academy members such as Professors Salim Abdool Karim, Quarraisha Abdool Karim, Glenda Gray, and Shabir Madhi are already working tirelessly in advising the South African government on effective ways of dealing with the pandemic. There are in fact many other Academy members doing important scientific research on the pandemic including Professor Helen Rees who is leading the South African component of a global vaccine trial to identify treatments for COVID-19.

Beyond these influential contributions of our members, ASSAf believes that as the national academy it has a further responsibility to make public its position on COVID-19 at this historic moment in South African and world history.

The respiratory disease COVID-19 is caused by the novel coronavirus (SARS-CoV-2) and has spread across the world at an alarming rate with more than 4.63m confirmed cases and more than 311,000 dead at the time of writing (mid-May 2020). There are as yet no effective treatments and no available vaccines to provide immunity against infection and prevent disease.

Under these circumstances, the primary response of national governments has been to mitigate the risk of the spread of the disease through measures such as physical distancing, travel restrictions and national lockdowns.

The success of this complex of public health actions has been uneven around the globe and dependent on many factors such as how quickly governments respond (through testing and contact tracing, for example) and how efficiently life-saving equipment is deployed.

The South African government's response has been effective and rightly acknowledged both at home and abroad. At the moment, the rate of growth of infections and death rates in South Africa is among the lowest in the world and also among BRICS countries.

ASSAf recognises and applauds the South African government for underlining the fact that the national strategy has been based on scientific evidence and guided by the advice of scientists. This was achieved despite uncertainties resulting from limited and evolving epidemiological and medical evidence and the pressure that comes with responding to new and emerging scientific information.

In such fast-moving and uncertain contexts, it is perhaps inevitable that different views will result among scientists themselves – such as how, when and where to ease the lockdown. Yet what cannot be denied is that strong, science-based governmental leadership has saved many lives for which South Africa can be thankful.

The leadership of ASSAf believes, however, that the government's response to the pandemic must be further strengthened by attending to the following **three concerns**.

First, it is crucial that the National Coronavirus Command Council, and the structures reporting to it, such as the Ministerial Advisory Committee on COVID-19, include in its advisory bodies scientists from a much broader range of disciplines. While it is important to have epidemiologists, vaccinologists and infectious disease experts on these bodies, we believe that the pandemic is not simply a medical problem but a social problem as well. This means that social scientists and humanities scholars should also form part of these advisory structures as the following examples illustrate.

Psychologists need to advise on the far-reaching mental health costs of the pandemic following extreme forms of isolation. Sociologists need to advise on the efficacy of social distancing in human settlements marked by inequality. Anthropologists need to advise on meaningful rituals of mourning when numbers are restricted for funeral attendance and family members cannot touch loved ones in their final moments. Economists must advise on how to enfranchise workers such as the self-employed.

Social work academics are needed to advise on managing family distress including the rise in domestic violence and the social effects of lockdown on children and the elderly. Political scientists must advise on the norms that should govern the relationship between government and its citizens in emergency conditions. And historians of pandemics can advise on lessons learnt that could be invaluable in making sense of the crisis and its likely course – for example, what happened with the so-called Spanish flu of 1918 when some cities or countries opened prematurely.

In partnership with the medical scientists, government would benefit from such an inclusive, multidisciplinary approach to science advice that can only strengthen the leadership response to the pandemic.

Second, it is critical that the National Coronavirus Command Council expands its focus to the regional African context. A virus, especially this rapid transmission coronavirus, does not respect national borders. In normal times, thousands of Africans travel back and forth every month between South Africa and the other SADC states and beyond. It is vital that the regional connectedness of our African neighbours is accounted for in the deliberations of the National Coronavirus Command Council. We should do so not only because of the regional, integrated character of the public health crisis but as a statement of solidarity with African neighbour states with even more precarious national health systems. Through the exchange of scientific ideas and the sharing of support, the pandemic offers a strategic opportunity for bolstering regional co-operation in fighting the pandemic.

ASSAf enjoys valuable and productive partnerships with African scientists through, for example, the Network of African Science Academies (NASAC), the African Academy of Sciences (AAS) and The World Academy of Sciences Sub-Saharan Africa Regional Partner (TWAS-SAREP). In fact, ASSAf proudly hosts the International Science Council's Regional Office for Africa that brings together African scientists to influence the international science agenda *guided by regional priorities*. ASSAf believes that the collective expertise of leading scientists from across the African region would fortify a continental response to the pandemic in line with the vision of the African Union.

Third, while it is understandable that the work of the National Coronavirus Command Council deals with managing the immediate crisis, it is not too soon for a broad range of scientific advice to be drawn on to address urgent concerns such as the future of the economy, business, education, human settlements, the environment and, of course, health care reform. The novel coronavirus has laid bare the deep inequalities in our society. We dare not reset as a country without addressing the dangerous fault lines exposed by the pandemic.

In this task, government must use the best available scientific evidence that is being generated across South Africa such as the highly significant Coronavirus Rapid Mobile Survey (CRAM). CRAM is a nationally representative sample of 10,000 people who are surveyed in six waves to collect evidence on child hunger, household welfare and social behaviour in relation to COVID-19. This kind of scientific data, as one example, will be invaluable in reconstructing South African society in a post-COVID-19 world.

In conclusion, there is an indispensable connection between science and the public trust. In the face of a pandemic, with all the fear and uncertainty of a novel virus, the credibility of governmental authority depends more than usual on winning the trust of the public. And there is no better way of maintaining that public trust than by speaking with

HOW TO CITE:
ASSAf. Public Statement on COVID-19. S Afr J Sci. 2020;116(7/8), Art. #sta0320, 2 pages. https://doi.org/10.17159/sajs.2020/sta0320

one voice on the authority of evidence-based science and employing remedies in the pandemic that uphold the values of our Constitution.

To this urgent task, ASSAf commits the multidisciplinary expertise of its membership in the ongoing service of advising government in these challenging times.

The Council of the Academy of Science of South Africa (signatories)

Professor Jonathan Jansen (President), Professor Barney Pityana (Vice President), Professor Brenda Wingfield (Vice President), Professor Sabiha Essack (General Secretary), Professor Eugene Cloete (Treasurer), Professor Stephanie Burton, Professor Norman Duncan, Professor Zeblon Vilakazi, Professor Mary Scholes, Professor Johann Mouton, Professor Wim de Villiers, Professor Refilwe Phaswana-Mafuya, Professor Wieland Gevers (Advisor), Professor Sunil Maharaj (Advisor), Professor Evance Kalula (Advisor), Professor Himla Soodyall (Executive Officer)

Additional Expert Inputs:

Professor Jimmy Volmink

Issued:

18 May 2020

COVID-19 Statement: The unanticipated costs of COVID-19 to South Africa's quadruple disease burden

As of early July 2020, the full force of COVID-19 has yet to strike in South Africa, though it has already impacted the economy further and disrupted the healthcare system. While much has been said about the former, with few exceptions, little has been discussed about the disruption to routine, essential healthcare services. The pandemic brings threats previously unknown and has reordered priorities for health. Hospitals have reprogrammed care units to accommodate COVID-19 patients, while others have temporarily closed[1].

To date, government leadership has done a remarkable job of trying to limit the spread of SARS-CoV-2 infection by promoting the most effective prevention toolkit currently available – social-behavioural measures such as social distancing, handwashing and ensuring that transportation and workplaces apply safety protocols.

The health leadership of ASSAf believes, however, that government's leadership, response and guidance to the pandemic can be further strengthened by attending to the following concerns. It is crucial that the National Coronavirus Command Council, and the structures reporting to it, ensure that by focusing on COVID-19, we do not lose sight of the opportunity costs of shifting priorities. The advisory bodies need to engage with practitioners and researchers who:

1. are focused on health services to address both the supply side and the demand side issues in innovative ways to reach the majority of the population
2. are focused on priority setting to ensure that we think more carefully about how many lives we are losing or saving from COVID-19 in contrast to lives that could be lost by disruption of essential services
3. are able to understand both the indirect and direct effects of the pandemic on co-morbidities
4. are able to help to understand the social and psychological ramifications of the pandemic and efforts to contain it, and can engage with experts who can think about ways to use this crisis to transform health services

These individuals should not be confined to clinicians and epidemiologists; to ensure the best outcome there is a need to also engage communicators, as well as behavioural and social scientists.

It is acknowledged that even in the best of times, efforts to address South Africa's quadruple disease burden is challenging. This is generated by the confluence of HIV and TB with obesity and non-communicable diseases (NCDs), continuing poor outcomes for child and maternal mortality as well as injury and violence, all the in the context of the inequities produced by our social determinants of health. In particular, the associated loss of income as a result of the lockdown has also increased hunger, and will have an impact on malnutrition. The full extent of the potentially drastic knock-on effect of COVID-19 on these critical issues is however, not appreciated. We already know that the nation-wide lockdown has resulted in both the demand side and the supply side shifts. On the supply side, human resources shift to COVID-19 has resulted in limited services for diagnosis, treatment and prevention of other health issues. On the demand side, the public has avoided health facilities, and has been affected by transport restrictions. For example, since the beginning of the lockdown, the National Institute for Communicable Diseases (NICD) has shown a 48% decline in testing for TB, as well as a 33% reduction in newly diagnosed positive cases over a 5-week period compared to the 6-week period preceding the level-5 lockdown[2]. In related data, since the beginning of the lockdown, 27 March 2020, the Gauteng Department of Health estimated that around 1,090 patients with TB, untreated HIV, and chronic conditions such as diabetes are more vulnerable to severe COVID-19. For example, COVID-19 may indirectly impact management of patients with diabetes, but more directly when diabetics are infected ensuing in poorer health outcomes. This exacerbation of the multiple-disease burden puts further strain on the national healthcare system. Not only does it have cost implications but more importantly, the human resources for delivering routine care during a pandemic, are limited[1,3].

Until we have a vaccine or a cure, COVID-19 might be amongst us for a while. Finite resources cannot be diverted solely to the pandemic. Careful priority setting, taking into consideration the costs and benefits of basic health interventions and services, are critical to the success and sustainability of public health gains of the past decades, while simultaneously addressing the COVID-19 pandemic.

There is also an urgent need for transparent and explicit decision-making that takes into account the losses and gains of shifting resources and simultaneously fostering equity in the distribution of health expenditure and subsequent outcomes.

When developing guidance for the health system and weighing different options, policy makers must consider the potential effects of COVID-19 on South Africa's complex disease burden. We need to ensure the continuity of health promotion, disease prevention and treatment services in order to avert excess death from the top four conditions and to prevent increases in their incidence during and after the COVID-19 pandemic. Even if the resources at health facilities were not crowded out by focus on COVID-19, the economic impact of the pandemic, such as increased unemployment, has the potential to erode spending power of those who can no longer afford to pay their transport costs to the clinic.

Possible impacts of privileging COVID-19 on the quadruple burden of disease include the following examples:

HIV: A six-month interruption of supply of ARVs across the whole population of HIV patients on treatment in South Africa would lead to an approximately 2-fold increase in HIV-related deaths over a one-year period. This amounts to an excess of between 83,800 and 140,900 adult HIV deaths should such a high level of disruption occur[4].

***Maternal and child health*:** Disruption to maternal and child health (MNCH) services could have a similarly devastating impact. A 9%-18% reduction on MNCH coverage over a 6-month period would lead to an additional 2,160 child deaths at a minimum in South Africa[5], despite children being at extremely low risk for severe COVID-19 illness. Of all MNCH services, sustaining routine childhood immunisation is particularly important. Measles in particular is a highly contagious disease that mostly affect children under the age of 5 years. The basic reproductive number of measles in a susceptible population is between 12 and 18; i.e. number of people that could be infected on average by every one person with measles. In contrast, while we do not know with certainty, the reproductive number of symptomatic cases of SARS-CoV-2 is thought to be approximately 2.5[6]. Previous South African research in 2010 shows what can happen when health workers are diverted to focus on a single issue - in this case a catch-up campaign for measles. In 2010, over this same three week period in 52 districts there was a 30% decrease in children completing the primary course of immunisation, a 10% decrease in antenatal visits and a 12%-17% decrease in use of injectable contraceptives[7].

Non-communicable diseases: South Africa has high rates of type 2 diabetes (12%), obesity and overweight (68% of women and 31% of men age 15+), and hypertension (35%)[8], which may actually be

HOW TO CITE:
ASSAf Standing Committee on Health. COVID-19 Statement: The unanticipated costs of COVID-19 to South Africa's quadruple disease burden. S Afr J Sci. 2020;116(7/8), Art. #sta0321, 2 pages. https://doi.org/10.17159/sajs.2020/sta0321

underestimates of the burden of these co-morbidities[9,10]. Control of these NCDs is critical since individuals with these co-morbidities are at greater risk for severe COVID-19 illness and death[11]. Any pandemic response needs to ensure that disruption to routine medical appointments and tests are minimised to prevent interruptions in NCD management and continuity of care.

Violence and injury: During the original lockdown, with alcohol sales restrictions, trauma admissions and motor vehicle injuries were reduced. Subsequently and following the opening of alcohol sales there has been a surge of both intentional and unintentional harm. At the same time, officially reported cases of gender-based violence (GBV) seemed to decrease, the number of GBV distress calls increased from 12,000 to almost 80,000 by week three of the lockdown[12], suggesting that women could not access services. A 2014 study estimated that GBV, and in particular violence against women, costs the South African economy a minimum of between R28.4 billion and R42.4 billion per annum, or between 0.9% and 1.3% of gross domestic product (GDP) in the year 2012/2013[13]. A loss of focus on the fight against GBV during the pandemic will be extremely detrimental for women and children.

In Summary:

The COVID-19 pandemic has revealed pre-existing weaknesses in our healthcare system that have been exacerbated. Efficient and equitable allocation of resources are thus critical now more than ever. Unless we prioritise interventions that that are cost-effective and address the major challenges from both the demand side and the supply side, South Africa will experience increased mortality and morbidity from diseases that have been side-lined in favour of COVID-19. This outcome will obliterate hard-won improvements in life expectancy over the past decade, thwarting any chance of South Africa reaching its SDG 2030 targets. To avert this scenario, the ASSAf Standing Committee on Health urges the National Department of Health:

1. **to engage a broad-spectrum of stakeholders without delay**
2. **to request evidence of the potential trade-offs and the consequent resource implications and**
3. **to promote a coordinated and collaborative funded research programme that encompasses multiple disciplines for both understanding the health burden complexity and for breakthrough innovations in public health and healthcare**

Authors in alphabetical order (ASSAf Standing Committee on Health Members):

Prof John Ataguba, Prof Olalekan Ayo-Yusuf , Prof Minrie Greeff, Prof Karen Hofman, Dr Elizabeth Lutge, Prof Shabir Madhi, Prof Angela Mathee, Prof Neil McKerrow, Prof Mosa Moshabela, Prof Shane Norris, Prof Solomon Rataemane, Prof Steve Reid, Prof Linda Richter, Prof Heidi van Rooyen, and Dr Caradee Wright

Issued:

3 July 2020

References:

1. Steyn D. Covid-19: SA's TB response has already been affected [Internet]. CityPress. 2020 [cited 2020 May 24]. Available from: https://city-press.news24.com/News/covid-19-sas-tb-response-has-already-been-affected-20200427
2. NICD. Impact of COVID-19 intervention on TB testing in South Africa [Internet]. 2020 May [cited 2020 Jun 3]. Available from: https://www.nicd.ac.za/wp-content/uploads/2020/05/Impact-of-Covid-19-interventions-on-TB-testing-in-South-Africa-10-May-2020.pdf
3. Almost 11,000 HIV-positive patients in Gauteng have skipped ARV collection during lockdown [Internet]. SowetanLIVE. [cited 2020 May 24]. Available from: https://www.sowetanlive.co.za/news/south-africa/2020-05-19-almost-11000-hiv-positive-patients-in-gauteng-have-skipped-arv-collection-during-lockdown/
4. Jewell BL, Mudimu E, Stover J, Kelly SL, Phillips A. Potential effects of disruption to HIV programmes in sub-Saharan Africa caused by COVID-19: results from multiple mathematical models [Internet]. 2020 [cited 2020 May 23]. Available from: https://figshare.com/articles/Potential_effects_of_disruption_to_HIV_programmes_in_sub-Saharan_Africa_caused_by_COVID-19_results_from_multiple_mathematical_models/12279914
5. Roberton T, Carter ED, Chou VB, Stegmuller AR, Jackson BD, Tam Y, et al. Early estimates of the indirect effects of the COVID-19 pandemic on maternal and child mortality in low-income and middle-income countries: a modelling study. The Lancet Global Health [Internet]. 2020;8:e901–080. https://doi.org/10.1016/S2214-109X(20)30229-1
6. Lai C-C, Shih T-P, Ko W-C, Tang H-J, Hsueh P-R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. International Journal of Antimicrobial Agents. 2020 Mar 1;55(3):105924. https://doi.org/10.1016/j.ijantimicag.2020.105924
7. Verguet S, Jassat W, Bertram MY, Tollman SM, Murray CJL, Jamison DT, et al. Supplementary immunization activities (SIAs) in South Africa: comprehensive economic evaluation of an integrated child health delivery platform. Glob Health Action [Internet]. 2013 Mar 1 [cited 2020 Jun 3];6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3587392/
8. South Africa Demographic and Health Survey 2016 [Internet]. Pretoria: National Department of Health, Statistics. Statistics South Africa; 2017 [cited 2017 Nov 8]. Available from: http://www.mrc.ac.za/bod/SADHS2016.pdf
9. Stokes A, Berry KM, Mchiza Z, Parker W, Labadarios D, Chola L, et al. Prevalence and unmet need for diabetes care across the care continuum in a national sample of South African adults: Evidence from the SANHANES-1, 2011-2012. PLoS ONE. 2017 Oct 2;12(10):e0184264. https://doi.org/10.1371/journal.pone.0184264
10. Berry KM, Parker W, Mchiza ZJ, Sewpaul R, Labadarios D, Rosen S, et al. Quantifying unmet need for hypertension care in South Africa through a care cascade: evidence from the SANHANES, 2011-2012. BMJ Glob Health [Internet]. 2017 Aug 16 [cited 2020 May 25];2(3). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5656122/
11. Kluge HHP, Wickramasinghe K, Rippin HL, Mendes R, Peters DH, Kontsevaya A, et al. Prevention and control of non-communicable diseases in the COVID-19 response. The Lancet [Internet]. 2020 May 8 [cited 2020 May 23];0(0). Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)31067-9/abstract
12. Gender-based violence cases rose by 500% since start of lockdown - Lifeline [Internet]. [cited 2020 May 25]. Available from: https://www.iol.co.za/the-star/news/gender-based-violence-cases-rose-by-500-since-start-of-lockdown-lifeline-48193496
13. Bafana Khumalo, Sisonke Msimang, Katie Bollbach. Too Costly to Ignore - The Economic Impact of Gender-Based Violence in South Africa [Internet]. Johannesburg: KPMG Human and Social Services; 2014 [cited 2020 May 25]. Available from: http://www.ci.uct.ac.za/overview-violence/reports/too-costly-to-ignore-the-economic-impact-of-GBV-in-SA

AUTHORS:
Shaun Mowat[1]
Bruce Rhodes[1]

AFFILIATION:
[1]Department of Economics, School of Accounting, Finance and Economics, University of KwaZulu-Natal, Durban, South Africa

CORRESPONDENCE TO:
Shaun Mowat

EMAIL:
mowats@ukzn.ac.za

DATES:
Received: 02 Sep. 2019
Revised: 24 Apr. 2020
Accepted: 09 May 2020
Published: 27 July 2020

HOW TO CITE:
Mowat S, Rhodes B. Identifying and assigning values to the intangible cultural benefits of ecosystem services to traditional communities in South Africa. S Afr J Sci. 2020;116(7/8), Art. #6970, 6 pages. https://doi.org/10.17159/sajs.2020/6970

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Bettine van Vuuren

KEYWORDS:
cultural ecosystem services, traditional communities, neoclassical economic valuation, deliberative approaches, types of value

FUNDING:
National Research Foundation (South Africa; grant no. 105679)

Identifying and assigning values to the intangible cultural benefits of ecosystem services to traditional communities in South Africa

Cultural ecosystem services make an important and valuable contribution to human well-being. However, research efforts in relation to ecosystems do not reflect this value, with the majority focusing on provisioning service contributions in developed countries, with cultural services largely neglected. Consideration of the contribution and importance of these services in South Africa focuses on the more tangible cultural ecosystem services such as recreational and educational benefits, with a paucity of research on the more intangible aspects such as sense of identity, belonging and worship of the ancestors. This lack of research is out of keeping with evidence of an intimate and profound relationship between the land and traditional communities in South Africa. Here we reflect on the available evidence of the nature of cultural ecosystem services to traditional communities in South Africa, and consider one aspect of the global debate on cultural ecosystem services by analysing the suitability of two predominant methods of ascertaining their value – neoclassic economic valuation and deliberative approaches – in a South African context. The types and nature of the values associated with cultural ecosystems, and the way of life of traditional communities, suggest the use of deliberative approaches is better suited to this task. It is hoped that these discussions will encourage researchers from a range of disciplines to engage in furthering research efforts in this area, and improve the evidence base on identifying, assessing and valuing these services, which are of significant importance and value to many of the most marginalised and vulnerable members of South African society.

Significance:

- Evidence from the literature suggests that cultural ecosystem services demonstrate a range of value types and ranges. The presence of a range of values puts cultural ecosystem services beyond the reach of neoclassical economic valuation methods.
- Deliberative approaches are the most suitable method for eliciting the range and dimensions of value associated with cultural ecosystem services. There is a need for research in a South African context to develop frameworks and methods to identify, assess and measure the range of values associated with cultural ecosystem services.

Introduction

Cultural ecosystem services (CES) make an important and valuable contribution to human well-being, yet extant research efforts are lacking.[1,2] In addition, there is under-representation of studies within an African context.[3] This research gap needs to be filled because traditional communities have a more intimate and profound relationship with the land and poor policy decisions raises the level of risk to these already vulnerable groups.[4-7] Through this paper, we aim to encourage a greater level of engagement and research of CES in South Africa and provide an overview of valuing cultural services and their importance to local traditional communities.

Extensive research shows that the history of a people or community is closely interwoven with the land they inhabit.[8-10] The CES benefits enjoyed by a given community depend on their history, culture and relationship with the land. Therefore, similarities and differences will exist both within and between different countries. The CES benefits will be location dependent and non-transferable. This suggests that research on CES needs to be tailored to the specific location and community in question.

This paper is concerned with issues surrounding CES and traditional communities in South Africa. Whilst all may enjoy the benefits of CES, there is evidence to suggest that, globally, they are of more significance and importance to traditional communities – a situation recognised in the literature, and by the World Bank, the Convention on Biodiversity, and many national governments, amongst others.[4-7,11,12] This is especially true where CES relate to important intangible aspects of people's lives, such as a sense of community, place and identity, which has been constructed by a community through living and interacting with the same environment over many generations.[4-7,11] Moreover, they may be relatively more important to the poorest and most marginalised communities within a country.[5] In a South African context, these two groups are often the same.

There may be similarities with the CES benefits enjoyed by traditional communities in other countries, although the way in which CES are experienced will be unique to a particular community. Amongst traditional communities in South Africa, much of this culture relates to the land,[13,14] which is more than just a productive economic asset. Land is important in many cultural practices and rituals, and the often-intangible cultural aspects of individual and community life. The significance and the cultural nature of the relationship between traditional communities and the land in South Africa is recognised in the literature.[9,10,13,14] Moreover, it is recognised and accepted as significant by the courts, including the Constitutional Court.[15] This relationship with the land helps explain why the land issue in South Africa is such an emotive one[16], and why there is often such robust opposition from affected communities whose land is threatened by development, such as the on-going case of the proposed mining operation at Xolobeni in the Eastern Cape[16,17].

Categorisation and definition

There have been numerous attempts to categorise the benefits arising from the interaction of people and ecosystems into CES.[18-20] Despite some differences, it is agreed that CES relate to the contributions ecosystems make to the material and non-material benefits to humans.[21] Tangible benefits include those relating to: physical and experiential interactions, such as educational, recreation and eco-tourism, and more intangible benefits relate to aesthetic, artistic, spiritual, sense of place, identity and social cohesion. This distinction is important as research efforts tend to concentrate on the former at the expense of the latter.[1-3] Moreover, the degree of substitutability between CES is likely to decrease to zero as they move towards the more intangible aspects related to benefits, such as identity, sense of place, and the worshipping of the ancestors. These are important considerations as the more intangible CES are the most important to traditional communities in South Africa.

A necessary step in accounting for CES values in decision-making is the identification and assessment of the relevant services, including the range and types of benefits and values they provide. This is a significant, and often difficult, undertaking due to the intangible element of CES.[5,7,21,22] A related issue is the manner in which CES have hitherto been conceptualised, with a tendency to conflate values, services and benefits.[5,22,23] In response, Fish et al.[22] call for a more explicit connection between benefits and ecosystems, to locate specific CES and their benefits within specific ecosystems. Taken collectively, these issues often leave CES as something somewhere out there, everywhere and therefore nowhere.

Current research into cultural ecosystem services

Available evidence suggests that traditional communities enjoy a more significant, experiential and tacit relationship with the land.[4,5,7] Yet research efforts on CES tend to focus on the more tangible aspects such as recreation and tourism in developed countries.[1,2] South Africa is no exception, and whilst there is some engagement with CES in South Africa,[24,25] the majority of the research reflects the global picture in being more concerned with the tangible aspects of CES. For example, there is consideration of tourism and protected areas[26-28], recreation[29] and education CES[30], but little evidence on the intangible aspects of CES in a South African context, and none that attempts to systematically identify specific CES, their associated benefits, and relate these to a specific landscape or community. This is both understandable and regrettable. Evidence on the nature of the relationship and hence the value that traditional communities ascribe to the land is therefore, at best, fragmented in a South African context. What can be pieced together suggests a significant relational value associated with the intangible aspects of CES, relating to community, sense of place and identity, and the worship of the ancestors.[9,10,13,14,31]

Identifying and accounting for CES value in decision-making is essential, especially if they relate to values that are important to marginalised and vulnerable communities.[5,32,33] The controversy over the granting of the mining licence at Xolobeni in the Eastern Cape serves as an illustration of the consequences of excluding values that are of importance to a local community. The socio-economic impact assessment[34] undertaken as part of the environmental impact assessment that accompanied the mining application, suggested that the value of the land to the local community was entirely instrumental. There was no attempt to consider other values, including CES values. If the land held only instrumental values, it could be assumed that the community would be able to substitute these for compensation, either by moving to similar land elsewhere or through financial payment. However, the robust and sustained community opposition to the proposed mining operation suggests that the land holds values other than instrumental – values that cannot be substituted. The fragmented evidence that does exist suggests that these values relate to CES.[13,15,31,35] For example, a member of the Xolobeni community explained the importance of their land thus:

> *My family are buried on this land. My father, brother and grandchild are all here, as well as many others. In Pondo culture we cannot move them. If the mine comes we will have to leave and they will stay behind. This land is sacred to us. Maybe others don't understand but it is very important.*[31]

These cultural benefits have value. For instance, a belief in the ancestors amongst many traditional communities in South Africa imbues the land with significant cultural importance, suggesting a strong sense of belonging and attachment to the land.[9,13,14,36] The sense of place attachment so engendered may have a number of benefits for individuals, improving their well-being and welfare.[37,38] Ecosystem services, which have a positive effect on the utility and well-being of at least one person in a relevant population, have an economic value, even when this value occurs outside of a market.[39]

Types of values

Identifying the type and nature of CES related to the land in a South African context is a necessary and important step for informed policy yet it may be insufficient if there is no indication of their relative importance and/or value. Moreover, how to ascertain such values depends on their type and dimensions. The literature identifies a number of different types of values. Although different terminology is often used, the literature makes a distinction between what could be termed 'held', and assigned or instrumental values.[21,40] Pascual et al.[41,42] suggest a third category of value, namely relational values, which refer to how individuals relate to each other and the natural world, and will depend in part on their held values. The more tangible aspects of CES are likely to demonstrate instrumental values whilst the more intangible aspects are likely to be more relational in value.[41-43] The available evidence suggests that it is the more intangible, relational values that are of importance to traditional communities in South Africa.[9,10,13,14,31]

Values people assign to ecosystem services are affected by the social context. Value orientation could be for oneself (self-oriented value) or for others (other-oriented value). The values are place specific and based on people's life experiences, the use and non-use of an ecosystem and its services, cultural characteristics and the economic and political setting. Moreover, values, rules and perceptions are often a social construct. Group values are thus unlikely to be the aggregate of the values of the individuals making up the group. Individuals may have a different response to questions, including questions of value depending on whether they are questioned in a group or individual setting.[2,41] Farber et al.[40] consider the ultimate origins of values lie within shared goals, the shared value system of a community or society, and hence values may be individually or group held. It may be sufficient to elicit individual values if enjoyment of the object in question is on an individual basis. However, such values are not appropriate if the object in question has a shared dimension, such as land or forest that is important to the culture of a group or community, as is often the case amongst traditional communities in South Africa.[9,10,13,14,31] The formation of such values and preferences is through social interactions, involve shared knowledge and are communal, especially in South Africa where traditional communities share the land, giving rise to shared or group values, in addition to more individually held values.[2,21,41,44]

Criticisms of Neoclassical valuation methods

Rooted in individual consumer sovereignty, the conventional neoclassical approach to capturing such values and thereby guiding policy is to estimate all costs and benefits, including non-market values, in monetary terms and compare them in a cost–benefit analysis. If the social benefits (as an increase in social welfare) of a project or policy are greater than its social costs (decrease in social welfare), the cost–benefit test is passed. Social value is the sum of the values expressed by the individuals making up the society under consideration.[45]

The simplification of reducing everything to a money metric comes at a cost, often through the neglect of values difficult to estimate in monetary terms. Attempts to overcome this difficulty and thereby add to the values measurable in a cost–benefit analysis process (including ecosystem services) is achieved through the use of various demand-based techniques. Notable examples include contingent valuation and the travel cost method.[46] Importantly, these techniques only elicit individual values. Although the academic and policymaking communities[19,47] have now generally accepted environmental valuation studies, valuation is not without its critics[48,49]. Significantly, the nature of these criticisms is especially pertinent to employing such methods to value CES.[46]

Neoclassical economics ascertains values for non-market goods through individualistic utilitarian values, with the assumption that these are pre-formed and stable, and reflect all possible values, which a utility-maximising individual will consider. The value of any good or service to society is then an aggregation of these individual values, suggesting that in maximising their own utility, individuals will ensure that society's utility is maximised.

There is little evidence to support the assumption that preferences are pre-formed and stable, especially in terms of often complex, unfamiliar non-market goods like CES.[50-52] Having to spontaneously attribute monetary values to these poorly formed preferences in a stated preference exercise is problematic and the process is unlikely to elicit meaningful and robust values.[51-54]

By definition, CES have a shared dimension. However, most stated preference studies present them in social isolation and seek individual values thereof.[52,53] This ignores important values an individual may hold as a citizen, including shared societal or community values, held both for themselves and for other members of the community.[55] This is relevant because CES are likely to have a plurality of values, including shared values – especially amongst traditional communities in South Africa[9,10,13,14,31] sharing communal land – characterised by incommensurability and missed by a process only focused on individual values.[44,49,51,52]

If values are plural, it may not be possible to represent this as a continuous utility function, implying that improvements in one dimension may not compensate for decreases in another. It is contested here that the implicit assumption of perfect substitutability of these dimensions is not a reasonable one. This has implications for the assumptions underlying cost–benefit analysis that various dimensions of value are comparable and, where necessary, can be traded off and compensated.[45] cost–benefit analysis has no established way of trading off more than one dimension of value, making it impossible to establish which outcome would in fact deliver the highest net value to society. Such an outcome would only be possible if all stakeholders could agree on how the different dimensions and types of value could be traded off against each other. Arrow's impossibility theorem showed that there is no single way that an aggregation of individual preferences can lead to the derivation of a sensible social ranking of choice.[54]

The elicitation of such values is further complicated in situations in which respondents are involved in subsistence farming and so generally are unfamiliar with money, making it difficult to express values for complex non-market goods and services in such a metric.[56] Such situations characterise the lives of many rural communities in South Africa.[57,58] Moreover, a reliance on monetary metrics may dilute the voice of the poor and the vulnerable by virtue of their limited ability to pay because of budget constraints – an important component in stated preference approaches.[5,56] Monetary aggregation can be used as what Ani[15,59] calls an aggregation weapon, often to the benefit of the rich and the detriment of the poor. This discounting of traditional communities' values through monetary metrics may under-estimate their cultural values.

Despite concerns, stated preference methods still have a role in estimating values of (some) CES. Where the values are likely to be individually held, self-regarding and relate to the more tangible aspects of CES, they may well be a useful tool in the decision- and policy-making tool bag. There are several ways in which the method can be used creatively or in conjunction with more participatory approaches, to ascertain the importance and value of CES.[23,43] However, for traditional communities in South Africa, the complex interactions between ecosystems, the services they provide and the contribution these services make to human welfare suggest that a pluralistic approach that utilises deliberative approaches be taken to ascertain their importance and value.[2,41]

Deliberative approaches

Deliberative processes, based in political, psychological and social theory, call for increased participation by citizens in the decision-making process, as public and social deliberation is fundamental to the political legitimacy of decision-making.[50] They are intended to overcome many of the shortcomings of stated preference methods in estimating values for cultural ecosystem values, allowing for a more logical, meaningful and comprehensive capturing of such values. In particular, they may address the lack of pre-formed values, the plurality of values, and issues of legitimacy and equity in decision-making.[33,50,52,53,55]

Advocates of the method suggest it is capable of improving preference and value elicitation, especially where these are not pre-formed, such as those provided by CES. A well-designed deliberative approach, in which preference formation emerges from the deliberations, should be integral to the process of valuation.[51,55] This technique explicitly recognises that certain, often deeply held values, are difficult to trade off with other values without recourse to discussion and negotiation, and that it is difficult to isolate valuation from the process of decision-making, especially where people believe there to be important moral or ethical issues at stake that warrant debate.[60]

All stakeholders, through increased participation, are more likely to view the deliberative process as fair and democratic. This legitimacy is enhanced if some affected groups consider issues other than simply economic efficiency.[51-53,55] The legitimacy and transparency of the process is important to ensure widespread acceptance of the final decision, especially from groups that may have suffered loss and/or considerable trade-offs.[23]

South African context

Whilst there is an increasing call for the use of more deliberative approaches in valuing CES, the evidence in the literature suggests that the call either is in general terms, or is implicitly considered within a developed world context.[33,43,54,61] However, there are a number of reasons why the general approach may be particularly well suited to the South African context, where there is a tradition of such processes. Discussion, negotiation, the participation of the community, and the seeking of consensus in resolving problems with a mix of different values and issues has a long history amongst many of the indigenous people of Africa, including South Africa.[13,14,59,62,63] The basis and process of the traditional resolution of conflict is one that bears many similarities with deliberative approaches. Consideration of the workings of the traditional court system provides important evidence of this.[62]

In contrast, approaches transposed from more developed countries are generally technical and scientific top-down methods, that rely on expert input and bear little resemblance to either the reality or experiences of most traditional communities in South Africa. Moreover, in a South African context, there is often a lack of capacity among relevant actors – including the state – to use such approaches, rendering these approaches ill-suited to the task for which they were designed.[7,64] South African society is multicultural and multi-ethnic; its citizens hold a myriad of different beliefs, opinions and values. Deliberative approaches have the ability to address many of the complex issues of such a society, including in ascertaining and including the value and importance of CES in decision-making.[64]

Additionally, the use of more deliberative approaches to CES has the potential to allow local and/or traditional knowledge to have a voice, and may allow for the greater inclusion of local knowledge and experience, and be more likely to capture these values, which are important to local communities.[64] As suggested above in relation to the granting of a mining licence at Xolobeni, the disregarding of values from decision-making, such as those related to CES that are important to people or communities, may result in such decisions being considered illegitimate. Deliberative processes, if utilised correctly, confer legitimacy on any decision arrived at.[21,33,64-66] This is an important consideration in general but perhaps especially so in a South African context, where the majority of the population have long been denied a voice, and where the poor and marginalised continue to struggle to be heard.[65,66]

It is beyond the scope of this review to give a detailed criticism of deliberative approaches, yet it is acknowledged that they are not a panacea. A small group vulnerable to dysfunctional power dynamics and peer pressure is unlikely to be representative. People may be excluded from the group, or not participate fully due to poor education or lack of confidence. Class, age, race, culture and gender may all affect participation.[67] Outcomes may be complex or nuanced and far less clear than a single number generated under cost–benefit analysis.[51] Inequalities in power, knowledge and communication abilities may be mirrored in inequalities in the deliberative process.[51,66] These issues need to be managed, particularly in a South African context with its great historical disparities in wealth, power, knowledge and education.[64,65] While legislation in South Africa actively encourages the use of participatory approaches to decision-making, the process is often a 'box-ticking' exercise with little real engagement or consideration of the needs or the concerns of the most marginalised members of society,[64] thus undermining any legitimacy underlying the approach.

Traditional deliberative process in South Africa are often patriarchal and autocratic.[63,67] Traditional leaders are unelected, sometimes unpopular and can have a vested interest in the status quo. They may wish to preserve the benefits that come with their position, preventing them from being an honest broker in community affairs, including the outcomes of deliberation.[67] Relying on the state for their source of power, they may have an incentive to support policies favoured by the government, rather than seeking what is in the best interests of their communities.[68]

Conclusions

Identifying, assessing and ascertaining CES values is important and necessary as these services and their associated benefits make important and valuable contributions to the well-being and welfare of individuals and natural resource dependent communities. Ignorance of the nature and magnitude of these benefits in South Africa suggests that current decisions on projects or policies that may significantly impact negatively on CES are at best sub-optimal.

Research on developing and applying frameworks to identify and assess the nature of any benefits associated with ecosystems is required. Such research should seek to ensure that there are clear links identified between the source of the CES, the types and nature of the benefits enjoyed, and the recipient of the CES. While some work has been done in developed countries to explicitly link CES benefits and landscapes,[11,22,69] such evidence is lacking in any South African context, including for traditional communities. A systematic understanding of these values and associated benefits is required. This understanding would be a precursor to any attempts to ascertain the value of such benefits, either in monetary terms or in some other form of value indicator. Continued ignorance of these issues will result in the continued loss of ecosystems and their accompanying cultural services, which is potentially harmful to those most vulnerable traditional communities. In South Africa, most CES are likely to be relational and intangible in nature, and to display shared plural and incommensurable values. Taken collectively, the nature and range of likely values put them beyond the reach of neoclassical valuation methods.

Although there is increasing interest in the use of deliberative approaches in ascertaining the importance of CES, an increase in examples of their application, and an increase in guidance of best practice in their use, the vast majority of this is considered and undertaken in a developed world context.[33,43,54,61] Indeed the importance of CES to traditional communities in South Africa is likely to differ fundamentally from the same or similar processes in a more developed world context. Such a research gap needs to be filled, to help build on and adapt a relatively rich history of the use of deliberation amongst traditional communities in South Africa. There is a need to 'Africanise' the approach to ensure its suitability for local conditions. To exclude these values is to continue to marginalise the poor and make decisions that are to their significant detriment.

Acknowledgements

We thank the anonymous reviewers for their thoughtful comments, which served to strengthen the article. This work is based on the research supported in part by the National Research Foundation of South Africa (grant no. 105679). The opinions, findings, conclusions, and recommendations expressed in this article are those of the authors alone.

Authors' contributions

This paper stems from PhD work by Shaun Mowat. Dr Bruce Rhodes was the supervisor. As such the intellectual input and conceptualisations were, overall, jointly constructed but principally driven by Shaun Mowat who as such is first author. The drafts were worked on jointly until a final version was agreed upon.

References

1. Milcu AI, Hanspach J, Abson D, Fischer J. Cultural ecosystem services: A literature review and prospects for future research. Ecol Soc. 2013;18(3), Art. #44. https://doi.org/10.5751/es-05790-180344
2. Scholte SSK, Van Teeffelen AJA, Verburg PH. Integrating socio-cultural perspectives into ecosystem service valuation: A review of concepts and methods. Ecol Econ. 2015;114:67–78. https://doi.org/10.1016/j.ecolecon.2015.03.007
3. Van der Ploeg S, De Groot R, Wang Y. The TEEB Valuation Database: Overview of structure, data and results. Wageningen: Foundation for Sustainable Development; 2010.
4. Fazey I, Proust K, Newell B, Johnson B, Fazey JA. Eliciting the implicit knowledge and perceptions of on-ground conservation managers of the Macquarie Marshes. Ecol Soc. 2006;11(1), Art. #25. https://doi.org/10.5751/es-01665-110125
5. Satz D, Gould RK, Chan KMA, Guerry A, Norton B, Satterfield T, et al. The challenges of incorporating cultural ecosystem services into environmental assessment. Ambio. 2013;42(6):675–684. https://doi.org/10.1007/s13280-013-0386-6
6. Mauro F, Hardison PD. Traditional knowledge of indigenous and local communities: International debate and policy initiatives. Ecol Appl. 2000;10(5):1263–1269. https://doi.org/10.1890/1051-0761(2000)010-[1263:tkoial]2.0.co;2
7. Kenter JO, Hyde T, Christie M, Fazey I. The importance of deliberation in valuing ecosystem services in developing countries – Evidence from the Solomon Islands. Glob Environ Chang. 2011;21:505–521. https://doi.org/10.1016/j.gloenvcha.2011.01.001
8. Fish R, Church A, Willis C, Winter M, Tratalos JA, Haines-Young R, et al. Making space for cultural ecosystem services: Insights from a study of the UK nature improvement initiative. Ecosyst Serv. 2016; 21(August):329–343. https://doi.org/10.1016/j.ecoser.2016.09.017
9. Cocks ML, Dold T, Vetter S. 'God is my forest' – Xhosa cultural values provide untapped opportunities for conservation. S Afr J Sci. 2012;108(5–6), Art. #880, 8 pages. https://doi.org/10.4102/sajs.v108i5/6.880
10. Cundill G, Bezerra JC, De Vos A, Ntingana N. Beyond benefit sharing: Place attachment and the importance of access to protected areas for surrounding communities. Ecosyst Serv. 2017;28:140–148. https://doi.org/10.1016/j.ecoser.2017.03.011
11. Klain SC, Chan KMA. Navigating coastal values: Participatory mapping of ecosystem services for spatial planning. Ecol Econ. 2012;82:104–113. https://doi.org/10.1016/j.ecolecon.2012.07.008
12. Sobrevila C. The role of indigenous peoples in biodiversity conservation: The natural but often forgotten partners. Washington DC: World Bank Group; 2008.
13. Hunter M. Reaction to conquest. Lansdowne: Citadel Press; 1979.
14. Hammond-Tooke WD. The Buntu-speaking peoples of southern Africa. 2nd ed. London: Routledge and Kegan Paul Ltd; 1974.
15. SAFLII. Baleni and Others v Minister of Mineral Resources and Others (73768/2016) [2018] ZAGPPHC 829; [2019] 1 All SA 358 (GP) (22 November 2018). Cape Town: Southern African Legal Information Institute; 2018. p. 40. Available from: http://www.saflii.org/za/cases/ZAGPPHC/2018/829.html
16. Kepe T Ntsebeza L, editors. Rural resistance in South Africa: The Mpondo revolts after 50 years. Leiden: Brill; 2011.

17. De Wet J. Collective agency and resistance to imposed development in rural South Africa. Working Papers in Development Sociology and Social Anthropology no. 373. Bielefeld: Bielefeld University; 2013. Available from: https://www.uni-bielefeld.de/soz/ab6/ag_sozialanthropologie/working_paper/WP373.pdf

18. Costanza R, Arge R, De Groot R, Farber S, Grasso M, Hannon B, et al. The value of the world's ecosystem services and natural capital. Nature. 1997;387:253–260. https://doi.org/10.1038/387253a0

19. Millennium Ecosystem Assessment: Ecosystems and human well-being: Synthesis report. Washington DC: Island Press; 2005.

20. Haines-young R, Potschin MB. Common International Classification of Ecosystem Services (CICES) V5.1: Guidance on the application of the revised structure. Nottingham: Fabis Consulting; 2018. Available from: https://cices.eu/content/uploads/sites/8/2018/01/Guidance-V51-01012018.pdf

21. Chan KMA, Satterfield T, Goldstein J. Annex A: Rethinking ecosystem services to better address and navigate cultural values. Ecol Econ. 2012;74:1–4. https://doi.org/10.1016/j.ecolecon.2011.11.011

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

23. Chan KMA, Guerry AD, Balvanera P, Klain S, Satterfield T, Basurto X, et al. Where are cultural and social in ecosystem services? A framework for constructive engagement. Bioscience. 2012;62(8):744–756. https://doi.org/10.1525/bio.2012.62.8.7

24. De Vos A, Cumming GS, Roux DJ. The relevance of cross-scale connections and spatial interactions for ecosystem service delivery by protected areas: Insights from southern Africa. Ecosyst Serv. 2017;28:133–139. https://doi.org/10.1016/j.ecoser.2017.11.014

25. Wisely SM, Alexander K, Mahlaba T, Cassidy L. Linking ecosystem services to livelihoods in southern Africa. Ecosyst Serv. 2018;30(Part C):339–341. https://doi.org/10.1016/j.ecoser.2018.03.008

26. Roques KG, Jacobson SK, McCleery RA. Assessing contributions of volunteer tourism to ecosystem research and conservation in southern Africa. Ecosyst Serv. 2018;30:382–390. https://doi.org/10.1016/j.ecoser.2017.12.014

27. Swemmer L, Mmethi H, Twine W. Tracing the cost/benefit pathway of protected areas: A case study of the Kruger National Park, South Africa. Ecosyst Serv. 2017;28:162–172. https://doi.org/10.1016/j.ecoser.2017.09.002

28. Clements HS, Cumming GS. Manager strategies and user demands: Determinants of cultural ecosystem service bundles on private protected areas. Ecosyst Serv. 2017;28:228–237. https://doi.org/10.1016/j.ecoser.2017.02.026

29. Smith MKS, Roux DJ, Hayes J. Adventure racing enables access to cultural ecosystem services at multiple scales. Ecosyst Serv. 2017;28:149–161. https://doi.org/10.1016/j.ecoser.2017.05.017

30. Smit IPJ, Roux DJ, Swemmer LK, Boshoff N, Novellie P. Protected areas as outdoor classrooms and global laboratories: Intellectual ecosystem services flowing to-and-from a National Park. Ecosyst Serv. 2017;28:238–250. https://doi.org/10.1016/j.ecoser.2017.05.003

31. Pierce T. Postcards from Xolobeni [webpage on the Internet]. c2018 [cited 2018 Dec 28]. Available from: https://thompierce.com/xolobeni

32. Klain SC, Satterfield TA, Chan KMA. What matters and why? Ecosystem services and their bundled qualities. Ecol Econ. 2014;107:310–320. https://doi.org/10.1016/j.ecolecon.2014.09.003

33. Orchard-Webb J, Kenter JO, Bryce R, Church A. Deliberative democratic monetary valuation to implement the ecosystem approach. Ecosyst Serv. 2016;21:308–318. https://doi.org/10.1016/j.ecoser.2016.09.005

34. Barbour T. Social impact assessment for Xolobeni Heavy Mineral Sands Project. Vol. 1. Sun Valley: Groundwater Consulting Services (GCS) (Pty) Ltd.; 2007.

35. Baleni D. Baleni and Others Vs Minister of Mineral Resources. Heads of Argument. 2018.

36. Costanza R, De Groot R, Sutton P, Van der Ploeg S, Anderson SJ, Kubiszewski I, et al. Changes in the global value of ecosystem services. Glob Environ Chang. 2014;26(1):152–158. https://doi.org/10.1016/j.gloenvcha.2014.04.002

37. Casakin HP, Kreitler S. Place attachment as a function of meaning assignment. Open Environ Sci. 2008;2:80–87. https://doi.org/10.2174/1876325100802010080

38. Collins D, Kearns R. Place attachment and community activism at the coast: The case of Ngunguru, Northland. N Z Geog. 2013;69(1):39–51. https://doi.org/10.1111/nzg.12005

39. Hanley N, Hynes S, Jobstvogt N, Paterson D. Economic valuation of marine and coastal ecosystems: Is it currently fit for purpose? J Ocean Coast Econ. 2015;2(1):1–38. https://doi.org/10.15351/2373-8456.1014

40. Farber SC, Costanza R, Wilson MA. Economic and ecological concepts for valuing ecosystem services. Ecol Econ. 2002;41:375–392. https://doi.org/10.1016/s0921-8009(02)00088-5

41. Pascual U, Balvanera P, Dı S, Roth E, Stenseke M, Watson RT, et al. Valuing nature's contributions to people : The IPBES approach. Curr Opin Environ Sustain. 2017;26:7–16.

42. Chan KM, Gould RK, Pascual U, Hirons M, Comberti C, Dunford R, et al. Editorial overview: Relational values: What are they, and what's the fuss about? Ecol Soc. 2018;35(1):A1–7. https://doi.org/10.1016/j.cosust.2018.11.003

43. Hirons M, Comberti C, Dunford R. Valuing cultural ecosystem services. Annu Rev Environ Resour. 2016;41(1):545–574. https://doi.org/10.1146/annurev-environ-110615-085831

44. Chee YE. An ecological perspective on the valuation of ecosystem services. Biol Conserv. 2004;120(4):549–565.

45. Pearce D, Atkinson G, Mourato S. Cost-benefit analysis and the environment: Recent developments. Paris: OECD Publications; 2006.

46. Bateman I, Carson R, Day B, Hanemann M, Hanley N, Hett T, et al. Economic valuation with stated preference techniques. Cheltenham, England: Edward Elgar; 2002.

47. South African National Treasury. A framework for considering market-based instruments to support environmental fiscal reform in South Africa – Draft policy paper. Pretoria: National Treasury; 2006.

48. Arrow K, Solow R, Portney PR, Learner EE, Radner R, Schuman H. Report of the NOAA Panel on contingent valuation. Washington DC: NOAA; 1993.

49. Norton BG, Noonan D. Ecology and valuation: Big changes needed. Ecol Econ. 2007;63(4):664–675.

50. Bunse L, Rendon O, Luque S. What can deliberative approaches bring to the monetary valuation of ecosystem services? A literature review. Ecosyst Serv. 2015;14:88–97. https://doi.org/10.1016/j.ecoser.2015.05.004

51. Spash CL. Deliberative monetary valuation (DMV): Issues in combining economic and political processes to value environmental change. Ecol Econ. 2007;63:690–699.https://doi.org/10.1016/j.ecolecon.2007.02.014

52. Lo AY, Spash CL. Deliberative monetary valuation: In search of a democratic and value plural approach to environmental policy. J Econ Surv. 2013;27(4):768–789. https://doi.org/10.1111/j.1467-6419.2011.00718.x

53. Lienhoop N, Bartkowski B, Hansju B. Informing biodiversity policy: The role of economic valuation, deliberative institutions and deliberative monetary valuation. Environ Sci Policy. 2015;54:522–532. https://doi.org/10.1016/j.envsci.2015.01.007

54. Kenter JO, Reed MS, Irvine K, O'Brien L, Brady E, Bryce R, et al. UK National Ecosystem Assessment Follow-on. Work Package Report 6: Shared, plural and cultural values of ecosystems – Summary. Cambridge, UK: UNEP-WCMC LWEC, UK; 2014. p. 48.

55. Sagoff M. Aggregation and deliberation in valuing environmental public goods: A look beyond contingent pricing. Ecol Econ. 1998;24:213–230. https://doi.org/10.1016/s0921-8009(97)00144-4

56. Christie M, Fazey I, Cooper R, Hyde T, Kenter JO. An evaluation of monetary and non-monetary techniques for assessing the importance of biodiversity and ecosystem services to people in countries with developing economies. Ecol Econ. 2012;83(2012):67–78. https://doi.org/10.1016/j.ecolecon.2012.08.012

57. Baiphethi MN, Jacobs PT. The contribution of subsistence farming to food security in South Africa. Agrekon. 2009;48(4):459–482. https://doi.org/10.1080/03031853.2009.9523836

58. Aliber M, Hart TGB. Should subsistence agriculture be supported as a strategy to address rural food insecurity? Agrekon. 2009;48(4):434–458. https://doi.org/10.1080/03031853.2009.9523835

59. Ani EI. Africa and the prospects of deliberative democracy. South African J Philos. 2013;32(3):207–219.

60. Kenter JO. Editorial: Shared, plural and cultural values. Ecosyst Serv. 2016;21:175–183. https://doi.org/10.1016/j.ecoser.2016.10.010

61. Bryce R, Irvine KN, Church A, Fish R, Ranger S, Kenter JO. Subjective well-being indicators for large-scale assessment of cultural ecosystem services. Ecosyst Serv. 2016;21:258–269. https://doi.org/10.1016/j.ecoser.2016.07.015

62. Skelton A. Tapping indigenous knowledge: Traditional conflict resolution, restorative justice. Acta Juridica. 2007;1:228–246.

63. LiPuma E, Koelble T. Deliberative democracy and the politics of traditional leadership in South Africa: A case of despotic domination or democratic deliberation? J Contemp Afr Stud. 2009;27(2):201–223. https://doi.org/10.1080/02589000902867287

64. Oelofse C, Scott D, Oelofse G, Houghton J. Shifts within ecological modernization in South Africa: Deliberation, innovation and institutional opportunities. Local Environ. 2006;11(1):61–78. https://doi.org/10.1080/13549830500396214

65. Williams JJ. Community participation: Lessons from post-apartheid South Africa. Policy Stud. 2006;27(3):197–217. https://doi.org/10.1080/01442870600885982

66. Hamilton C. Uncertain citizenship and public deliberation in post-apartheid South Africa. Soc Dyn. 2009;35(2):355–374. https://doi.org/10.1080/02533950903116091

67. Bentley KA. Are the powers of traditional leaders in South Africa compatible with women's equal rights? Three conceptual arguments. Hum Rights Rev. 2005;6(4):48–68.https://doi.org/10.1007/s12142-005-1010-3

68. De Kadt D, Larreguy HA. Agents of the regime? Traditional leaders and electoral politics in South Africa. J Polit. 2018;80(2):382–399. https://doi.org/10.1086/694540

69. Brown G, Raymond CM, Corcoran J. Mapping and measuring place attachment. Appl Geogr. 2015;57:42–53.

Check for updates

AUTHORS:
Nosiphiwe P. Ngqwala[1]
Petros Muchesa[2]

AFFILIATIONS:
[1]Environmental Health and Biotechnology Research Group, Division of Pharmaceutical Chemistry, Faculty of Pharmacy, Rhodes University, Makhanda, South Africa
[2]Water and Health Research Centre, Faculty of Health Sciences, University of Johannesburg, Johannesburg, South Africa

CORRESPONDENCE TO:
Nosiphiwe Ngqwala

EMAIL:
n.ngqwala@ru.ac.za

DATES:
Received: 31 Oct. 2018
Revised: 28 Feb. 2020
Accepted: 13 Mar. 2020
Published: 29 July 2020

HOW TO CITE:
Ngqwala NP, Muchesa P. Occurrence of pharmaceuticals in aquatic environments: A review and potential impacts in South Africa. S Afr J Sci. 2020;116(7/8), Art. #5730, 7 pages. https://doi.org/10.17159/sajs.2020/5730

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:
water pollution, water systems, impact of pharmaceuticals, environmental pollution

FUNDING:
National Research Foundation (South Africa)

Occurrence of pharmaceuticals in aquatic environments: A review and potential impacts in South Africa

The carbon footprint of pharmaceuticals through manufacturing, distribution, the incineration of unwanted pharmaceuticals as well as the packaging of pharmaceutical waste is an emerging and enormous challenge. Pharmaceuticals are major contributors to water pollution in aquatic environments that include surface water and groundwater. These pollutants arise not only from waste products but also from pharmaceutical products that have not been properly disposed of. The continuous exposure to unspecified sub-therapeutic doses of antibiotics presents risks to humans and other animals. Due to their extensive use and incomplete elimination, antibiotics have been detected in various environmental waters. The persistence of antibiotics in the environment and chronic exposure of organisms to these chemical stressors has also proven to have ecotoxicological effects. The prevailing emergence of antimicrobial resistance amongst bacteria is an area of primary concern, especially with regard to the release of antibiotics into the environment. Resistance is the acquired ability of bacterial populations to render an antibiotic ineffective as a result of a change in bacterial DNA which occurs when bacteria are subjected to an antibiotic concentration that will not kill them. A sub-lethal concentration possibly exerts a selective pressure that can result in the development of antimicrobial resistance in bacteria. It is clear that there is a need for extensive research to improve regulations and guidance on pharmaceutical waste management, pharmaceutical take-back programmes and consumer awareness.

Significance:

Pharmaceuticals are major contributors to water pollution in aquatic environments that include surface water and groundwater. This review examines the potential sources of pharmaceuticals in aquatic environments, their occurrence in South Africa, and public and environmental health implications posed by their presence. This information will provide a baseline for research and development to optimise water treatment technologies and to improve national, provincial, regional and municipal regulations and legislation.

Introduction

Pharmaceuticals are among the prime examples of contaminants that have recently been detected in water systems, with up to 90% of oral drugs that pass through the human body ending up in the water supply.[1] These emerging pollutants, therefore, present a new global water quality challenge with potentially serious implications to human health and ecosystems. Some of the challenges include the development of antibiotic-resistant bacteria and genes[2], persistence of endocrine disrupting compounds in aquatic systems, and other deleterious ecotoxicological effects[3,4]. The increased consumption, disposal and presence of human pharmaceuticals in the environment, particularly in aquatic systems, has raised concerns worldwide due to their continued introduction into the environment mainly via hospital effluents, agricultural activities and waste-water treatment plants (WWTP). Inadequate removal efficiencies of pharmaceuticals in WWTP leads to contamination of surface water, groundwater and treated drinking water.[5,6] Therefore, legislative measures need to be in place to prevent and manage any possible risks that these compounds pose to aquatic systems. In places such as Australia, Canada, the USA and some European countries where preventative measures have been implemented, the regulations are stringent and mainly apply to controlled substances and cytotoxic drugs other than pharmaceuticals and still preclude the release in sewage. However, most low- and middle-income countries, including South Africa, do not have regulations pertaining to pharmaceutical traces as pollutants in aquatic systems. This absence has resulted in very little or no environmental monitoring of these chemical stressors.[7]

The presence of pharmaceutically active compounds at low concentrations in surface water, groundwater, seawater, sediments or drinking water has been reported worldwide for more than 50 years.[8] The rapidly growing pharmaceutical industry has been pushed by the high consumption of pharmaceuticals, resulting in high frequency of detection of these contaminants in aquatic environments.[9] Pharmaceuticals such as antibiotics, analgesics, anti-inflammatories, hormones, non-steroidal anti-inflammatory drugs, beta-blockers, blood lipid regulators, and anti-epileptics have been detected in aqueous environments in concentrations ranging from nanograms to micrograms per litre.[10,11] We review the potential sources of pharmaceuticals in the aquatic environments, their occurrence in South Africa and public and environmental health implications posed by their presence.

Pathways of pharmaceuticals in aquatic environment

The number of pharmaceutical compounds that are prescribed for medical and veterinary use varies with region and country. Studies by Caldwell et al.[12] and Boxall et al.[13] have reported 3500 and 4000 pharmaceutical compounds, respectively, that are consumed globally on a daily basis. Other studies in Europe and the USA have also reported a daily intake of 5000 and 10 000, respectively.[14,15] However, the number of pharmaceuticals and their consumption in developing countries such as South Africa has not been quantified. The large consumption of pharmaceuticals by humans and other animals provides different pathways in which pharmaceuticals can enter aquatic environments. Pharmaceutical substances are biologically active and hydrophilic so that the human body can take them up easily. In the body, they are persistent, to avoid degradation before they have a curing effect. Depending on the pharmacology of a medical substance, it can be excreted as a mixture of metabolites, as unchanged substance, or conjugated with an inactivating compound attached to the molecule.[16,17] Generally, pharmaceutical compounds are metabolised by the body. Although some pharmaceuticals are completely degraded by the body, some are partially excreted through body wastes. These pharmaceuticals end up in the sewage system and eventually enter the environment through sewage leakages or discharge of waste water from sewage treatment plants which deposit into the aquatic systems (Figure 1).[18,19]

The main sources of pharmaceuticals are represented by improper disposal of medicines at domestic sites, hospital discharges, aquaculture facilities, animal farming activities, municipal and industrial WWTPs (Figure 1).[20,21] Contamination of water sources by these pharmaceuticals can occur via various pathways, which include surface run-off or leaching of human and other animal waste, and waste-water effluent discharges.[22,23] The conventional WWTPs consist of mechanical and chemical processes followed by biological treatment to remove, precipitate, and biodegrade the organic compounds based on their physicochemical characteristics. However, the majority of pharmaceuticals are not usually completely mineralised in conventional WWTPs equipped with primary and secondary processes.[21,24] Although tertiary treatments using technologies such as reverse osmosis, ultrafiltration, nanofiltration, ozonation and photolysis are more efficient in the removal of pharmaceuticals[21,25], their application in developing countries such as South Africa is relatively expensive.[26] In the conventional treatment, they are either partially retained in the sludge, or metabolised to a more hydrophilic but still persistent form and, therefore, pass through the WWTP and enter surface water and groundwater. Their removal from waste water is variable and depends on the properties of the substance and process conditions such as sludge retention time, hydraulic retention time and temperature.[27] Levels of many pharmaceutically active compounds barely reduce in waste-water treatment, and therefore are detected in treated effluents. In addition to agricultural activities, domestic and hospital waste water as main sources of pharmaceutical pollutants, and poor or no sanitation facilities in rural African communities can contribute to pharmaceutical contamination of water resources, as faecal matter is washed from the ground into the surface water during rainy seasons.[28]

Pharmaceuticals in the aquatic environment of South Africa

Pharmaceutical compounds are released mostly unchanged or as metabolites mostly conjugated to polar molecules which can be easily re-transformed to the original active compound before being diluted by the large volumes of water to sub-therapeutic concentration.[28] Once in the aqueous environment, pharmaceuticals can be distributed to various facets of the environment. Most pharmaceutical products have low volatility and therefore spread throughout the environment mixed with water as aqueous solutions or suspensions; some may get adsorbed onto soil particles and enter a food chain.[29] The presence of pharmaceuticals in aquatic environments has been well established in developed countries due to their potential environmental and health impacts. Identified pharmaceuticals belong to the following groups: antibiotics, lipid regulators, beta-blockers, steroids and related hormones, cancer drugs, diuretics, anti-epileptics, antidepressants, tranquillisers, non-steroidal anti-inflammatory drugs, anxiolytics, proton pump inhibitors and analgesics.[30] These pharmaceuticals have been identified from environmental samples, with most being commonly used for the treatment of conditions related to the central nervous, cardiovascular and digestive systems (Table 1).[31,32] Despite their wide use, relatively few studies have been done in Africa, including South Africa, to investigate the behaviour and fate of these emerging pollutants in the aquatic environment. In this review, we discuss the pharmaceuticals commonly detected in South African waters: analgesics, anti-inflammatories, antibiotics and antiretrovirals.

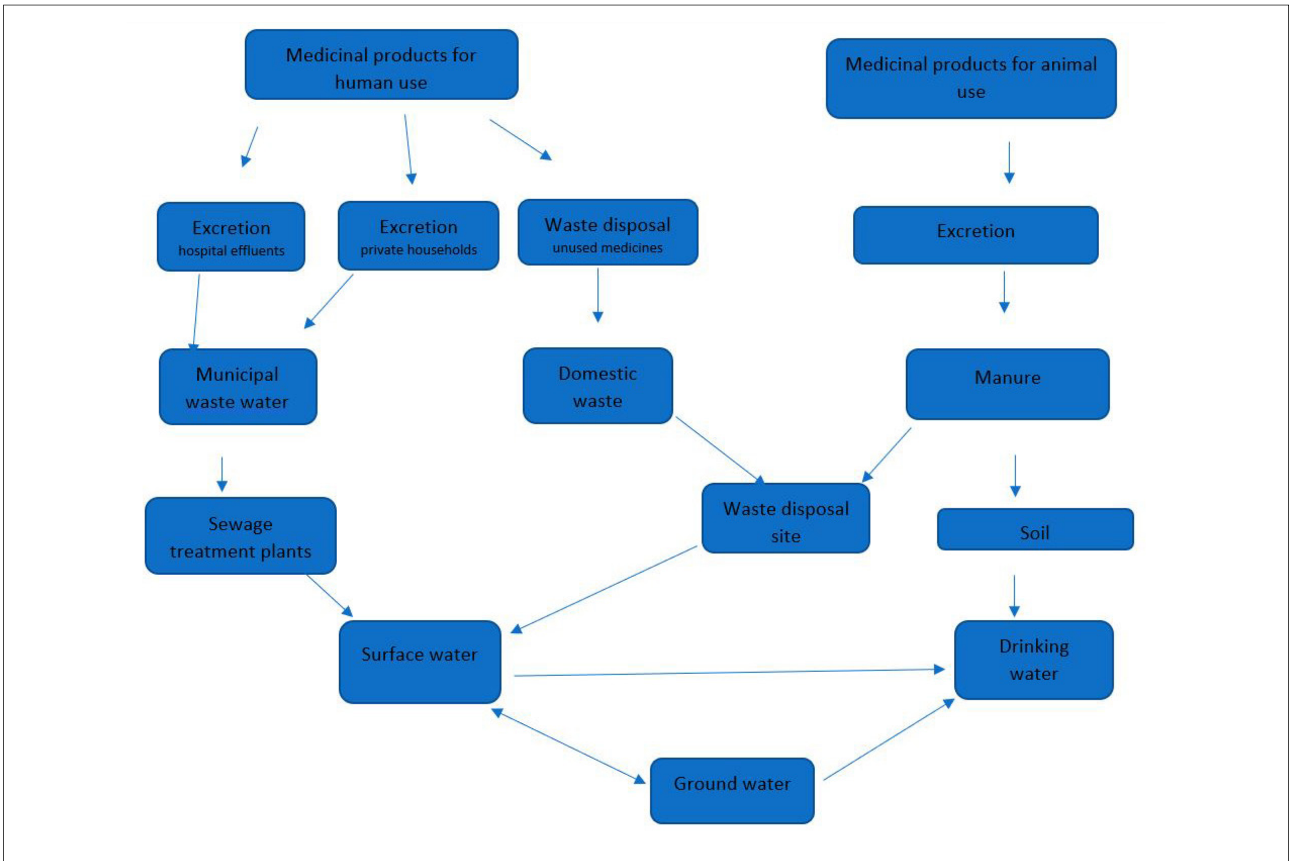


Figure 1: Sources and pathways by which pharmaceuticals may enter water resources.

Table 1: Pharmaceuticals that have been found in aqueous ecosystems (adapted from Nikolaou et al.[31] and Gros et al.[32])

Therapeutic groups	Pharmaceuticals
Cancer drugs	Cyclophosphamide and ifosfamide
Diuretics	Furosemide
Anti-epileptics	Carbamazepine
Antidepressants	Mianserin, fluoxetine, citalopram, venlafaxine, nordiazepam, oxazepam, 7-aminoflunitrazepam and paroxetine
Anxiolytics	Alprazolam, bromazepam, oxazepam
Tranquillisers	Diazepam
Steroids and related hormones	17-beta-oestradiol, oestrone, 17-alpha-ethinyloestradiol, diethylstilboestrol and diethylstilboestrol acetate
Beta-blockers	Metoprolol, propranolol, nadolol, atenolol, sotalol and betaxolol
Antibiotics	Erythromycin, ofloxacin, chlortetracyline, oxytetracycline, streptomycin, ciprofloxacin, trimethoprim, sulfamethoxazole, lincomycin, nalidixic acid, amoxicillin and azithromycin
Analgesics	Aspirin, ibuprofen, paracetamol, metamizole, codeine, indomethacin, acetaminophen, propyphenazone and phenazone
Non-steroidal anti-inflammatory drugs	Diclofenac, naproxen, ketoprofen and mefenamic acid
Lipid regulators	Bezafibrate, gemfibrozil, clofibric acid, mevastatin, pravastatin and fenofibrate
Anti-ulcer agents	Loratadine, famotidine and ranitidine
Proton pump inhibitors	Lansoprazole

Analgesics and anti-inflammatory drugs

Analgesics are a class of pharmaceuticals that are used to relieve pain whereas anti-inflammatories are a class of drugs that are used to treat or reduce inflammation and swelling. Analgesics can have both anti-inflammatory and antipyretic properties.[33] The most commonly used analgesics and anti-inflammatories in South Africa include diclofenac, naproxen, ibuprofen, acetaminophen (paracetamol), aspirin and ketoprofen.[34] Studies in South Africa have shown that ibuprofen is usually detected in aquatic environments in higher concentrations than other non-steroidal anti-inflammatory drugs. A study on the occurrence of selected pharmaceuticals in water and sediment of Umgeni River (KwaZulu-Natal, South Africa) by Matongo et al.[35] reported concentrations of ibuprofen of up to 2.94 μg/L in effluent samples at the WWTP connected to the river. Their study further noted a higher concentration of ibuprofen (12.94 μg/L) in the effluent samples compared to influent samples, showing the ineffectiveness of the WWTP in removing ibuprofen.[35] Therefore the receiving river could potentially be contaminated over time. Another study in the same province monitored naproxen, ibuprofen and diclofenac in a river and WWTP located around the city of Durban (KwaZulu-Natal, South Africa). In that study, maximum concentrations of 6.84 μg/L, 19.2 μg/L and 9.69 μg/L were reported for naproxen, ibuprofen and diclofenac, respectively. The corresponding maximum concentrations detected in WWTP effluent samples for naproxen, ibuprofen and diclofenac were 14.4 μg/L, 67.9 μg/L and 23.5 μg/L, respectively.[36] A maximum concentration of 221 μg/L was recorded in the same study from a WWTP influent sample. A study in the North West Province reported a maximum concentration of 13.7 μg/L for ibuprofen in effluent samples at a WWTP.[37] In surface water, the highest concentration of ibuprofen (62.0 μg/L) was detected in KwaZulu-Natal at the point where a tributary, Msunduzi, joins the Umgeni River (Table 2).[36]

Antibiotics

Antibiotics are chemical entities constituting an integral part of modern medicine and are an essential line of defence against pathogenic bacteria and fungi by eradicating or inhibiting their growth. Approximately 200 000 tonnes of antibiotics are produced every year, worldwide.[38,39] Although antibiotics are useful, their overuse, inappropriate use and unregulated use have contributed to the presence of their metabolites and residues in the environment. Based on their biological and physicochemical properties, antibiotics can persist for extended periods of time in the environment, where they can contaminate water sources.[40] There has been increasing concern over the past two decades about pollution of water sources by antibiotics, with several studies focusing on WWTP effluent discharges as potential hotspots of contamination.[41,42] A huge proportion of antibiotics that are ingested end up in waste water and are amongst the most detected pharmaceuticals in WWTPs around the world.[43] They are classified as recalcitrant bio-accumulative environmental pollutants, with the rate of degradation of most of these compounds such that it cannot offset their accumulation.[44] Most antibiotics prescribed and dispensed cause side effects which differ in severity according to the agent used; however, in most cases, these side effects are outweighed by the therapeutic benefits and are therefore overlooked during chemotherapy. The major concern about the presence of antibiotics in the environment is the proliferation of antimicrobial resistance genes and antimicrobial resistance bacteria, which reduce the therapeutic potential against human and non-human animal bacterial pathogens.[43]

Several antibiotics have been reported in environmental waters due to their extensive use in humans and other animals to treat infections. The global increase in the use of these antibiotics has been observed in low- and middle-income countries, including South Africa, where there is a high incidence of diseases such as cholera, typhoid, meningitis, gonorrhoea, tuberculosis (TB) and malaria that are associated with drug-resistant pathogens.[45,46] In addition, other factors such as the discharge of untreated waste, the low cost of individual antibiotics, the discharge through urine and faeces of animals and the lack of drug return programmes have been cited as contributors to the presence of antibiotics in most African waters.[46] The occurrence of antibiotics in South African waters has been investigated in WWTPs as well as in surface waters.[35,47,48,49] Among the investigated antibiotics in South African aquatic systems, a high concentration of sulfamethoxazole (59.28 μg/L) was detected at the influent of the WWTP of the Northern water works in Durban.[35] A study by Agunbiade and Moodley[50] detected ciprofloxacin at concentrations of 27 μg/L and 14 μg/L in the influent and effluent of a WWTP in KwaZulu-Natal. However, no ciprofloxacin has been reported in South African surface waters (Table 2). A more recent study to determine the concentrations of antibiotics and other pharmaceuticals at two hospital WWTPs in the North West Province reported a maximum of 45.38 μg/L and 3.22 μg/L in influent and effluent samples, respectively, for tetracycline.[37] Other pharmaceuticals – namely azithromycin, ofloxacin, norfloxacin and erythromycin – have also been detected in WWTPs, but in lower concentrations (<10 μg/L or 10 ng/L).[35,48,49,51,52] To date, there are no studies that have reported the occurrence of antibiotics in potable drinking water in South Africa.

Antiretroviral drugs

Antiretroviral (ARV) treatment comprises a combination of nucleoside reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors and protease inhibitors that act to inhibit multiple, viral targets and approved integrase inhibitors in patients with viral resistance, usually people with HIV. A combination of drugs is also used to prevent mother-to-child transmission.[52] In South Africa, at least 6.2 million people are HIV positive, according to statistics released in 2016.[53] From that number, half of these individuals are enrolled for ARV therapy, with the number increasing every year.[54] A study by Schoeman et al.[55] estimated that around 162 883 kg of ARV drugs could reach the aquatic systems in South Africa, based on high consumption figures of 2 500 000 and 2 150 880 people on ARV therapy in 2011 and 2012, respectively.[56]

The detection of ARV drugs in water resources of South Africa has been reported for the provinces of KwaZulu-Natal[57,58], Gauteng[59] and the Western Cape[60]. The ARV drugs abacavir, atazanavir, darunavir, didanosine, efavirenz, emtricitabine, indinavir, lamivudine, lopinavir, maraviroc, nevirapine, raltegravir, ritonavir, saquinavir, tenofovir, zalcitabine and zidovudine have been reported in South African aquatic environments. Among ARV drugs, emtricitabine has been detected at the highest concentration of 172 μg/L in influent samples of a WWTP in the Western Cape.[60] Low concentrations of up to 0.013 μg/L were reported for emtricitabine in a separate study in surface water samples. In South African aquatic environments, the most prevalent ARV drug detected has been efavirenz, which has been detected in Durban WWTP influent samples at high concentrations of up to 140 μg/L (Table 2).[58] Low concentrations of efavirenz, ranging between 0.002 μg/L and 2.45 μg/L, have been reported in surface waters.[58,61] Zidovudine and darunavir have also been detected at relatively high concentrations (53 μg/L and 43 μg/L, respectively) in the influent samples of WWTPs in KwaZulu-Natal (Table 2).[57]

Table 2: Maximum concentrations for pharmaceuticals quantified in South African water bodies

Pharmaceuticals	Maximum concentration (mg/L)			References
	WWTP Influent	WWTP Effluent	Surface water	
Analgesics and anti-inflammatories				
Ibuprofen	221	67.9	62.0	Kanama et al.[37]; Madikizela and Chimuka[36]
Naproxen	109.3	14.4	6.8	Madikizela and Chimuka[36]
Diclofenac	115.1	23.5	9.7	Madikizela and Chimuka[36]
Fenoprofen	80	47		Madikizela et al.[34]
Antibiotics				
Sulfamethoxazole	59.28	0.0803	6.01	Matongo et al.[35]; Nyamukamba et al.[49]
Ciprofloxacin	27	14		Agunbiade and Moodley[50]
Tetracycline	45.38	3.22		Kanama et al.[37]
Antiretroviral drugs				
Emtricitabine	172	41.7	0.013	Mosekiemang et al.[60]; Rimayi et al.[61]
Efavirenz	140	93.1	2.45	Mtolo et al.[58]
Zidovudine	53	0.5		Abafe et al.[57]
Darunavir	43	17		Abafe et al.[57]

WWTP, waste-water treatment plant

Impacts of pharmaceuticals in aquatic systems

Assessment of the environmental risk posed by pharmaceuticals and their metabolites has become a major focus in recent years because of their continuous introduction into aquatic systems. In the South African context, the fact that the country is classified as a water scarce country, makes the situation even more critical as there are relatively few water resources. Although pharmaceuticals may be present in aquatic environments in low concentrations, their extensive use, high reactivity with biological systems, continuous release and relatively low degradation makes them pseudo-persistent in aquatic environments. The potential effects to the environment and public health are chronic rather than acutely toxic, and depend on exposure, that is, bioavailability, susceptibility to the compound in question, and the degradability of the compound.[62] Globally, relatively few studies have focused on determining the toxicity of pharmaceuticals in the environment, putting humans and other animals at risk as there is a lack of information on the potential toxicity of these biologically active compounds.[10]

South Africa has been battling with epidemics of both TB and HIV/Aids over the last two decades. The country has the third worst TB epidemic worldwide (after China and India) and is among the six countries that account for 60% of the global TB burden. South Africa has the highest burden of the disease in Africa, with an estimated incidence of 454 000 cases, at a rate of 834 cases per 100 000 population as reported in 2015.[63] Prevalence of HIV is the highest in the world, with 12.7% of the population in 2016 reported to be infected.[63] In addition, South Africa reported the highest number of HIV-associated TB cases worldwide, with 59.9% of TB patients co-infected with HIV in 2017. Furthermore, patients co-infected with TB and HIV have increased susceptibility to other infections such as Pneumocystis pneumonia.[64] The prevalence of these two epidemics in South Africa means that tonnes of pharmaceuticals (ARV, antitubercular and antimicrobial drugs) are consumed to control the two epidemics. This presents a new problem concerning the presence and fate of these drugs in the aquatic environment. Pharmaceuticals can therefore pose potential environmental and public health issues that are of importance to South Africa.

Environmental impacts

Pharmaceuticals are designed to interfere with specific metabolic, enzymatic, or cell-signalling mechanisms at low concentrations through a specific mode of action in humans. The persistence of pharmaceuticals in the environment and chronic exposure to these chemical stressors can have ecotoxicological effects on non-target organisms.[62,65] The nature of the aqueous environment, together with the physicochemical properties of the pharmaceuticals, also play an important role as they determine whether the pharmaceuticals will succumb to the processes (including the employed treatment) or persist in the environment.[66] For example, fluoroquinolones, sulfonamides, trimethoprim and cephalosporins are resistant to microbial biodegradation and tend to persist in WWTP and other environmental compartments.[67] Fluoroquinolones also have strong adsorptive properties and tend to accumulate on sediments and other organic matter thus elevating their persistence in environmental matrices.[65] In addition, the presence of antimicrobial compounds in the waste water at particular levels can reduce and/or inhibit the growth of sludge bacteria that are involved in biotransforming drugs and degrading organic matter. This inhibition can decrease the efficiency of the WWTP and may result in contamination of receiving water bodies.[68]

Toxicity studies of fish, daphnia and algae have been used to predict environmental concentrations and ecological risk of most pharmaceuticals.[67,69] The biological activity of pharmaceuticals released in aquatic systems has been observed in nature and laboratory investigations have shown that they cause both acute and chronic effects. For example, the antibiotics clarithromycin sulfamethoxazole, ofloxacin, lincomycin, enrofloxacin and ciprofloxacin have been reported to be toxic to freshwater algae.[20] Low concentrations (in nanograms/litre) of the synthetic oestrogen 17-alpha-ethinyloestradiol often used in contraceptive pills have been shown to enlarge fish livers and affect the sexual characteristics of male fish in surface water. The anti-inflammatory drug diclofenac also seems to be cause for concern for aquatic organisms.[70] A study done by Fent et al.[71] reported that diclofenac was associated with the disappearance of the Orient white-backed vulture in India and Pakistan. In mammals, diclofenac has been reported to affect the liver and kidneys. Furthermore, propranolol (a beta-blocker) detected in northeastern Spain was reported to have toxic effects on zooplankton and benthic organisms.

Public health impacts

Drinking water and consumption of aquatic organisms are two ways in which humans can be exposed to pharmaceuticals that pollute the aquatic environment. Therefore, possible risks of exposure for human health are a subject of concern, especially for the countries that use surface water as their main source of drinking water. Several quantitative pharmaceutical risk assessment studies on exposure to trace levels of pharmaceuticals in drinking water, conducted in different parts of the world, have shown very low risks to human health based on toxicological data.[72,73] However, these studies do not rule out possible effects on human health as some studies are often been based on limited sets of monitoring data which do not consider long-term effects of exposure and have limited knowledge on the mixed effects of pharmaceuticals in drinking water consumed by humans.[73] In addition, some studies focus on pharmaceutical concentrations in surface water only, and not drinking water, to assess human health risk, assuming that drinking water treatment plants do not remove any of the pharmaceuticals.[74,75]

In addition to introducing toxins to drinking water, the development of resistance to antimicrobial compounds is another risk that pharmaceuticals in aquatic environments can pose to public health. The overuse and misuse of antibiotics may cause a risk to human health by promoting antibiotic-resistant bacteria and antibiotic resistance genes in aquatic environments.[2,5] This occurs as a result of the high selective pressure imposed by antibiotics on bacteria. The bacterial community that can withstand this antimicrobial pressure will survive and multiply, leading to more resistant strains in the aquatic environments.[76] The resistant genes can be horizontally transferred from animal to human pathogens and also across different classes of antibiotics used in veterinary and medical contexts, especially when the antibiotics have the same mechanism of action.[38] Horizontal gene transfer is a mechanism by which bacteria can disseminate novel traits, including acquired antimicrobial resistance. It is usually carried out by mobile DNA elements such as plasmids, transposons and integrons.[77] When bacteria acquire resistance genes to protect them against various antibiotics, they are capable of employing several biochemical types of resistance mechanisms such as antibiotic inactivation, target modification, alteration in permeability and bypass metabolic pathway.[78] Inadequate management of waste water may therefore release antibiotics, antibiotic-resistant bacteria and antibiotic resistance genes into the environment, thus presenting a potential environmental health risk. Antibiotic resistance is a major health concern; the presence of antibiotics in treated waste water is increasing and will lead to higher mortality and morbidity as untreatable infectious diseases increase.[70]

Antimicrobial resistance has become a great challenge in clinical therapy mainly because it compromises the effectiveness of antibiotics, resulting in therapeutic failure, elevated health costs, and increased morbidity and mortality rates.[79] For example, pathogens such as multidrug resistant *Klebsiella pnemoniae* cannot be treated with any antibiotic currently on the market.[76]

Research gaps and future perspectives in South Africa

The presence, persistence and toxicity of pharmaceuticals in the aquatic environment is an important subject that needs to be extensively investigated to help prevent effects on the environment and human health. There is a lack of baseline studies in South Africa to counter any effects that can be caused by the presence of pharmaceuticals in South African water systems. South Africa has particular challenges, such as a high burden of HIV/Aids and TB, with resistant strains of TB prevalent in the population. This points to a high use of ARV drugs and antibiotics, resulting in relatively high concentrations being released into aquatic environments. Therefore, there is a need to quantify and determine their fate, and extrapolate their possible long-term effects on the environment and public health. There have been few studies conducted in South Africa on pharmaceutical drugs in water and their biodegradation profile, even though South Africa has an estimated 7.7 million people living with HIV, of which 62% of all people living with HIV in South Africa are on ARV treatment. Therefore, it is expected that the concentration of ARV drugs will be considerably high in waste water and surface water. This can be determined by comprehensive quantitative pharmaceutical risk assessments that are unique to the needs of South Africa. These studies will help to analyse different aspects of pharmaceutical exposure to the environment and humans, and the toxicity and associated health risks.

Removal of pharmaceuticals by WWTPs is important in reducing disposal to aquatic environments Currently, treatment processes in South Africa cannot remove pharmaceuticals completely, resulting in their discharge into water bodies. There is a need for research to determine how these sewage treatments are efficient in removing different types of pharmaceuticals. This can also be applied to drinking water plants that use potentially contaminated surface water as their source. This will help redesign treatment plants that can exhaustively remove pharmaceuticals that can be toxic or harmful to the environment and humans. Most quantitative pharmaceutical risk assessments have focused on urban areas, neglecting the rural populations that often utilise unpurified water for drinking. Therefore, studies in rural areas in South Africa will provide relevant information on the occurrence and fate of pharmaceuticals in the environment that can be compared to studies in urban areas. Not only do these pharmaceuticals negatively affect human life but they also affect aquatic life negatively. Several studies suggest diverse negative effects on aquatic life that are exposed to these trace amounts of pharmaceuticals in their habitats. South Africa is a semi-arid area, and pharmaceutical remains are especially harmful to aquatic life during drought, because the concentration automatically becomes higher due to low volumes of water.

Waste management is a pressing issue in South Africa, and improper dumping of dangerous health-care waste is a serious concern. The groups that are responsible for this misconception are nurses, pharmacists and other health-care professionals. It also places an unacceptably high financial and human resources burden on health authorities to manage the problem. Health-care waste entering the normal domestic waste stream will end up being disposed of in municipal landfill sites. When health-care waste is placed in landfills or buried, contamination of groundwater may occur and may result in the spread of *Escherichia coli*.

If landfills are insecure, expired drugs may come into contact with children and scavenging animals. Evidence suggests that the presence of antibiotics in waste water may be contributing to antibiotic resistance, and if these antibiotics are present in waste water for a longer period, they may cause genetic effects in humans and marine life. It is therefore essential that health-care facilities dispose of all waste in accordance with national, provincial, regional and municipal regulations and legislation. Hence, it is essential to raise public awareness and encourage consumers to adopt proper disposal practices for unwanted pharmaceuticals.

Conclusion

Pharmaceuticals are used in human and veterinary medicines, aquaculture, animal husbandry, and also in agriculture for the treatment of diseases. The uptake and excretion of these pharmaceuticals pose a risk for the contamination of aquatic systems. This is a good reason for intensive research into the fate of pharmaceuticals, which covers their detection, distribution, transformation, and impact on microorganisms in the surrounding environment, the development of resistance of microorganisms to pharmaceuticals, and the possible harmful effects on the treatment of animal and human bacterial disease. In addition, there should also be continued research and development to optimise water treatment technologies and to improve national, provincial, regional and municipal regulations and legislation. There is a need for cooperation of stakeholders, manufacturers, regulators, veterinarians, pharmacists and consumers to agree on ways to decrease the harmful levels of pharmaceuticals in the environment. It is also important that the pharmaceutical industry be responsible for the development and implementation of pollution control measures and monitoring thereof.

Acknowledgements

We acknowledge the National Research Foundation of South Africa and Rhodes University for their financial support.

Authors' contributions

N.P.N. conceptualised the review; both authors wrote the manuscript.

References

1. Pothitou P, Voutsa D. Endocrine disrupting compounds in municipal and industrial wastewater treatment plants in Northern Greece. Chemosphere. 2008;73(11):1716–1723. https://doi.org/10.1016/j.chemosphere.2008.09.037

2. Berendonk TU, Manaia CM, Merlin C, Fatta-Kassinos D, Cytryn E, Walsh F, et al. Tackling antibiotic resistance: The environmental framework. Nat Rev Microbiol. 2015;13(5):310–317. https://doi.org/10.1038/nrmicro3439

3. Saravanan M, Hur JH, Arul N, Ramesh M. Toxicological effects of clofibric acid and diclofenac on plasma thyroid hormones of an Indian major carp, *Cirrhinus mrigala* during short and long-term exposures. Environ Toxicol Pharmacol. 2014;38(3):948–958. https://doi.org/10.1016/j.etap.2014.10.013.

4. Peltzer PM, Lajmanovich RC, Attademo AM, Junges CM, Teglia CM, Martinuzzi C, et al. Ecotoxicity of veterinary enrofloxacin and ciprofloxacin antibiotics on anuran amphibian larvae. Environ Toxicol Pharmacol. 2017;51:114–123. https://doi.org/10.1016/j.etap.2017.01.021

5. Houeto P, Carton A, Guerbet M, Mauclaire A, Gatignol C, Lechat P, et al. Assessment of the health risks related to the presence of drug residues in water for human consumption: Application to carbamazepine. Regul Toxicol Pharmacol. 2012;62(1):41–48. https://doi.org/10.1016/j.yrtph.2011.11.012

6. Moreno-González R, Rodriguez-Mozaz S, Gros M, Barceló D, León V. Seasonal distribution of pharmaceuticals in marine water and sediment from a Mediterranean coastal lagoon (SE Spain). Environ Res. 2015;138:326–344. https://doi.org/10.1016/j.envres.2015.02.016

7. He K, Soares A, Adejumo H, McDiarmid M, Squibb K, Blaney L. Detection of a wide variety of human and veterinary fluoroquinolone antibiotics in municipal wastewater and wastewater-impacted surface water. J Pharmaceut Biomed. 2015;106:136–143. https://doi.org/10.1016/j.jpba.2014.11.020

8. González Alonso S, Catalá M, Maroto R, Gil J, De Miguel Á, Valcárcel Y. Pollution by psychoactive pharmaceuticals in the Rivers of Madrid metropolitan area (Spain). Environ Int. 2010;36(2):195–201. https://doi.org/10.1016/j.envint.2009.11.004

9. Bayer A, Asner R, Schüssler W, Kopf W, Weiß K, Sengl M, et al. Behavior of sartans (antihypertensive drugs) in wastewater treatment plants, their occurrence and risk for the aquatic environment. Environ Sci Pollut Res. 2014;21(18):10830–10839. https://doi.org/10.1007/s11356-014-3060-z

10. Golet E, Alder A, Hartmann A, Ternes T, Giger W. Trace determination of fluoroquinolone antibacterial agents in urban wastewater by solid-phase extraction and liquid chromatography with fluorescence detection. Anal Chem. 2001;73(15):3632–3638. https://doi.org/10.1021/ac0015265

11. Yang Y, Ok Y, Kim K, Kwon E, Tsang Y. Occurrences and removal of pharmaceuticals and personal care products (PPCPs) in drinking water and water/sewage treatment plants: A review. Sci Total Environ. 2017;596:303–320. https://doi.org/10.1016/j.scitotenv.2017.04.102

12. Caldwell DJ, Mastrocco F, Anderson PD, Länge R, Sumpter JP. Predicted-no-effect concentrations for the steroid estrogens estrone, 17β-estradiol, estriol, and 17α- ethinylestradiol. Environ Toxicol Chem. 2012;31:1396–1406. https://doi.org/10.1002/etc.1825

13. Boxall AB, Rudd MA, Brooks BW, Caldwell DJ, Choi K, Hickmann S, et al. Pharmaceuticals and personal care products in the environment: What are the big questions? Environ Health Perspect. 2012;120(9):1221–1229. https://doi.org/10.1016/j.envint.2013.06.012

14. Dong Z, Senn DB, Moran RE, Shine JP. Prioritizing environmental risk of prescription pharmaceuticals. Regul Toxicol Pharmacol. 2013;65(1):60–67. https://doi.org/10.1016/j.yrtph.2012.07.003

15. Hughes SR, Kay P, Brown LE. Global synthesis and critical evaluation of pharmaceutical data sets collected from river systems. Environ Sci Technol. 2013;47(2):661–677. https://doi.org/10.1021/es3030148

16. Halling-Sørensen B, Nors Nielsen S, Lanzky P, Ingerslev F, Holten Lützhøft H, Jørgensen S. Occurrence, fate and effects of pharmaceutical substances in the environment – A review. Chemosphere. 1998;36(2):357–393. https://doi.org/10.1016/S0045-6535(97)00354-8

17. Clara M, Kreuzinger N, Strenn B, Gans O, Kroiss H. The solids retention time – a suitable design parameter to evaluate the capacity of wastewater treatment plants to remove micropollutants. Water Res. 2005;39(1):97–106. https://doi.org/10.1016/j.watres.2004.08.036

18. Madikizela L, Tavengwa N, Chimuka L. Status of pharmaceuticals in African water bodies: Occurrence, removal and analytical methods. J Environ Manag. 2017;193:211–220. https://doi.org/10.1016/j.jenvman.2017.02.022

19. Ayman Z, Işık M. Pharmaceutically active compounds in water, Aksaray, Turkey. CLEAN – Soil Air Water. 2015;43(10):1381–1388. https://doi.org/10.1002/clen.201300877

20. Klatte S, Schaefer HC, Hempel M. Pharmaceuticals in the environment – a short review on options to minimize the exposure of humans, animals and ecosystems. Sustain Chem Pharm. 2017;5:61–66. https://doi.org/10.1016/j.scp.2016.07.001

21. Yang X, Flowers RC, Weinberg HS, Singer PC. Occurrence and removal of pharmaceuticals and personal care products (PPCPs) in an advanced wastewater reclamation plant. Water Res. 2011;45(16):5218–5228. https://doi.org/10.1016/j.scitotenv.2017.04.102

22. Zorita S, Mårtensson L, Mathiasson, L. Occurrence and removal of pharmaceuticals in a municipal sewage treatment system in the south of Sweden. Sci Total Environ. 2009;407:2760–2770. https://doi.org/10.1016/j.scitotenv.2008.12.030

23. Rodriguez-Mozaz S, Chamorro S, Marti E, Huerta B, Gros M, Sànchez-Melsió A, et al. Occurrence of antibiotics and antibiotic resistance genes in hospital and urban wastewaters and their impact on the receiving river. Water Res. 2015;69:234–242. https://doi.org/10.1016/j.watres.2014.11.021

24. Galán MJG, Díaz-Cruz MS, Barceló D. Removal of sulfonamide antibiotics upon conventional activated sludge and advanced membrane bioreactor treatment. Anal Bioanal Chem. 2012;404(5):1505–1515. https://doi.org/10.1007/s00216-012-6239-5

25. Peake BM, Braund R, Tong A, Tremblay LA. The life-cycle of pharmaceuticals in the environment. Cambridge: Elsevier; 2015.

26. Li Y, Zhu G, Ng WJ, Tan SK. A review on removing pharmaceutical contaminants from wastewater by constructed wetlands: Design, performance and mechanism. Sci Total Environ. 2014;468:908–932. https://doi.org/10.1016/j.scitotenv.2013.09.018

27. Vieno NM, Tuhkanen T, Kronberg L. Determination of pharmaceutical residues in fish bile by solid phase microextraction couple with liquid chromatography–tandem mass spectrometry (LC/MS/MS). Environ Sci Technol. 2005;39:8220–8226. https://doi.org/10.1021/es051124k

28. Segura PA, Takada H, Correa JA, El Saadi K, Koike T, Onwona-Agyeman S, et al. Global occurrence of anti-infectives in contaminated surface waters: Impact of income inequality between countries. Environ Int. 2015;80:89–97. https://doi.org/10.1016/j.envint.2015.04.001

29. Comerton AM, Andrews RC, Bagley DM. Practical overview of analytical methods for endocrine-disrupting compounds, pharmaceuticals and personal care products in water and wastewater. Philos T R Soc A. 2009;367(1904):3923–3939. https://doi.org/10.1098/rsta.2009.0111

30. Ngumba E, Gachanja A, Tuhkanen T. Occurrence of selected antibiotics and antiretroviral drugs in Nairobi River Basin, Kenya. Sci Total Environ. 2016;539:206–213. https://doi.org/10.1016/j.scitotenv.2015.08.139

31. Nikolaou A, Meric S, Fatta D. Occurrence patterns of pharmaceuticals in water and wastewater environments. Anal Bioanal Chem. 2007;387(4):1225–1234. https://doi.org/10.1007/s00216-006-1035-8

32. Gros M, Petrović M, Barceló D. Development of a multi-residue analytical methodology based on liquid chromatography–tandem mass spectrometry (LC–MS/MS) for screening and trace level determination of pharmaceuticals in surface and wastewaters. Talanta. 2006;70(4):678–690. https://doi.org/10.1016/j.talanta.2006.05.024

33. Truter I. Prescribing patterns of anti-migraine medicines in South Africa using a claims database. Int J Clin Pharm. 2015;37(3):447–451. https://doi.org/10.1007/s11096-015-0089-0

34. Madikizela LM, Ncube S, Chimuka L. Analysis, occurrence and removal of pharmaceuticals in African water resources: A current status. J Environ Manage. 2020;253:109741.https://doi.org/10.1016/j.jenvman.2019.109741

35. Matongo S, Birungi G, Moodley B, Ndungu P. Occurrence of selected pharmaceuticals in water and sediment of Umgeni River, KwaZulu-Natal, South Africa. Environ Sci Pollut Res. 2015;22(13):10298–10308. https://doi.org/10.1007/s11356-015-4217-0

36. Madikizela LM, Chimuka L. Occurrence of naproxen, ibuprofen, and diclofenac residues in wastewater and river water of KwaZulu-Natal Province in South Africa. Environ Monit Assess. 2017;189(7):348. https://doi.org/10.1007/s10661-017-6069-1

37. Kanama KM, Daso AP, Mpenyana-Monyatsi L, Coetzee MA. Assessment of pharmaceuticals, personal care products, and hormones in wastewater treatment plants receiving inflows from health facilities in North West Province, South Africa. J Toxicol. 2018;2018, Art. #3751930, 15 pages. https://doi.org/10.1155/2018/3751930

38. Davies J, Davies D. Origins and evolution of antibiotic resistance. Microbiol Mol Biol Rev. 2010;74(3):417–433. https://doi.org/10.1128/MMBR.00016-10

39. Rehman M, Rashid N, Ashfaq M, Saif A, Ahmad N, Han J. Global risk of pharmaceutical contamination from highly populated developing countries. Chemosphere. 2015;138:1045–1055. https://doi.org/10.1016/j.chemosphere.2013.02.036

40. Gothwal R, Shashidhar T. Antibiotic pollution in the environment: A review. CLEAN – Soil Air Water. 2014;43(4):479–489. https://doi.org/10.1002/clen.201300989

41. Costanzo S, Murby J, Bates J. Ecosystem response to antibiotics entering the aquatic environment. Mar Pollut Bull. 2005;51(1–4):218–223. https://doi.org/10.1016/j.marpolbul.2004.10.038

42. Leung H, Minh T, Murphy M, Lam J, So M, Martin M, et al. Distribution, fate and risk assessment of antibiotics in sewage treatment plants in Hong Kong, South China. Environ Int. 2012;42:1–9. https://doi.org/10.1016/j.envint.2011.03.004

43. Ncibi M, Sillanpää M. Optimized removal of antibiotic drugs from aqueous solutions using single, double and multi-walled carbon nanotubes. J Hazard Mater. 2015;298:102–110. https://doi.org/10.1016/j.jhazmat.2015.05.025

44. Chen K, Zhou J. Occurrence and behavior of antibiotics in water and sediments from the Huangpu River, Shanghai, China. Chemosphere. 2014;95:604–612. https://doi.org/10.1016/j.chemosphere.2013.09.119

45. Ndihokubwayo JB, Yahaya AA, Desta AT, Ki-Zerbo G, Odei EA, Keita B, et al. Antimicrobial resistance in the African region: Issues, challenges and actions proposed. African Health Monitor. 2013;16:27–30.

46. Laxminarayan R, Duse A, Wattal C, Zaidi AK, Wertheim HF, Sumpradit N, et al. Antibiotic resistance – the need for global solutions. Lancet Infect Dis. 2013;13(12):1057–1098. https://doi.org/10.1016/s1473-3099(13)70318-9

47. Fekadu S, Alemayehu E, Dewil R, Van der Bruggen B. Pharmaceuticals in freshwater aquatic environments: A comparison of the African and European challenge. Sci Total Environ. 2019;654:324–337. https://doi.org/10.1016/j.scitotenv.2018.11.072

48. Faleye AC, Adegoke AA, Ramluckan K, Fick J, Bux F, Stenström TA. Concentration and reduction of antibiotic residues in selected wastewater treatment plants and receiving waterbodies in Durban, South Africa. Sci Total Environ. 2019;678:10–20. https://doi.org/10.1016/j.scitotenv.2019.04.410

49. Nyamukamba P, Moloto MJ, Tavengwa N, Ejidike IP. Evaluating physicochemical parameters, heavy metals, and antibiotics in the influents and final effluents of South African wastewater treatment plants. Pol J Environ Stud. 2019;28(3):1305–1312. https://doi.org/10.15244/pjoes/85122

50. Agunbiade FO, Moodley B. Occurrence and distribution pattern of acidic pharmaceuticals in surface water, wastewater, and sediment of the Msunduzi River, KwaZulu-Natal, South Africa. Environ Toxicol Chem. 2015;35:36–46. https://doi.org/10.1002/etc.3144

51. Swartz CD, Genthe B, Chamier J, Petrik LF, Tijani JO, Adeleye AP. Emerging contaminants in wastewater treated for direct potable re-use: The human health risk priorities in South Africa. WRC project no. K5/2369. Pretoria: Water Research Commission; 2016.

52. Shah AS, Stelzle D, Lee KK, Beck EJ, Alam S, Clifford S, et al. Global burden of atherosclerotic cardiovascular disease in people living with HIV: Systematic review and meta-analysis. Circulation. 2018;138(11):1100–1112. https://doi.org/10.1161/circulationaha.117.033369

53. Statistics South Africa. Mid-year population estimates. Pretoria: Statistics South Africa; 2016. Available from: http://www.statssa.gov.za/publications/P0302/P03022016.pdf

54. NDOH. National consolidated guidelines for the prevention of mother-to-child transmission of HIV and management of HIV in children, adolescents and adults [webpage on the Internet]. c2015 [cited 2020 Feb 28]. Available from: https://www.health-e.org.za/2015/

55. 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), Art. #272. https://doi.org/10.4172/2157-7064.1000272

56. World Health Organization. Global update on HIV treatment 2013: Results, impact and opportunities [webpage on the Internet]. c2013 [cited 2020 Feb 28]. Available from: www.who.int/hiv/pub/progressreports/update2013/en

57. Abafe OA, Späth J, Fick J, Jansson S, Buckley C, Stark A, et al. LC-MS/MS determination of antiretroviral drugs in influents and effluents from wastewater treatment plants in KwaZulu-Natal, South Africa. Chemosphere. 2018;200:660–670. https://doi.org/10.1016/j.chemosphere.2018.02.105

58. Mtolo SP, Mahlambi PN, Madikizela LM. Synthesis and application of a molecularly imprinted polymer in selective solid-phase extraction of efavirenz from water. Water Sci Technol. 2019;79(2):356–365. https://doi.org/10.2166/wst.2019.054

59. Schoeman C, Dlamini M, Okonkwo OJ. The impact of a wastewater treatment works in Southern Gauteng, South Africa on efavirenz and nevirapine discharges into the aquatic environment. Emerg Contam. 2017;3(2):95–106. https://doi.org/10.1016/j.emcon.2017.09.001

60. Mosekiemang TT, Stander MA, De Villiers A. Simultaneous quantification of commonly prescribed antiretroviral drugs and their selected metabolites in aqueous environmental samples by direct injection and solid phase extraction liquid chromatography-tandem mass spectrometry. Chemosphere. 2019;220:983–992. https://doi.org/10.1016/j.chemosphere.2018.12.205

61. Rimayi C, Odusanya D, Weiss JM, De Boer J, Chimuka L. Contaminants of emerging concern in the Hartbeespoort Dam catchment and the uMngeni River estuary 2016 pollution incident, South Africa. Sci Total Environ. 2018;627:1008–1017. https://doi.org/10.1016/j.scitotenv.2018.01.263

62. Fabbri E. Pharmaceuticals in the environment: Expected and unexpected effects on aquatic fauna. Ann N Y Acad Sci. 2015;1340:20–28. https://doi.org/10.1111/nyas.12605

63. World Health Organization (WHO). Global tuberculosis report 2016. Geneva: WHO; 2016. Available from: https://apps.who.int/iris/bitstream/handle/10665/250441/9789241565394-eng.pdf;jsessionid=71484E5429BB1D586F26C66F517D740F

64. World Health Organization (WHO). Global tuberculosis report 2018. Geneva: WHO; 2018. Available from: https://apps.who.int/iris/handle/10665/274453

65. Mitani K, Kataoka H. Determination of fluoroquinolones in environmental waters by in-tube solid-phase microextraction coupled with liquid chromatography–tandem mass spectrometry. Anal Chim Acta. 2006;562(1):16–22. https://doi.org/10.1016/j.aca.2006.01.053

66. Michael I, Rizzo L, McArdell C, Manaia C, Merlin C, Schwartz T, et al. Urban wastewater treatment plants as hotspots for the release of antibiotics in the environment: A review. Water Res. 2013;47(3):957–995. https://doi.org/10.1016/j.watres.2012.11.027

67. Ebele A, Abou-Elwafa Abdallah M, Harrad S. Pharmaceuticals and personal care products (PPCPs) in the freshwater aquatic environment. Emer Cont. 2017;3(1):1–16. https://doi.org/10.1016/j.emcon.2016.12.004

68. Magwira CA, Aneck-Hahn N, Taylor MB. Fate, occurrence and potential adverse effects of antimicrobials used for treatment of tuberculosis in the aquatic environment in South Africa. Environ Pollut. 2019;254, Art. #112990. https://doi.org/10.1016/j.envpol.2019.112990

69. Godoy A, Kummrow F, Pamplin P. Occurrence, ecotoxicological effects and risk assessment of antihypertensive pharmaceutical residues in the aquatic environment – A review. Chemosphere. 2015;138:281–291. https://doi.org/10.1016/j.chemosphere.2015.06.024

70. Yao D, Xi X, Huang Y, Hu H, Hu Y, Wang Y, et al. A national survey of clinical pharmacy services in county hospitals in China. PLoS ONE. 2017;12(11), e0188354. https://doi.org/10.1371/journal.pone.0188354

71. Fent K, Weston A, Caminada D. Erratum to "Ecotoxicology of human pharmaceuticals". Aquat Toxicol. 2006;78(2):207. https://doi.org/10.1016/j.aquatox.2005.09.009

72. De Jesus Gaffney V, Almeida C, Rodrigues A, Ferreira E, Benoliel M, Cardoso V. Occurrence of pharmaceuticals in a water supply system and related human health risk assessment. Water Res. 2015;72:199–208. https://doi.org/10.1016/j.watres.2014.10.027

73. Houtman C, Kroesbergen J, Lekkerkerker-Teunissen K, Van der Hoek J. Human health risk assessment of the mixture of pharmaceuticals in Dutch drinking water and its sources based on frequent monitoring data. Sci Total Environ. 2014;496:54–62. https://doi.org/10.1016/j.scitotenv.2014.07.022

74. Cunningham V, Perino C, D'Aco V, Hartmann A, Bechter R. Human health risk assessment of carbamazepine in surface waters of North America and Europe. Regul Toxicol Pharmacol. 2010;56(3):343–351. https://doi.org/10.1016/j.yrtph.2009.10.006

75. Vulliet E, Cren-Olivé C. Screening of pharmaceuticals and hormones at the regional scale, in surface and groundwaters intended to human consumption. Environ Pollut. 2011;159(10):2929–2934. https://doi.org/10.1016/j.envpol.2011.04.033

76. Marti E, Variatza E, Luis Balcazar J, Balcazar JL. The role of aquatic ecosystems as reservoirs of antibiotic resistance. Trends Microbiol. 2014;22:36–41. https://doi.org/10.1016/j.tim.2013.11.001

77. Jain H, Mulay S, Mullany P. Persistence of endodontic infection and *Enterococcus faecalis*: Role of horizontal gene transfer. Gene Rep. 2016;5:112–116. https://doi.org/10.1016/j.genrep.2016.09.010

78. Giedraitienė A, Vitkauskienė A, Naginienė R, Pavilonis A. Antibiotic resistance mechanisms of clinically important bacteria. Medicina. 2011;47(3), Art. #19. https://doi.org/10.3390/medicina47030019

79. Munier A, De Lastours V, Barbier F, Chau F, Fantin B, Ruimy R. Comparative dynamics of the emergence of fluoroquinolone resistance in staphylococci from the nasal microbiota of patients treated with fluoroquinolones according to their environment. Int J Antimicrob Agents. 2015;46(6):653–659. https://doi.org/10.1016/j.ijantimicag.2015.09.004

Check for updates

AUTHORS:
Ridhwannah Gangat[1]
Heidi van Deventer[1,2]
Laven Naidoo[2]
Elhadi Adam[1]

AFFILIATIONS:
[1]Archaeology and Environmental Studies, Department of Geography, University of the Witwatersrand, Johannesburg, South Africa
[2]Council for Scientific and Industrial Research, Pretoria, South Africa

CORRESPONDENCE TO:
Heidi van Deventer

EMAIL:
HvDeventer@csir.co.za

DATES:
Received: 08 July 2019
Revised: 13 Mar. 2020
Accepted: 18 Mar. 2020
Published: 29 July 2020

HOW TO CITE:
Gangat R, Van Deventer H, Naidoo L, Adam E. Estimating soil moisture using Sentinel-1 and Sentinel-2 sensors for dryland and palustrine wetland areas. S Afr J Sci. 2020;116(7/8), Art. #6535, 9 pages. https://doi.org/10.17159/sajs.2020/6535

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Yali Woyessa

KEYWORDS:
soil moisture content, volumetric water content, hydroperiod, random forest regression, machine learning regression models

FUNDING:
Water Research Commission (South Africa), Council for Scientific and Industrial Research (South Africa)

Estimating soil moisture using Sentinel-1 and Sentinel-2 sensors for dryland and palustrine wetland areas

Soil moisture content (SMC) plays an important role in the hydrological functioning of wetlands. Remote sensing shows potential for the quantification and monitoring of the SMC of palustrine wetlands; however, this technique remains to be assessed across a wetland–terrestrial gradient in South Africa. The ability of the Sentinel Synthetic Aperture Radar (SAR) and optical sensors, which are freely available from the European Space Agency, were evaluated to predict SMC for a palustrine wetland and surrounding terrestrial areas in the grassland biome of South Africa. The percentage of volumetric water content (%VWC) was measured across the wetland and terrestrial areas of the Colbyn Wetland Nature Reserve, located in the City of Tshwane Metropolitan Municipality of the Gauteng Province, using a handheld SMT-100 soil moisture meter at a depth of 5 cm during the peak and end of the hydroperiod in 2018. The %VWC was regressed against the Sentinel imagery, using random forest, simple linear and support vector machine regression models. Random forest yielded the highest prediction accuracies in comparison to the other models. The results indicate that the Sentinel images have the potential to be used to predict SMC with a high coefficient of determination (Sentinel-1 SAR = R^2>0.9; Sentinel-2 optical = R^2>0.9) and a relatively low root mean square error (Sentinel-1 RMSE =<17%; Sentinel-2 optical = RMSE <21%). Predicted maps show higher ranges of SMC for wetlands (> 50%VWC; $p<0.05$) compared to terrestrial areas, and therefore SMC monitoring may benefit the inventorying of wetlands, as well as monitoring of their extent and ecological condition.

Significance:

- The freely available and space-borne Sentinel sensors show potential for the quantification of surface soil moisture across a wetland–terrestrial gradient.
- Significant differences between the surface soil moisture of palustrine wetlands and terrestrial areas, imply that inventorying and monitoring of the extent and hydroperiod of palustrine wetlands can potentially be done.

Introduction

Globally, it is estimated that more than 85% of wetlands have been transformed, primarily owing to the loss of natural habitat resulting from land conversion, but also as a consequence of other pressures including changes to the hydrological regime, water pollution and invasive species.[1,2] In South Africa, the extent of natural and transformed wetlands is unknown[3,4], although sub-national studies have shown that 58% of wetlands in the Umfolozi secondary catchment (in the KwaZulu-Natal Province), had already been transformed irreversibly already by the 1990s[5]. Regional and automated inventorying and monitoring of wetland extent is critical for their conservation and management.

Remote sensing has played an important role in the detection and monitoring of the inundated sections of wetlands, both internationally and in South Africa. For example, the Global Inundation Extent from Multi-Satellites (GIEMS) and Global Surface Water Explorer products have been generated from coarse-scale satellite imagery which reflect the extent of inundation of larger artificial and natural wetlands.[6,7] In South Africa, the national land-cover products include open water classes[8], and, more recently, monitoring of the monthly extent of inundation[9] has improved our ability to characterise the hydroperiod of wetlands. Yet early estimations of the extent of wetland cover showed that nearly 89% of wetlands are either arid or covered with vegetation (palustrine) in nature, whereas only 11% may be inundated.[4] The development of indices which would characterise the extent and nature of palustrine wetlands is, therefore, a gap and top priority for South Africa.

Surface soil moisture is an important variable of palustrine wetlands that could potentially inform on the extent, hydroperiod and ecological condition of wetlands.[10] Wetlands are areas where the soil becomes intermittently (±3 months in a year or less), seasonally (3–9 months per annum) or permanently (>9 months per annum) saturated within 50 cm from the soil surface.[11,12] Traditional in-situ methods to measure percentage volumetric water content (%VWC), which use the dielectric properties within the soil (for example, the gravimetric method), are limited in representing the spatial and temporal variations of soil moisture[13,14], and are also labour intensive, time consuming and costly[15]. Regional prediction of soil moisture content (SMC) through surface hydrological models has been stymied by the availability of coarse-scale data sets, in the range of ~10–100 km spatial resolution[16-19], which prohibits the prediction of SMC for the small wetland features of an arid to semi-arid country such as South Africa. As precipitation is highly variable across South Africa[20], frequent temporal updates of SMC would be essential to improve the understanding of soil saturation periods for wetlands. The prediction of SMC using space-borne sensors offers several advantages to traditional measurements and other modelling methods, the most important being that these sensors are able to estimate SMC frequently at regional scale.[21]

Several studies have investigated the capability of space-borne sensors (including both Synthetic Aperture Radar (SAR) and optical sensors) in estimating surface SMC (see review by Filion et al.[22]). Some of the most recent studies done using the latest available sensor technology managed to achieve high coefficients of determination ($R^2 = >0.72$) and low root mean square errors (RMSEs $<13\%$).[23-26] In general, the active SAR sensors were restricted to C-band sensors, which can penetrate between 5 cm and 10 cm into the canopy or soil, at a spatial resolution ranging from 10 m to 100 m. Passive L-band SAR sensors, on the other hand, can penetrate to a depth of 30 cm, and have spatial resolutions ranging from 3 km to 35 km[27]; however, they are costly and not suitable for monitoring small wetlands. SAR sensors have the advantage over optical sensors in that they are not affected by cloud cover, yet scattering of the signal on highly textured areas and dense vegetation may reduce the accuracy of prediction. Optical sensors, in contrast to SAR sensors, cannot penetrate soil depth to estimate SMC, but infer SMC from the total reflectance of soil, vegetation and water across the visible/near infrared (VNIR: 400 nm–1200 nm) and the short-wave infrared (SWIR: 1200 nm–2500 nm) regions of the electromagnetic spectrum.[28] Unlike SAR sensors, detection of SMC using optical sensors is affected by cloud cover, cover texture and the density of vegetation.[29] The incorporation of vegetation indices, multiple phenological periods and different incident angles in the SMC predictions decreased the RMSE, and in this way, reduced the impact of vegetation on the prediction.[26,30,31] However, a study done by Hornacek et al.[32] showed that vegetation $\leq$ 1 kg/m^2 had very little influence on the estimation of SMC in terrestrial systems. In general, critical limitations of these SAR and optical sensors for the estimation and monitoring of SMC are the spatial resolution of detection and cost. In an arid to semi-arid country such as South Africa, wetlands are small in extent, and palustrine wetlands often are composed of a mosaic of soil and vegetation cover. SMC therefore offers the advantage of a single variable to monitor across the landscape to inform wetland extent, hydroperiod and ecological condition.

The Sentinel SAR and optical sensors were launched between 2014 and 2017 by the European Space Agency (ESA) and consist of twin satellites each: Sentinel SAR Sentinel-1A and Sentinel-1B (S1A, S1B) and optical Sentinel-2A and Sentinel-2B (S2A, S2B). Images from these sensors were made freely available to the public and consequently offered several new opportunities for testing the capabilities of space-borne sensors in the quantification and monitoring of features on earth, including SMC of palustrine wetlands. These operational space-borne sensors hold promise for predicting SMC at a regional scale for palustrine wetlands in South Africa, but these sensors are yet to be assessed for their capabilities in the temperate regions of the southern hemisphere. In addition, none of the studies assessed SMC across a terrestrial–wetland gradient, and whether thresholds can be selected for determining the maximum extent of a wetland. Average %VWC values measured in terrestrial systems abroad ranged from 24% to 45%, while in-situ SMC measured in wetlands was generally approximately above 50%.[26,31,33]

The aim of this study was therefore to determine whether the Sentinel-1 and Sentinel-2 sensors have potential to be used in estimating SMC across a gradient of palustrine wetlands to terrestrial areas in the grassland biome of South Africa. The grassland biome extends to approximately a third of the land mass of South Africa[34], hosts a number of palustrine wetlands, and is one of the biomes most threatened by multiple pressures such as land conversion to urban areas and mining[35]. Our objectives were to (1) assess the capability of the SAR and optical Sentinel sensors in estimating SMC and (2) determine whether there are significant differences between SMC values between terrestrial areas and the wetlands, to inform a proposed threshold of wetland extent. We hope the outcome will contribute to South Africa's National Wetland Monitoring Programme.[36]

Methods

Study area

The study area (approximately 70 ha) was situated in the Colbyn Wetland Nature Reserve (CWNR), located in the City of Tshwane Metropolitan Municipality of the Gauteng Province, South Africa (25°44'21.67"S; 28°15'15.35"E) (Figure 1). The study area falls within the grassland biome and experiences a temperate climate with the rainfall season between September and March and the dry season between April and August. This ecoregion experiences an average summer rainfall of between 650 mm and 750 mm per annum and an evapotranspiration of 524 mm annually.[37]

The Hartbeesspruit River drains the catchment, flowing in a northeasterly direction up to a dolerite dyke. The dolerite dyke forms a barrier on the northern side of the study area, and even though it has been breached by the river, water backs up southwest of the dyke, resulting in the formation of a channelled valley-bottom wetland. The adjacent hillside slopes contribute seepage towards the wetland, and groundwater also contributes interflow towards the channel.[38] Most of the wetland remains

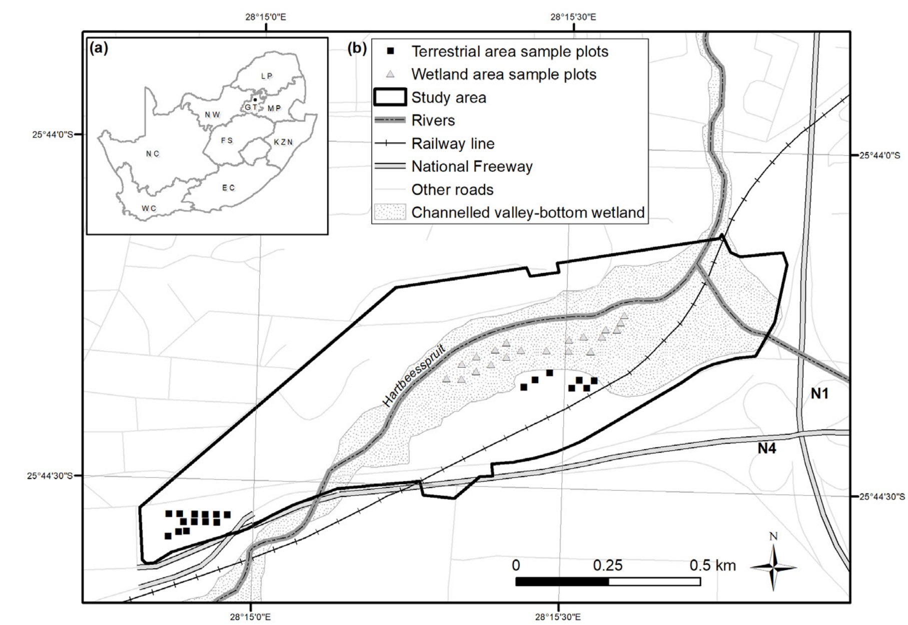


Figure 1: (a) The location of the study area, the Colbyn Wetland Nature Reserve, within the Gauteng Province of South Africa. (b) A channelled valley-bottom wetland forms where the Hartbeesspruit (River) backs up southwest of a dolerite dyke. The locations of sample plots in the wetland and terrestrial areas are indicated.

permanently saturated throughout the year, and peat has been found in the centre part of the wetland near the channel (extent estimated at 4.68 ha).[39] The CWNR is considered a palustrine wetland, with the full cover of grasses and sedges dominating in the temporary saturated zones of the wetland (e.g. cotton wool grass or *Imperata cylindrica*) and reeds (*Phragmites australis),* bulrush *(Typha capensis)* and the lesser pond sedge (*Carex acutiformis)* in the permanently saturated zones.[40] Only the river channel has open water partly visible on satellite imagery, lined by exotic tree species, such as the weeping willow (*Salix babylonica*) and poplars (*Populus × canescense*), with the latter also found in a part of the permanently saturated zone of the wetland.

The CWNR is exposed to a number of pressures and impacts. Drainage has been disrupted by numerous roads, resulting in high energy run-off leading to erosion of the wetland.[41] Weirs have been built along the channel to alleviate the effects of erosion in order to prevent further degradation of the wetland.[42] The Koedoespoort Railway line crosses through the CWNR, causing soil compaction which results in an increase in soil moisture saturation in parts of the study area.

Data collection

Image acquisition and pre-processing

Sentinel-1 SAR data acquisition and pre-processing

The S1A and S1B SAR C-band images were downloaded as Ground Range Detected (GRD) data (10x10 m spatial resolution) from the Copernicus website (https://scihub.copernicus.eu/dhus/#/home) (Table 1). The data were acquired in Interferometric Wide (IW) for both the vertical-transmit, vertical-receive (VV) and vertical-transmit, horizontal-receive (VH) polarisation modes. S1A and S1B GRD data were pre-processed using ESA's Sentinel Application Platform (SNAP) software version 6.0 (2018) for radiometric calibration, multi-looking and terrain correction. Multi-looking was applied to each Sentinel-1 image (applying 2 and 2 multi-looking factors for range and azimuth, respectively) to reduce the speckle present in the images which subsequently converts the 10-m spatial resolution image to a 20-m spatial resolution image. Radiometric calibration of the SAR images converts the data from a digital number format to backscatter in *sigma naught* or *sigma dB*. Backscatter signal errors associated with terrain, orientation and geo-referencing of the imagery were corrected with the Range Doppler Terrain Correction using the Shuttle Radar Topography Mission 3 arc-seconds 30-m digital elevation model.[43]

Sentinel-2 data acquisition and pre-processing

Sentinel-2 optical images (S2A and S2B), processed to Level 1C, were acquired as close as possible to the SAR images, but avoiding imagery with >20% cloud coverage (Table 1). The images were downloaded from the United States Geological Survey (USGS) Earth Explorer website[44] as 10 individual spectral bands (Table 2). Bands 2, 3, 4 and 8 are provided by ESA at a 10-m spatial resolution while bands 5, 6, 7, 8a, 11 and 12 are at 20-m spatial resolution. Bands with a 60-m spatial resolution (bands 1, 9 and 10) are mainly used in atmospheric correction and cirrus-cloud screening and were not required for estimating the percentage SMC (%SMC). Three procedures were necessary for pre-processing the Sentinel-2 satellite images: (1) atmospheric correction; (2) resampling the 20-m multispectral images to 10 m using the Sen2Cor algorithm using the default settings in SNAP; and (3) sub-setting to extract the study area.

In-situ soil moisture collection

Prior to sampling, several field visits were made to plan sampling positions in the wetland and terrestrial areas. The extent of the wetland was guided by the National Wetlands Map 5[4], and identified through characterising the nature of the soil, extracted from the ground using a soil auger, as well as vegetation species as indicator plants. Subsequent to this first scoping field visit, the boundaries of the channelled valley-bottom wetland were adjusted to match field observations of terrestrial and wetland areas, resulting in an extent of 28.7 ha (Figure 1). The sampling period was selected to coincide with the peak hydroperiod, which would help to detect the maximum level and extent of soil moisture in the wetland for wetland inventorying. A stratified random sampling method was chosen to collect in-situ*,* %VWC measurements in the wetland and terrestrial areas. Stratified random sampling ensured that the point sample measurements were well distributed in order to represent the

Table 1: Acquisition dates and times of the Sentinel-1A/1B (S1A, S1B) and Sentinel-2A/2B (S2A, S2B) images as well as the dates of in-situ observations

Sensor	Scene ID no.	Date (2018)	Time of overpass of sensor (GMT+ 2 h)	Hydroperiod
S1A	S1A_IW_GRDH_1SDV_20180326T164655_20180326T164720_021188_0246E	26 March	18:44	Peak
S1B	S1B_IW_GRDH_1SDV_20180328T033428_20180328T033453_010226_012958_A3E7	28 March	05:33	Peak
S2A	L1C_T35JPM_A014432_20180328T081650	28 March	09:45	Peak
S2B	L1C_T35JPM_A006024_20180502T081534	02 May	09:45	End

Table 2: Spectral bands and associated wavelength ranges of the optical Sentinel-2A and 2B images (adapted from the European Space Agency Sentinel online, 2019)

Spatial resolution (m)	Band number	S2A		S2B		Use
		Central wavelength (nm)	Bandwidth (nm)	Central wavelength (nm)	Bandwidth (nm)	
10	2	442.7	21	442.2	21	Aerosol correction, land measurement
	3	492.4	66	492.1	66	Land measurement
	4	559.8	36	559.0	36	Land measurement
	8	664.6	31	664.9	31	Land measurement, water vapour correction
20	5	704.1	15	703.8	16	Land measurement
	6	740.5	15	739.1	15	Land measurement
	7	782.8	20	779.7	20	Land measurement
	8a	832.8	106	832.9	106	Land measurement, water vapour
	11	864.7	21	864.0	22	Land measurement
	12	945.1	20	943.2	21	
60	1	1373.5	31	1376.9	30	Aerosol correction
	9	1613.7	91	1610.4	94	Water vapour correction
	10	2202.4	175	2185.7	185	Cirrus detection

wetland and terrestrial sampling areas. Available Sentinel images were downloaded and used to determine suitable positions for the sampling plots which were positioned to fit both the Sentinel SAR and optical image pixels. A sampling plot the size of 10 m x 10 m was positioned within a 20 m x 20 m pixel of the Sentinel-1 and Sentinel-2 image pixels. A total of 40 sampling plots was planned for the field survey, with 20 located in the wetland area and 20 located in the terrestrial area (Figure 1). For each sample plot, five replicate measurements of %VWC were recorded in order to capture the variation of the observed %SMC within the top layer of the soil surface. This yielded a total of 200 readings for the terrestrial and wetland areas during each sampling campaign.

Near-surface volumetric SMC was acquired using a handheld SMT-100 soil moisture and temperature probe.[45] The probe measures the %VWC at a depth of 5 cm. The centre and corners of each sample plot were mapped in ArcGIS version 10.5[46] and then uploaded to several e-Trex 30 Global Positioning System (GPS) devices.[47] The GPS devices were then used to navigate to the same location for successive sampling campaigns. Previous studies recommended that ground measurements should be made within a 2-h window period around the sensor overpass time so as to minimise diurnal variation in SMC and vegetation on radar backscatter.[48,49] Therefore, three probes were used by three teams to record the %VWC within a 2-h time period around the satellite overpass, which included the hour before and after the time of overpass of each Sentinel sensor.

To determine the impact of vegetation on the regression, the vegetation height in various zones was randomly measured and recorded during the field campaigns. In general, the density and height of the vegetation in both zones varied little for the duration of %VWC data collection between March and May of 2018. Regardless, sample plots were planned at least 2 m away from macrophytes and trees where the height of the vegetation was >2 m high. In general, the canopy height in sample plots of similar grassland palustrine wetland sites (Chrissiesmeer, Mpumalanga Province) was 1.5–2 m, with an estimated biomass ≤850 g/m^2.[50] The Normalised Difference Vegetation Index (NDVI)[51,52] is often used to compensate for the influence of vegetation in estimating SMC[24,49]. However, according to Hornacek et al.[32], vegetation and texture have very little impact on the %SMC modelling if grass vegetation is ≤1 kg/m^2. Consequently, no adjustments were made for vegetation in this paper, because the above-ground biomass and sedges in the study area are likely <850 g/m^2.

Data analysis

In order to assess the Sentinel sensors' capability to estimate the SMC, backscatter from S1A and S1B and reflectance values from S2A and S2B were extracted from the respective images and regressed against the average of the five in-situ %VWC measurements taken for each plot. The centre point recorded for each sample plot in shapefile format was used to extract backscatter values for VV, VH polarisation modes as well as VV+VH as a modelling scenario, in ArcMap 10.5.[46] Similarly, the spectral reflectance values of the optical sensors were extracted for all the bands, excluding bands 1, 9 and 10 (60-m resolution bands) for the same points.

The Sentinel-1 backscatter values and Sentinel-2 reflectance band values were regressed against the %VWC values (in-situ measurements) using both a parametric (the simple linear regression model or SLR) and two non-parametric (support vector machine – SVM and random forest – RF) algorithms in the Waikato Environment for Knowledge Analysis (Weka) software version 3.8.[53] These regression models are commonly used in the remote sensing of environmental variables and the capabilities of these models were compared in this study (see review by Gangat[26]). In principal, parametric models assume normal distribution of the data, and hence they depend on mean and standard deviation statistics. These models are less complex in terms of tuning and require a fixed number of input variables.[54] Non-parametric models do not assume normal distribution and, as spectral data are often not normally distributed, non-parametric methods have been found to outperform parametric methods in remote sensing classification and prediction. A data split was used with 30% data for the training data set and 70% for the validation data set to test the best model for regressing the observed %VWC to the estimated %SMC. Individual polarisations and bands as well as a combination of the polarisations and all bands were evaluated for each sensor in predicting %SMC. The best model to predict %SMC from the radar and optical images was selected where the highest coefficient of determination (R^2) and lowest RMSE was attained. In order to estimate the %SMC, backscatter from S1B and reflectance values from optical S2B were extracted from the respective images. A Shapiro–Wilk test was used to test the differences between the wetland and terrestrial areas, for both the in-situ %VWC and predicted %SMC, to assess whether thresholding would be possible for wetland mapping. A $p<0.05$ was used to identify significant differences.

Results

Ability of Sentinel-1 and Sentinel-2 to estimate soil moisture content

Of the various modelling scenarios, the Sentinel images were capable of predicting the %SMC with the majority of coefficients of determination (R^2)>0.7 and RMSEs<21% (Table 3). Of the four sensors which used RF, S1B produced a high R^2 of 0.92–0.94 and the lowest RMSE of 10%. S2B achieved the second-highest results with an R^2 of 0.92–0.94 and RMSE of 12–14%. S2A produced slightly better results (R =0.70–0.86; RMSE=13–20%) than S1A, which resulted in the lowest R^2 of 0.58–0.72 and highest error at RMSE=19–24%.

Table 3: Comparison of the different modelling approaches and validation models, using coefficient of determination (R^2) and root mean square error (RMSE) between the percentage volumetric water content (%VWC) and predicted percentage of soil moisture content (%SMC), across the four Sentinel sensors evaluated using simple linear regression (SLR), support vector machine (SVM) and random forest (RF) modelling algorithms

		SLR		SVM		RF	
		R^2	RMSE	R^2	RMSE	R^2	RMSE
S1A	VV	0.01	40	0.01	50	0.58	24
	VH	0.10	34	0.10	35	0.72	19
	VV+VH	0.01	40	0.03	48	0.69	23
S1B	VV	0.05	39	0.05	32	**0.92**	**10**
	VH	0.12	36	0.12	37	**0.94**	**10**
	VV+VH	0.16	34	0.16	36	**0.94**	**10**
S2A	2-Blue	0.09	34	0.30	37	0.82	13
	3-Green	0.25	32	0.50	33	0.86	18
	4-Red	0.28	31	0.53	32	0.86	18
	5-VRE	0.23	32	0.48	33	0.86	18
	6-VRE	0.1	37	0.14	41	0.72	19
	7-VRE	0.25	37	0.25	41	0.7	19
	8-NIR	0.7	35	0.7	38	0.74	19
	11-SWIR	0.11	35	0.11	39	0.72	20
	12-SWIR	0.45	27	0.45	27	0.74	19
	All bands	0.53	25	0.6	25	0.75	19
S2B	2-Blue	0.41	29	0.45	30	**0.94**	12
	3-Green	0.40	29	0.40	29	**0.94**	13
	4-Red	0.34	31	0.34	30	**0.94**	12
	5-VRE	0.30	31	0.30	31	**0.92**	13
	6-VRE	0.25	32	0.25	33	**0.92**	13
	7-VRE	0.18	34	0.18	34	**0.92**	14
	8-NIR	0.09	36	0.09	36	**0.94**	14
	11-SWIR	0.36	30	0.36	31	**0.92**	14
	12-SWIR	0.42	28	0.42	29	**0.94**	13
	All bands	0.36	30	0.45	30	**0.94**	12

S1A, Sentinel-1A; S1B, Sentinel-1B; S2A, Sentinel-2A; S2B, Sentinel-2B; V, vertical-receive vertical-transmit; VH, vertical-receive horizontal-transmit; VRE, vegetation red edge; SWIR, short-wave infrared

Of the three polarisation modes (VV, VH and VH+VV) associated with the two Sentinel-1 (SAR) sensors, the VH polarisation mode and the VH+VV modelling scenario yielded higher accuracies ($R^2>0.7$) and lower errors (RMSE<19%) than the VV polarisation (Table 3). S1B showed the highest coefficient of determination ($R^2>0.9$) when the VH polarisation and VH+VV modelling scenario were used, with an RMSE of 10% in both instances. The results for S1A were slightly lower at $R^2=0.72$, with a slightly higher RMSE of 19% for VH and 23% for VH+VV. The single VV polarisation showed the lowest coefficient of determination and highest error ($R^2=>0.58$; RMSE = >24%) for both of the Sentinel-1 sensors where the RF algorithm was used. The VH polarisation mode, however, contributed more to the accuracies of the combined VH+VV modelling scenario inputs than the single polarisation (VV) mode.

A combination of all the bands for the optical sensors S2A and S2B, in general, resulted in high accuracies ($R^2>0.7$ and RMSE<20%) when the RF algorithm was used, whereas the SLR and SVM showed an $R^2<0.6$ and RMSE were 30–50% (Table 3). Some exceptions are evident where the use of the blue, green, red and vegetation red edge (VRE) bands produced comparable results to the combined bands, most noticeably when the RF algorithm was used ($R^2=0.7$ for S2A, or even higher for S2B ($R^2>0.9$). Five of the S2B individual bands resulted in the highest coefficients of determination in predicting %SMC ($R^2=>0.94$; RMSE=13%), namely: blue (band 2:496–492 nm), green (band 3: 560–559 nm), red (band 4: 664–665 nm), NIR (band 8: 833–835 nm) and SWIR (band 12: 2185–2204 nm).

When comparing the regression model scenarios, the RF algorithm outperformed the SLR and SVM. RF achieved ranges of the coefficients of determination from $R^2=0.58$ to $R^2=0.94$ and RMSE values between 10% and 24% (Table 3). In contrast, the non-parametric SVM had lower accuracies ranging from $R^2=0.01$ to $R^2=0.7$ and RMSE values between 25% and 50%. The parametric SLR algorithm showed similar ranges of coefficients of determination to that of the non-parametric SVM algorithm (from R^2 = 0.01 to R^2 = 0.7) and RMSE values ranging from 25% to 40%. Consequently, the RF algorithm was applied to the S1B SAR, using only VH polarisation which contributed most to the backscatter values, and all the bands from the S2B image, to predict %SMC for the CWNR.

Comparison of observed %VWC and predicted %SMC in wetland and terrestrial areas

In-situ observed %VWC ranged from 16% to 100% in the wetland areas and from 1% to 37% for the terrestrial area (Table 4). In comparison, the S1B predicted %SMC ranged from 30% to 100% for wetlands and 11% to 39% for terrestrial areas, and the S2B predicted %SMC ranged from 4% to 78% in the wetlands and 4% to 57% in the terrestrial areas. The mean soil moisture values for in-situ observed %VWC in the wetlands and terrestrial areas were higher on 28 March 2018 (mean±standard deviation = 91%±21 and 20%±8, respectively) compared to those measured a month later on 2 May 2018 (mean±standard deviation = 74%±27 and 6%±3, respectively). Similar trends were visible in the predicted %SMC mean values for both the wetland and terrestrial areas. In contrast, the percentage of coefficient of variance (%COV) showed an increase over the month (from 28 March 2018 to 2 May 2018), for both the in-situ %VWC and %SMC predicted from both sensors, within the wetland and terrestrial areas. In general, the %COV was much lower in the terrestrial area for both dates, considering the %VWC and %SMC, compared to the wetland area (Figure 2).

Table 4: Descriptive statistics for in-situ observed percentage of volumetric water content (%VWC) and predicted percentage of soil moisture content (%SMC) at the time of Sentinel sensors overpass on the 28 March 2018 for Sentinel-1B (S1B) and on 2 May 2018 for Sentinel-2B (S2B)

		S1B: 28 March 2018		S2B: 2 May 2018	
		Observed %VWC	Predicted %SMC	Observed %VWC	Predicted %SMC
Wetland areas	Minimum	16.2	30.1	29.2	3.9
	Maximum	100.0	100.0	100.0	77.6
	Mean	90.7	80.7	74.8	35.7
	Median	100.0	92.7	88.8	39.0
	s.d.	20.8	21.4	27.3	23.3
	%COV	23.0	26.5	36.4	65.3
Terrestrial areas	Minimum	4.5	11.4	1.3	3.9
	Maximum	36.9	39.4	16.9	56.9
	Mean	20.3	23.4	5.9	16.8
	Median	20.6	21.8	5.5	13.7
	s.d.	8.0	7.4	3.0	13.2
	%COV	39.3	31.8	50.8	78.4
Terrestrial and wetland areas	Minimum	4.5	11.4	1.3	3.9
	Maximum	100.0	100.0	100.0	77.6
	Mean	57.4	52.2	42.7	29.4
	Median	37.9	35.5	36.3	21.5
	s.d.	38.7	32.9	39.9	22.4
	%COV	67.5	63.0	93.3	76.0

s.d., standard deviation; %COV, coefficient of variation

Soil moisture values, whether in-situ %VWC or predicted %SMC, were significantly different between the wetland and terrestrial areas for both

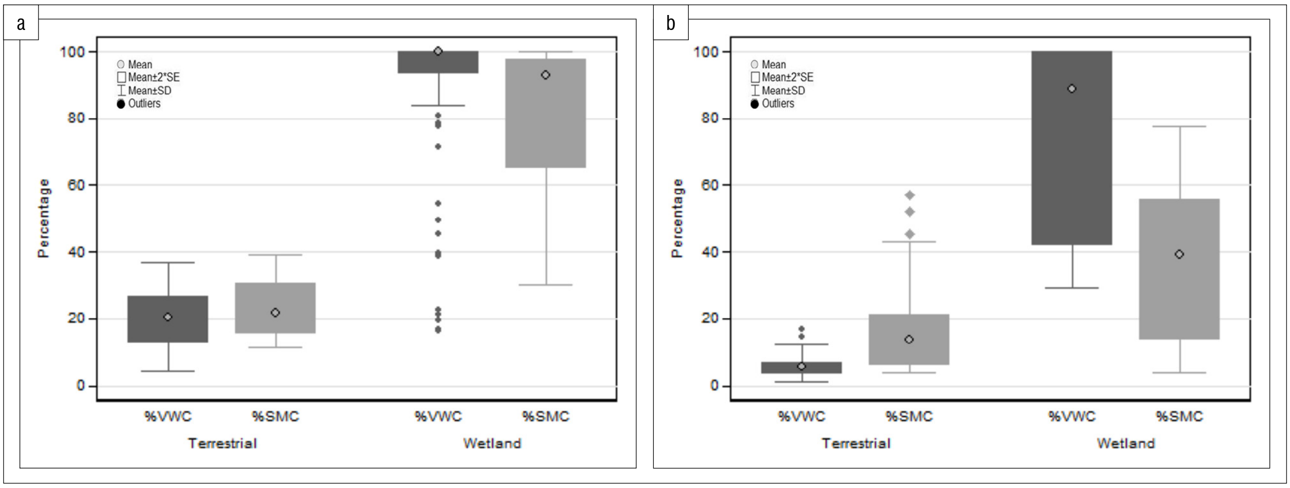


Figure 2: Percentage volumetric water content (%VWC) and predicted percentage of soil moisture content (%SMC) values between terrestrial and wetland areas for (a) Sentinel-1B on 28 March 2018 and (b) Sentinel-2B on 2 May 2018.

dates (Table 5). The overlap in the minimum %VWC and %SMC with the maximum %VWC and %SMC (Table 4) showed that there is still confusion in the soil moisture values between the wetland and terrestrial areas, which makes it difficult to identify a certain threshold for wetland mapping. An average threshold of 50%VWC and/or 50%SMC was therefore used to distinguish wetland extent from the terrestrial areas.

Table 5: Differences of significance (*p*-value resulting from Shapiro–Wilk's test) between the wetland and terrestrial areas for in-situ percentage volumetric water content (%VWC) and the predicted percentage soil moisture content (%SMC) resulting from the Sentinel-1B (S1B) and -2B (S2B) predictions

Date	In-situ (observed) measurements	Sensor	Predicted soil moisture content
28 March 2018	0.0000000000000023	**S1B**	0.00000000000000022
2 May 2018	0.0000000000000024	**S2B**	0.00000000000000025

Predicted soil moisture maps for the study area

The predicted %SMC maps from the S1B and S2B sensors show a variation in the extent of soil saturation (Figure 3a and 3b). Although differences are visible between the S1B and S2B predictions on the edges of the study area, both maps show a higher level of soil saturation in the centre of the channelled valley-bottom wetland, southeast of the Hartbeesspruit River's channel. These areas read nearly 100% of %VWC during the sampling for both dates.

Using the 50% threshold, nearly 47% (32.7 ha) of the extent of the study area could be wetland from the S1B prediction (Figure 3a), whereas approximately 23% (15.7 ha) would be predicted as wetland when using the S2B %SMC map (Figure 3b).

The standard error graphs illustrating the observed in-situ %VWC against the predicted %SMC, represent the level of overestimation and underestimation of the model (Figure 4a and 4b). The results displayed show a coefficient of determination for S1B of R^2=0.91 and for S2B of R^2=0.86. Both S1B and S2B showed that observed %VWC values of under 50% are expected to be underpredicted, whereas above this threshold, predicted %SMC is overestimated, in each case by 16% and 21%, respectively (Figure 4).

Discussion

This study showed that the freely available Sentinel SAR and optical sensor data have potential for estimating soil moisture in palustrine wetlands. These sensors were used to predict SMC for a palustrine wetland in the grassland biome of South Africa with a RF algorithm resulting in a high coefficient of determination (R^2=>0.7) and a low RMSE= <24% using various modelling scenarios. The results are comparable to those of other studies done within wetland and terrestrial areas in temperate climates of Poland, Germany and Italy.[23-26] Further work is required to test the capabilities of these Sentinel sensors across other palustrine wetlands in South Africa, to assess the potential for upscaling the sensors for wetland inventorying and monitoring.

Soil moisture ranges showed significant differences between the wetland and terrestrial areas. Although the wetland–terrestrial gradient and a threshold for wetland mapping have not been explored in other studies,

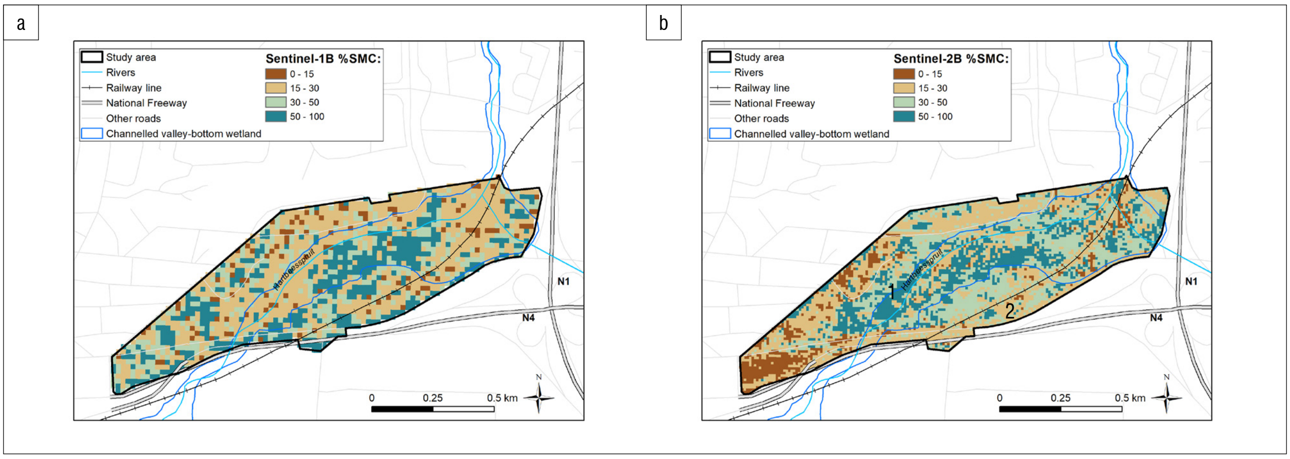


Figure 3: Predicted percentage soil moisture content (%SMC) map showing the variation in soil moisture derived from (a) Sentinel-1B for 28 March 2018 and (b) Sentinel-2B for 2 May 2018 sampling campaigns. The numbers 1 and 2 in Figure 3b refer to points of interest mentioned in the Discussion.

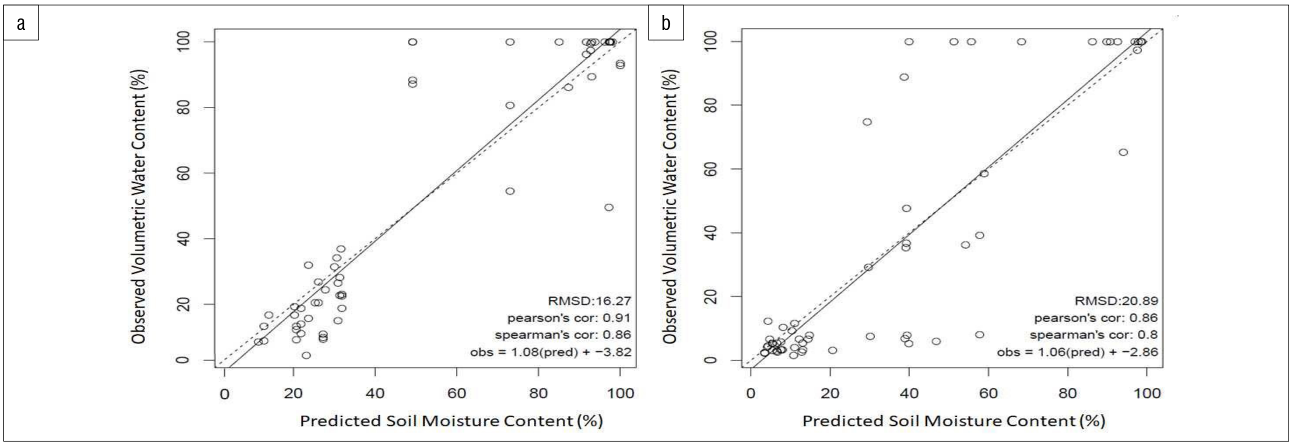


Figure 4: Standard error regression graphs displaying how well (a) Sentinel-1B and (b) Sentinel-2B captured the variability of the soil moisture content across the study area, using the random forest regression model. The solid line is the predictive model and the dotted line is the 1:1 line.

their measured SMC values were compared to this study to assess such a threshold. In Germany, the %VWC was 30–99.1%VWC with a mean of 50%VWC over a floodplain with a grass canopy of up to 2 m in height, measured at the end of the growth period.[26] In Italy, the mean %VWC in a dense terrestrial grassland area was 45%VWC at the end of the growth period.[31] A study conducted on the coastal plains of Washington District Capital (USA) attained an average of 59%VWC for their wetland areas and 24%VWC in the terrestrial areas.[33] Our findings show mean soil moisture levels for the in-situ measured data in the wetland of >75%VWC – significantly higher ($p<0.05$) compared to the mean %VWC measured in the terrestrial area (20% and 6% on the 28 March and 2 May 2018, respectively). An overlap in %VWC and %SMC values still makes it difficult to identify a threshold for mapped wetlands with high confidence. Consequently, we propose an interim 50% threshold for measured %VWC or predicted %SMC to distinguish the extent of the wetland in the CWNR for these two dates. Continuous assessment of the different soil moisture ranges over multiple time periods and hydrological regimes would be critical to confirm this proposed threshold and would provide insight into determining the maximum extent of wetlands.

The site showed a variation in the magnitude and areal extent of soil moisture between 28 March 2018 and 2 May 2018. The magnitude of the mean observed %VWC decreased from 28 March to 2 May by 6%. The predicted %SMC also showed a decline – by 45% – which may be explained by differences in prediction of %SMC by the radar and optical sensors. Regardless, it appears as if the sensors would have the potential to detect the variation in the degree of soil saturation. The predicted %SMC maps of 28 March 2018 (using S1B) and 2 May 2018 (using S2B) showed differences in the areal extent of soil saturation across the wetland, which could be attributed to a variation in rainfall and the subsequent infiltration and interflow of water into the soil. If a 50%SMC threshold is used to extract the areal extent of the wetland, the extent of the wetland covered 47% or 32.7 ha on 28 March and 23% or 15.7 ha on 2 May. Interestingly, the sampling campaign on 28 March 2018 took place shortly after an intense rainfall storm (22 March 2018) whereas the sampling campaign on 2 May 2018 took place approximately 2 weeks after a less intense rainfall period (Figure 5). This could possibly explain the higher mean values of 91%VWC and 23%COV recorded directly after the rain spell within the top layer of the soil, which subsequently may have infiltrated to deeper soils a month later, resulting in a lower mean of 75%VWC and 6%COV. Theoretically, the continuous rainfall events over summer (rainfall started approximately mid-February of 2018)[37] would lead to progressively accumulated water in the main part of the channelled valley-bottom wetland through surface run-off and groundwater accumulation. The accumulation of water in soil of the palustrine wetland would have resulted in an increase in the dielectric constant, resulting in higher backscatter and reflectance values. These results are evident from the comparison between the predicted SMC maps from S1B and S2B, which suggests that changes in soil saturation could potentially be used to detect soil saturation across a wetland's hydroperiod. We recommend that both sensors be used to test whether a more refined threshold for wetland mapping can be derived from time-series data.

In modelling SMC with the use of satellite imagery, our methods show that the RF non-parametric algorithm outperformed the SLR and SVM algorithms. This finding is in line with other studies in which non-parametric algorithms outperformed parametric algorithms.[54,55] Our study showed a minor difference in model performance for the radar data when using either VV, VH or VV+VH polarisation ($R^2>0.92$ and RMSE=10%). This differs from the findings of Dabrowska-Zielinska et al.[24], who indicated that the SAR VH outperforms VV and dual polarisation. For the optical data, the use of all bands achieved the highest R^2 (0.94) and the lowest RMSE (12%), compared to the use of individual bands. The use of bands across the visible, NIR and SWIR all contributed to the optimisation of the prediction of %SMC in the RF algorithm.

When comparing the parametric and non-parametric predictive models (SLR, RF and SVM), the comparative evaluation results show that RF outperformed the other parametric SLR and non-parametric SVM. RF resulted in the highest coefficients of determination (R^2 0.7–0.94) when either VV+VH or all optical bands were used. In contrast, the SLR and SVM show a poorer relationship between %VWC and %SMC when all bands were used (R^2 ranged from 0.4–0.6), while for the radar, SLR showed no relation ($R^2<0.2$). In general, non-parametric models are more suitable for reflectance and backscatter data, because these are often not normally distributed.[50,55] In addition, the non-parametric models are able to predict values using data sets with fewer observations over large regions. Although RF and SVM often show comparative results in remote sensing prediction of continuous variables, SVM requires more customisation to obtain the optimum model performance, whereas RF is easier to use and does not require reiterative testing.

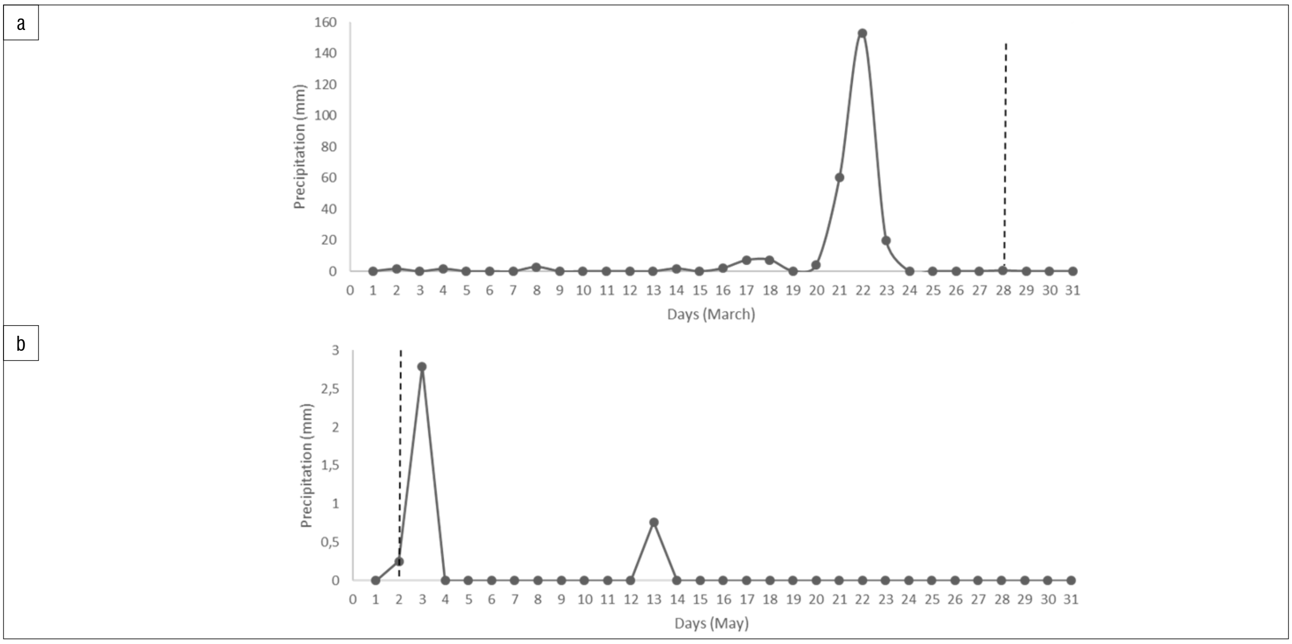


Figure 5: Total amount of daily precipitation (mm) for the month[37] in which the sampling was done. Graph (a) shows a rainfall peak shortly before the acquisition for Sentinel-1B on the 28 March 2018 (indicated by the dotted vertical line) and graph (b) shows the Sentinel-2B acquisition on 2 May 2018 (indicated by the dotted vertical line). Precipitation scale ranges on the *y*-axes are different to account for differences in the maximum precipitation of the two sample dates.

Despite the concern of the influence of vegetation on backscatter or reflectance data, our results show that it was likely not a significant influence on the prediction of %SMC in the case of CWNR. Above-ground biomass values for palustrine wetlands in the grassland biome of South Africa are estimated to be <850 g/m$^{2(50)}$ and <1 kg/m$^{2(32)}$, and are considered to have very little influence on the estimation of SMC. The inclusion of NDVI in the S2B regression showed a minor increase in model performance ($R^2=0.9$ and RMSE of 0.2; results not shown). Further work is required to assess whether vegetation and texture indices improve the estimation of SMC in other sites and climatic regions.

Regional monitoring of soil moisture would contribute greatly to South Africa's National Wetland Monitoring Programme, particularly because it is a common variable across palustrine wetland type, whereas vegetation communities would vary. To date, the SAR sensors available for the estimation of SMC are limited to C-band sensors, which can penetrate only to a depth of 5 cm. L-band sensors, on the other hand, have the advantage of deeper penetration through canopy cover or into bare soils. The Advanced Land Observing Satellite (ALOS)-2 L-band SAR sensor of the Japan Aerospace Exploration Agency (JAXA) has a high spatial resolution (10 m) and a temporal resolution of 46 days; however, these data were not freely available for use in monitoring at the time of this study, but the Japanese ministry announced in November 2019 that the ALOS archive would be made openly accessible. Several L-band sensors are planned to be launched from 2020 onwards, potentially for public use, including ALOS-3 from JAXA, NASA ISRO Synthetic Aperture Radar (NISAR) from NASA/ISRO and TanDEM-L from the German Space Agency.[56]

Conclusion

This study proves that the freely available Sentinel-1 (SAR) and Sentinel-2 (optical) sensors have potential in the estimation of the extent and degree of soil saturation in palustrine wetlands in the grassland biome of South Africa. The Sentinel-1 SAR and Sentinel-2 optical sensors were able to predict %SMC with a high coefficient of determination ($R^2>0.9$) and low RMSE ($<21\%$) along a wetland–terrestrial gradient in the grassland biome of the Gauteng Province of South Africa. The non-parametric RF algorithm outperformed the parametric SLR and non-parametric SVM algorithms in predicting %SMC for the CWNR during the peak of the hydroperiod in March and May 2018. Significant differences between the surface soil moisture of the palustrine wetland and the surrounding terrestrial areas imply that inventorying and monitoring of the extent and hydroperiod of palustrine wetlands can potentially be done. An SMC threshold of $\geq 50\%$ was used as a potential threshold to determine the extent of the wetland area; however, owing to uncertainties resulting from the overlap in the measured %VWC, further work would be required to confirm whether this threshold is relevant across the hydroperiod and other grassland sites. The predictions for the two months (March and May) showed a potential accumulation of soil saturation over the period, which may have resulted from interflow and groundwater accumulation. Although other studies suggest that vegetation has an influence on the prediction of soil moisture over an area, the incorporation of the NDVI in predicting %SMC from the optical Sentinel-2B image, showed only a minor improvement in the prediction. The prediction of SMC in the grassland biome of South Africa can play a significant role in improving the representation of the natural variation in soil saturation values of palustrine wetlands and enables the detection of outlier seasons or years associated with the impacts of global and changing climate.

Acknowledgements

This work was funded by the South African Water Research Commission (WRC) under the project K5/2545 'Establishing remote sensing toolkits for monitoring freshwater ecosystems under global change' as well as the Council for Scientific and Industrial Research (CSIR) through the project titled 'Common Multi-Domain Development Platform (CMDP) to Realise National Value of the Sentinel sensors for various land, freshwater and marine societal benefit areas'. Several fieldwork assistants have kindly contributed time, including Razeen Gangat, Mohamed Gangat, Taskeen Gangat, Zahra Varachia, Salman Mia, Yonwaba Atyosi and Jason le Roux. A special thanks to Tamsyn Sherwill from the WRC for including us in some of the Colbyn Wetland Nature Reserve projects and conservation initiatives.

Authors' contributions

R.G.: Student who undertook the literature search, planned the fieldwork, coordinated the fieldwork, did the analysis, and drafted the research output. H.v.D.: Primary supervisor who provided guidance to the student on the formulation of the research question, hypothesis, and execution of the fieldwork, participated in the fieldwork, and revised and edited the research output. L.N.: Co-supervisor who provided guidance to the student on the execution of the fieldwork, participated in the fieldwork, analysed the radar data and provided input and edits to the draft and revised research output. E.A.: Co-supervisor who provided a review of the research output.

References

1. Dudgeon D, Arthington AH, Gessner MO, Kawabata ZI, Knowler DJ, Lévêque C, et al. Freshwater biodiversity: Importance, threats, status and conservation challenges. Biol Rev. 2006;81(2):163–182. https://doi.org/10.1017/S1464793105006950
2. Díaz S, Settele J, Brondízio E, Ngo HT, Guèze M, Agard J, et al. Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services [document on the Internet]. c2019 [cited 2020 Mar 13]. Available from: https://www.ipbes.net/sites/default/files/downloads/spm_unedited_advance_for_posting_htn.pdf
3. Van Deventer H, Smith-Adao L, Mbona N, Petersen C, Skowno A, Collins NB, et al. South African Inventory of Inland Aquatic Ecosystems (SAIIAE). Report number CSIR/NRE/ECOS/IR/2018/0001/A. Pretoria: South African National Biodiversity Institute; 2018. http://hdl.handle.net/20.500.12143/5847
4. Van Deventer H, Van Niekerk L, Adams J, Dinala MK, Gangat R, Lamberth SJ, et al. National Wetland Map 5 – An improved spatial extent and representation of inland aquatic and estuarine ecosystems in South Africa. Water SA. 2020;46(1):66–79. https://doi.org/10.17159/wsa/2020.v46.i1.7887
5. Begg GW. The wetlands of Natal (Part 2): The distribution, extent and status of wetlands in the Mfolozi catchment. Report 71. Pietermaritzburg: Natal Town and Regional Planning Commission; 1988.
6. Fluet-Chouinard E, Lehner B, Rebelo LM, Papa F, Hamilton SK. Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data. Remote Sens Environ. 2015;158:348–361. http://dx.doi.org/10.1016/j.rse.2014.10.015
7. Pekel J-F, Cottam A, Gorelick N, Belward AS. High-resolution mapping of global surface water and its long-term changes. Nature. 2016;540:418–422. https://doi.org/10.1038/nature20584
8. GeoTerraImage (GTI). Technical report 2013/2014: South African national land cover dataset version 5. Pretoria: GTI; 2015.
9. Thompson M, Hiestermann J, Moyo L, Mpe T. Cloud-based monitoring of SA's water resources. Position IT Magazine. 2018;38–41.
10. Erwin KL. Wetlands and global climate change: The role of wetland restoration in a changing world. Wetl Ecol Manag. 2008;17:71–84. https://doi.org/10.1007/s11273-008-9119-1
11. Burton TM, Tiner RW. Ecology of wetlands. In: Gene E, Likens GE, editors. Encyclopedia of inland waters. Oxford: Academic Press; 2009. p. 507–515. https://doi.org/10.1016/B978-012370626-3.00056-9
12. Ollis D, Snaddon K, Job N, Mbona N. Classification system for wetlands and other aquatic ecosystems in South Africa: User manual: Inland systems. SANBI Biodiversity Series 22. Pretoria: South African National Biodiversity Institute; 2013.
13. Engman ET. Applications of microwave remote-sensing of soil-moisture for water-resources and agriculture. Remote Sens Environ. 1991;35:213–226. https://doi.org/10.1016/0034-4257(91)90013-V
14. Wood EF, Lettenmaier DP, Zartarian VG. A landsurface hydrology parameterization with subgrid variability for general circulation models. J Geophys Res. 1992;97:2717–2728. https://doi.org/10.1029/91JD01786
15. Santi E, Paloscia S, Pettinato S, Notarnicola C, Pasolli L, Pistocchi A. Comparison between SAR soil moisture estimates and hydrological model simulations over the Scrivia Test Site. Remote Sens. 2013;5(10):4961–4976. https://doi.org/10.3390/rs5104961
16. McDonnell JJ, Sivapalan M, Vaché K, Dunn S, Grant G, Haggerty R, et al. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. Water Resour Res. 2007;43, W07301. https://doi.org/10.1029/2006WR005467

17. Crow WT, Yilmaz MT. The Auto-Tuned Land Assimilation System (ATLAS). Water Resour Res. 2014;50:371–385. https://doi.org/10.1002/2013WR014550

18. Riley W, Shen C. Characterizing coarse-resolution watershed soil moisture heterogeneity using fine-scale simulations. Hydrol Earth Syst Sci. 2014;18:2463–2483. https://doi.org/10.5194/hess-18-2463-2014

19. Tebbs E, Gerard F, Petrie A, De Witte E. Emerging and potential future applications of satellite-based soil moisture products. In: Petropoulos GP, Srivastava P, Kerr Y, editors. Satellite soil moisture retrieval: Techniques and applications. Amsterdam: Elsevier; 2016; p. 379–400. https://doi.org/10.1016/B978-0-12-803388-3.00019-X

20. Schulze RE, Lynch SD. Annual precipitation. In: Schulze RE, editor. South African atlas of climatology and agrohydrology. WRC Report no. 1489/1/06, Section 6.2. Pretoria: Water Research Commission; 2007.

21. Wang L, Qu J. Satellite remote sensing applications for surface soil moisture monitoring: A review. Fron Earth Sci. 2009;3(2):237–247. https://doi.org/10.1007/s11707-009-0023-7

22. Filion R, Bernier M, Paniconi C, Chokmani K, Melis M, Soddu A, et al. Remote sensing for mapping soil moisture and drainage potential in semi-arid regions: Applications to the Campidano plain of Sardinia, Italy. Sci Total Environ. 2015;543:862–876. https://doi.org/10.1016/j.scitotenv.2015.07.068

23. Sadeghi M, Babaeian E, Tuller M, Jones S. The optical trapezoid model: A novel approach to remote sensing of soil moisture applied to Sentinel-2 and Landsat-8 observations. Remote Sens Environ. 2017;198:52–68. https://doi.org/10.1016/j.rse.2017.05.041

24. Dabrowska-Zielinska K, Musial J, Malinska A, Budzynska M, Gurdak R, Kiryla W, et al. Soil moisture in the Biebrza Wetlands retrieved from Sentinel-1 imagery. Remote Sens. 2018;10:1979. https://doi.org/10.3390/rs10121979

25. Holtgrave A, Forster M, Greifeneder F, Notarnicola C, Kleinschmit B. Estimation of soil moisture in vegetation-covered floodplains with Sentinel-1 SAR using support vector regression. Photogramm Eng Remote Sensing. 2018;86:85–101. https://doi.org/10.1007/s41064-018-0045-4

26. Gangat R. Assessing whether soil moisture content (SMC) can be estimated for wetlands in the grassland biome of South Africa using freely available space-borne sensors [MSc thesis]. Johannesburg: University of the Witwatersrand; 2019.

27. Klinke R, Kuechly H, Frick A, Förster M, Schmidt T, Holtgrave A, et al. Indicator-based soil moisture monitoring of wetlands by utilizing Sentinel and Landsat remote sensing data. Photogramm Eng Remote Sensing. 2018;86(2):71–84. https://doi.org/10.1007/s41064-018-0044-5

28. Lu N, Chen S, Wilske B, Sun G, Chen J. Evapotranspiration and soil water relationships in a range of disturbed and understood ecosystems in the semi-arid Inner Mongolia, China. J Plant Ecol. 2001;4(1–2):49–60. https://doi.org/10.1093/jpe/rtq035

29. Saalovaara K, Thessler S, Malik RN, Tuomisto H. Classification of Amazonian primary rain forest vegetation using Landsat ETM+ satellite imagery. Remote Sens Environ. 2005;97:39–51. https://doi.org/10.1016/j.rse.2005.04.013

30. Said S, Kothyari UC, Arora MK. Vegetation effects on soil moisture estimation from ERS-2 SAR images. Hydrol Sci J. 2012;57(3):517–534. https://doi.org/10.1080/02626667.2012.665608

31. Paloscia S, Pettinato S, Santi E, Notarnicola C, Pasolli L, Reppucci A. Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation. Remote Sens Environ. 2012;34:234–248. https://doi.org/10.1016/j.rse.2013.02.027

32. Hornacek M, Wagner W, Sabel D, Truong HL, Snoeij P, Hahmann T, et al. Potential for high resolution systematic global surfaces soil moisture retrieval via change detection using Sentinel-1. IEEE J Sel Top Appl Earth Obs Remote Sens. 2012;5(4):359–366. https://doi.org/10.1109/JSTARS.2012.2190136

33. Lang MW, Kasischke ES, Prince SD, Pittman KW. Assessment of C-band synthetic aperture radar data for mapping and monitoring coastal plain forested wetlands in the Mid-Atlantic Region, USA. Remote Sens Environ. 2007;122:4120–4130. https://doi.org/10.1016/j.rse.2007.08.026

34. O'Connor TG, Bredenkamp GJ. Grassland. In: Cowling RM, Richardson DM, Pierce SM, editors. Vegetation of South Africa. Cambridge, UK: Cambridge University Press; 1997.

35. South African National Biodiversity Institute (SANBI). Grasslands ecosystem guidelines: Landscape interpretation for planners and managers. Pretoria: SANBI; 2013.

36. Wilkinson M, Danga L, Mulders J, Mitchell S, Malia D. The design of a National Wetland Monitoring Programme. Consolidated technical report. Volume 1. WRC Report no. 2269/1/16. Pretoria: Water Research Commission; 2016.

37. Agricultural Research Council – Soil, Climate and Water (ARC-ISCW). National AgroMet Climate Databank. Pretoria: ARC-ISCW; 2018.

38. Grundling PL. Genesis and hydrological function of an African mire: Understanding the role of peatlands in providing ecosystem services in semi-arid climates [PhD thesis]. Ontario: University of Waterloo; 2015. http://hdl.handle.net/10012/9037

39. Delport L. A Holocene wetland: Hydrology response to wetland rehabilitation in Colbyn Valley, Gauteng, South Africa [MSc thesis]. Johannesburg: University of Johannesburg; 2016. http://hdl.handle.net/10210/235764

40. Venter I, Grobler LER, Delport L, Grundeling P. Colbyn Wetland monitoring and evaluation report – TIER 3B assessment. Report prepared for the Working for Wetlands Programme. Pretoria: South African Department of Environmental Affairs; 2016.

41. Sherwill T. Colbyn Valley: The ultimate urban wetland survivor. The Water Wheel. 2015;22–26.

42. South African Department of Environmental Affairs (DEA). Draft biodiversity management plan for the Colbyn Valley wetland. Pretoria: DEA; 2015.

43. United States Geological Survey (USGS). Shuttle Radar Topography Mission, 1 Arc Second scene SRTM_u03_n008e004, unfilled unfinished 2.0. College Park, MD: University of Maryland; 2004.

44. United States Geological Survey (USGS). Earth explorer [webpage on the Internet]. c2018 [cited 2020 Mar 13]. Available from: http://earthexplorer.usgs.gov

45. Sichuan Weinasa Technology Co Ltd. Takeme-10EC meter conductivity soil moisture and temperature sensor [webpage on the Internet]. c2018 [cited 2020 Mar 13]. Available from: https://weinasa.en.alibaba.com/product/60787926381-807021906/Takeme_10EC_meter_conductivity_soil_moisture_and_temperature_sensor.html

46. Environmental Systems Research Institute (ESRI). ArcGIS desktop 10.4. Redlands, CA: ESRI; 1999–2017.

47. GARMIN. Our most popular handheld GPS with 3-axis compass. c2011 [cited 2020 Mar 13]. Available from: https://buy.garmin.com/en-ZA/ZA/p/87774

48. Muller E, Décamps H. Modelling soil moisture – reflectance. Remote Sens Environ. 2001;76(2):173–180. https://doi.org/10.1016/S0034-4257(00)00198-X

49. Baghdadi N, El Hajj M, Zribi M, Fayad I. Coupling SAR C-band and optical data for soil moisture and leaf area index retrieval over irrigated grasslands. IEEE J Sel Top Appl Earth Obs Remote Sens. 2015;99:1–15. https://doi.org/10.1109/IGARSS.2016.7729919

50. Naidoo L, Van Deventer H, Ramoelo A, Mathieu R, Nondlazi B, Gangat R. Estimating above ground biomass as an indicator of carbon storage in vegetated wetlands of the grassland biome of South Africa. Int J Appl Earth Obs Geoinf. 2019;78:118–129. https://doi.org/10.1016/j.jag.2019.01.021

51. Rouse J, Hass R, Schell J, Deering D. Monitoring vegetation systems in the Great Plains with ERTS. Third Earth Resources Technology Satellite Symposium; 1973 December 10–14; Greenbelt, MD, USA. Washington DC: NASA; 1974. SP-3511: 309–317.

52. Tucker CJ. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ. 1979;8(2):127–150.

53. Eibe F, Hall MA, Witten IA. The WEKA Workbench. Online appendix for 'Data Mining: Practical Machine Learning Tools and Techniques'. 4th ed. Burlington, MA: Morgan Kaufmann; 2016.

54. Naidoo L, Mathieu R, Main R, Kleynhans W, Wessels K, Asner G, et al. Savannah woody structure modelling and mapping using multi-frequency (X-, C- and L-band) synthetic aperture radar data. ISPRS J Photogramm Remote Sens. 2015;105:234–250. https://doi.org/10.1016/j.isprsjprs.2015.04.007

55. Ali I, Greifeneder F, Stamenkovic J, Neumann M, Notarnicola C. Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. Remote Sens. 2015;7:16398–16421. https://doi.org/10.3390/rs71215841

56. The CEOS Database – Catalogue of satellite missions [webpage on the Internet]. c2019 [cited 2020 Mar 13]. Available from: http://database.eohandbook.com/database/missiontable.aspx

Characterisation of wet and dry summer seasons and their spatial modes of variability over Zimbabwe

AUTHORS:
Moven Manjowe[1]
Terence Mushore[1,2]
Electidom Matandirotya[1]
Emmanuel Mashonjowa[1]

AFFILIATIONS:
[1]Department of Physics, University of Zimbabwe, Harare, Zimbabwe
[2]Department of Geography, University of KwaZulu-Natal, Pietermaritzburg, South Africa

CORRESPONDENCE TO:
Moven Manjowe

EMAIL:
manjowem@gmail.com

DATES:
Received: 28 June 2019
Revised: 13 Jan. 2020
Accepted: 25 Feb. 2020
Published: 29 July 2020

HOW TO CITE:
Manjowe M, Mushore T, Matandirotya E, Mashonjowa E. Characterisation of wet and dry summer seasons and their spatial modes of variability over Zimbabwe. S Afr J Sci. 2020;116(7/8), Art. #6517, 7 pages. https://doi.org/10.17159/sajs.2020/6517

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Yali Woyessa

KEYWORDS:
long-term rainfall, Inter-Tropical Convergence Zone, ITCZ, empirical orthogonal function analysis

FUNDING:
None

In the face of a changing climate, knowledge on the characteristics of wet and dry summers and their modes of variability becomes of great importance to Zimbabwe so that proper policies and planning can be implemented to maximise the positive impacts of climate change while minimising its negative impacts. We used time series of rainfall anomaly index, empirical orthogonal function analysis and composite analysis to determine the temporal and spatial characteristics of summer rainfall in Zimbabwe between 1980 and 2013. Results indicate that there is a possible shortening in the length of the summer season, running from November to March. There were 14 anomalous summer seasons (7 dry and 7 wet) during the 33 summer seasons in the study period. Three dominant modes of variability were identified for Zimbabwean summer rainfall for the period under study: (1) an east–west gradient accounting for about 63% of the total variability, (2) a northeast–southwest oscillation accounting for about 18% of the variability and (3) a northwest–southeast oscillation accounting for about 6% of the variability. From the results, the high frequency of occurrence of above or below normal summer seasons is a major concern due to their far-reaching effects on Zimbabwe's economic and social well-being. Circulation mechanisms influencing such characteristics should be studied and possible predictors identified for the development of short-term climate prediction models. If the forecasting of extreme events is reliable and done with sufficient lead time, the information can be utilised to mitigate the adverse effects while maximising their positive impacts. The findings of this study are important for informing economic activities in sectors such as farming, energy, mining, and tourism, which rely heavily on summer rainfall. However further studies must be carried out to identify the atmospheric and oceanic circulations which lead to such rainfall variability. Such information is critical in the development of reliable rainfall forecasts and early warning systems.

Significance:

Rainfall affects a variety of socio-economic activities, especially in Zimbabwe where rain-fed agriculture is a major contributor to the gross domestic product. Therefore, understanding spatial and temporal variations in rainfall enhances the formulation of strategies and decision-making to ensure sustainable development in the country.

Introduction

Southern African countries including Zimbabwe have been plagued by recurrent droughts in recent years. The devastating drought of 1982–1984 caused the crop yields in southern Africa to decline to 10% of the average historical levels with many sources of water drying up.[1] At least 90% of the livestock in the region was wiped out in the 1982/1983 drought.[2] In the 1991/1992 summer season, rainfall was 50% of normal levels over southeastern Africa, surpassing the 1982/1983 drought in severity.[3] According to Zhakata[4], the 1982/1983 season was one of the worst on record for southern Africa since the 1901/1902 season. Zhakata[4] further reported that, in the 1991/1992 rainy season, 90% of inland dams dried up, crops withered, and livestock perished in the thousands in Zimbabwe. A recent study by Manatsa et al.[5] showed that there is a strong correlation between maize yield and annual rainfall in rainfed agricultural systems, indicating a strong dependency of maize production and other agricultural activities on rainfall patterns. Thus, characterisation of wet and dry summers over Zimbabwe and their modes of variability is of great importance to the region.

Previous studies on rainfall variability used low spatial resolution data from a sparse rainfall network.[2,6,7] These studies mainly focused on rain-bearing features like the Inter-Tropical Convergence Zone (ITCZ). Hence many studies have not addressed the variability which results from other rain-bearing features like westerly cloud bands and cyclones in the Indian Ocean. The Zimbabwean Department of Meteorological Services precipitation data set that is used, amongst others, by many researchers[2,8] fails to meet the World Meteorological Organization (WMO) standards for rainfall station density. According to WMO regulations for areas with a variable topography like Zimbabwe, for adequate rainfall analysis, station density must be one rainfall station in every 240 km^2 of land.[9] However, high-resolution rainfall data from the Global Precipitation Climatology Center (GPCC) at 0.5° latitude by 0.5° longitude grid squares are freely available from NOAA/OAR/ESRL PSD, Boulder, Colorado, USA[10] and have the potential to be used to adequately map rainfall variability characteristics for Zimbabwe[9]. The use of higher-resolution rainfall data sets allows for a wide range of required spatial and temporal details[10] which have not yet been fully utilised in Zimbabwe. This is particularly important for Zimbabwe where the topography suppresses or enhances some rainfall triggering systems. Research outputs from using high-resolution data sets would be beneficial, for example in informing localised adaptation strategies, especially in view of potential impacts of the current and anticipated climatic changes on rainfed agriculture, hydropower generation and tourism.

Many studies carried out on variability characteristics of summer rainfall over Zimbabwe have focused on external forcing mechanisms and covered large areas like the whole of southern Africa, thus generalising to large homogeneous regions.[11,12] Mamombe et al.[8] concentrated on the effects of the El Niño Southern Oscillation and observed that the spatial variability varies in east–west and north–south gradients. However, such results fail to explain observed historical rainfall events which show a northeast–southwest oscillation and a northwest–southeast oscillation. Ogwang et al.[12] and Kabanda and Jury[13] concentrated on the roles played by the South Atlantic and Indian Oceans in the variability of summer rainfall over Zimbabwe and the rest of the southern African region. This had the effect of smoothing out the small localised variabilities due to the use of low-resolution data sets such as the GPCC precipitation data set with a resolution of 1.0º latitude by 1.0º longitude by Nicholas and Mwafulirwa[14] and Ogwang et al.[11] and also data sets built from a low density network of rainfall stations[6]. Similarly, Mamombe et al.[8] used data from coarse station networks interpolated between stations to produce a high-resolution data set. Such a data set fails to capture the influence of topography which plays a major role in determining the characteristics of precipitation patterns for regions with variable topography like Zimbabwe.[8,14] Due to the use of low-resolution rainfall data sets, these studies failed to fully explain the observed historical variations in Zimbabwean summer rainfall. With the availability of high-resolution data sets from GPCC and other advanced mathematical tools, it has become important to update existing knowledge on the temporal and spatial variability of Zimbabwean summer rainfall. The results can then be compared with those obtained by other authors in the past. The knowledge gained is of great value to the farming community, water resource managers, the energy sector and government for the proper planning and effective management of water resources.

The objectives of this study were to enhance the understanding of the spatial pattern of inter-annual rainfall variability, quantify the number of extreme (dry/wet) summer events which occurred between 1980 and 2013 and establish spatial modes of variability of summer rainfall in Zimbabwe using a high-resolution GPCC data set at 0.5° latitude by 0.5° longitude grid squares. The main hypothesis was that inter-annual variability of summer rainfall over southern Africa, including Zimbabwe, results from the fluctuations in the behaviour of rain-bearing systems during the summer period. These major systems hypothesised to influence rainfall variability are the ITCZ, Angola low pressure system, westerly waves, easterly waves and cyclones or lows in the Mozambican channel.

Data sources and methodology

Description of the study area

Zimbabwe is a landlocked country in southern Africa (Figure 1). Its total area is 309 580 km^2, of which a small portion (3910 km^2) is covered by water. Despite its small size, the terrain ranges from high altitudes in the eastern and central parts of the country to very low altitudes in the extreme northern, western and southern areas. The highest point (Mount Nyanga) is 2592 m above sea level while the lowest (Beit Bridge) is about 400 m above sea level. The eastern border is only 200 km from the Indian Ocean, so the country is influenced to a large extent by the atmospheric systems in the Mozambican channel. In the low-altitude areas (southwestern regions), the amount of rainfall is very low, resulting in semi-arid conditions, while the high-altitude areas (northeastern regions) enjoy significantly high amounts of precipitation.

The high-altitude areas along the eastern margin of the country also coincide with areas which receive the highest rainfall annually (Figure 1). The eastern highlands also receive rainfall throughout the year due to proximity to the Indian Ocean, which brings in moist air and orographic uplift which enhances activity. In the rest of the country, rainfall is confined to the summer period. Zimbabwe's summer season runs from end October through to late April, with the highest precipitation occurring late December to early February.[12] The start and cessation dates of the summer season are quite variable and thus inter-annual variability is high. At times, the season starts in early October and stretches into late April. Long-term mean seasonal rainfall shows north–south and west–east gradients. Such a phenomenon is in response to the country's topography and airflow that approaches Zimbabwe from the northwest and northeast, promoting moist warm air advection.[15,16] Cumulative seasonal rainfall exhibits a high degree of inter-annual variability. According to Buckle[17], the variability is higher to the south and lower to the north.

Rainfall variability responds to the weakening or strengthening of synoptic scale features, namely: the South Atlantic Ocean high pressure system, the South Indian Ocean high pressure system[18], the Angola low pressure system[19], the ITCZ[20] and finally the number and strength of tropical cyclones passing in the Mozambican channel. These features control the summer mid-tropospheric westerly or easterly wave behaviour, times and areas of surface convergence or divergence, and moist or dry air advection into different parts of the country, hence affecting rainfall totals and intensities at different times of the season as well as in different areas of the country.

Rainfall data collection and processing

The gridded monthly rainfall data were obtained from 0.5 x 0.5 gridded data from the GPCC monthly precipitation data set for the period 1975–2013. The GPCC precipitation data are provided freely by the NOAA/OAR/ESRL PSD (http://www.esrl.noaa.gov/psd).[10] The high-resolution data set which merges station and model data sets is suitable for our study area where topography plays a crucial role in modulating precipitation events; due to their low density, rainfall stations alone do not provide sufficient data to depict rainfall variability.

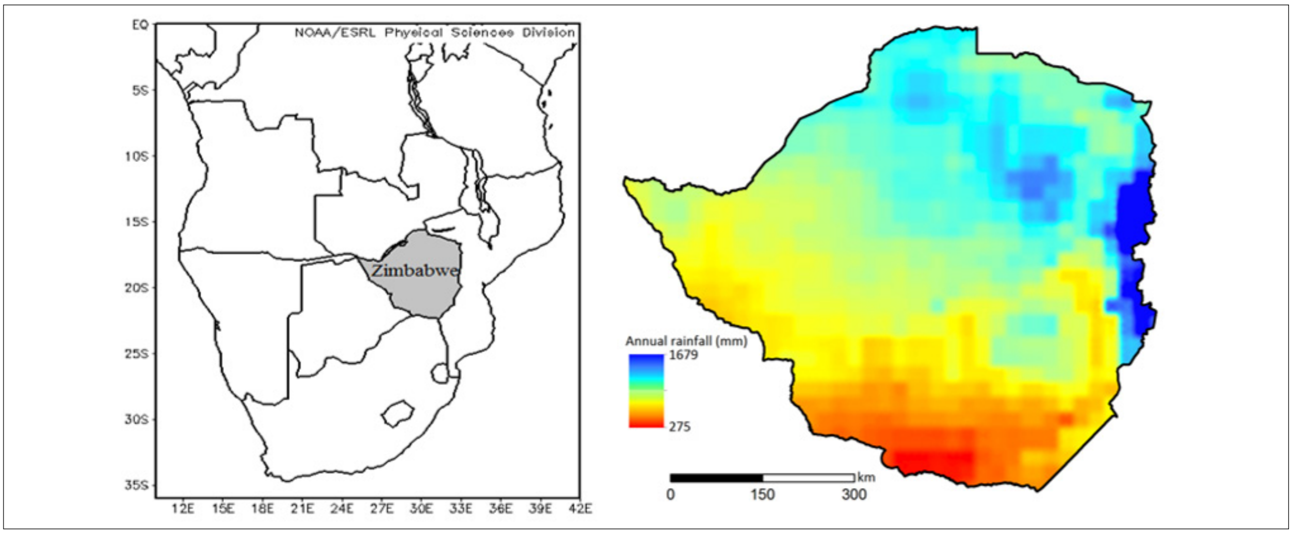


Figure 1: Map of southern Africa showing (a) the geographical position of Zimbabwe and (b) a map of the annual rainfall distribution in Zimbabwe.

Monthly spatial rainfall analysis using the standardised precipitation index

Monthly rainfall data from GPCC were formulated into their standardised departures using the historical means and standard deviations for each month. Hence, the seasonal cycle (based on averages calculated for the period 1980 to 2013) was removed, such that the standardised data had a mean of zero and a variance of one. In this way, diverse grid points were combined into indices and compared. The time series of monthly and seasonal precipitation was evaluated. The dry and wet years were selected by filtering the years with a standardised anomaly greater than 1 (wet years), those with a standardised anomaly less than -1 (dry years) and those with a standardised anomaly between -1 and 1 (normal years). This selection criterion was successfully used for identifying wet and dry years for Malawi.[14] The grid point's precipitation indices were averaged to create one precipitation index for Zimbabwe. This area index was used to identify the dry and wet years.

The normalisation method is used to produce averages of observations from areas with different characteristics, especially in the tropics where rainfall is convective. Normalisation minimises differences in received rainfall owing to terrain and elevation. An index created in this manner has unit variance and zero mean. The creation of the Standardised Rainfall Index for Zimbabwe was based on grid points at a 1.0° latitude × 1.0° longitude resolution.

Data for each grid point for each month was normalised with respect to the individual means and standard deviation according to Equation 1:

$$Z_i = \frac{(X_i - U_i)}{\sigma}, \qquad \text{Equation 1}$$

where Z_i represents the individual standardised departure of values (monthly data over k years); X_i (i=1, 2..., k) is the observed individual monthly value, U_i is the long-term monthly mean (or seasonal etc.) rainfall, and σ is the long-term standard deviation. In this way, three monthly totals are obtained from monthly values for each grid point starting for December to February (summer) from 1980 to 2013.

The ZRI index was used to select dry years using the following criterion: for all dry years, $Z_i \leq -1$; for all wet years, $Z_i \geq +1$; and for normal years, $-1 < Z_i < +1$.

Retrieval of modes of rainfall variability

In order to retrieve the rainfall variability modes for Zimbabwe, the empirical orthogonal function (EOF) analysis method was performed on the GPCC monthly rainfall data. EOF is a very commonly applied statistical technique in the field of atmospheric science. This technique was known as principal component analysis before Lorenz[21] introduced the term EOF into the literature as another name for the eigenvectors of a principal component analysis. EOF is one of the most common methods to extract important patterns from measurements of atmospheric variables.[22] Ogwang et al.[11] used this technique to show the dominant modes of variability of rainfall over Uganda. Ogallo[23] performed an EOF analysis to identify homogeneous regions of climate variability for Eastern Africa. Indeje et al.[24] used EOF analysis on the October–December season rainfall from 1961 to 1990 over East Africa. The EOF function defined by the formula of both space and time distributed data is computed using Equation 2:

$$Z\,(x\,y\,z) = \sum_{i=1}^{N} PC\,(t) \times EOF(x,y), \qquad \text{Equation 2}$$

where Z (xyz) is the original time series as a function of time (t) and space (x,y). EOF (x,y) shows the spatial maps (x,y) of Zimbabwean summer rainfall of each summer season and its variation from year to year. PC(t) is the principal component that shows how the amplitude of each EOF varies with time. The technique enables the capture of a new set of variables of most observed variance from the data through the linear combination of the original variable.

The eigenvector with the highest eigenvalue is the first principal component of the data set. The December–February average rainfall data are for the period 1980–2013 inclusive. A number of studies have been carried out on the EOF method and its significance.[25,26] In this study, it was used to show the dominant modes of variability of summer rainfall over Zimbabwe.

Results and discussions

Mean rainfall patterns

The long-term mean annual cycle of rainfall over Zimbabwe is presented in Figure 2. Figure 2 shows that significant rainfall (mean monthly totals above 60 mm) starts in November and ends around March. It can be inferred that there is a possible shortening of the summer season, which is a shift from the results of Buckle[17], but in agreement with the results of Mamombe et al.[8] These results also show that the peak rainfall occurs in January when the effects of the ITCZ are strongest (mean monthly totals above 140 mm); this finding is in agreement with the results of previous studies.[18,22] About 65% of the summer rainfall in Zimbabwe occurs during the period from December to February.

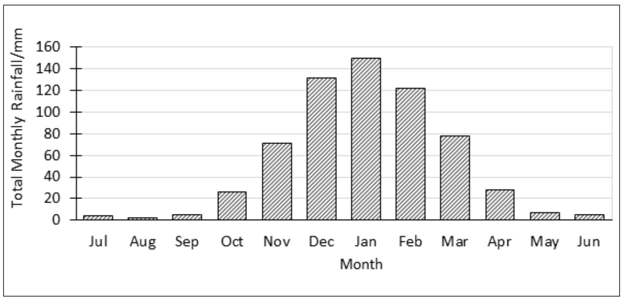


Figure 2: Annual cycle of rainfall (mm/month) in Zimbabwe for the period 1980 to 2013.

Figure 3 shows the spatial distribution for the long-term mean monthly rainfall for Zimbabwe for the months November to March. The maps in 3a–e show a gradual increase in precipitation to a peak in January, declining in March. From Figure 3, it can be seen that rainfall decreases in north–south and east–west gradients for Zimbabwe. This observation can be explained by the dominant influence of moisture advection into the country due to northeast and southeast airflow which brings moisture from the equatorial regions and from the Indian Ocean. Such results agree with those of Mamombe et al.[8] and Buckle[17].

Figure 4 shows the long-term mean rainfall maps (1980–2013) for Zimbabwe. Figure 4a is the long-term mean map for the period November to March while Figure 4b is for the period December to February. The northeastern part of Zimbabwe receives mean monthly rainfall in the range of 100 mm to 200 mm in the November to March period and 125 mm to 200 mm in the December to February period. For the two seasons, the southwestern areas receive rainfall in the range of 0 to 100 mm. The spatial rainfall distribution for the two seasons is similar. By choosing the December to February season to study the characteristics of Zimbabwean summer rainfall, the results from Figure 4 show that this objective can be accomplished. Figure 4a and 4b show that the northeastern areas receive more rainfall while the southwestern areas receive less rainfall – in agreement with results shown in Figure 3. From past studies, the rainfall system responsible for rainfall in the months of October to March was found to be westerly cloud bands.[17,21] From the results of this study, the amount of rainfall received in the months of October and April has decreased, leading to a possible reduction in the length of the Zimbabwean summer season. As such, it can be concluded that, due to climate change, the contribution of Zimbabwean summer rainfall due to the westerly cloud bands is decreasing while the contribution from the ITCZ remains dominant.

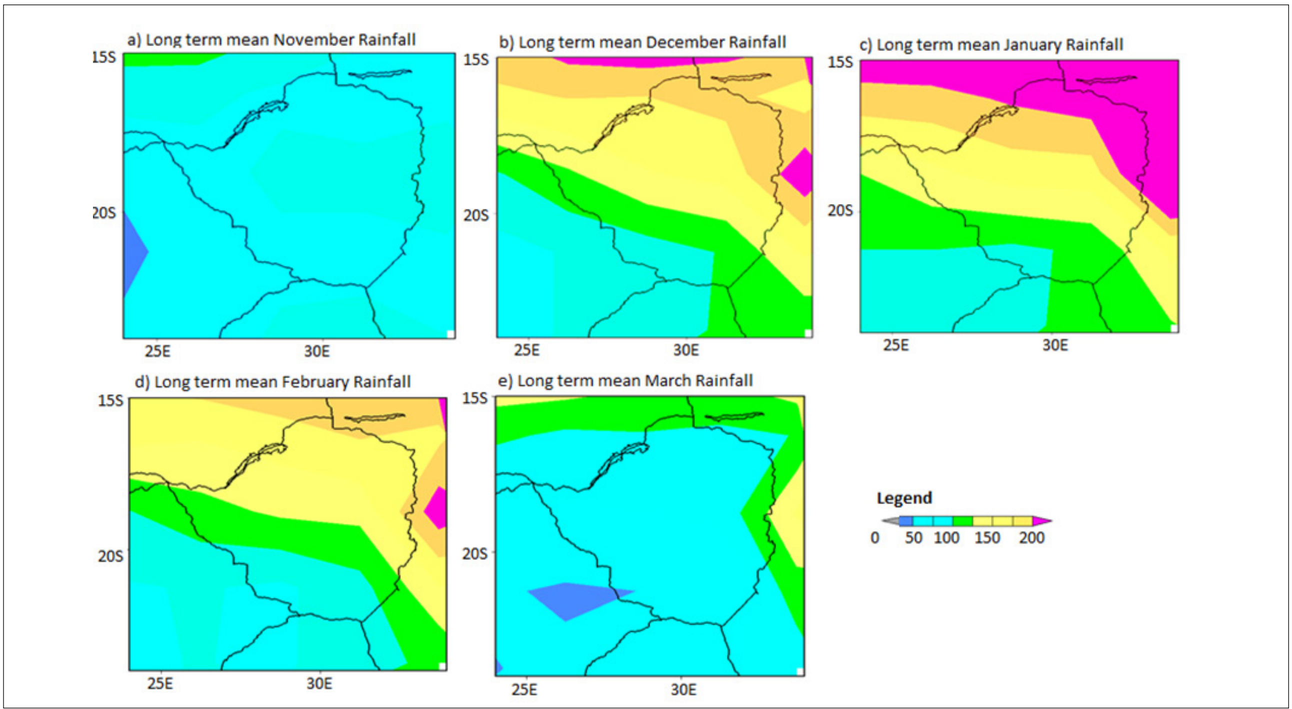


Figure 3: Long-term monthly mean rainfall (mm/month) maps for Zimbabwe for the summer months, November to March.

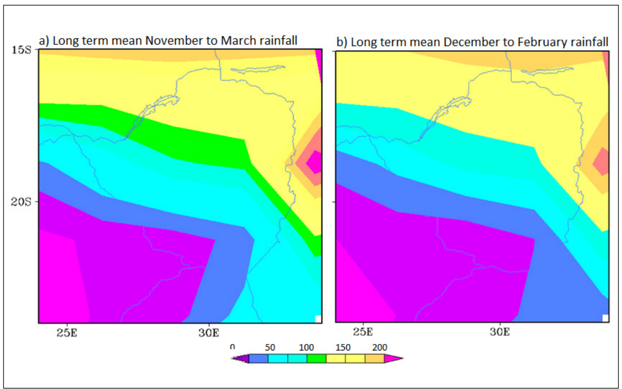


Figure 4: Long-term mean maps for Zimbabwe summer rainfall (in mm) for the period 1980 to 2013 for the November to March and December to February periods.

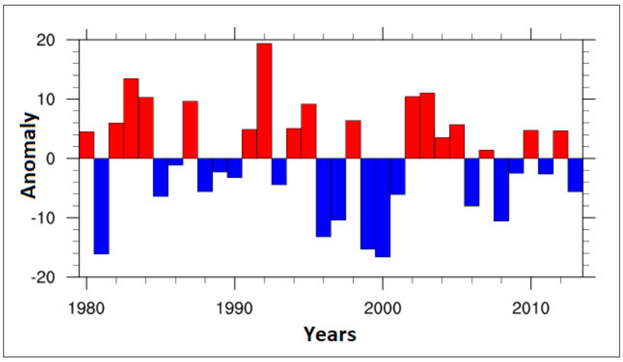


Figure 5: The Standardised Rainfall Index for Zimbabwe time series for the period 1980 to 2013.

Spatio-temporal characteristics of Zimbabwe rainfall, 1980–2013

Temporal rainfall analysis using Standardised Rainfall Index time series

From the Standardised Rainfall Index for Zimbabwe time series for the period 1980 to 2013 (Figure 5), seven wet years (1980/1981, 1995/1996, 1996/1997, 1998/1999, 1999/2000, 2005/2006 and 2007/2008) and seven dry years (1982/1983, 1983/1984, 1986/1987, 1991/1992, 1994/1995, 2001/2002 and 2002/2003) are identified. About 42% of the summer seasons between 1980 and 2013 were extreme rainfall events (wet/dry seasons). It can be deduced that the first period, from the 1981 season to the 1996 season, was relatively dry (five dry seasons) followed by a relatively wet period between 1996 and 2013 (six wet seasons). The 10-year cycle in which successive events of extreme dryness occur is also evident from the findings of previous studies.[2,3,6] An extreme wet period precedes or succeeds an extreme dry period in Zimbabwe, as the atmosphere tries to establish a balance. During the study period, rainfall is characterised by oscillating periods of wet and dry years.

Modes of variability from EOF analysis

Figure 6 shows the time series and spatial pattern of the first three eigenvectors (EOF 1–3) of the December–February seasonal rainfall. The data for all 396 months (December–February for 33 years) were used to build the eigenvectors. The first three eigenvectors (PC) explain about 87% of the total variance. Positive loading follows the central watershed (high altitude areas of the country). Weak loading is found in the northeast, southeast, southwest and western parts of the country, corresponding to the low altitude areas. These results cement previous results of other authors which emphasise the dominant role of topography in determining the spatial characteristics of summer rainfall in Zimbabwe.[8,15]

From EOF 1, 2 and 3, summer rainfall in Zimbabwe has three dominant spatial modes of variability. The largest mode accounts for about 62.7% of summer rainfall variability which occurs in a northeast to southwest gradient. Such a pattern can be explained by moisture advection into the country from the northeast as the ITCZ approaches from the north. The second mode accounts for 17.7% of this variability, emphasising an oscillating pattern in which the northeast/southwest parts are wetter/drier in some years and drier/wetter in other years. This pattern shows evidence of rain-bearing features which approach the country from the southwest and which are stronger in some years and weaker in other years, competing with the ITCZ approaching from the north. Such a feature

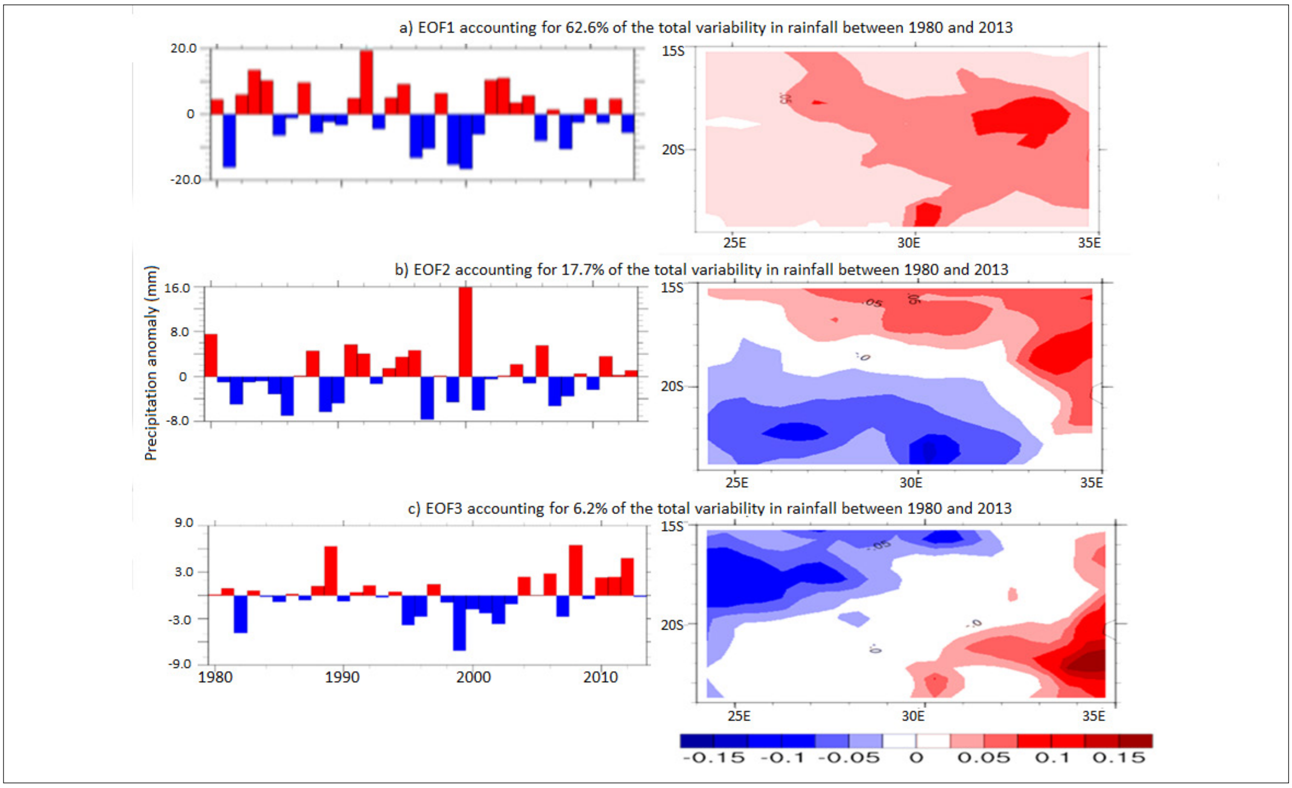


Figure 6: Time series and eigenvectors for the variability of Zimbabwean summer rainfall, 1980 to 2013.

is the westerly cloud bands which are active at the onset and cessation of summer in southern Africa. These results differ from those of Mamombe et al.[8] and Manatsa and Mukwada[7] who found an east to west oscillation. The difference could be because the previous studies focused on the effect of El Niño Southern Oscillation and Indian Ocean Dipole Zonal Mode, while the current study looked at other localised synoptic scale features. The smallest mode of variability identified is the northwest to southeast oscillation which results in the northwest/southeast areas being wetter/drier in some years but becoming drier/wetter in other years. This mode accounts for only 6.5% of the total variability. This once more shows the presence of rain-bearing features which control moisture advection into the country – one located to the northwest of the country and another to the southeast. Such features can be identified as the Angola low pressure system and cyclones in the Mozambican channel. The results show the dominance of the ITCZ in influencing both temporal and spatial characteristics of Zimbabwean summer rainfall from EOF 1, while EOF 3 shows the role of the Angola low pressure system and cyclones in the Mozambican channel.

Spatial distribution of rainfall follows a northeast to southwest bias with areas to the northeast receiving more rainfall than areas to the southwest. However, areas to the southwest receive more rainfall in the November to January period than in the January to March period while areas to the northeast receive less rainfall in the November to January period and more rainfall in the January to March period. During dry years, southern and western areas are wetter than normal compared to areas to the north and eastern areas. In years where westerly/easterly waves are weak/strong, the westerly cloud bands/ITCZ are weak/strong which results in drier/wetter southwest/northeast areas. The reverse occurs in other years, resulting in the northwest to southwest fluctuation. A lesser see saw occurs in a northwest to southeast direction, meaning that when cyclone activity is strong/weak in the Mozambican channel, the Angola low pressure system is respectively weak/strong.

Spatial characteristics for December–February rainfall

From the EOF analysis, three dominant modes of variability were obtained in terms of spatial distribution of Zimbabwean summer rainfall (Figure 6). These were the northeast to southwest variation, northeast to southwest oscillation and the smaller mode which is the southeast to northwest oscillation. The results in this section show which variation is dominant for wet and dry years. It is noted that during dry years, the areas usually dry become wetter while those usually wet become drier.

Spatial distribution of summer rainfall in Zimbabwe for wet years

From Figure 7a and 7b, it can be deduced that the rainfall for Zimbabwe during wet years follows the pattern for the long-term mean with a northeast to southwest orientation with rainfall decreasing in the same pattern. The highest amount of precipitation and precipitation anomaly is in the Eastern Highlands in the Mukandi area (250–280 mm monthly mean and >100 mm above long-term mean rainfall). The lowest rainfall anomaly is in the extreme southwest parts around the Beit Bridge to Plumtree area (120–160 mm monthly mean and 50–60 mm above the long-term mean rainfall). The EOF analysis results show the dominating role of the ITCZ on Zimbabwean summer rainfall. From the rainfall anomaly map (Figure 7b), it can be seen that all areas have positive rainfall anomalies. The anomaly belts are oriented in a southeast to northwest direction with anomalies decreasing on either side of the highest anomaly belt, corresponding to the country's topography.

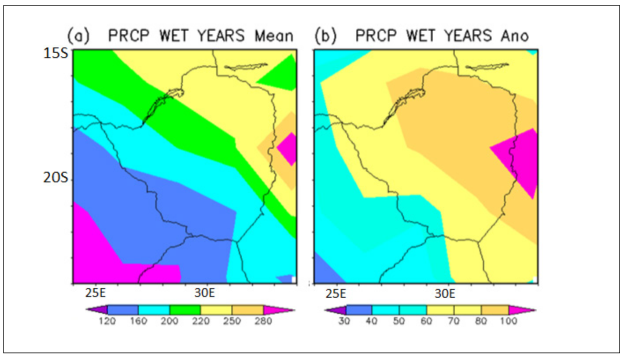


Figure 7: (a) Rainfall composite mean and (b) rainfall anomaly maps for Zimbabwe for the November to January season (in mm) from 1980 to 2013.

Figure 8 shows the spatial distribution of rainfall for six chosen wet years (1980/1981 season, 1995/1996 season, 1996/1997 season, 1998/1999 season, 1999/2000 season and 2007/2008 season). The maps in Figure 8 show two modes of variability. The dominant rain-bearing factor responsible for wet years is possibly tropical cyclones or a very deep low-pressure system in the Mozambican channel. Tropical cyclones can move into the interior of the subregion either through the Limpopo River to the southeast or through the Zambezi River to the northeast of the country. These cyclones promote moisture advection into the southeast, southern and southwest areas of the country.

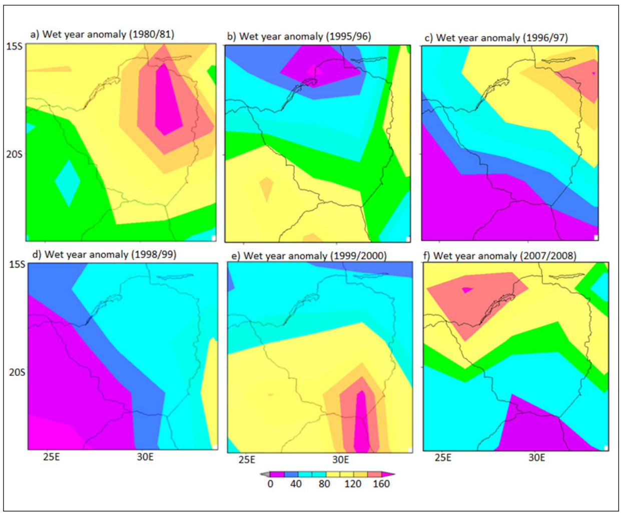


Figure 8: Rainfall distribution maps (in mm) for Zimbabwe for six wet years over the December–February months.

Spatial distribution of summer rainfall in Zimbabwe for dry years

Figure 9a and 9b show the composite maps for mean rainfall and anomalies in dry years. It can be deduced that the rainfall for Zimbabwe during dry years also follows the pattern for the long-term mean with a northeast to southwest gradient. The highest amount of precipitation is in the extreme northeast around Mutoko and Muzarabani areas (about 120 mm monthly mean). The lowest rainfall is in the southwest and southeast regions in the Chiredzi, Masvingo, Beit Bridge and Plumtree areas (<60 mm monthly mean).

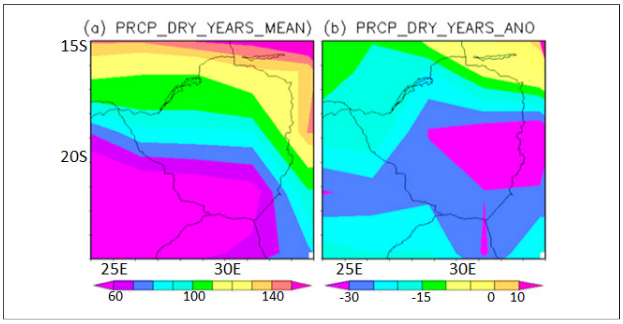


Figure 9: (a) Mean and (b) anomaly maps for rainfall distribution (in mm) over Zimbabwe for six dry years for the December–February months.

However, the anomaly map (Figure 9b) shows that all areas receive below normal rainfall, with the largest deficit being in the area that usually receives the highest rainfall, the Eastern Highlands (anomaly of -30 mm below long-term monthly mean). Smaller deficits are further to the north, meaning that the tropical rain-bearing system (ITCZ) is still active even in dry years. The low-pressure system in the Mozambican channel is weak during dry years, resulting in large deficits of rainfall in the Eastern Highlands around the Mukandi area.

Figure 10 shows the spatial distribution of rainfall for six chosen dry years: (10a) 1981/1982 season, (10b) 1983/1984 season, (10c) 1986/1987 season, (10d) 1991/1992 season, (10e) 1994/1995 season and (10f) 2001/2002 season. Regions usually drier in wet years (northwest and southwest parts) become wetter in dry years. For dry years, the effects of the ITCZ, westerly waves and easterly waves are revealed. From these maps, the dominant effects of each of the three factors can be shown by the areas with positive or negative anomalies. From the anomaly maps, the swings from northeast to southwest and southeast to northwest emerge, showing the effects of the five synoptic scale features which control rainfall characteristics over Zimbabwe. The scale of contribution is also seen by analysing the level of anomaly during such oscillations.

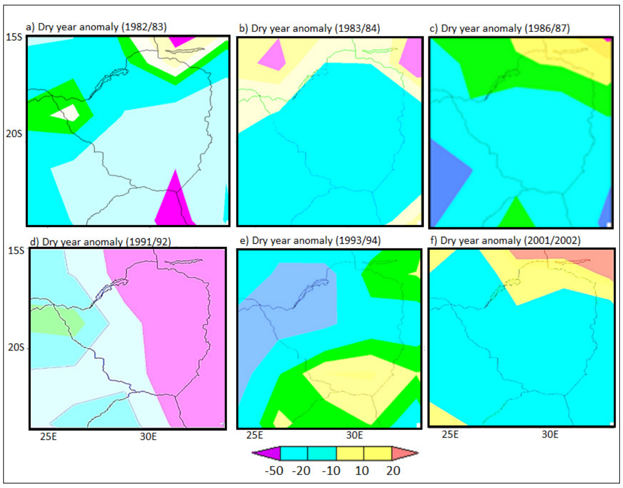


Figure 10: Rainfall distribution maps (in mm) for Zimbabwe for six dry years over the December–February months.

Conclusion

We characterised wet and dry summers and their modes of variability on an inter-annual time scale for Zimbabwe using GPCC model data. The Standardised Rainfall Index, EOF and composite analysis were applied on monthly data covering the period 1980–2013. Based on our findings, the comparatively high spatial resolution GPCC model data set is useful for rainfall studies in regions like Zimbabwe that have a variable topography. Our findings are in agreement with the results of Mwafulrirwa[14] who utilised model rainfall data to study rainfall variability over Malawi. This conclusion can be deduced from the capability to accurately identify known historical extreme rainfall events in Zimbabwe from the literature. It was also concluded that there is a possible decrease in the length of the rainfall summer season in Zimbabwe, although more research needs to be done to ascertain and verify these results. Also, the frequency of extreme summer events in both wet and dry seasons was high – at approximately 42% of the 33 years used in the study. Based on the EOF analysis, the spatial distribution of rainfall for the different modes of variability show evidence of the five previously postulated rain-bearing systems being the drivers of inter-annual variability of rainfall in Zimbabwe, as opposed to the results of many researchers e.g. Manatsa et al.[5,7] and Mamombe et al.[8] who argued that rainfall variability over southern Africa is attributed to the El Niño Southern Oscillation and the Indian Ocean Dipole. The ITCZ, westerly cloud bands, tropical cyclones, easterly waves and Angola low pressure systems are the major systems that contribute to characteristics of wet and dry summers over Zimbabwe for the period under study – this is in agreement with the findings of Jury et al.[1,21] and Buckle[17] who also report the importance of the ITCZ but downplay the role of tropical cyclones and westerly cloud bands on variability of southern Africa rainfall. Future studies must be done to further analyse these five features, especially their changes in view of climate change. This study focused on Zimbabwe, therefore similar analysis should be done for other regions in southern Africa to verify and ascertain these findings. Resultant findings would then greatly enhance the production of accurate and reliable seasonal forecasts for efficient early warning systems for the region.

Acknowledgements

We thank Mr Collen Mutasa, Dr Vitalis Musara and Mr Innocent Masukuwedza for their guidance, support and immense contributions towards the successful production of this paper; their suggestions and corrections were greatly valuable for the successful completion of this study. Lastly, we thank the Government of Zimbabwe, Government of the Republic of China, World Meteorological Organization and the responsible authorities at NUIST. We gratefully acknowledge the GPCC and NOAA/NCDC for providing data sets.

Authors' contributions

M.M. was responsible for the conceptualisation of the article; identifying the appropriate methodology; data collection, processing and analysis; writing the first draft; and writing revisions. T.M. was involved in the identification of appropriate data sources; validation of data; writing the first draft; and student supervision. E.Matandirotya was instrumental in writing and revising the article; student supervision; and project leadership. E.Mashonjowa was the overall supervisor for the whole team and provided leadership and advice critical for the successful completion of the work.

References

1. Jury MR, Levey K. The climatology and characteristics of drought in the eastern Cape of South Africa. Int J Climatol. 1993;13(6):629–641. https://doi.org/10.1002/joc.3370130604
2. Makarau A. Intra-seasonal oscillatory modes of the southern Africa summer circulation [PhD thesis]. Cape Town: University of Cape Town; 1995.
3. Jury MR. The structure and possible forcing mechanisms of the 1991-1992 drought in southern Africa. S Afr Tydskr Natuurw Tegnol. 1993;12(1):8–16. https://doi.org/10.4102/satnt.v12i1.551
4. Zhakata W. Impacts of climate variability and forecasting on agriculture (Zimbabwe experience). In: Proceedings of the Workshop on Reducing Climate-Related Vulnerability in Southern Africa; 1996 October 1–4; Victoria Falls, Zimbabwe. Silver Spring, MD: NOAA; 1996. p. 131–139.
5. Manatsa D, Mushore T, Lenouo A. Improved predictability of droughts over southern Africa using the standardized precipitation evapotranspiration index and ENSO. Theor Appl Climatol. 2017;127(1–2):259–274. https://doi.org/10.1007/s00704-015-1632-6
6. Dlangezwa K, Africa S. Predictability of Zimbabwe summer rainfall 1. Int J Climatol. 1997;17(13):1421–1432. https://doi.org/10.1002/(SICI)1097-0088(19971115)17:13%3C1421::AID-JOC202%3E3.0.CO;2-Z
7. Manatsa D, Mukwada G. Rainfall mechanisms for the dominant rainfall mode over Zimbabwe relative to ENSO and/or IODZM. Sci World J. 2012;2012, Art. #926310, 15 pages. https://doi.org/10.1100/2012/926310
8. Mamombe V, Kim WM, Choi YS. Rainfall variability over Zimbabwe and its relation to large-scale atmosphere–ocean processes. Int J Climatol. 2017;37(2):963–971. https://doi.org/10.1002/joc.4752
9. Awadallah AG. Selecting optimum locations of rainfall stations using kriging and entropy. Int J Civ Environ Eng. 2012;12(1):36–41.
10. Levitus S, Boyer T. Data set. Boulder, CO: NOAA/OAR/ESRL PSD; 1998.
11. Ogwang BA, Chen H, Li X, Gao C. Evaluation of the capability of RegCM4.0 in simulating East African climate. Theor Appl Climatol. 2016;124(1–2):303–313. https://doi.org/10.1007/s00704-015-1420-3
12. Ogwang BA, Guirong T, Haishan C. Diagnosis of September-November drought and the associated circulation anomalies over Uganda. Pakistan J Meteorol. 2012;9(17):11–24.
13. Kabanda TA, Jury MR. Inter-annual variability of short rains over northern Tanzania. Clim Res. 1999;13(3):231–241. https://doi.org/10.3354/cr013231
14. Mwafulirwa ND. Climate variability and predictability in tropical southern Africa with a focus on dry spells over Malawi [MSc thesis]. KwaDlangezwa: University of Zululand; 1999.
15. Muchuru S, Landman WA, De Witt DG. Seasonal rainfall predictability over the Lake Kariba catchment area. Water SA. 2014;40(3):461–470. https://doi.org/10.4314/wsa.v40i3.9
16. Unganai LS, Mason SJ. Long-range predictability of Zimbabwe summer rainfall. Int J Climatol. 2002;22(9):1091–1103. https://doi.org/10.1002/joc.786
17. Buckle C. Weather and climate in Africa. Harlow: Longman; 1996.
18. Zimbabwean Department of Meterological Services. Climate handbook of Zimbabwe. Harare: Zimbabwean Department of Meteorological Services; 1981. p. 48.
19. Matarira CH. Theoretical and applied climatology frequency and tracks of anticyclones and their effect on rainfall patterns over Zimbabwe. Theor Appl Climatol. 1990;66(1):53–66. https://doi.org/10.1007/BF00865526
20. Makarau A, Jury MR. Seasonal cycle of convective spells over southern Africa. Int J Climatol. 1997;17(12):1317–1332. https://doi.org/10.1002/(SICI)1097-0088(199710)17:12%3C1317::AID-JOC197%3E3.0.CO;2-A
21. Lorenz EN, Project MI of TSF. Empirical orthogonal functions and statistical weather prediction. Boston, MA: Department of Meteorology, Massachusetts Institute of Technology; 1956.
22. Hannachi A. A primer for EOF analysis of climate data [document on the Internet]. c2004 [cited 2020 Jan 13]. Available from: http://www.o3d.org/eas-6490/lectures/EOFs/eofprimer.pdf
23. Ogallo LJ. Spatial and temporal patterns of East African seasonal rainfall derived from the principal component analysis. Int J Climatol. 1989;9(22):145–167. https://doi.org/10.1002/joc.3370090204
24. Indeje M, Semazzi FHM, Ogallo LJ. ENSO signals in East African rainfall seasons. Int J Climatol. 2000;20(1):19–46. https://doi.org/10.1002/(SICI)1097-0088(200001)20:1%3C19::AID-JOC449%3E3.0.CO;2-0
25. Beilfuss R. A risky climate for southern African hydro. Berkeley, CA: International Rivers; 2012. Available from: http://climenews.com/letoltes/zambezi_climate_report_final.pdf
26. Behera SK, Luo JJ, Masson S, Delecluse P, Gualdi S, Navarra A, et al. Erratum: Paramount impact of the Indian Ocean dipole on the East African short rains: A CGCM study (Journal of Climate vol. 18 (21) (4514–4530)). J Clim. 2006;19(7):1361. https://doi.org/10.1175/JCLI3541.1

Check for updates

Identification of solar periodicities in southern African baobab $\delta^{13}C$ record

AUTHOR:
Pieter B. Kotzé[1,2,3]

AFFILIATIONS:
[1]South African National Space Agency, Space Science, Hermanus, South Africa
[2]Centre for Space Research, North-West University, Potchefstroom, South Africa
[3]Physics Department, Stellenbosch University, Stellenbosch, South Africa

CORRESPONDENCE TO:
Pieter Kotzé

EMAIL:
pkotze@sansa.org.za

DATES:
Received: 13 Aug. 2019
Revised: 28 Jan. 2020
Accepted: 10 Mar. 2020
Published: 29 July 2020

HOW TO CITE:
Kotzé PB. Identification of solar periodicities in southern African baobab $\delta^{13}C$ record. S Afr J Sci. 2020;116(7/8), Art. #6813, 5 pages. https://doi.org/10.17159/sajs.2020/6813

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Yali Woyessa

KEYWORDS:
tree ring, carbon isotope data, spectral analysis, solar influences, climate variability

FUNDING:
None

Spectral analysis using wavelet, Lomb–Scargle and maximum entropy techniques of the proxy rainfall record of northeastern South Africa based on annual carbon isotope ($\delta^{13}C$) data obtained from baobab trees for the period 1600 AD – 2000 AD show clear evidence of the presence of characteristic solar periodicities. Solar periodicities that were identified above the 95% confidence level include the ~11-year Schwabe cycle, the ~22-year Hale cycle as well as the 80–110-year Gleissberg cycle. A Morlet wavelet analysis of the $\delta^{13}C$ data between 1600 AD and 1700 AD shows the effect of the Maunder sunspot minimum on both the Schwabe and Hale cycles during this time.

Significance:

A spectral analysis of $\delta^{13}C$ baobab tree ring data for southern Africa between 1600 AD and 2000 AD is presented. The results show – for the first time – that the $\delta^{13}C$ time series contains the 11-year Schwabe, 22-year Hale as well as the 80–110-year Gleissberg solar periodicities. In addition, the influence of the Maunder sunspot minimum between ~1650 AD and 1715 AD could also be clearly identified in the $\delta^{13}C$ data for the first time. These findings are of significant importance to investigations of solar influences on climate variability.

Introduction

Carbon isotope ratios have been used in the literature[1] by several palaeoclimate researchers as a proxy for rainfall in a specific region where data were made available by tree ring investigations. During the process of photosynthetic absorption of atmospheric CO_2 by trees and plants, active discrimination occurs against $^{13}CO_2$ in favour of $^{12}CO_2$.[2] During periods of high rainfall, rapid exchange of CO_2 with the atmosphere and trees takes place. During periods of low rainfall, on the other hand, the amount of CO_2 available for photosynthesis is restricted to only that inside leaves. Therefore, in regions of low rainfall, like southern Africa, lower $\delta^{13}C$ values are associated with wetter conditions, and higher $\delta^{13}C$ values are associated with drier conditions during the growth of the wood.[3] In order to link carbon isotope ratios to rainfall, one is restricted to trees that only rely on precipitation as a water source. Carbon ratios from the African baobab (*Adansonia digitata* L.) have shown strong correlation with summer rainfall patterns in southern Africa in a previous study.[4]

The sun is a variable star that exhibits changes on multiple spatial and temporal scales, and is the most important source of energy for the earth's climate system (see for example a review on this topic by Gray et al.[5] and the references therein). Solar activity is characterised by periodicities and cycles that vary on a variety of timescales. The most important periodicities are the ~11-year sunspot cycle (also known as the Schwabe cycle), the ~22-year magnetic cycle (also known as the Hale cycle), the 80–110-year Gleissberg cycle as well as the ~205–year De Vries cycle. These periodicities of solar origin have been identified in several tree ring records and cosmogenic data such as ^{10}Be and ^{14}C (see for example the review by Usoskin[6] and the references therein). Convincing evidence for solar influences on weather and climate has been found in the last ~150 years (Hoyt and Schatten[7] and references therein). Souza Echer et al.[8] found a clear correlation in the 22-year periodicity between global air surface temperature and sunspot number in the interval 1880–2000. Although solar signals have been detected in climate records (e.g. Beer et al.[9]), one should keep in mind that climate and rainfall is a non-linear system and that several forces are involved in the process. Apart from the direct and indirect effects of solar irradiance (total solar irradiance and ultraviolet), particle radiation effects (solar energetic particles and cosmic rays) also have an effect on the troposphere (see Figure 21 in Gray et al.[5]). Rigozo et al.[10] investigated the solar activity signals in tree ring data in Chile over a period of ~400 years, employing spectral analysis techniques. They found clear evidence for the presence of solar activity: Schwabe (~11-year), Hale (~22-year), fourth-harmonic of the 208-year De Vries (~52-year) and Gleissberg (~80-year) cycles. Several investigations have indicated that solar variations had an effect on pre-industrial climate throughout the Holocene (the warm interval since the last ice age). These studies have been done in several parts across the globe and utilised a large variety of palaeoclimate proxies. The reader is referred to the recent review by Lockwood[11]. An interesting finding by McKinnell and Crawford[12] is the long-term lunar-induced 9.3-year and 18.6-year cycles in the temperature records in some regions over a 400-year period. A wavelet analysis of centennial (1895–1994) rainfall data in southern Brazil by Souza Echer et al.[13] confirmed the existence of strong 8.9-, 11.7- and 24.9-year periodicities. ^{14}C time series studies on tree growth rings from Bashkiria (Russia) by Kocharov et al.[14] revealed periods close to 21, 9.4 and 13.5 years using advanced spectral analysis methods. Damon et al.[15], who also studied a time series of $\delta^{14}C$ in tree rings, for the time interval between 1065 AD and 1250 AD, confirmed these results. They also detected a period of ~52 years and attributed it to the fourth harmonic of the Suess or De Vries cycle. The climate of southern Africa is to a large extent semi-arid.[16] Except for the southern coastal area, the southern African region receives almost all its rainfall in summer. A characteristic feature of the southern African rainfall record is the pronounced ~80-year cycle that can be traced back 3500 years.[16]

In this investigation, $\delta^{13}C$ measurements obtained from the Pafuri 1000-year-old baobab trees[17] were analysed between the years 1600 AD and 2000 AD to identify the presence of possible solar periodicities. Spectral analysis methods such as wavelets[18], Lomb–Scargle[19,20] and maximum entropy[21] were used to obtain statistically significant results above the 95% confidence level. In addition, a Morlet[22] wavelet analysis of $\delta^{13}C$ between 1600 AD and

1700 AD was performed to investigate the variation in the strengths of, particularly, the Schwabe and Hale cycles during the Maunder sunspot minimum period of 1645–1715 AD (Eddy).[23]

Data and method of analysis

Woodborne et al.[17] did a carbon isotope analysis of four South African baobab trees located in the Pafuri area (22°22.925'S, 31°13.145'E; 22°26.647'S, 31°04.745'E; 23°15.765'S, 31°33.309'E; 22°24.352'S, 31°16.676'E) in order to obtain a proxy for rainfall in this region. A map of southern Africa showing the location of the trees is given in Figure 1.

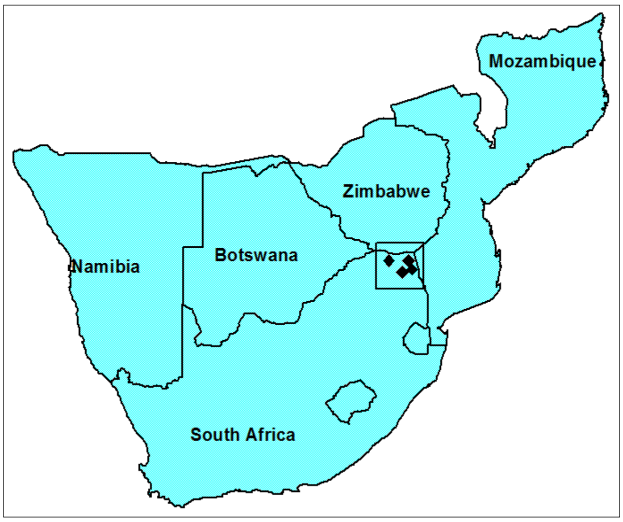


Figure 1: A map of southern Africa showing the location of the four baobab trees (inside the square area as indicated) whose $\delta^{13}C$ data were used in this study.

Samples obtained were large enough to separate individual ring structures at 1-year intervals. Woodborne et al.[17] subsequently performed an accelerator mass spectrometry radiocarbon analysis of these samples and calibrated the dates using well-known standards and procedures. They used the extracted $\delta^{13}C$ data to confirm abnormal dry conditions during the Little Ice Age, particularly during 1635 and 1695. The $\delta^{13}C$ data were obtained from Woodborne et al.[24] (http://www.ncdc.noaa.gov/paleo/study/17995). Figure 2 shows the data between AD 1600 and 2000 utilised in this investigation.

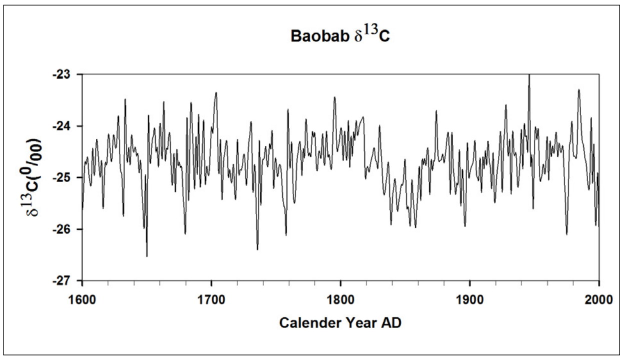


Figure 2: A plot showing the $\delta^{13}C$ tree ring data at annual intervals between 1600 AD and 2000 AD as obtained by Woodborne et al.[17] from baobab trees in the Pafuri region of southern Africa.

Three spectral analysis approaches were employed in this investigation to perform a spectral analysis of $\delta^{13}C$ time series, namely the wavelet, the Lomb–Scargle periodogram method, and the maximum entropy method, in order to identify characteristic periodicities above the 95% confidence level. Wavelet analysis is used extensively to decompose a time series into time–frequency space and to establish the modes of variability, particularly when the time series under investigation contains non-stationary power at different periods. As the wavelet transform is a localised transform in both space (time) and frequency, this particular property can be usefully applied to extract information from the signal, which is not possible with Fourier techniques (Daubechies).[25] All graphs and plots in this paper were produced using the SigmaPlot (www.systat.com) plotting package, while the contour plots were generated using the Interactive Data Language wavelet applet (http://www.exelisvis.com/ProductsServices/IDL.aspx).

Results and discussion

Morlet wavelet power spectra of order 6 were obtained by analysing the $\delta^{13}C$ time series at a cadence of 1 year between 1600 AD and 1700 AD. This was done particularly to investigate the behaviour of the Schwabe (~11-year sunspot) cycle as well as the Hale (~22-year magnetic) cycle with time during this interval. The results can be seen in Figure 3 below.

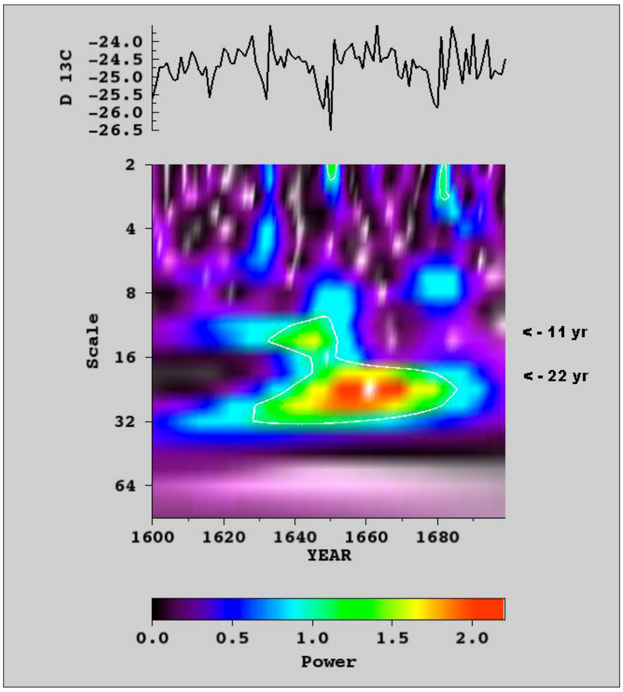


Figure 3: A Morlet wavelet power spectrum of the Pafuri $\delta^{13}C$ data between 1600 AD and 1700 AD, showing the behaviour of the Schwabe (~11-year) and Hale (~22-year) solar periodicities during this interval. The white contour line represents the 95% confidence level. The top plot shows the annual $\delta^{13}C$ data.

From Figure 3, the 11-year Schwabe sunspot cycle as well as the 22-year Hale solar magnetic cycle can be clearly identified. It is interesting to note that the 11-year sunspot cycle could only be observed between ~1630 AD and 1650 AD above the 95% confidence level, while between ~1650 AD and 1700 AD it disappeared completely. This particular interval coincides with the sunspot Maunder Minimum (1645–1715 AD) when no sunspots could be observed on the surface of the sun and was also known as the Little Ice Age. In the northern hemisphere, extreme cold spells were experienced during the Little Ice Age. It is quite extraordinary that a disappearance of the 11-year Schwabe cycle on the sun did indeed coincide with a similar pattern in the $\delta^{13}C$ data obtained from the summer rainfall region of southern Africa. The purpose of this investigation is not to show that rainfall is influenced by the sun, but that during extreme sunspot minimum conditions like the Maunder Minimum, we observe that solar periodicities like the 11-year Schwabe period in rainfall proxy parameters such as $\delta^{13}C$ can indeed disappear or become statistically insignificant. On the other hand, due to the sub-dominant role of the 11-year periodicity during the Maunder Minimum, we observe a dominant 22-year periodicity in this time, supporting observations by Mursula et al.[26] Similar observations during the Maunder Minimum in solar proxies showing a dominant 22-year periodicity accompanied by a less-dominant Schwabe cycle have been reported by Silverman[27] in aurora data and by Miyahara et al.[28] in ^{14}C data.

We proceeded with the analysis of $\delta^{13}C$ by performing a Lomb–Scargle spectral analysis between 1600 AD and 2000 AD at each 100-year interval in order to identify other prominent periodicities in the data. Results obtained are displayed in Figure 4.

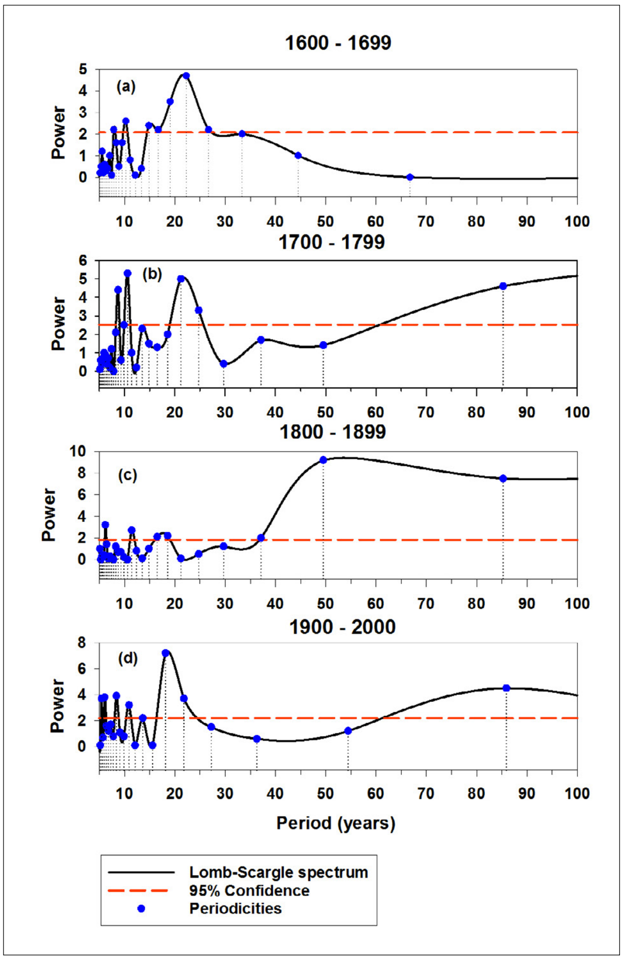


Figure 4: Lomb–Scargle spectral analysis of Pafuri baobab $\delta^{13}C$ data between (a) 1600 and 1699, (b) 1700 and 1799, (c) 1800 and 1899 and (d) 1900 and 2000. The black curves are spline fits of the periodicities revealed by the Lomb–Scargle method.

From Figure 4a (and Table 1) a Lomb–Scargle analysis revealed periodicities in the $\delta^{13}C$ data at ~18 and 22 years above the 95% confidence level, while the 10.3-year period is marginally statistically significant, confirming our wavelet analysis regarding the 11-year periodicity. The ~18-year period fits in with the general conclusion by Tyson et al.[16] that climate variability in the 16–20-year range is a characteristic feature of southern Africa. However, one cannot exclude the possibility that this ~18-year periodicity can be linked to the lunar-induced 18.6-year cycle as reported by McKinnell and Crawford[12]. This ~18-year cycle has been identified as a characteristic oscillation in southern African rainfall and has been extensively documented in the meteorological records. It appears predominantly in the summer rainfall tree-ring data, but could not be identified during 1700–1799 AD. Figure 4b, representing periodicities observed between 1700 AD and 1799 AD, shows a resurgence of both the 11-year as well as the 22-year cycles, while we also notice with interest the appearance of an 85-year period which represents the solar Gleissberg cycle. Variability at about 80 years has been observed on a global scale. A review by Hoyt and Schatten[7] summarises its presence in temperature and precipitation data sets as well as tree-ring investigations. During 1700–1799 AD, we also observed a 24.7-year as well as a 8.7-year periodicity (Table 1). These periods have also been observed in the rainfall records of southern Brazil[13] during 1894–1995. The 8.7-year period can possibly be a result of long-term, lunar-induced, 9.3-year modulation of sea-surface temperatures as reported by McKinnell and Crawford[12]. Periodicities identified between 1800 AD and 1899 AD, exceeding the 95% confidence level, can be seen in Figure 4c and Table 1. Of particular interest is the 11-year Schwabe cycle at 12.1 years as well as the 85-year Gleissberg cycle. In addition, we also see the appearance of a 50-year periodicity which is only observed during 1800–1899 AD. This particular periodicity has been identified by Cohen and Lintz[29] as a 'shoulder' to the Gleissberg solar cycle. During 1900–2000 (Figure 4d), our Lomb–Scargle analysis not only revealed the presence of the Schwabe (at 10.9 years), Hale (at 21.7 years) and Gleissberg (at 85.9 years) solar periodicities, but also the ~18-year and ~9-year periods to be statistically significant above the 95% level. Of particular interest is that the ~18-year period has the highest amplitude/strength during 1900–2000.

As a next step to verify results obtained by wavelet and Lomb–Scargle methods, we performed a maximum entropy spectral analysis of the Pafuri baobab $\delta^{13}C$ data for the time interval 1600–2000 AD. Maximum entropy results are shown in Figure 5 and summarised in Table 1.

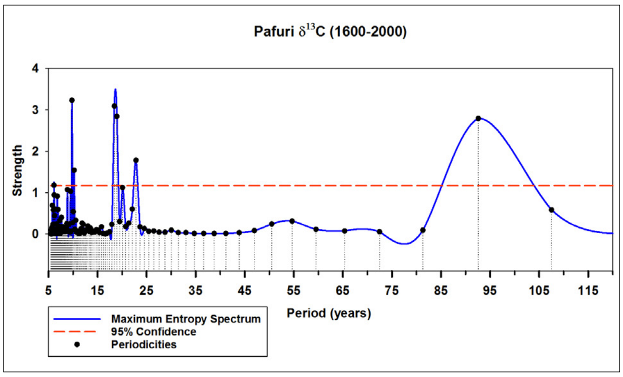


Figure 5: Maximum entropy spectral analysis of the Pafuri baobab $\delta^{13}C$ data for the time period between 1600 and 2000. The blue curve is a spline fit of the identified periodicities.

From our maximum entropy analysis, we could clearly identify the 11-year Schwabe (at 10.2 years), the 22-year Hale (at 22.8 years) as well as the Gleissberg solar period (at 92.5 years) above the 95% statistically significant level. In addition, the ~18-year period which is characteristic of southern African rainfall could also be verified. These results show, particularly, that the solar periodicities identified through our spectral analysis techniques are statistically significant.

Table 1: All periodicities identified above the 95% confidence level during each interval between 1600 AD and 2000 AD using Lomb–Scargle and maximum entropy spectral analysis techniques

Lomb–Scargle periodicities				Maximum entropy periodicities
1600–1699 AD	1700–1799 AD	1800–1899 AD	1900–2000 AD	1600–2000 AD
–	85.2	85.2	85.9	92.5
–	–	49.5	–	–
22.2	21.2, 24.7	–	21.7	22.8
15.5, 18.8	–	18.1	18.1	18.4
10.3	10.6	12.1	10.9	10.2
–	8.7	–	8.4	–

A cross-spectral analysis was subsequently performed to determine whether a statistically significant correlation exists between sunspot periodicities and the Pafuri $\delta^{13}C$ data between 1600 AD and 2000 AD. Cross-spectral analysis is a method to determine the common variability between two time series as a function of periodicity (frequency) and to determine the coherency between them. The Fortran computer program REDFIT-X[30] was used to determine the coherence between sunspot numbers and δ^{13}C observations. Results obtained are shown in Figure 6.

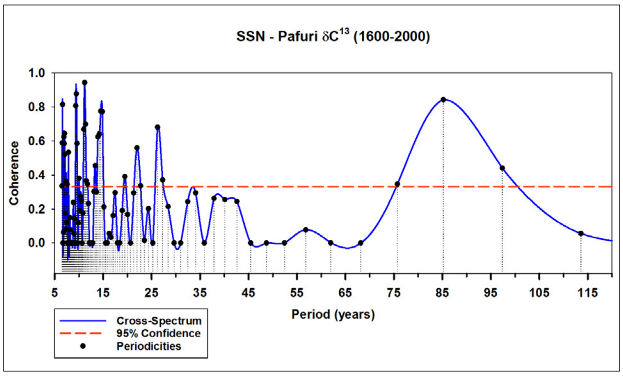


Figure 6: A cross-spectral analysis between annual mean sunspot numbers (ssn) and δ^{13}C data between 1600 AD and 2000 AD.

The cross-spectral results clearly show exceptionally strong coherence for the 11-year Schwabe cycle (>0.9) as well as the 85-year Gleissberg cycle (>0.8), while coherence for the 22-year Hale cycle was at a moderate level of 0.56. These results are all well above the 95% confidence level as displayed in Figure 6.

Conclusions

In this study, δ^{13}C measurements from the Pafuri baobab trees between the years 1600 AD and 2000 AD have been analysed in order to identify the presence of possible well-known solar periodicities. Spectral analysis methods such as wavelets, Lomb–Scargle and maximum entropy were employed to obtain statistically significant results above the 95% confidence level. Results obtained show clear evidence of the presence of the 11-year Schwabe sunspot cycle, the 22-year Hale magnetic cycle as well as the 80–100-year Gleissberg solar cycle. In addition, the ~18-year period which is characteristic of southern African rainfall, could also be identified in our spectral analysis, confirming previous investigations of the rainfall pattern across this region. Previous investigations linked this periodicity to lunar-induced effects[12] on sea-surface temperatures. A Morlet wavelet analysis of δ^{13}C data, particularly concentrating on the time interval between 1600 AD and 1700 AD, revealed, in particular, that the Schwabe cycle could only be observed between ~1630 AD and 1650 AD above the 95% confidence level, while between ~1650 AD and 1700 AD it disappeared completely. This particular period coincides with the sunspot Maunder Minimum period when hardly any sunspots could be identified on the surface of the sun. This period is also known as the Little Ice Age when exceptionally low temperatures were recorded in Europe. To the best of our knowledge, this is the first identification of the presence of the Maunder Minimum in δ^{13}C tree-ring data in southern Africa. In order to verify our results, we further performed a cross-spectral analysis of sunspot data and δ^{13}C measurements between 1600 AD and 2000 AD. Evidence obtained from this analysis showed exceptionally strong coherence, particularly for the Schwabe and Gleissberg cycles, while a moderate coherence exists for the 22-year Hale cycle. Waple et al.[31] investigated solar irradiance forcing on climate over a 200-year period between 1650 AD and 1850 AD preceding the industrial era with its anthropogenic influence. This era is characterised by a wide range of solar irradiance values and includes the Maunder Minimum. The authors conclude that on multi-decadal to century timescales, solar forcing is dominated by the ~90-year Gleissberg cycle, while at decadal timescales the 11-year and 22-year solar cycles dominate.

Sundqvist et al.[32] analysed a 350-year δ^{18}O stalagmite record from the summer rainfall region in South Africa and reported a positive correlation with regional air surface temperatures at interannual timescales. They found that during parts of the Little Ice Age, the regional temperature could have been ~1.5 °C colder than today. Prominent 22- and 11-year cycles also indicated that both the solar magnetic as well as the El Nino–Southern Oscillation cycle could be important drivers of multi-decadal to interannual climate variability in southern Africa. A surface water temperature reconstruction from Antarctica[33] shows similar features and a cooling of more than 2 °C between 1300 AD and 1800 AD. These records show clear evidence that the Little Ice Age, originally identified in the northern hemisphere, also had a counterpart in the southern hemisphere. Peristykh and Damon[34] did an interesting spectral analysis study of δ^{14}C tree-ring data divided into intervals before, during and after the Maunder Minimum. They found that the ~22-year Hale cycle is most prominent during the Maunder Minimum, while the 11-year cycle is totally suppressed. In addition, they also reported that *before* the Maunder Minimum the Hale cycle had a period of 22.9 years, whereas *during* the minimum the cycle had two periodicities at 24 and 15.8 years, and *after* the minimum the Hale cycle had a period of 26 years. Our investigation also confirmed the existence of a 15.5-year periodicity during 1600–1699 AD, but periodicities observed between ~25 and 30 years were found to be below the 95% confidence level. They were therefore regarded as not statistically significant.

The results obtained in this investigation, in particular δ^{13}C measurements from the Pafuri baobab trees in southern Africa between the years 1600 AD and 2000 AD, provide strong evidence of the presence of solar cycles. These δ^{13}C isotope measurements are regarded by palaeoclimatologists as a proxy for rainfall in a particular region. Results obtained in this investigation can be regarded as the first identification of prominent solar cycles in δ^{13}C tree-ring data in southern Africa. Of particular interest is the behaviour of the Schwabe cycle during the Maunder Minimum period between ~1650 AD and 1700 AD.

Acknowledgements

The author thanks the NOAA National Center for Environmental Information for archiving and making the tree ring data available. The assistance and constructive comments of two anonymous reviewers in evaluating this paper are appreciated.

Conflicts of interest

The author declares that no competing interests exist.

References

1. Norstrom E, Holmgren K, Morth CM. Rainfall-driven variations in d13C composition and wood anatomy of *Breonadia salicina* trees from South Africa between AD 1375 and 1995. S Afr J Sci. 2005;101(3–4):162–168.
2. McCarroll D, Loader NJ. Stable isotopes in tree rings. Quat Sci Rev. 2004;23(7):771–801. https://doi.org/10.1016/j.quascirev.2003.06.017
3. Farquhar GD, O'Leary MH, Berry JA. On the relationship between carbon isotope discrimination and the intercellular carbon dioxide concentration in leaves. Funct Plant Biol. 1982;9(2):121–137. https://doi.org/10.1071/PP9820121
4. Robertson I, Loader NJ, Froyd CA, Zambatis N, Whyte I, Woodborne S. The potential of the baobab (*Adansonia digitata* L.) as a proxy climate archive. Appl Geochem. 2006;21(10):1674–1680. https://doi.org/10.1016/j.apgeochem.2006.07.005
5. Gray LJ, Beer J, Geller M, Haigh JD, Lockwood M, Matthes K, et al. Solar Influences on Climate. Rev Geophys. 2010;48, RG4001. https://doi.org/10.1029/2009RG000282
6. Usoskin IG. A history of solar activity over millennia. Living Rev Sol Phys. 2013;10:1–94. https://doi.org/10.12942/lrsp-2013-1
7. Hoyt DV, Schatten KH. The role of the sun in climate change. New York: Oxford University Press; 1997.
8. Souza Echer MP, Echer E, Nordemann DJR, Rigozo NR. Multi-resolution analysis of global surface air temperature and solar activity relationship. J Atmos Solar-Terrestrial Phys. 2009;71:41–44. https://doi.org/10.1016/j.jastp.2008.09.032
9. Beer J, Mende M, Stellmacher R. The role of the Sun in climate forcing. Quat Sci Rev. 2000;19:403–415. https://doi.org/10.1016/ S0277-3791(99)00072-4

10. Rigozo NR, Nordemann DJR, Souza Echer MP, Echer E, Silva HE, Prestes A, et al. Solar activity imprints in tree ring width from Chile (1610–1991). J Atmos Solar-Terrestrial Phys. 2007;69(9):1049–1056. https://doi.org/10.1016/j.jastp.2007.03.010

11. Lockwood M. Solar induced climate effects. In: Encyclopaedia of sustainability science and technology. Berlin: Springer; 2012.

12. McKinnell SM, Crawford WR. The 18.6-year lunar nodal cycle and surface temperature variability in the northeast Pacific. J Geophys Res. 2007;112, C02002. https://doi.org/10.1029/2006JC003671

13. Souza Echer MP, Echer E, Nordemann DJR, Rigozo NR, Prestes A. Wavelet analysis of a centennial (1895-1994) southern Brazil rainfall series. Clim Change. 2008;87:489–497.

14. Kocharov GE, Ostryakov VM, Peristykh AN, Vasil'ev VA. Radiocarbon content variations and Maunder Minimum of solar activity. Sol Phys. 1995;159:381–391. https://doi.org/10.1007/BF00686539

15. Damon PE, Eastone CJ, Hughes MK, Kalin RM, Long A, Peristykh AN. Secular variation of $\delta^{14}C$ during the Medieval Solar Maximum: A progress report. Radiocarbon. 1998;40(1):343–350. https://doi.org/10.1017/S003382220001821X

16. Tyson PD, Cooper GRJ, McCarthy TS. Millennial to multi-decadal variability in the climate of southern Africa. Int J Climatol. 2002;22:1105–1117. https://doi.org/10.1002/joc.787

17. Woodborne S, Hall G, Robertson I, Patrut A, Rouault M, Loader NJ, et al. A 1000-year carbon isotope rainfall proxy record from South African baobab trees (*Adansonia digitata* L.). PLoS ONE. 2015;10(5), e0124202. https://doi.org/10.1371/journal.pone.0124202

18. Torrence C, Compo GP. A practical guide to wavelet analysis. Bull Am Meteorol Soc. 1998;79:61–78.

19. Lomb NR. Least-squares frequency analysis of unequally spaced data. Astrophys Space Sci. 1976;39:447–462. https://doi.org/10.1007/BF00648343

20. Scargle JD. Studies in astronomical time series II – Statistical aspects of spectral analysis of unevenly spaced data. Astrophys J. 1982;263:835–853. https://doi.org/10.1086/160554

21. Ables JG. Maximum entropy spectral analysis. Astron Astrophys Suppl. 1974;15:383–393.

22. Morlet J, Arens G, Forgeau I, Giard D. Wave propagation and sampling theory. Geophysics. 1982;47(2):203–221. https://doi.org/10.1190/1.1441328

23. Eddy JA. The Maunder Minimum. Science. 1976;192:1189–1202. https://doi.org/10.1126/science.192.4245.1189

24. Woodborne S, Hall G, Robertson I, Patrut A, Rouault M, Loader NJ, et al. NOAA/WDS Paleoclimatology – Pafuri, South Africa 1000 year baobab tree carbon isotope data [data set]. NOAA National Centers for Environmental Information. 2015 May 13 [cited 2020 Jan 28]. Available from: http://www.ncdc.noaa.gov/paleo/study/17995. https://doi.org/10.25921/k336-r342

25. Daubechies I. The wavelet transform, time-frequency localization and signal analysis. IEEE Trans Info Theory. 1990;36(5):961–1005. https://doi.org/10.1109/18.57199

26. Mursula K, Usoskin IG, Kovaltsov GA. Persistent 22-year cycle in sunspot activity: Evidence for a relic solar magnetic field. Sol Phys. 2001;198:51–56. https://doi.org/10.1023/A:1005218414790

27. Silverman SM. Secular variation of the aurora for the past 500 years. Rev Geophys. 1982;30:333–351. https://doi.org/10.1029/92RG01571

28. Miyahara H, Masuda K, Muraki Y, Furuzawa H, Menjo H, Nakamura T. Cyclicity of solar activity during the Maunder Minimum deduced from radiocarbon content. Sol Phys. 2004;224:317–322. https://doi.org/10.1007/s11207-005-6501-5

29. Cohen TJ, Lintz PR. Long-term periodicities in the sunspot cycle. Nature. 1974;250:398–400. https://doi.org/10.1038/250398a0

30. Ólafsdóttir KB, Schultz M, Mudelsee M. REDFIT-X: Cross-spectral analysis of unevenly spaced paleoclimate time series. Comput Geosci. 2016;91:11–18. https://doi.org/10.1016/j.cageo.2016.03.001

31. Waple AM, Mann ME, Bradley RS. Long-term patterns of solar irradiance forcing in model experiments and proxy based surface temperature reconstructions. Clim Dyn. 2002;18:563–578. https://doi/10.1007/s00382-001-0199-3

32. Sundqvist HS, Holmgren K, Fohlmeister J, Zhang Q, Bar Matthews M, Spötl C, et al. Evidence of a large cooling between 1690 and 1740 AD in southern Africa. Sci Rep. 2013;3:1767. https://doi.org/10.1038/srep01767

33. Berter NA, Mayewski NPA, Carter L. Cold conditions in Antarctica during the Little Ice Age – Implications or abrupt climate change mechanisms. Earth Planet Sci Lett. 2011;308:41–51. https://doi.org/10.1016/j.epsl.2011.05.021

34. Peristykh AN, Damon PE. Modulation of atmospheric ^{14}C concentration by the solar wind and irradiance components of the Hale and Schwabe solar sycles. Solar Phys. 1998;177:343–355. https://doi.org/10.1023/A:1004982321191

South African log resource availability and potential environmental impact of timber construction

AUTHORS:
Philip L. Crafford[1]
C. Brand Wessels[1]

AFFILIATION:
[1]Department of Forest and Wood Science, Stellenbosch University, Stellenbosch, South Africa

CORRESPONDENCE TO:
Philip Crafford

EMAIL:
pcrafford@sun.ac.za

DATES:
Received: 30 May 2019
Revised: 01 Apr. 2020
Accepted: 01 Apr. 2020
Published: 29 July 2020

HOW TO CITE:
Crafford PL, Wessels CB. South African log resource availability and potential environmental impact of timber construction. S Afr J Sci. 2020;116(7/8), Art. #6419, 8 pages. https://doi.org/10.17159/sajs.2020/6419

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Tania Douglas

KEYWORDS:
wood-based building systems, global warming potential, embodied energy, residential housing, South Africa

FUNDING:
Hans Merensky Foundation (South Africa)

We investigated the South African log resource availability and the potential global warming impact of an increasing wood-based residential building market. We have shown that, with the use of wood resources currently exported as chips, as well as planting trees in areas that have been earmarked for afforestation, a sustainable residential building market, where all constructions are wood-based, is possible. However, in the short term, imports of wooden building components might be necessary if rapid growth in wood-based building occurs. Basic modelling analyses show that if the market share of wood-based buildings increases to 20% of new constructions, the embodied energy and global warming potential of the residential building sector could decrease by 4.9%. If *all* new constructions were wood based, the total embodied energy and global warming potential of the residential building sector could decrease by up to 30%.

Significance:

- A novel finding of this paper is that sufficient local log resource options exist to realise a sustainable all-wood residential construction market in South Africa.
- The likely implications in terms of embodied energy and potential global warming impact of using wood-based materials for residential buildings compared to conventional brick and mortar or reinforced concrete buildings were also analysed and found to be favourable.

1. Introduction

Numerous studies have shown that timber is not only renewable, but is also the best performer across most environmental impact factors when compared to building material alternatives such as steel and concrete, with particularly good performance in terms of greenhouse gas emissions.[1-6] Trees absorb carbon dioxide during the photosynthetic process to form wood, which is a largely carbon-based material. Timber structures effectively store a similar mass of carbon that was removed from the atmosphere by the tree and fixed as wood.

Approximately 70% of local residential roof truss systems are wood based.[7] However, only 1% of new residential housing structures in South Africa can be described as wood-based structures. [Note: wood-based building systems in this paper comprise timber frame, cross-laminated timber and other wood-based materials such as orientated strand board or plywood.]

The rest are brick and mortar or cement block with timber roof truss systems (Slabbert W 2017, email communication, November 15). In some countries, such as the USA, Canada and Australia, well over 90% of residential housing is timber frame.[8] According to Palmer[8], timber frames account for about 70% of all housing stock in developed countries, representing close to 150 million homes.

According to Beradi[9], the building sector in developed countries produces up to 40% of their total greenhouse gases (GHG). In South Africa, it is estimated that the energy used in the construction of buildings is responsible for about 27% of South Africa's total anthropogenic carbon dioxide emissions.[10] The environmental footprint of residential buildings in South Africa can be reduced in various ways. Firstly, the traditional brick and mortar building materials and constructions can be replaced by lower environmental impact systems such as timber frame or even new timber panel systems (i.e. cross-laminated timber). Secondly, various strategies to decrease the operating impact of buildings can be introduced (solar energy, insulation, LED technology, etc.).

Currently, operational life-cycle energy requirements of conventional buildings are higher than the embodied energy.[11] However, as low-energy and near-zero-energy buildings (and employing energy-saving technology) become more prevalent, embodied energy will become a larger part of the total building energy requirements.[12,13] The objective of this study was to determine whether local forest resources would be able to supply the required wood for substantial growth in wood-based residential development in South Africa by an analysis of (1) the residential housing footprint in South Africa, (2) available log resources for wood-based buildings and (3) likely building-system environmental impacts.

2. Background and literature review

2.1 Residential housing footprint in South Africa

The annual South African population growth rate decreased steadily from 2.8% in 1972 to 1.3% in 2017, and is expected to continue to decrease in the near future.[14] Data on national completed residential buildings demonstrate a rise and fall development curve during the period 2000–2016. Compiled building data from Statistics SA[15] were selected as background data for further scenario modelling. These data include all completed residential buildings reported by South African municipalities (Figure 1). Take note that not all government-subsidised low-cost housing units were included as, in many cases, these units are reported and financed separately. These data were not available and could not be included in this study.

According to Statistics SA, over the period 2000–2016, the average house in South Africa was 114 m^2 and an average of 54 111 houses were constructed annually. On average, 1 040 651 m^2 of houses smaller than 80 m^2, 3 436 302 m^2 of houses bigger than 80 m^2 and 1 665 624 m^2 of flats and townhouses were completed annually.

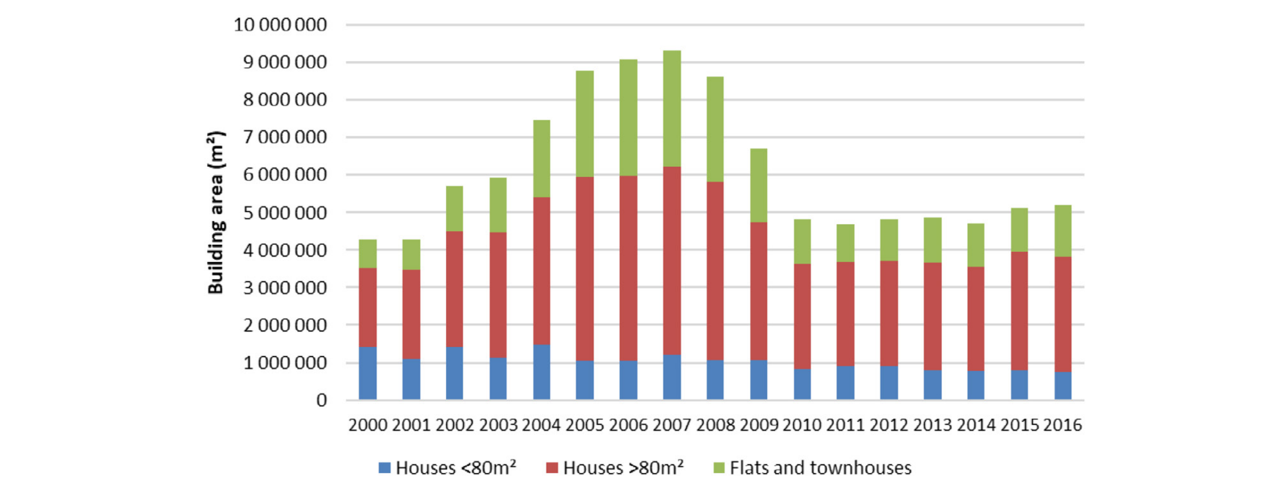


Figure 1: South Africa's completed residential building area.[15]

2.2 Market growth potential of wood-based residential buildings

The extent to which new building systems can increase its market share in a country is dependent on many factors. Cost, resource availability, legislation, building culture, user's perception of a building method and type, skills availability and the perception of the environmental credentials of the building system, can play a role also.

In Germany, the number of new, single family and two-family houses built with wood has tripled in the past 25 years from 6% of the market share in the early 1990s to 18% in 2017.[16] The UK timber frame housing share of all new buildings reached 27.6% in 2015 and was predicted to rise to around 32% by 2018.[17]

South Africa is generally perceived as a country with limited forest resources. However, the South African plantation forestry industry is very productive. Despite having only about 1.8 million ha covered with closed canopy plantations and forests, the annual national industrial roundwood production was 17.5 million m^3 in 2015.[7]

Wood resources for future houses can come from either (2.2.1) a change in forest resource use, (2.2.2) new forest plantings, or (2.2.3) imports. South Africa's industrial roundwood production is used mainly for the production of pulp and board products (51%), sawn lumber (24%) and chip exports to Asia (17%) (Figure 2). In 2016, sawn timber production was 2.3 million m^3 of which 70% was used in construction, mainly for roof truss material.[7] Sawn timber resources are already oversubscribed and mainly used in house construction (roof trusses), therefore it is not likely that any additional timber could be sourced for future house construction from the current sawmilling resource. Table 1 provides estimates of potential future log resources available for timber-based housing components such as sawn timber and board products.

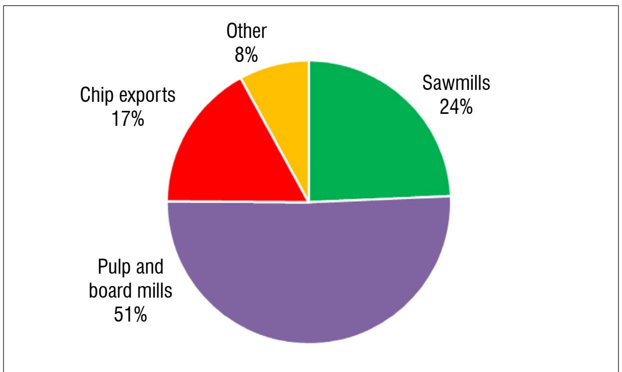


Figure 2: South African industrial roundwood consumption by different sectors.[7]

Chip exports is the most likely available resource which could potentially be used for future housing elements. Export chips are either from eucalypt or wattle trees. Recently developed new technology such as green-gluing of eucalypt timber enable the manufacture of engineered, high-grade structural timber from fast-grown pulp wood resources.[20] Cross-laminated timber (CLT) would offer another product solution for young eucalypt or trees grown for pine pulp or wood chip. According to Guo et al.[21], CLT is a relatively new product. CLT is a European developed product, with an 80% installed market share in Europe, but countries like Canada, the USA and Australia are also showing rapid market growth.[21] Globally, production increased from 25 000 m^3 in 1996 to 600 000 m^3 in 2014 and was estimated to reach 1 000 000 m^3 in 2016. Other housing components that could potentially be manufactured from young pulp tree resources include products such as oriented strand board, and possibly parallel strand lumber.

2.2.1 Wood chips

Over the past 10 years, an average of 3.5 million tons of wood chips was exported from South Africa annually.[7] A slight decrease in chip exports was evident in 2015 with only 2.3 million tons exported. Depending on chip moisture content and using a sawmill volume recovery rate of 40%, an average of 2.3 million tons of chips would result in 2.6 million m^3 sawlogs or 1.04 million m^3 of sawn timber. The national average sawmill volume recovery rate of softwood sawmills in South Africa is 47.4%.[22] Generally, smaller diameter logs such as pulp logs will result in lower volume recovery rates.[23] Some processors of small diameter eucalypt logs into green, unseasoned sawn timber obtain volume recovery rates of 50% but do not include shrinkage loss as they sell products wet off saw.[24] A volume recovery rate of 40% was assumed to be a reasonable estimation of dry sawn timber that will be recoverable. For board products such as oriented strand board or reconstituted lumber such as parallel strand lumber, the volume recovery rates will depend on the process and final product, but could vary between 70% and 80%.[25] In this study, a conservative 40% recovery rate was assumed for a timber product such as CLT and 55% for a combination of sawn timber and board products (i.e. for timber-frame building).

2.2.2 Afforestation

In South Africa, afforestation with fast growing plantation species is also a possibility. Although available land considered suitable for plantations is limited in South Africa, communal areas of 100 000 ha were earmarked by the government for afforestation in the Eastern Cape. There is also about 40 000 ha private farmland available in KwaZulu-Natal for afforestation.[26] If successful, these afforestation plans have the potential to produce an additional annual sustainable supply of 2.07 million m^3 roundwood or about 1 million m^3 of timber within about 24 years of establishment, if destined for sawlogs only. These figures were calculated using a mean annual increment of 14.8 m^3/ha/year for softwood sawlogs[27] and the national average sawmill volume recovery rate of 47.4%[22].

There is also potential for afforestation in areas previously not considered suitable for plantation forestry. Recent research shows the potential of dryland forestry in the Western Cape coastal areas.[28] Von Doderer[19] identified 175 000 ha of potential dryland forest plantation area in the Western Cape. This area could result in a potential annual yield of 738 255 m^3 of timber (based on a mean annual increment of 8.9 m^3/ha/year and a 47.4% volume recovery rate) within about 30 years of establishment. In addition, research by Wessels et al.[29] showed that some species grown on the dry west coast of southern Africa could produce high-value sawn timber. Undoubtedly there are other areas in the country where trees can be grown in dry areas previously not considered suitable for forestry. However, research is required to quantify this potential.

2.2.3 Import

Although it is not always the preferable option from a socio-economic perspective, import of sawn timber is also a possibility. Research from other countries has shown that where shipping is over short land transport distances, the environmental impact of timber imports can be relatively low.[30] In 2016, South Africa only imported 2% of its annual structural timber.[22] Currently, the three major import countries include Brazil, Chile and Zimbabwe. Past trade and most likely future countries for import include Argentina, New Zealand, Germany, Zambia and Mozambique (Stears A 2020, email communication, February 1). Board products such as oriented strand board, the preferred option for timber frame housing wall covering, are currently only available from imported sources. However, research will be required to quantify the environmental impacts of importing these materials into South Africa.

Table 1: Potential future log resources available for timber-based housing components such as sawn timber and board products

Description	Log volume (m^3/year)	Availability (years)†	Data source
Current chip export resources: eucalypt and wattle logs	2 600 000	Immediate	Forestry South Africa[7]
Current pulp, board, and other log resources: eucalypt, wattle and pine	11 850 000	Immediate	Forestry South Africa[7]
Import logs or wood products	–	Immediate	
Afforestation Eastern Cape / KwaZulu-Natal: 140 000 ha	2 070 000	24 (8)	South African Department of Environmental Affairs[18]
Dryland afforestation Western Cape: 175 000 ha	1 557 500	30 (10)	Von Doderer[19]

†Values in parentheses indicate availability for pulpwood rotations and thinnings.

From the data in Table 1, it can be seen that – excluding imports and current pulp, board, and other log resources – there could be an estimated 6.23 million m^3 of log resources available for wood house components in the future. This amount could be processed into between 2.9 and 4.9 million m^3 of products depending on the product type and recovery rates. If timber frame construction requires on average 0.3 m^3 of processed wood-based products per square metre (similar in volume to CLT according to Table 2), it means that between 9.6 and 16.3 million m^2, or between 84 210 and 142 982 houses of 114 m^2, can be built sustainably per annum. This is nearly double the amount of formal annual residential development at present. This clearly indicates the resource potential for an increased wood-based construction market in South Africa.

2.3 Water availability

The South African forest industry has informally standardised on Forestry Stewardship Council (FSC) certification and, already in 2007, 97.8% of all industrial roundwood produced in South Africa was FSC and ISO certified.[31]

Numerous research projects on water use efficiency and stream flow reduction of plantation species in South Africa have been completed.[32,33] Most of the research indicates that water balance is a location- and species-specific issue, and is likely to be a constraint in future, for example, on high water use agricultural food crops.

According to Von Doderer[19], the introduction of selected plantation species may have a positive effect on the water balance, for example when replacing intensive agriculture under irrigation or when establishing short rotation plantations on land that is covered with so-called alien invader plants.

Although it is likely to be a constraining factor in some instances, water availability for afforestation in this study has been considered, as (earmarked dryland) areas not meeting the minimum water requirements were excluded in the land availability assessment by applying the so-called aridity index.[19]

2.4 Building system impacts

In 2014, cement-based building products such as mortar, screed, plaster, concrete and paving accounted for 3.59 million tons carbon dioxide equivalent ($mtCO_2$ eq.) GHG emissions or 29.4% of the emissions of the major building product groups in South Africa. An additional 3.36 $mtCO_2$ eq. (27.6%) of the emissions of the major building product groups was caused by masonry wall elements. More specifically, concrete hollow blocks and clay brick production contributed 60% and 40%, respectively, of masonry GHG impact.[13] Concrete stock blocks require a considerable amount of GHG-generating cement. Clay stock brick production requires energy intensive processes and the major GHG emissions arise from fossil fuel burning to fire brick kilns.[34] For South Africa, no data on embodied energy (EE), global warming potential (GWP) or life-cycle analysis for timber frame or wood-based building systems could be found.

Until now, the world has relied heavily on CO_2-intensive concrete development for building structures.[35] On the other hand, wood-based

Table 2: Research study results on building system embodied energy and global warming potential impacts

Building system	Description	MJ/m^2	CO^2 eq./m^2	Wood (m^3/m^2)	Gross floor area (m^2)	Life cycle (years)	Country	Year	Source
Brick	Low energy	5588	527.17[a]	0.1[b]	231	50	Australia	2017	Thomas and Ding[39]
Timber frame	Low energy	4717	445.00[a]	0.3[b]	231	50	Australia	2017	Thomas and Ding[39]
Reinforced concrete	Conventional	1541	308.2	–	4 floors	50	China	2017	Guo et al.[21]
CLT	Low energy	847	-84	–	4 floors	50	China	2017	Guo et al.[21]
Reinforced concrete	Conventional	3095.2[a]	292	–	4 floors	50	Sweden	2014	Dodoo et al.[40]
CLT	Conventional	1208.4[a]	114	0.27[c]	4 floors	50	Sweden	2014	Dodoo et al.[40]
Brick	Conventional	5400	509.43[a]	–	192	70	Italy	2010	Blengini and Di Carlo[41]
Brick	Conventional	6132	578.49[a]	–	150	30	Spain	2006	Casals[42]
Timber frame	Standard light	2212	208.68[a]	–	94	100	New Zealand	2004	Mithraratne and Vale[43]

[a]These results were obtained by multiplying a factor from the South African primary energy production and greenhouse gas emissions ratio in 2014.[13]
[b]Brick and timber-frame wood volume per square metre was obtained from Pajchrowski et al.[44]; the timber frame house had an all wood-based ground floor, first floor and roof structure.
[c]CLT (cross-laminated timber) showed a slightly lower volume of wood per square metre, most likely due to a reinforced concrete foundation and ground floor.

systems have been gaining market share in some areas[16,17] and have a comparatively lower CO_2 and EE impact. Bribian et al.[36] report that laminated wood absorbs 582 kg CO_2 per m^3 (not incinerated at end of life), while reinforced concrete emits 458 kg CO_2 and steel 12 087 kg CO_2 per m^3. In the same way, Ferguson[37] reported that rough-sawn timber produces 750 MJ/m^3, concrete 4800 MJ/m^3 and steel 266 000 MJ/m^3. Although these differences are quite large, material quantities do not enable any realistic building system comparisons in terms of building area because not all building systems require the same amount and format of materials per unit area.

A building system review by Cuchí et al.[38], performed in Spain, showed overall average GWP emissions of 500 kg CO_2 and EE of 5754 MJ for all building materials considered per building area (m^2). In another life-cycle energy study of brick and timber residential buildings, Thomas and Ding[41] compared 10 standard Australian brick buildings to similar thermal and structural performing timber designs. Three life-cycle stages were analysed, including materials and construction, maintenance and end-of-life over a 50-year life cycle. Compared to Cuchí et al.'s[38] findings, the material and construction phase resulted in similar EE and GWP impacts per square metre (Table 2).

Embodied energy carries an increasing importance in residential life-cycle impacts. Chastas et al.[11] performed an in-depth literature review which considered 90 life-cycle energy analysis case studies of residential buildings over a 50-year life cycle and constructed in the past decade. The results showed an increasing percentage of EE in the transformation from conventional to passive, low-energy and near-zero-energy buildings. EE dominates in low-energy and near-zero-energy buildings with a share of 26–57% and 74–100%, respectively.

Embodied energy and GWP of buildings, particularly in residential dwellings, can be very complex to determine. Studies based on life-cycle analysis methodology and newly developed product category rules[45] for buildings, were selected as the most valid data sources from which to derive the normalised building impacts. Table 2 summarises the best available literature results for building system EE and GWP impact per square metre, compiled from multiple international sources. Mean volume of timber (including wood-based panels) per building system was also included.

It is important to note that operational, maintenance or end-of-life energy were not included and were assumed equal for all systems. End-of-life energy contributed on average less than 2% of total life-cycle energy for both timber frame and brick cladded homes.[39] In the same way, no notable differences between timber frame and brick home maintenance energy over 50 years was evident. In terms of CLT and reinforced concrete demolition energy demands, due to lack of CLT system demolition energy data, it was assumed equal.[21]

3. Methodology

3.1 Log resource analysis

The findings in Section 2 clearly show the timber resource potential for wood-based residential development in South Africa. Based on these findings, we continued with the environmental impact analysis by comparing selected scenarios as seen in Section 3.2. Table 3 presents a summary of the (potential) annual sustainable log resource supply and associated wood-based development coefficients for the South African sawmill industry. The values and volume/building recovery coefficients in Table 3 originate from Figure 1, Tables 1 and 2 and Section 2.2. These data (Tables 2 and 3) formed the basis for the final analysis as depicted in Figures 3 and 4.

A conservative approach was again followed and the 73% of houses (bigger and smaller than 80 m^2) was selected as timber frame systems and 27% (flats and townhouses) as CLT. Compared to timber frame, CLT is a very new building system, therefore an even higher timber frame market could have been assumed. In that case, a higher percentage (i.e. 73%+) of timber frame systems would realise a higher log to product recovery and ultimately equate to more potential homes. It is important to note, in terms of resource availability, that wood chips result in an immediate potential log resource, whereas afforestation could take up to 30 years to supply in log demand.

3.2 Building system impact

Four potential residential building scenarios were selected based on the existing international examples of growth in wood-based development, available building technology and local potential log resources. Table 4 presents these scenarios and input values for South Africa: current (1% residential wood-based buildings), 10%, 20% and 100% residential wood-based buildings. The 10% and 20% growth scenarios were based on market growth values in wood-based buildings experienced in western European countries such as Germany and England over a period of about two decades. The 100% scenario is an extreme value to illustrate the environmental impact of constructing only wood-based residential buildings. Mean building area values for houses smaller than 80 m^2, houses larger than 80 m^2 as well as flats and townhouses are indicated in Figure 1. Most applicable building system impacts (from the grey shaded areas in Table 2), i.e. brick and timber frame building, were assigned to all houses smaller and bigger than 80 m^2 whereas reinforced concrete and CLT system impacts were assigned to the remaining flats and townhouses portion. In each case, the carefully selected building system (i.e. brick, timber frame, reinforced concrete and CLT) with its impacts, either best represented South African building and climate conditions or provided the most conservative analyses in terms of GWP.

Building system impact values here represent EE impacts for all processes required to produce and construct each building, such as foundations, walls, roof, windows, and doors. These impacts include a wide range of materials and processes, for example, the brick and mortar system includes on average 0.1 m^3 of wood per square metre – mostly due to the roof structure.

End of life and maintenance energy was not included and assumed equal for all direct system comparisons. However, wood in buildings can be reused or used for heat or bio-energy, which both have positive climate effects. According to the literature[46,47], treated wood can be landfilled (as municipal solid waste), incinerated (waste to energy) and recycled (cleared from CCA treatment), of which proper incineration

Table 3: Annual sustainable log resource potential and wood-based development coefficients for South Africa for timber frame (TF) and cross-laminated timber (CLT) buildings

Resource	Building type	%	Log volume m^3	TF recovery†	TF m^3/m^2	CLT recovery†	CLT m^3/m^2
Wood chips	Total	100	2 600 000	0.55	0.3	0.4	0.27
	Houses	73	1 898 000	0.55	0.3	–	–
	Flats	27	702 000	–	–	0.4	0.27
Afforestation	Total	100	3 627 500	0.55	0.3	0.4	0.27
	Houses	73	2 648 075	0.55	0.3	–	–
	Flats	27	979 425	–	–	0.4	0.27
Total	Total	100	6 227 500	0.55	0.3	0.4	0.27
	Houses	73	4 546 075	0.55	0.3	–	–
	Flats	27	1 681 425	–	–	0.4	0.27

†TF and CLT recovery are the timber frame and cross-laminated timber construction material recovery coefficients as discussed in Section 2.2.1.

Table 4: Four projected development scenarios with minimum and maximum impact values

Building system		Brick / reinforced concrete			Timber frame / cross-laminated timber		
System EE and GWP impacts		MJ/m²	kg CO_2 eq./m²	m²	MJ/m²	kg CO_2 eq./m²	m²
Current (1% wood)	<80 m² and >80 m²	5400/5588	509/527	4 432 183	2212/4717	208/445	44 770
	Flats and townhouses	1541/3095	308/292	1 648 968	847/1208	-84/114	16 656
10% wood	<80 m² and >80 m²	5400/5588	509/527	4 029 257	2212/4717	208/445	447 695
	Flats and townhouses	1541/3095	308/292	1 499 061	847/1208	-84/114	166 562
20% wood	<80 m² and >80 m²	5400/5588	509/527	3 581 562	2212/4717	208/445	895 391
	Flats and townhouses	1541/3095	308/292	1 332 499	847/1208	-84/114	333 125
100% wood	<80 m² and >80 m²	5400/5588	509/527	0	2212/4717	208/445	4 476 953
	Flats and townhouses	1541/3095	308/292	0	847/1208	-84/114	1 665 624

EE, embodied energy; GWP, global warming potential

technology and methodology according to US EPA does not emit GHGs. However, incineration is not a viable/available option across South Africa yet and was not included in the South African Wood Preservers Association guidelines.

3.3 Sensitivity analyses and limitations

This study focused on log resource availability for an increased wood-based residential building market. Important impacts such as GDP generation and job creation per development scenario were not in the scope of this study. Although elements such as water quality, air pollution and economy are critical in building system comparison, these were not included in the study due to scope and resource constraints. This limits the impact of this research.

Population growth was not included as a direct parameter in the analyses, because the South African population growth is already fairly low, and declining. However, many other factors such as political instability of neighbouring countries and subsequent immigration to South Africa, diseases such as HIV and malaria, and economic growth influence population growth and building rates. Excluding growth also allowed the model to be time independent and therefore easier to apply. All these factors introduce uncertainty in the analysis – which should be taken into consideration by the user.

Simple cradle to gate (material and construction) system boundaries were selected to evaluate likely residential GWP impact comparisons. No local life-cycle assessment building system impacts were available for timber frame and CLT residential houses. Therefore, best available literature for total building system EE and GWP impacts was used to explain the likely system impacts and the expected variation between conventional and low energy technologies.

4. Results and discussion

4.1 Log resource potential

The number of potential wood-based residential houses per annum are shown in Figure 3. This number was computed from data in Table 3 and the average house footprint of 114 m² as per Section 2.1. The total number of potential wood-based houses per annum, considering only current available wood chips (2.2.1) as resource, equates to 39 646 houses (30 523 houses and 9123 flats).

The total number of potential wood-based houses per annum, considering only afforestation (2.2.2) resources, equates to 55 314 houses (42 586 houses and 12 728 flats). That is 1203 more than the average new builds in the past 17 years. Considering both wood chips and afforestation resource potential, close to 95 000 wood-based houses (172% of current supply) could be realised each year.

Import of wood-based materials offers a further resource to consider, if a very large wood-based residential market in South Africa develops. Import allows for a immediate supply option of all required wood-based materials. In this case, economics, quality and environmental considerations will determine the potential housing units, which in theory, is extensive.

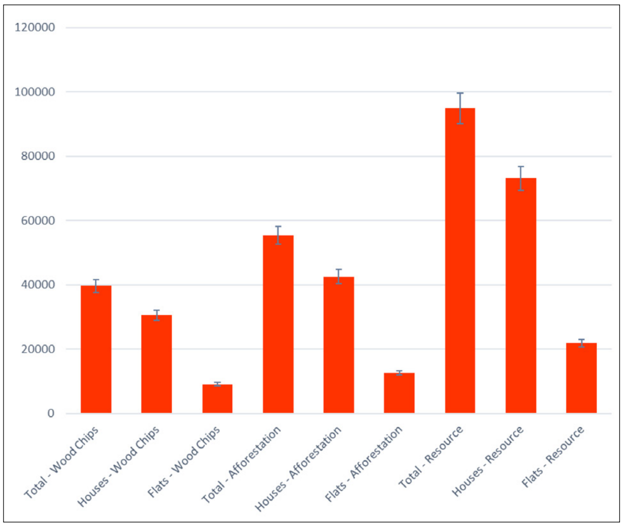


Figure 3: Potential number of 114-m² wood-based houses per annum, with 5% error bars.

4.2 Building system impact

The results discussed here are based on the development scenarios as defined in Table 4 and, more specifically, the maximum impact values (to best represent South African development practices). Impact values comprise a range of annual building system EE and GWP impacts per building system, translated in mean area (m²) as seen in Table 2. Output values in Figure 4a and 4b are minimum and maximum annual residential EE (MJ) and GWP (kg CO_2 eq.) per development scenario. Each impact bar in Figure 4b for EE consists of at least three major categories – energy for construction (0.2–1%), transport (0.1–7%) and material production (92–99.7%) – and, depending on the system, varies considerably in mean EE contribution.[11]

Building material and material production, thus building system choice, represent by far the biggest EE quantity, with transportation of goods being second. These findings support the rationale for an increased wood-based system introduction, as it is the best performer across most environmental impact factors – especially in terms of GWP, compared to building material alternatives such as steel and concrete, with particularly good performance in terms of greenhouse gas emissions.[1-6] It is important to note that, in this study, the total embodied building *system* impact was selected as output values. If comparing purely building *structures* (excluding furnishing, painting, plumbing, insulation, etc.), in relation, an even greater difference would be expected between wood-based and other systems.

Brick and mortar residential homes (<80 m² and >80 m²) comprise the bulk (73%) of the formal residential housing market in South Africa. However, the mean EE and GWP impact from residential homes (<80 m² and >80 m²) contribute 83% of the total annual South African footprint. This proportion is mostly due to the smaller scale and subsequent

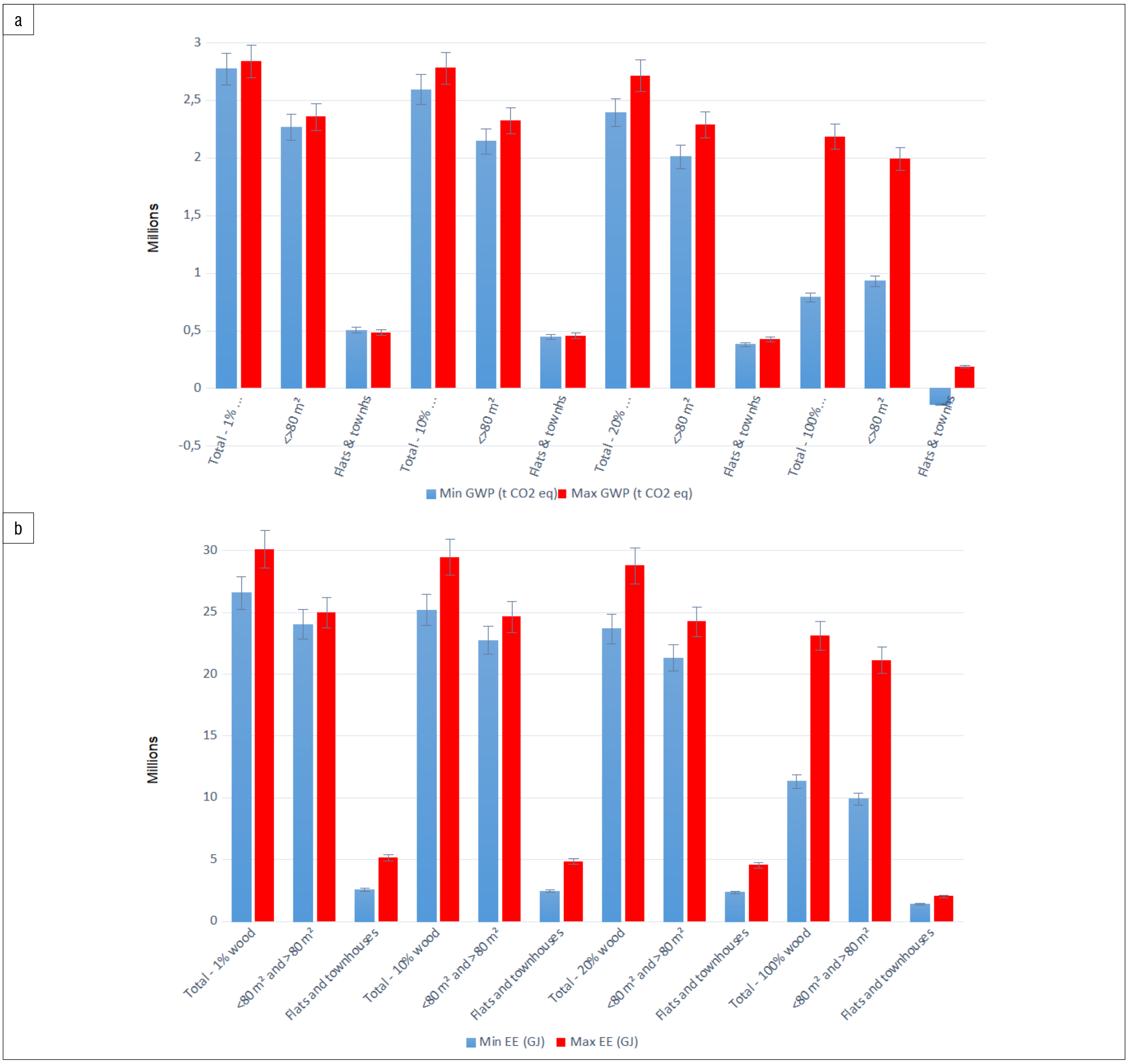


Figure 4: South African minimum and maximum annual residential building (a) global warming potential (GWP) and (b) embodied energy (EE) impacts, with 5% error bars.

inefficiencies as well as the building system difference compared to multi-storey flats and townhouses.

A 10% wood residential market increase will amount to a 2.4% savings in mean annual EE and GWP compared to the current scenario. The 20% market increase will amount to a 4.9% savings in mean annual EE and GWP compared to the current scenario. Finally, an all wood market would amount to a 30% savings in mean annual EE and GWP compared to the current scenario.

South Africa had an estimated total GWP of 590 million ton CO_2 eq. in 2014 – an extraordinary 243 million ton more than in 2006.[7] The major building products amounted to 12.2 million ton CO_2 eq. in 2014 and represented only 2.1% of the total national GHG impact. These major building material impacts include all industries, i.e. roads, commercial, government and industrial sectors. Figure 4 presents minimum and maximum annual residential GWP and EE building impacts, respectively, with 5% error bars to explain likely variability for total development, normal houses and townhouses and flats. It is evident that, if selected minimum (low-energy technology) impact values were considered, much greater GWP and EE savings could be anticipated for wood-based development.

As mentioned earlier, we evaluated only residential EE and not operational energy impacts. Recent studies show that EE for conventional buildings contributes as little as 10% of total building life-cycle energy impacts compared to operational energy impacts.[11] Although not considered in this study, wood-based buildings generally also perform well in terms of operational energy efficiency. Wood is 400 times better than steel and 10 times better than concrete (per volume) in resisting the flow of heat due to its low conductivity and good insulating ability, which can lead to considerable energy savings.[48] However, EE can contribute up to 100% in modern near-zero-energy buildings and, therefore, plays an ever-increasing role in total life-cycle energy.

This modelling study showed that, with market growth of wood-based residential buildings similar to those in Germany and England (i.e. 10–20% of new buildings), there will be a moderate reduction in EE and GWP emissions of less than 5% of total residential building values. If all new residential buildings were wood based, the total reduction in EE and GWP could be a substantial 30%. Even though the potential to reduce EE and GWP in the short to medium term seems to be moderate, it will still be an important contribution to climate change mitigation. The Wedges Theory of Pacala and Socolow[49] showed that it was not possible to reduce GWP to acceptable target levels with a single initiative or technology. Many different industries, sectors and technologies will all have to contribute to combat global warming. If the effects of climate change result in more severe weather events, it could also be that more dramatic changes in consumer behaviour or even government intervention will result in faster and more dramatic changes in building methods and materials, such as the 100% wood-based residential building scenario modelled here.

Only residential housing construction was considered in this study, as it has traditionally been the market segment of choice for wood-based building in other countries. New technologies and products such as CLT also make it possible to build medium-rise buildings from wood-based materials. An 18-storey building was built in Vancouver (Canada) in 2017 from mainly CLT and glulam beams.[50] The commercial and industrial building sectors might therefore in future become adopters of wood-based building.

Due to the limited forest cover in South Africa, the perception is often that significant increases in the market share of wood-based buildings are not possible (at least from local wood resources). This study showed that this perception is not correct. Current resources, available in large volumes such as eucalyptus for chip export, could potentially support considerable growth in wood buildings. In the longer term, however, new afforestation will be required if wood-based buildings become the norm in South Africa. In the short term, supply gaps of wood building components could potentially be alleviated by imports using shipping with short land transport distances. However, research is required to quantify the environmental transport impacts from such imports.

Apart from the environmental advantages of building with wood, wood-based development also has many other positive spin-offs, such as job creation, technological advancement and development of other ecosystem services.[17]

5. Conclusion and recommendations

5.1 Log resource potential

It was shown that with the use of wood resources currently exported as chips, as well as planting trees in areas that have been earmarked for afforestation, it will be possible (in the long term) to sustain a future residential building market where all constructions are wood based. However, in the short term, imports of wood building components might be necessary if rapid growth in the wood-based building market occurs in residential development.

5.2 Building system impact

The basic impact modelling showed that incremental 10% and 20% increases in residential wood-based buildings market share show a moderate environmental benefit, compared to current national GHG impacts of the residential building sector. Further, we demonstrated that (based on maximum impact values), a 100% increase in local wood-based development could result in a substantial 30% GWP saving in residential building impact. However, if selected minimum (low-energy technology) impact values were considered, far greater GWP and EE savings can be expected for wood-based development.

5.3 Further consideration

Contrary to Australia and the USA, South Africa does not have a culture of designing and building with wood. Therefore, further research that includes other impacts such as social and economic comparisons with regard to an increase in wood-based building, is recommended. Finally, the interaction of operational energy and embodied energy of wood-based buildings compared to conventional buildings in South Africa should be investigated – a life-cycle analysis approach is recommended.

Acknowledgements

We gratefully acknowledge the Hans Merensky Foundation for providing study sponsorship that enabled this research to be undertaken.

Authors' contributions

P.L.C. was responsible for the article design, technical analyses and write-up. C.B.W. supervised the research and assisted with the write-up and language editing.

References

1. Petersen AK, Solberg B. Environmental and economic impacts of substitution between wood products and alternative materials: A review of micro-level analyses from Norway and Sweden. Forest Policy Econ. 2005;7(3):249–259. https://doi.org/10.1016/S1389-9341(03)00063-7
2. Werner F, Richter K. Wooden building products in comparative LCA: A literature review. Int J Life Cycle Ass. 2007;12(7):470–479. https://doi.org/10.1007/s11367-007-0317-5
3. Upton B, Miner R, Spinney M, Heath LS. The greenhouse gas and energy impacts of using wood instead of alternatives in residential construction in the United States. Biomass Bioenerg. 2008;32(1):1–10. https://doi.org/10.1016/j.biombioe.2007.07.001
4. Sathre R, O'Connor J. Meta-analysis of greenhouse gas displacement factors of wood product substitution. Environ Sci Policy. 2010;13(2):104–114. https://doi.org/10.1016/j.envsci.2009.12.005
5. Wang L, Toppinen A, Juslin H. Use of wood in green building: A study of expert perspectives from the UK. J Clean Prod. 2014;65(1):350–361. https://doi.org/10.1016/j.jclepro.2013.08.023
6. Crafford PL, Blumentritt M, Wessels CB. The potential of South African timber products to reduce the environmental impact of buildings. S Afr J Sci. 2017;113(9/10), Art. #2016-0354. https://doi.org/10.17159/sajs.2017/20160354
7. Forestry South Africa (FSA). South African forestry and forest products industry [webpage on the Internet]. c2015 [cited 2020 May 15]. Available from: http://www.forestry.co.za/statistical-data/
8. Palmer S. Sustainable homes: Timber frame housing [homepage on the Internet]. c2015 [cited 2020 May 15]. Available from: www.sustainablehomes.co.uk
9. Beradi U. A cross-country comparison of the building energy consumptions and their trends. Resour Conserv Recylc. 2017;123(1):230–241. https://doi.org/10.1016/j.resconrec.2016.03.014
10. Milford R. Greenhouse gas emission baseline and reduction potentials from buildings in South Africa [document on the Internet]. No date [cited 2020 May 15]. Available from: http://www.cidb.org.za/publications/Documents/Greenhouse%20Gas%20Emission%20Baselines%20and%20%20Reduction%20Potentials%20from%20Buildings%20in%20South%20Africa.pdf
11. Chastas P, Theodosiou T, Bikas D. Embodied energy in residential buildings – towards the nearly zero energy building: A literature review. Build Environ. 2016;105(1):267–282. http://dx.doi.org/10.1016/j.buildenv.2016.05.040
12. Sartori I, Hestnes AG. Energy use in the life cycle of conventional and low-energy buildings: A review article. Energy Build. 2007;39(3):249–257. https://doi.org/10.1016/j.enbuild.2006.07.001
13. Ampofo-Anti NL, Dumani N, Van Wyk L. Potential for reducing greenhouse gas emissions in the South African construction sector. DBSA Green Fund Research and Policy Development to advance a green economy in South Africa: Research report [document on the Internet]. c2015 [cited 2017 Oct 12]. Available from: www.sagreenfund.org.za/research
14. UN Department of Economic and Social Affairs. Population Division [webpage on the Internet]. No date [cited 2020 May 15]. Available from: https://esa.un.org/unpd/wpp/
15. Statistics SA. P5041.1 – Selected building statistics of the private sector as reported by local government institutions. [homepage on the Internet]. No date [cited 2018 Nov 17]. Available from: http://www.statssa.gov.za/?page_id=1854&PPN=P5041.1
16. Alfter D, Lüdtke J, Maack C. Mitigating climate change. Creating value. Utilising resources efficiently. Charter for Wood 2.0 [document on the Internet]. No date [cited 2020 May 15]. Available from: http://www.timber.co.za/uploads/charter_for_wood_2.0_web.pdf
17. Adamson J, Browne R. Spotlight UK forestry market. Savills [document on the Internet]. No date [cited 2020 May 15]. Available from: http://www.savills.co.uk/research_articles/141557/215817-0

18. South African Department of Environmental Affairs (DEA). Working for forest [homepage on the Internet]. No date [cited 2017 Aug 15]. Available from: https://www.environment.gov.za/projectsprogrammes/workingforforest

19. Von Doderer CCC. Determining sustainable lignocellulosic bioenergy systems in the Cape Winelands District Municipality, South Africa [unpublished doctoral dissertation]. Stellenbosch: Stellenbosch University; 2012. http://scholar.sun.ac.za/handle/10019.1/71838

20. Crafford PL, Wessels CB. The potential of young, green finger-jointed *Eucalyptus grandis* lumber for roof truss manufacturing. South Forests. 2016;78(1):61–71. http://dx.doi.org/10.2989/20702620.2015.1108618

21. Guo H, Liu Y, Meng Y, Huang H, Sun C, Shao Y. A comparison of the energy saving and carbon reduction performance between reinforced concrete and cross-laminated timber structures in residential buildings in the severe cold region of China. Sustainability. 2017(9):1426. https://doi.org/10.3390/su9081426

22. Crickmay and Associates. South African Lumber Index. Unpublished technical report; August 2013.

23. Wessels CB. Cant sawing log positioning optimization: A simulation study. Forest Prod J. 2009;59(4):17–22.

24. De Kock SL. The economic feasibility of producing green glued eucalyptus CLT in South Africa [BSc report]. Stellenbosch: Stellenbosch University; 2016.

25. Puettmann M, Kaestner D, Taylor A. Corrim report – Module E life cycle assessment of oriented strandboard (OSB) production [document on the Internet]. No date [cited 2020 May 15]. Available from: https://corrim.org/wp-content/uploads/Module-E-OSB-Final-w-survey.pdf

26. Chamberlain D, Essop H, Hougaard C, Malherbe S, Walker R. Part 1: The contribution, costs and development opportunities of the forestry, timber, pulp and paper industries in South Africa [document on the Internet]. No date [cited 2020 May 15]. Available from: https://www.environment.gov.za/sites/default/files/reports/part1contributionscost_forestrytimberpulpindustries.pdf

27. Kotze H, Kassier H, Fletcher Y, Morley T. Growth modelling and yield tables. In: Bredenkamp S, Upfold B, editors. South African forestry handbook. Pietermaritzburg: South African Institute for Forestry; 2012. p. 175–210.

28. Du Toit B, Malherbe GF, Kunneke A, Seifert T, Wessels CB. Survival and long-term growth of eucalyptus in semi-arid sites in a Mediterranean climate, South Africa. South Forests. 2017;79(3):235–249. https://doi.org/10.2989/20702620.2016.1254914

29. Wessels CB, Crafford PL, Du Toit B, Seifert T. Variation in physical and mechanical properties from three drought tolerant *Eucalyptus* species grown on the dry west coast of southern Africa. Eur J Wood Prod. 2016;74:563–575. https://doi.org/10.1007/s00107-016-1016-3

30. Liao C, Tseng P, Lu C. Comparing carbon dioxide emissions of trucking and intermodal container transport in Taiwan. Transp Res Rec. Part D 14. 2009;14(7):493–496. https://doi.org/10.1016/j.trd.2009.05.002

31. Government Communications (GCIS). Forestry. In: Tibane E, Vermeulen A, editors. South Africa Yearbook 2013/14. Pretoria: GCIS; 2013. p. 49–56. Available from: http://www.gcis.gov.za/sites/www.gcis.gov.za/files/docs/resourcecentre/yearbook/2013-4Agriculture.pdf

32. Dvorak WS. Water use in plantations of eucalypts and pines: A discussion paper from a tree breeding perspective. Int Forest Rev. 2012;14(1):110–119. https://doi.org/10.1505/146554812799973118

33. Albaugh J, Dye P, King J. Eucalyptus and water use in South Africa. J For Res. 2013; Art. #852540. http://dx.doi.org/10.1155/2013/852540

34. United States Environmental Protection Agency. Background document for life cycle greenhouse gas emission factors for clay brick reuse and concrete recycling [document on the Internet]. c2003 [cited 2020 May 15]. Available from: https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P10088LL.txt

35. Miller SA, John VM, Pacca SA, Horvath A. Carbon dioxide reduction potential in the global cement industry by 2050. Cem Concr Res. 2018;114:115–124. https://doi.org/10.1016/j.cemconres.2017.08.026

36. Bribian IZ, Capilla AV, Uson AA. Life cycle assessment of building materials: Comparative analysis of energy and environmental impacts and evaluation of the eco-efficiency improvement potential. Build Environ. 2011;46(5):1133–1140. https://doi.org/10.1016/j.buildenv.2010.12.002

37. Ferguson I. Environmental properties of timber. Research paper commissioned by the FWPRDC [document on the Internet]. c1996 [cited 2020 May 15]. Available from: https://drive.google.com/file/d/1C2OW729e2dBQ6aUx678GdxgHAjHNwUBi/view?usp=sharing

38. Cuchí A, Wadel G, Lopez F, Sagrera A. Guía de la eficiencia energética para losadministradores de fincas [Energy efficiency guide for property administrators]. Barcelona: Natural Gas Foundation; 2007. Available from: http://fabregaspere.com/blog2/wp-content/uploads/2014/09/2007.02.07.Guia-de-la-eficiencia-energetica-para-Administradores-de-Fincas.pdf

39. Thomas D, Ding G. Comparing the performance of brick and timber in residential buildings – The case of Australia. Energy and Buildings. 2018;159(1):136–147. https://doi.org/10.1016/j.enbuild.2017.10.094

40. Dodoo A, Gustavsson L, Sathre R. Life cycle carbon implications of conventional and low-energy multi-storey timber building systems. Energy Build. 2014;82(1):194–210. http://dx.doi.org/10.1016/j.enbuild.2014.06.034

41. Blengini GA, Di Carlo T. The changing role of life cycle phases, subsystems and materials in the LCA of low energy buildings. Energy Build. 2010;42(6):869–880. https://doi.org/10.1016/j.enbuild.2009.12.009

42. Casals XG. Analysis of building energy regulation and certification in Europe: Their role, limitations and differences. Energy Build. 2006;38(5):381–392. https://doi.org/10.1016/j.enbuild.2005.05.004

43. Mithraratne N, Vale B. Life cycle analysis model for New Zealand houses. Build Environ. 2004;39(4):483–492. https://doi.org/10.1016/j.buildenv.2003.09.008

44. Pajchrowski G, Noskowiak A, Lewandowska A, Strykowski W. Wood as a building material in the light of environmental assessment of full life cycle of four buildings. Constr Build Mater. 2014;52(1):428–436. https://doi.org/10.1016/j.conbuildmat.2013.11.066

45. Tyrens A. Product category rules according to ISO 14025:2006 product group 2014:02, in UN CPC 531 buildings [webpage on the Internet]. c2014 [cited 2020 May 15]. Available from: https://www.iso.org/standards.html

46. Jambeck J, Weitz K, Solo-Gabriele H, Townsend T, Thorneloe S. CCA-treated wood disposed in landfills and life-cycle trade-offs with waste-to-energy and MSW landfill disposal. J Waste Manag. 2007;*27*(8):21–28. https://doi.org/10.1016/j.wasman.2007.02.011

47. Ribeiro AB, Mateus EP, Ottosen LM, Bech-Nielsen G. Electrodialytic removal of Cu, Cr, and As from chromated copper arsenate-treated timber waste. Environ Sci. 2000;34(5):784–788. https://doi.org/10.1021/es990442e

48. Stalnaker J, Harris EC. Structural design in wood. 2nd ed. New York: Springer Science and Business Media; 2013.

49. Pacala S, Socolow R. Stabilization wedges: Solving the climate problem for the next 50 years with current technologies. Science. 2004;305(5686):968–972. https://doi.org/10.1126/science.1100103

50. Connolly T, Loss C, Iqbal A, Tannert T. Feasibility study of mass-timber cores for the UBC Brock Commons Tallwood Building. Paper presented at: 40th IABSE Symposium: Tomorrow's Megastructures; 2018 September 19–21; Nantes, France.

Check for updates

Sequence-related amplified polymorphism markers – a tool for litchi breeders in Africa

AUTHORS:
Elliosha Hajari[1]
Dzunisani Nonyane[1]
Regina Cronje[1]

AFFILIATION:
[1]Agricultural Research Council – Tropical and Subtropical Crops, Nelspruit, South Africa

CORRESPONDENCE TO:
Elliosha Hajari

EMAIL:
HajariE@arc.agric.za

DATES:
Received: 08 Oct. 2019
Revised: 13 Mar. 2020
Accepted: 18 Mar. 2020
Published: 29 July 2020

HOW TO CITE:
Hajari E, Nonyane D, Cronje R. Sequence-related amplified polymorphism markers – a tool for litchi breeders in Africa. S Afr J Sci. 2020;116(7/8), Art. #7461, 6 pages. https://doi.org/10.17159/sajs.2020/7461

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:
SRAP markers, *Litchi chinensis*, genetic diversity, molecular characterisation, plant breeding

FUNDING:
Agricultural Research Council (South Africa), South African Litchi Growers' Association

Litchi represents an economically important crop in South Africa – however, the local industry is based on only five cultivars. In order to expand the gene pool and to extend the harvest season, new cultivars have been imported. Currently, cultivars are identified based on morphological characteristics, but these are not always reliable. Molecular markers provide a tool to supplement morphological characterisation, particularly in cases in which confusion exists. The present study reports on the application of sequence-related amplified polymorphism (SRAP) markers in litchi for assessment of genetic relationships and molecular characterisation. The results provide evidence for separation of cultivars based on maturation period and fruit characteristics. The SRAP markers provide a tool for molecular characterisation that can be readily used by researchers with limited budgets, which is common in many developing countries.

Significance:

- We report on the application of SRAP markers as a tool for litchi breeders in resource constrained countries.
- The tested molecular markers allowed for genotyping (molecular characterisation) of litchi cultivars and selections.
- The markers also revealed relationships between genetic and morphological (phenotypic) characteristics.

Introduction

Litchi (*Litchi chinensis* Sonn.) belongs to the Sapindaceae family and is a commercially important fruit tree in tropical and subtropical regions. It is indigenous to southern China, northern Vietnam and the Malay peninsula but is also cultivated in other countries including India, Taiwan, Thailand, Madagascar and South Africa.[1,2] There is a long history of litchi cultivation globally, and accordingly, a large number of selections have been developed over time. In China, more than 500 litchi accessions are stored at the National Litchi Germplasm Gene Bank at the Institute of Fruit Tree Research in the Guangdong Academy of Agricultural Sciences. Litchi was introduced into South Africa in 1876 and cultivation currently occurs in subtropical, frost-free regions in Limpopo, Mpumalanga and KwaZulu-Natal Provinces.[3] During the 2017/2018 season, 5545 tons of litchi were produced in South Africa, of which 65% was exported, 24% sold to the local market and 11% processed into products.[3] Hence, litchi cultivation makes an important contribution to the gross domestic product of the country and contributes towards job creation in the agricultural and processing sectors.

Despite the availability of a variety of litchi cultivars, the South African industry is based on just five: 'Mauritius' (89.8% of plantings) and 'McLeans Red' (6.4%) with 'Wai Chee', 'Fay Zee Siu' and 'Third Month Red' making up the rest (3.8%). As a consequence, the local litchi industry is characterised by a short production season with a limited range of cultivars that have a narrow genetic base.[4,5] In recent years, there has been increasing interest in expanding the gene pool of cultivars available in South Africa in order to extend the harvesting season (beyond the current window) to exploit the export market and to increase the genetic diversity available for breeding.[6] The Agricultural Research Council'sTropical and Subtropical Crops (ARC-TSC) in Nelspruit has an active breeding programme aimed at developing new cultivars for the litchi industry. In order to expand the gene pool, 30 cultivars from other litchi-producing countries have been imported and are currently being evaluated under South African conditions. These cultivars (as well as others present in the gene bank) have to be accurately identified and characterised in order to verify the identity of cultivars and to ensure that inadvertent mix-ups are avoided, which can have devastating effects on the livelihoods of growers and costly legal implications (in terms of Intellectual Property Rights).

Currently, identification of litchi is based on morphological traits (including vegetative, floral and fruit characteristics). The International Union for the Protection of New Varieties of Plants (UPOV) outlines a set of morphological descriptors for identification of litchi cultivars based on the distinctness, uniformity and stability of physical traits.[7] However, there are limitations to the use of these descriptors; for example, they can be inaccurate during the juvenile phase and can be influenced by environmental and other factors, which leads to misidentifications.[8] This situation is further complicated by confusion surrounding the naming of litchi cultivars which occurred as cultivar names were translated from Cantonese and Mandarin during distribution to other countries.[9,10] Furthermore, it is not uncommon for the same cultivar to have different names in various Chinese dialects.[9] This confusion has led to homonymies (cultivars having the same name but with different genetic profiles) and synonymies (cultivars having different names but with identical genetic profiles) which are prevalent in litchi cultivation programmes globally. Dissemination of cultivars to other countries has amplified this problem.[9] This confusion in litchi nomenclature is exacerbated by misidentification of seedlings and the observation that the same cultivar grown in different climates may produce fruit that appear morphologically different from that expected.[11]

An example of the confusion surrounding litchi nomenclature is exemplified by the Chinese cultivar 'Sanyuehong' which is known by its English translation of 'Third Month Red' in South Africa (the fruit of this cultivar mature in the third month of the lunar calendar in China, hence its name). This same cultivar is referred to by a third name 'Sum Yee Hong' in Australia.[5] Other cultivars have slightly different spellings in different countries, as is the case for 'Feizixiao'/'Fay Zee Siu' and 'Nuomici'/'No Mai Chee' while others are pronounced similarly but spelled differently such as 'Huaizhi' and 'Wai Chee'.[5]

Considering the above, there is a need for a supportive tool that can overcome some of the limitations imposed by morphological identification. Molecular markers can be used for this purpose and to unravel the genetic relationships between cultivars. The advantages of molecular characterisation are that markers are stable, thereby allowing plants to be sampled at any developmental stage and they are not influenced by environmental, pleiotropic or epistatic effects.[12] A range of molecular markers have been applied to litchi, each with advantages and limitations.[13] Applications relating to the use of markers for genetic diversity assessment and verification of cultivar identity have dominated the literature, which is not surprising considering the confusion in litchi nomenclature.[14] These include studies on isozymes[10,15], random amplified polymorphic DNA (RAPD)[16-20], amplified fragment length polymorphism (AFLP)[21,22], simple sequence repeat (SSR)[23-25], expressed sequence tagged SSRs (EST-SSRs)[26,27], inter simple sequence repeat (ISSR)[9] and single nucleotide polymorphism (SNP)[28] markers.

However, none of the markers mentioned above specifically targets coding regions of the genome.[29] One specific class of markers, the sequence-related amplified polymorphism (SRAP) markers, are novel because they target open reading frames. Furthermore, the unique construction of the primers, combined with the optimised polymerase chain reaction (PCR) running conditions, ensures efficient reproducibility.[29] The SRAP markers have been used for a range of purposes, for example, assessment of genetic relationships in apple and related species in which separation of genotypes was found to be based largely on geographic distribution.[30] Similarly, SRAP markers were used to analyse the genetic diversity of pomegranate[31] and passion fruit[32] with the former study reporting a low degree of genetic variation amongst genotypes while the latter reported a high degree of genetic variability. Other researchers have used a different approach, by linking SRAP markers to specific traits such as fruit shape in cucumber[33] and the colour around the stone of peach[34].

The aim of the present study was to investigate the suitability of SRAP markers for molecular characterisation and investigation of the genetic relationships between litchi cultivars maintained at the ARC-TSC (cultivars maintained in the gene bank and newly imported ones). An additional advantage of using SRAP markers is that their use requires simple laboratory equipment (the most complicated being a PCR machine), which makes this technology suitable for resource constrained researchers on the African continent and in other countries.

Materials and methods

Source material

Leaf samples (mature, hardened leaves from the most recent flush that were still soft at the time of collection) of 52 litchi cultivars (including all imported cultivars as well as those available in the gene bank) were collected at the ARC in Nelspruit, South Africa (-25.4884, 30.4028). Table 1 provides a summary of all the cultivars tested as well as their country of origin. Leaf samples were brought to the laboratory and immediately processed for analysis.

DNA extraction

Genomic DNA was extracted from leaf material using the Macherey-Nagel Kit (Düren, Germany) as per the manufacturer's instructions. Cell lysis was performed using a Precellys homogeniser with zinc zirconium beads (Bertin Technologies, France).

Polymerase chain reaction

All SRAP primers were first screened to ensure that clear peaks and non-ambiguous scoring data were obtained. The SRAP primers were chosen according to literature[29], i.e. 16 combinations of ME1 – EM4 (details of the primers tested are provided in Table 2). Genomic DNA was amplified in 15 μL reactions containing 25 ng DNA, TaKaRa EmeraldAmp Max HS PCR master mix (TaKaRa, Shiga, Japan), and 0.2 μM forward and 0.2 μM reverse primers using a G-Storm thermocycler. The PCR amplification conditions were as follows: hot start denaturation at 98 °C for 1 min, followed by five cycles of 1 min of additional denaturation at 94 °C, 1 min of annealing at 35 °C and 1 min of elongation at 72 °C. The initial amplification was followed by 35 cycles of denaturation at 94 °C for 30 s, the annealing temperature was increased to 50 °C for 30 s and elongation occurred for 1 min at 72 °C followed by a final elongation for 5 min at 72 °C. The PCR products were visualised via capillary electrophoresis (Qiaxcel Advanced, Qiagen, Hilden, Germany) using OM500 running conditions. All reactions were repeated to verify data.

Table 1: Litchi cultivars used in the current study and country of origin

Cultivar	Country of origin	Cultivar	Country of origin
Baitangying	China	Kwai Mai Pink	China
Bidum	India	Kwai May Red	China
Brewster	USA, but originally from China	Late Large Red	India
Casino	South African selection	Late Seedless	India
Chakrapat	Thailand	Madras19	India
Chompogo	China	Maguili	China
Early Delight	South Africa	Maskells	Unknown
Early Large Red	India	Mauritius 1-5	Imported from Mauritius, but originally from China
Emmerson	South African selection	McLeans Red	India
Erdon Lee	China	Miller Kent	South African selection
Fay Zee Siu	China	Mooragusha	India
Floridian	USA/China	Muzaffarpur	India
Garnet	Israel	Nuomici	China
Goose Egg	China	Rose Scented	India
Groff	Hawaii	Saharanpur	India
Haak Yip	China	Sahkeng	Taiwan
Hazipur	India	Salathiel	Australia
Hung Long	Vietnam	Shang Shu Huai	China
Jean Hang	Thailand	Shuijingqui	China
Johnstones Favourite	South African selection	Souey Tung	China
Kafri	India	Third Month Red	China
Kaimana	Hawaii	TS-LIT-049	South Africa
Kim Cheng Meesa	Thailand	Wai Chee	China
Kontand	South African selection	Yellow Red	China

Table 2: Primer sequences tested

Primer	Sequence
ME1	TGAGTCCAAACCGGATA
ME2	TGAGTCCAAACCGGAGC
ME3	TGAGTCCAAACCGGAAT
ME4	TGAGTCCAAACCGGACC
EM1	GACTGCGTACGAATTAAT
EM2	GACTGCGTACGAATTTGC
EM3	GACTGCGTACGAATTGAC
EM4	GACTGCGTACGAATTTGA

Data analysis

The sizes of the PCR products were determined using Screengel Software (Qiagen, Hilden, Germany) and the data were used to compile a database. A genetic distance matrix was created using GenAlEx 6.3, which was subjected to UPGMA (unweighted pair group method with arithmetic mean) cluster analysis. The genetic distance matrix was generated in a pairwise manner for each cultivar using the method of Huff et al.[35] The information generated from the distance matrix provided a calculation of the differences between each pair of compared genotypes. The cluster analysis was validated through calculation of the cophenetic correlation coefficient (CCC) with branch lengths denoting genetic distances between samples. The final tree was constructed using MEGA (version 5.05).

Results and discussion

The present study reports on the use of SRAP markers to characterise litchi germplasm at the ARC-TSC, as well as newly imported cultivars that have not previously been investigated and promising selections arising from the local breeding programme. Examination of published literature indicates that there are only two reports on the use of SRAP markers in litchi.[6,36] The first study considered the development of a core collection for breeding as the primary goal and the second study used different cultivars from the present investigation (cultivars of relevance to the Chinese industry). Furthermore, the current analysis included recently imported cultivars such as 'Baitangying', 'Jean Hang', 'Chompogo', 'Shujinqui', 'Maguili', 'Erdon Lee', 'Goose Egg', 'Yellow Red', 'Kwai May Red', 'Kim Cheng Meesa' and 'Hung Long', which have not been previously described in South Africa.

A summary of the band sizes attained with each of the primer pairs is provided in Table 3. All primer pairs were polymorphic to differing extents, as similarly reported by Bhatt et al.[37] using SRAP markers in cumin. A total of 3736 bands were scored across all 16 primers. This figure is similar to the 3939 bands that were scored by Zhou et al.[36] using 32 litchi genotypes and nine primer pairs. The total number of bands scored for each primer pair ranged from 77 with ME3/EM4 to 383 with ME2/EM1. Table 3 also highlights the most common band size present across all cultivars for each primer pair. This ranged from 252 bp with ME3/EM1 (present in 89% of the tested samples) to 1518 bp with ME3/EM2 (present in 60% of the tested samples).

The polymorphism information content (PIC) was used to indicate the ability of polymorphic loci to reveal genetic diversity amongst genotypes and is also presented in Table 3.[31] The PIC was lowest for ME1/EM2 (0.14) – this primer pair also generated fewer bands (94) than most of the other tested primers (>200). The highest PIC value of 0.30 was obtained with ME4/EM3. The average PIC in the present study (0.23) is similar to that reported by other researchers working on SRAP markers in various crops, e.g. 0.34 in cumin[37], 0.28 in coffee[38], 0.28 in pomegranate[31] and 0.23 in passion fruit[32]. According to Xie et al.[39], PIC values may be classified as revealing high (PIC>0.5), medium (0.5>PIC>0.25) or low (PIC<0.25) levels of polymorphism. In the present study, five primer pairs displayed a PIC value of >0.25 (indicative of medium levels of polymorphism between cultivars) while 11 had a PIC value of <0.25 (indicative of low levels of polymorphism between cultivars, Table 3). The low levels of polymorphism revealed by 11 of the tested primer pairs could be attributed to the high degree of genetic uniformity present among certain genotypes, for example, the 'Mauritius' selections and 'Madras' group (discussed below). Indeed, it is acknowledged that the commercially available litchi cultivars comprise a narrow genetic base[23] as a result of selection for desirable fruit traits[5]. Furthermore, other commercially important tropical and subtropical crops (including citrus[40] and mango[41]) are also characterised by having limited genetic variation.

Table 3: Summary of band characteristics per primer pair and polymorphism information content (PIC)

Primer pair	Most common band size (bp)	%Presence in cultivars	Total number of bands	PIC
ME1/EM1	375	42	258	0.22
ME1/EM2	491	52	94	0.14
ME1/EM3	253	94	274	0.22
ME1/EM4	415	54	293	0.20
ME2/EM1	1271	64	383	0.29
ME2/EM2	339	58	262	0.26
ME2/EM3	301	73	212	0.23
ME2/EM4	440	100	280	0.22
ME3/EM1	252	89	272	0.21
ME3/EM2	1518	60	180	0.22
ME3/EM3	1365	96	230	0.25
ME3/EM4	1269	85	77	0.18
ME4/EM1	1106	69	226	0.23
ME4/EM2	1213	71	293	0.27
ME4/EM3	281	65	266	0.30
ME4/EM4	381	100	136	0.17

The UPGMA dendrogram generated produced a CCC value of 0.945 with the Dice method indicating a good fit between the data and analysis method (Figure 1). The tested markers allowed for the separation of cultivars based on morphological characteristics and breeding history. For example, the five 'Mauritius' selections grouped together due to their high degree of genetic similarity. Similarly, 'Early Large Red' and 'Late Large Red' clustered together with a high degree of genetic similarity. 'Brewster' and 'Floridian' also clustered together and this is related to the fact that 'Floridian' is known to be the offspring of 'Brewster'. Further, these two cultivars display very similar morphological characteristics, which provides additional support for their close genetic relationship.[9] These and other similar observations provided an indication that the SRAP markers were functioning effectively as the relationships observed could be explained in terms of known characteristics.

The cluster analysis revealed a number of groupings based on physical characteristics (highlighted in Figure 1). The first was the 'Mauritius' group comprising a cluster of genotypes that produce fruit resembling the cultivar 'Mauritius', i.e. large, ellipse-shaped fruit with a pointed to round tip, characterised by dull red skin and medium-sized seed. This is true for all cultivars within this group apart from 'Kaimana' which produces fruit that are different to 'Mauritius' (i.e. round fruit with flat sides, a round tip and bright red skin). Similarly, Degani et al.[9] also reported the occurrence of a grouping of 'Mauritius'-like fruit using a different marker type, i.e. ISSR markers. In the present study, the occurrence of 'Kaimana' within the 'Mauritius' group might be explained by consideration of its parentage. 'Kaimana' is known to be an open pollinated seedling of 'Haak Yip' and it has been suggested that 'Mauritius' might be its other parent, thereby indicating a genetic relationship between 'Kaimana' and 'Mauritius'.[42] 'Early Delight' was also found to occur in close association with the 'Mauritius' group as the former was developed as the offspring from open pollinated 'Mauritius'[6] and 'TS-LIT-049' is also a selection of 'Mauritius'.

As mentioned above, there is uncertainty surrounding the naming of litchi, which has been exacerbated by exporting cultivars to different countries and subsequent translations of cultivar names.[9] Molecular markers can contribute towards dispelling some of this confusion.

In this regard, 'Muzaffarpur' and 'Late Large Red' have previously been reported to be the same cultivar.[43,44] However, Degani et al.[9] demonstrated that these were different cultivars using ISSR markers. The results presented in the current study corroborate the findings of Degani et al.[9] as differences were found between 'Muzaffarpur' and 'Late Large Red' (Figure 1). Furthermore, other cultivars often considered as synonyms (cultivars with different names but identical genetic profiles), e.g. 'Wai Chee'/'Salathiel' and 'Haak Yip'/'Souey Tung' were found to be different from each other in the present study, as evidenced from their positions on the dendrogram.[23] Further evidence to support this assertion is provided by the observation that 'Salathiel' shared only 27.3% genetic similarity with 'Wai Chee' and 'Haak Yip and 'Souey Tung' shared 31.1% genetic similarity (as per the genetic distance matrix). A similar finding was reported by Viruel and Hormaza[23] who used SSR markers. The above findings provide evidence that molecular markers can assist in efforts towards unravelling the confusion surrounding the naming of litchi cultivars. However, this needs to be a global effort involving all litchi-producing countries to ensure that standardised criteria are developed and applied consistently.

Examination of the cluster analysis revealed a second grouping of cultivars of Indian origin (also known as the 'Madras' group) which has been reported in other studies.[9,25] Cultivars within this group were characterised by relatively short branch lengths (0.08–0.10; Figure 1) indicating close genetic similarity.[45] These cultivars typically produce colourful, red fruit but under South African conditions, the fruit quality is often poor with soft, watery flesh and large seeds. In addition, alternate bearing can be a problem.[46] However, these cultivars could prove to be useful breeding parents to transfer desirable attributes such as fruit colour to potential offspring.

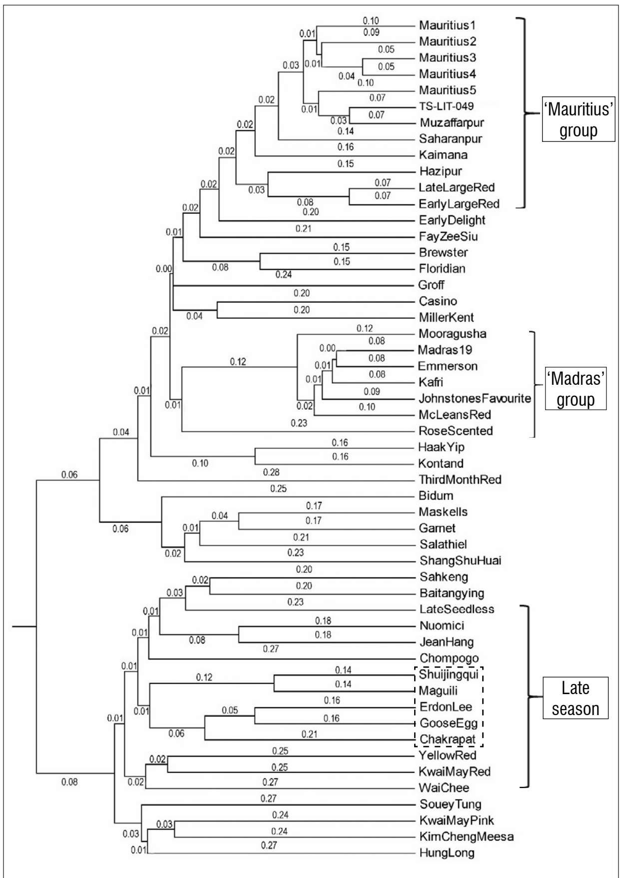


Figure 1: UPGMA (unweighted pair group method with arithmetic mean) dendrogram illustrating relationships between litchi cultivars and selections using sequence-related amplified polymorphism (SRAP) markers (Dice, CCC=0.945). Cultivars within the box with a dashed line produce typically large fruit with dark purple skin.

Furthermore, a common problem in litchi production is cracking of the pericarp of fruit, which results in major losses to growers. Cracked fruit cannot be sold as fresh fruit and can only be used for juicing (provided that fungal contamination has not set in) which results in less income than does the sale of fruit. A number of factors contribute to fruit cracking, with cultivar-specific differences being one, thus indicating a genetic link. For example, 'Nuomici' is known to be prone to cracking, with up to 80% of fruit lost to cracking disorders.[47] Considering this, cultivars within the 'Madras' group could also serve as a potential source of desirable traits for improved skin characteristics. In this context, the composition of the various cell layers comprising the pericarp is particularly important as this contributes significantly to the ability to resist the physical stresses associated with cracking.[47] In addition, the ability to withstand cracking also promotes better post-harvest storage characteristics in terms of the ability of fruit to withstand storage diseases.[47]

Some of the groupings apparent on the dendrogram can be explained by consideration of morphological characteristics of the cultivars. For example, 'Bidum' and 'Haak Yip' occur in close association on the dendrogram (although in adjacent groups) and these two cultivars are reported to look very similar to each other apart from a few differences, i.e. fruit of 'Bidum' are slightly smaller with more variation in size, have yellow red markings on the skin rather than the purple red colour typical of 'Haak Yip', have more chicken tongue (shrivelled) seeds and are not as sweet as 'Haak Yip'.[48] Chicken tongue seeds are a desirable trait as it means a larger percentage of flesh recovery than in fruit with larger seeds. 'Kwai May Red' and 'Kwai May Pink' also occur in close association on the dendrogram and the physical appearance of trees is also similar. It has been suggested that 'Kwai May Pink' might be a seedling of 'Kwai May Red'. Despite the high degree of genetic similarity (Figure 1) and similar tree morphological characteristics, there are slight differences in fruit characteristics of these two cultivars which allow for distinction between them. In this regard, 'Kwai May Red' has a red skin colour rather than the orange-pink colour of 'Kwai May Pink'. In addition, 'Kwai May Red' has firmer fruit with more small seeds, a higher flesh recovery and slightly better flavour than 'Kwai May Pink'. [5,48]

The current analysis also revealed instances where reported parents and offspring occurred in different groups on the dendrogram. Such findings are not uncommon, as discussed below. For example, 'Yellow Red' is assumed to be the offspring of 'Brewster'; however, these occur in different groups on the dendrogram – a finding that was also reported by Degani et al.[9] with ISSR markers. Hence, the parentage of 'Yellow Red' remains in question, as pointed out by those authors. Similarly, 'Sahkeng' is reported to be a seedling of 'Haak Yip' yet they occur in different groups. These two cultivars also display some differing morphological characteristics. For example, 'Sahkeng' produces short branches bearing large fruit with swollen skin segments and blunt protuberances, while 'Haak Yip' has long, fragile branches producing medium-sized fruit with smooth skin.[5,48] Another example is 'Kaimana', which has been reported to be a seedling of 'Haak Yip' but they do not occur in close association in the present study. Similarly, Madhou et al.[25] could not support 'Haak Yip' being the parent of 'Kaimana' when SSR markers were used. In the case of 'Salathiel', 'Nuomici' is reported to be its parent but they occur in adjacent groups on the dendrogram. These two cultivars also display fruit characteristics that are different from each other, for example, 'Salathiel' fruits are egg- to ball-shaped with thick, moderately rough skin while 'Nuomici' produces heart-shaped fruit with thin, smooth skin.[5] Similarly, the results presented by Degani et al.[15] using isozyme analysis could not support 'Nuomici' as the parent of 'Salathiel'.

Apart from a few exceptions, the general trend observed was that most of the cultivars that are harvested early and in the middle of the season occurred together at the top of the dendrogram while the late season cultivars occurred at the opposite end (Figure 1). Separation of cultivars based on maturation period has been reported using different markers, i.e. with RAPD and SNP markers.[17,18,28] Considering this, it has been suggested that fruit maturation period should be considered as one of the primary factors in litchi taxonomy.[9,23,25] A novel observation made in the present study is the presence of a sub-group within the late season cultivars which were characterised by production of large fruit with dark purple skin (Figure 1 – cultivars enclosed by the box with a dashed line). This is of interest locally, as these cultivars not only extend the harvest season, but also large fruit is a novel characteristic in the local market.

Conclusion

The current study reports on the suitability of SRAP markers for investigating the genetic relationships between litchi cultivars and selections. The tested markers allowed for separation of cultivars into groups based on similar fruit characteristics and fruit maturation period. Application of the SRAP markers allowed for creation of a molecular genotype reference database which will be a useful tool in the breeding programme in future. The imported cultivars expanded the gene pool available, particularly for the late season cultivars which were not previously available in South Africa (apart from 'Wai Chee'). Furthermore, many of the late season fruit are of a large size which is an additional benefit for exploitation in the local and international market. While other, more recently developed markers are available (e.g. SNP markers), they are significantly more expensive and require access to high-tech platforms for detection and analysis of data. Although the cost of DNA sequencing technologies is becoming more affordable in developed countries, this is not the case in developing countries where budgets for research are constrained. Hence, the SRAP markers provide an alternative tool to researchers lacking access to sophisticated platforms. The resource requirements are relatively simple and can be undertaken in a standard molecular biology laboratory.

Acknowledgements

The Agricultural Research Council (South Africa) and South African Litchi Growers' Association are acknowledged for funding and support.

Authors' contributions

All authors contributed to the conception and design of the study. Material preparation, data collection and sample analysis was performed by D.N. Data analysis, validation, data curation and writing was undertaken by E.H. R.C. contributed to data analysis and interpretation. All authors were involved in revisions of the manuscript and approved the final manuscript.

Conflicts of interest

The authors declare no conflict of interest.

Data availability

The data for this study are not available because they include breeding germplasm that have Intellectual Property restrictions.

References

1. Menzel CM, Simpson DR. A description of lychee cultivars. Fruit Var J. 1990;45:45–56.

2. Oosthuizen JH. Recommended cultural practices and research priorities for litchis in South Africa. S Afr Litchi Growers' Assoc Yearb. 1993;5:2–5.

3. National Agricultural Marketing Council. South African Fruit Trade Flow. 2018;30:1–13.

4. Begemann GJ. The South African Litchi Industry – an overview. Acta Hort. 2014;1029:35–39.

5. Menzel CM, Huang X, Liu C. Cultivars and plant improvement. In: Menzel CM, Waite GK, editors. Litchi and longan: Botany, production and uses. Cambridge: CABI Publishing; 2005. p. 59–86.

6. Froneman IJ, Nonyane D, Severn-Ellis A, Cronje RB, Sippel A. Expanding genetic diversity of the South African litchi germplasm collection to promote plant improvement. Acta Hort. 2015;1127:365–372. https://doi.org/10.17660/ActaHortic.2016.1127.57

7. International Union for the Protection of New Varieties of plants (UPOV). UPOV code: Litchi. Guidelines for conduction of tests for distinctness, uniformity and stability. Geneva: UPOV; 2012.

8. Arias RS, Borrone JW, Tondo CL, Kuhn DN, Irish BM, Schnell RJ. Genomics of tropical fruit tree crops. In: Schnell RJ, Priyadarshan PM, editors. Genomics of tree crops. New York: Springer; 2012. p. 209–239.

9. Degani C, Deng J, Beiles A, El-Batsri R, Goren M, Gazit S. Identifying lychee *(Litchi chinensis* Sonn.) cultivars and their genetic relationships using intersimple sequence repeat (ISSR) markers. J Am Soc Hort Sci. 2003;128:838–845.

10. Aradhya MK, Zee FT, Manshardt RM. Isozyme variation in lychee (*Litchi chinensis* Sonn.). Sci Hort. 1995;63:21–35. https://doi.org/10.1016/0304-4238(95)00788-U

11. Long Y, Cheng J, Mei Z, Zhao L, Wei C, Fu S, et al. Genetic analysis of litchi (*Litchi chinensis* Sonn.) in southern China by improved random amplified polymorphic DNA (RAPD) and inter-simple sequence repeat (ISSR). Mol Biol Rep. 2015;42:159–166. https://doi.org/10.1007/s11033-014-3755-8

12. Agarwal M, Shrivastava N, Padh H. Review: Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep. 2008;27:617–631. https://doi.org/10.1007/s00299-008-0507-z

13. Pandey DK, Dey A, Singh J. Biotechnological advances in lychee (*Litchi chinensis*) and their future implication in improvement of crop. In: Kumar M, Kumar V, Prasad R, Varma A, editors. The lychee biotechnology. Singapore: Springer Nature; 2017. p. 59–100. https://doi.org/10.1007/978-981-10-3644-6

14. Hajari E. Review: Molecular and related approaches to litchi improvement – historical perspective and future trends. J Hort Sci Biotech. 2019;94:693–702. https://doi.org/10.1080/14620316.2019.1624202

15. Degani C, Beiles A, El-Batsri R, Goren M, Gazit S. Identifying lychee cultivars by isozyme analysis. J Am Soc Hort Sci. 1995;120:307–312.

16. Anuntalabhochai S, Chundet R, Chiangda J, Apavatjrut P. Genetic diversity within lychee (*Litchi chinensis* Sonn.) based on RAPD analysis. Acta Hort. 2002;575:253–259. https://doi.org/10.17660/ActaHortic.2002.575.27

17. Liu C, Mei M. Construction of a lychee genetic linkage map based on RAPD markers. Acta Hort. 2003;625:131–136. https://doi.org/10.17660/ActaHortic.2003.625.13

18. Liu C, Mei M. Classification of lychee cultivars with RAPD analysis. Acta Hort. 2005;665:149–160. https://doi.org/10.17660/ActaHortic.2005.665.17

19. Kumar M, Gupta M, Shrivastava D, Prasad M, Prasad US, Sarin NB. Genetic relatedness among Indian litchi accessions (*Litchi chinensis* Sonn.) by RAPD markers. Int J Agric Res. 2006;1:390–400. https://doi.org/10.3923/ijar.2003.390.400

20. Monshi FI, Bhuiyan MSU, Tabassum R. Adaptability of litchi germplasm in hilly areas of Sylhet Agricultural University and screening their genetic variation by using RAPD markers. Int J Plant Breed Genet. 2015;9:218–227. https://doi.org/10.3923/ijpbg.2015.218.227

21. Yi G, Huo H, Chen D, Huang Z, Cai C, Qiu Y. Studies on genetic relationship among litchi varieties using AFLP. Acta Hort Sinica. 2003;30:399–403.

22. Pathak AK, Singh SP, Tuli R. Amplified fragment length polymorphism fingerprinting to identify genetic relatedness among lychee cultivars and markers associated with small-seeded cultivars. J Am Soc Hort Sci. 2014;139:657–668. https://doi.org/10.21273/JASHS.139.6.657

23. Viruel MA, Hormaza JI. Development, characterisation and variability analysis of microsatellites in lychee (*Litchi chinensis* Sonn., Sapindaceae). Theor Appl Genet. 2004;108:896–902. https://doi.org/10.1007/s00122-003-1497-4

24. Fan Q, Chen S, Zhou R, Shi S. Genetic variation of wild litchi (*Litchi chinensis* Sonn. subsp. *chinensis*) revealed by microsatellites. Conserv Genet. 2011;12:753–760. https://doi.org/10.1007/s10592-011-0182-4

25. Madhou M, Normand F, Bahorun T, Hormaza JI. Fingerprinting and analysis of genetic diversity of litchi *(Litchi chinensis* Sonn.) accessions from different germplasm collections using microsatellite markers. Tree Genet Genome. 2013;9:387–396. https://doi.org/10.1051/fruits/2010009

26. Xiang X, Chen DM, Ma SP, Ma WC, Fan J, Yang XY, et al. Core EST-SSR marker selection based on genetic linkage map construction and their application in genetic diversity analysis of litchi (*Litchi chinensis* Sonn.) germplasm resources. Acta Hort. 2014;1029:109–116. https://doi.org/10.17660/ActaHortic.2014.1029.12

27. Sun Q, Shuaipeng MA, Liu W, Jiang N, Fang J, Yang X, et al. Identification and analysis on true or false hybrids from hybridisation populations of litchi (*Litchi chinensis* Sonn.) using EST-SSR markers. Agric Biotech. 2015;4:38–43.

28. Liu W, Xiao Z, Bao X, Yang X, Fang J, Xiang X. Identifying litchi (*Litchi chinensis* Sonn.) cultivars and their genetic relationships using single nucleotide polymorphism (SNP) markers. PLoS ONE. 2015;10, e0135390. https://doi.org/10.1371/journal.pone.0135390

29. Li G, Quiros CF. Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: Its application to mapping and gene tagging in *Brassica*. Theor Appl Genet. 2001;103:455–461. https://doi.org/10.1007/s001220100570

30. Xu R, Hu D, Chen Z, Zhang C, Jiang X, Tang G. SRAP analysis on genetic relationships of genotypes in the genus *Malus* Mill. Biotech Biotech Equip. 2014;28:602–607. https://doi.org/10.1080/13102818.2014.948596

31. Soleimani MH, Talebi M, Sayed-Tabatabaei BE. Use of SRAP markers to assess genetic diversity and population structure of wild, cultivated and ornamental pomegranates (*Punica granatum* L.) in different regions of Iran. Plant Syst Evol. 2012;298:1141–1149. https://doi.org/10.1007/s00606-012-0626-4

32. Oluoch P, Nyaboga EN, Bargul JL. Analysis of genetic diversity of passion fruit (*Passiflora edulis* Sims) genotypes grown in Kenya by sequence-related amplified polymorphism (SRAP) markers. Ann Agrar Sci. 2018;16:367–375. https://doi.org/10.1016/j.aasci.2018.08.003

33. Meng H, Chen S, Chen Z, Chai D, Li Y. SRAP markers for fruit shape in cucumber. Pakistan J Bot. 2012;44:1380–1384.

34. Han J-C, Liu G-J, Chang R-F, Zhang X-Z. The sequence-related amplified polymorphism (SRAP) markers linked to the colour around the stone (Cs) locus of peach fruit. Afr J Biotech. 2012;11:9911–9914. https://doi.org/10.5897/AJB12.396

35. Huff, DR, Peakall R, Smouse PE. RAPD variation within and among natural populations of outcrossing buffalograss *Buchloe dactyloides* (Nutt) Engelm. Theor Appl Genet. 1993;86:927–934. https://doi.org/10.1007/BF00211043

36. Zhou J, Fu J-X, Wu Z-X, Huang S-S, Zhang Y-F, Wang Y, et al. Genetic diversity in litchi and longan germplasm as determined by SRAP markers. Acta Hort. 2011;918:799–806. https://doi.org/10.17660/ActaHortic.2011.918.105

37. Bhatt J, Kumar S, Patel S, Solanki R. Sequence-related amplified polymorphism markers based genetic diversity analysis of cumin genotypes. Ann Agrar Sci. 2017;15:434–438. https://doi.org/10.1016/j.aasci.2017.09.001

38. Jingrade P, Huded AK, Kosaraju B, Mishra MK. Diversity genotyping of Indian coffee (*Coffea arabica* L.) germplasm accessions by using SRAP markers. J Crop Improv. 2019;33:327–345. https://doi.org/10.1080/15427528.2019.1592050

39. Xie WG, Zhang XQ, Cai H, Liu W, Peng Y. Genetic diversity analysis and transferability of cereal EST-SSR markers to orchard grass (*Dactylis glomerata* L.). Biochem Syst Ecol. 2010;38:740–749. https://doi.org/10.1016/j.bse.2010.06.009

40. Sagawa CHD, Cristofani-Yaly M, Novelli VM, Bastianel M, Machado MA. Assessing genetic diversity of *Citrus* by DArT_seq™ genotyping. Plant Biosyst. 2018;152:593–598. https://doi.org/10.3389/fpls.2017.00577

41. Kuhn DN, Bally ISE, Dillon NL, Innes D, Groh AM, Rahaman J, et al. Genetic map of mango: A tool for mango breeding. Front Plant Sci. 2017;8:577. https://doi.org/10.1080/11263504.2017.1341438

42. Husselman JH, Froneman IJ. Cultivars. In: De Villiers EA, Joubert PH, editors. The cultivation of litchi. Nelspruit: Agricultural Research Council – Institute for Tropical and Subtropical Crops; 2002. p. 31–49.

43. Singh LB, Singh UP. The litchi. Uttar Pradesh, India: Superintendent, Printing and Stationery; 1954.

44. Pandey RM, Sharma HC. The litchi. New Delhi: Indian Council for Agricultural Research; 1989.

45. McGarigal K, Cushman S, Stafford S. Cluster analysis. Multivariate statistics for wildlife and ecology research. New York: Springer; 2000. p. 81–127.

46. South African Department of Agriculture, Forestry and Fisheries (DAFF). Litchi production guideline. Pretoria: DAFF; 2013. Available from: https://www.daff.gov.za/Daffweb3/Portals/0/Brochures%20and%20Production%20guidelines/Production%20Guidelines%20Litchi.pdf

47. Huang X. Fruit disorders. In: Menzel CM, Waite GK, editors. Litchi and longan: Botany, production and uses. Cambridge: CABI Publishing; 2005. p. 141–151.

48. Singh G, Nath V, Pandey SR, Ray PK, Singh HS. The litchi. New Delhi: Food and Agriculture Organisation of the United Nations; 2012.

Revised estimates of Taung's brain size growth

Cranial capacity, a proxy for the volume of the brain and associated cranial contents, is an important yardstick used to compare early hominin species because increasing brain size is a key characteristic of our lineage. In 1925, Raymond Dart claimed that a natural endocast found at the Buxton Limeworks near Taung, South Africa (which he named *Australopithecus africanus*), showed signs of neural reorganisation, but its juvenile status complicated comparison to other hominoid species. In an attempt to put its brain size and reorganisation into a comparative context, subsequent researchers have tried to estimate Taung's adult cranial capacity by comparison to coarse-grained hominoid growth data. In this study, we simulated brain growth in *A. africanus* using asymptotic growth models in known-age mountain gorillas, chimpanzees and modern humans, and show that, at just under 4 years old, Taung's brain had already finished or nearly finished growing according to hominoid developmental schedules. Percentage-growth remaining estimates are lower here than in previous studies using cross-sectional ontogenetic samples of unknown chronological age. Our new adult estimates (between 404 cm^3 and 430 cm^3 overall and 405–406 cm^3 for chimpanzee models) are smaller than previous estimates with a 'starting' cranial capacity of 404 cm^3, supporting the hypothesis that Taung's adult brain size would have fallen toward the lower end of the *A. africanus* range of variation and strengthening the case that Taung was female.

Significance:

- This is one of several recent studies to show that brain growth is completed in African apes and humans earlier than previously appreciated.
- New adult cranial capacity estimates for Taung are lower than previous estimates, supporting the hypothesis that Taung was female.
- Cessation of brain growth in hominoids at earlier ages than previously reported suggests that adult cranial capacities for hominin juvenile specimens have been overestimated.

Check for updates

AUTHORS:
Robert C. McCarthy[1]
Emily Zimel[2]

AFFILIATIONS:
[1]Department of Biological Sciences, Benedictine University, Lisle, Illinois, USA
[2]Department of Physical Sciences, Benedictine University, Lisle, Illinois, USA

CORRESPONDENCE TO:
Robert McCarthy

EMAIL:
rmccarthy@ben.edu

DATES:
Received: 31 Jan. 2019
Revised: 01 Apr. 2020
Accepted: 01 Apr. 2020
Published: 29 July 2020

HOW TO CITE:
McCarthy RC, Zimel E. Revised estimates of Taung's brain size growth. S Afr J Sci. 2020;116(7/8), Art. #5963, 7 pages. https://doi.org/10.17159/sajs.2020/5963

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:
brain growth, cranial capacity, brain ontogeny, growth curves, *Australopithecus africanus*

FUNDING:
Dr. Scholl Foundation, Benedictine University Natural Science Summer Research Program

Introduction

The type specimen of *Australopithecus africanus*, Taung, is a juvenile skull consisting of a partial face with fragmentary pieces of the basicranium attached, a mandible, and a natural hemi-endocast.[1] Taung has been the subject of intensive research focus because of its potential to resolve questions about hominin brain size and reorganisation[2-7] and *A. africanus* craniofacial growth[8-11], dental maturation[12-18] and brain ontogeny[19-21].

Raymond Dart originally estimated Taung's cranial capacity to be 520 cm^3 based on a reconstruction of the hemi-endocast.[2] Subsequent estimates have ranged between 382 cm^3 and 530 cm^3 (Tables 1–2). The most frequently cited estimate for Taung's cranial capacity – 404 cm^3 – was derived from an independent reconstruction of the hemi-endocast[22] and was recently corroborated by digital reconstruction of the endocast and endocranial cavity[23]. At 404 cm^3, Taung's cranial capacity is already at the lower end of the range of *A. africanus* variation (Table 2), even though the Taung juvenile died after gingival eruption of the first molars but before the they had moved into functional occlusion, and so still had several years remaining to reach adulthood.[12-19] Taung's importance to studies of hominin brain evolution and the scarcity of relatively complete crania and endocasts of adult *A. africanus* specimens have tempted researchers to estimate Taung's adult cranial capacity. Adult estimates range between 404 cm^3 and 785 cm^3 for 'starting' values $\geq$404 cm^3 (Table 1). Researchers working with different age estimates and differing ideas about *A. africanus* growth trajectories have sometimes modelled Taung with a large percentage-growth remaining, producing adult cranial capacity estimates >600 cm^3,[2,3,19,24-26] which are larger than any known adult *A. africanus* specimens. Uncertainty about brain growth parameters in fossil species makes it difficult to estimate how much growth Taung had already attained (and how much remained),[23,27-29] making it difficult to estimate adult cranial capacity. Size of the adult brain, which is approximated to some extent by cranial capacity, is an important parameter for understanding adaptive shifts in brain size and neural reorganisation early in human evolutionary history.

Previous adult predictions are based on overestimates of the amount of brain growth remaining in hominoids. At the time Taung was discovered in 1924, there were several misconceptions in the scientific literature about great ape growth and development. For example, it was thought that brain growth continued throughout the entire juvenile growth period until eruption of the third molars, or even beyond[19(and references therein),25,34,37,38], with humans reaching 81–88%[24,25,38] and great apes 85–92%[19,24] of adult brain size in the period just prior to eruption of the first molars; and that chimpanzees followed a tooth eruption schedule similar to that of modern humans, with the first permanent molar erupting at the end of the sixth year[19] or early in the fifth year[25]. Based in part on these misconceptions, Keith[24,25] and Dart[2,3] increased Taung's juvenile cranial capacity by 15–20% to produce adult cranial capacity estimates of 520–625 cm^3. Zuckerman[19] increased Taung's juvenile cranial capacity by 3% to as much as 57% based on the supposed amount of growth expected in chimpanzees and gorillas between eruption of the first molars and adulthood (Table 1). More recent estimates, while lower, have mostly relied on data from Ashton and Spence[33] who found that 92.5% of adult cranial capacity is attained on average prior to eruption of the first molar in cross-sectional samples of hominoids (including orangutans, gorillas, chimpanzees and modern humans).

Table 1: Previous juvenile and adult cranial capacity estimates for Taung presented in numerical order of adult cranial capacity. Table references are noted below or can be found in references[30,34]. C = chimpanzee; G = gorilla; H = modern human; A = australopith; gc = growth curve.

Reference	Year	Juv	Adult	% increase	Model
Falk and Clarke[28]	2007	382	390	2.1	C gc (98%)
Falk and Clarke[28]	2007	382	406	6.3	C gc (94%)
Falk[31]	1987	404[22]	412 (404–420)	2.0 (0–4.0)	C gc (96–100%)
Conroy et al.[32]	2000	405[22]	422	4.2	C (%, female)
Wolpoff	1996–1997	405[22]	425	4.9	95%
Neubauer et al.[23]	2012	403	428	6.2	A (Sts 71)
Conroy et al.[32]	2000	405[22]	431	6.4	C (94%, combined sex)
Holloway[22], Holloway et al.[27]	1970, 2004	405[22]	440[a]	8.6	C (92%)[33]
Conroy et al.[32]	2000	405[22]	455	12.4	C (%, male)
Keith[24]	1925	<450	≤520	15.6[b]	G, C (~86%)
Zuckerman[19]	1928	500	540 (515–566)	8.0 (3.0–13.2)	C
Tobias[34]	1971	500[19]	540	8.0	O, G, C, H (92.5%)[33]
Coon	1962	494	543	9.9	H ('unlikely')
Zuckerman[19]	1928	500	550	10.0	C, G ('likely')
Tobias	1965, 1967	520[2]	562	8.1	O, G, C, H (92.5%)[33]
Miller	1991	–	563	–	Mid-range estimate
Le Gros Clark	1947	500[19]–520[2]	570	9.6–14.0	C
Keith[22]	1931	500[19]	600	20.0	G
Le Gros Clark, Tobias	1955, 1963	520[2]	600	15.4	?
Zuckerman[19]	1928	500[19]	605[c]–638	21.0–27.5	C (max starting size/age; 'extremely unlikely')
Dart[3]	1956	520[2]	624	20.0	NR
Dart[2]	1926	520[2]	625	20.2[d]	NR
Schepers	1950	510–530	650	22.6–27.5	NR
Elliot Smith[26]	1925	520[2]	>650	>25[e]	NR
Zuckerman[19]	1928	520[2]	728	40.0	C (min starting size/age; 'extreme, ridiculously high')
Zuckerman[19]	1928	500[19]	769–785	54.0–57.0	C (min size/age; 'obviously ridiculous')

[a]Computed so that 405 cm³ is 92% of adult cm³.[33]
[b]Keith[24(p.234)] used somewhat imprecise language ('volume of the brain in the juvenile … must be less than 450 cc, and if we allow a 15 per cent. increase for the remaining stages of growth, the size of the adult brain will not exceed 520 cc.'), and so rounded this value to 15%.
[c]Miscalculated as 603 cm³ in Zuckerman[19].
[d]Dart[2] reported this value as 20%.
[e]Elliot Smith[26(p.235)] reported this as a 20% increase: '… brain would probably have increased in volume to the extent of a fifth had it attained the adult status.' He obtained this value by calculating per cent complete (520 x 100/650=80%) rather than percentage change ([650–520] x 100/520=25%).

Table 2: Cranial capacity values for *Australopithecus africanus* specimens. More information about ranges of estimates can be found in references[23,30,34]. Sex attributions follow Grine[35]. 3DR = three-dimensional digital reconstruction, including correcting distortion and reconstructing broken portions using chimpanzees and Sts 5 as models[23]; PERM = partial endocast reconstruction method[27].

Specimen	Sex	Cranial capacity	Range of estimates	Methods/notes
MLD 1	M	510[27]	500–650[3]	Missing frontal and temporal lobes; reconstructed via PERM[27]
MLD 37/38	F	440[23]	425–480	Undistorted posterior cranium filled with matrix, required reconstruction in rostral part of frontal lobes; 3DR
Sts 5 (skull no. 5)	F[a]	475[23]	473[23]–485[20]	Complete cranium with broken calotte that fits back on; 3DR
Sts 19/58 (skull no. 8)	F	436[5,27]	436–570	Posterior cranium, endocast made by hand, reconstructed via PERM[27]
Sts 25[b]	F	363[30,c]	350–375[30]	Estimated using regression on vault variables
Sts 60 (TM 1511)	M?	391[23]	384[23]–428	Natural endocast; 3DR
Sts 71 (skull no. 7)	F	412[23]	370–430[32]	Nearly complete but distorted cranium; 3DR
StW 505	M	568[23]	515–625	Relatively complete but distorted cranium; 3DR
Taung	F[d]	403[23]–404[22]	382[28]–525	Natural hemi-endocast, missing part estimated, mirror-imaged; 3DR
Type 2 (TM 1512)	F	457[27]	457–580	Distorted natural endocast missing basal portion

[a]Sts 5 in particular has been attributed as female and male at different times (see Grine[35] and Tawane and Thackeray[36] for references).
[b]We also considered a 400–425-cm³ estimate for Sts 17,[30] a partial face and associated cranial vault pieces that do not fit together perfectly; we followed Wolpoff's recommendation to exclude this specimen from consideration (Wolpoff M 2018, personal communication).
[c]Mid-point of range of estimated values.
[d]See discussion

Here we reconsider evidence for Taung's adult cranial capacity, taking into account the most up-to-date estimates of chronological age, cranial capacity and brain growth trajectories in African apes and modern humans; and use these new estimates to reassess brain size variability in *A. africanus*.

Materials and methods

We used mountain gorilla (*Gorilla beringei*), chimpanzee (*Pan troglodytes*) and modern human (*Homo sapiens*) cranial capacity growth curves to produce developmental simulations[10,11,23] of Taung's brain growth trajectory and to calculate the per cent changes that would have occurred between a given set of 'starting' ages and brain growth completion (Figure 1). We used starting age estimates of 3.73, 3.83 and 3.93 years based on a comparison of Taung's root length and crown development to *A. africanus* specimen Stw 402[17], but also considered 3.3-year[8,9,12] and 3.5-year[13,15] estimates (based on patterns of root and crown development) and 4-year[24,39], 4.5–5.5-year[40] and 5–7-year[41] estimates (based on varying interpretations of dental development and tooth eruption timing) in order to assess the impact of different starting ages on adult cranial capacity estimates.

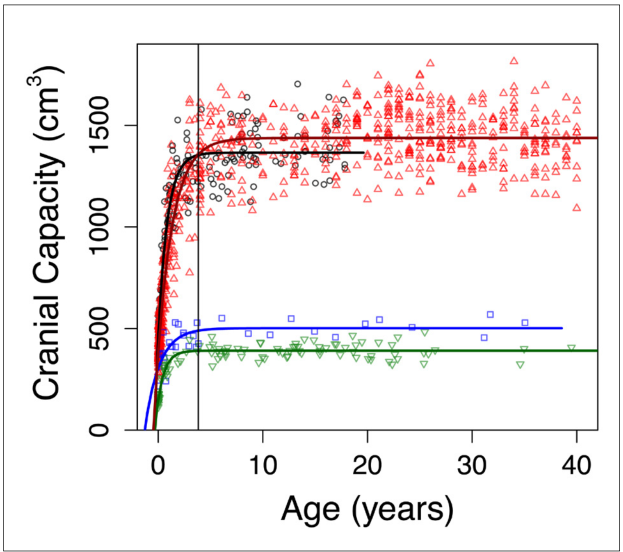


Figure 1: Distance growth curves for two samples of modern humans (black open circles, red triangles), chimpanzees (green triangles) and gorillas (blue squares) used to model Taung's cranial capacity growth. The vertical line indicates Taung's estimated age of 3.83 years.

We collated cranial capacities and/or brain masses for mountain gorilla, chimpanzee and modern human comparative samples from the literature[42-47] and set up each data set – including deleting outliers – following previously determined criteria[42,44,48,49]. Because raw data were not available in one instance[47], we digitised axes and data points from the original paper using WebPlotDigitizer[50]. We transformed brain mass data from previous studies[42-46] into cranial capacities following equations in Cofran and DeSilva[51]. As we are modelling the amount of growth that occurs between a given 'starting' age and asymptotic growth cessation, results would differ minimally (if at all) based on different methods for adjusting between brain mass (in grams) and cranial capacity (in cubic centimetres or millilitres). In this case, we prioritised maintaining all the data in the same units as the fossil data (cm^3) over leaving each data set in its original units. After visually examining growth curves for modern humans, chimpanzees and mountain gorillas, we decided to cap three data sets at 40 years of age in order to limit the known impact that brain shrinkage at older ages can have on curve-fitting.[43,46,48,52] The fourth (modern human) data set included only individuals up to 18 years.[47]

We fit the data in R (R Core Development) using a non-linear asymptotic regression model with a vertical offset (SSasymp), which is a 'self-starting' function that iteratively fits a model using initial estimates of the horizontal asymptote ('Asym'), the natural logarithm of the rate constant ('lrc'), and a numeric parameter corresponding to the response value when the input is zero ('R0'). We estimated cranial capacities from the regression fit at given starting ages for Taung and calculated percentage-change values for cranial capacities between each starting age and the maximum age for each data set (in each case the growth curve had reached asymptotic stability). We then estimated adult values for Taung by increasing Taung's juvenile cranial capacity by these percentage-change values. The R code for each of these operations is presented in the supplementary material. As noted above, we produced estimates for the most realistic 3.73–3.93-year-old starting values (Table 3), but also tested the effects of different juvenile starting cranial capacity estimates (Supplementary table 1) and different starting ages (Supplementary table 2). In this study, we prioritised testing the effects of different starting juvenile cranial capacity and age estimates over fitting confidence intervals, which adds a layer of complexity that will be addressed in a follow-up study.

To assess variability in the *A. africanus* sample, we calculated coefficients of variation (CV)[53] for the *A. africanus* hypodigm with different adult estimates for Taung (Table 4), and compared these values to bootstrapped samples of adult *Gorilla gorilla* + *G. berengei*, *P. troglodytes*, and *H. sapiens* matching the sample size of the *A. africanus* sample (n=10 including Sts 25, n=9 excluding Sts 25; see explanation below), as well as to historical CV values for *A. africanus* (Supplementary table 3). We did not include cranial capacity data for the recently published specimen StW 573 from the Silberberg Grotto at Sterkfontein because its taxonomic status is still under debate, and its endocast has not yet been virtually reconstructed to correct for distortion during the fossilisation process.[54]

We also simulated *A. africanus* cranial capacity growth by looking at per cent changes necessary to grow Taung to different *A. africanus* 'target' adults, thereby simulating different models for growth increases in Taung's cranial capacity[10,11] (Table 5). We compared these percentages to values for comparative samples to assess the likelihood that Taung's remaining brain growth would be sufficient to produce target adult cranial capacity values.

Results

Adult cranial capacity estimates for Taung fall within a relatively narrow range regardless of narrowly defined starting age and choice of species- or sex-specific models (Table 3). By 3.83 years of age, cranial capacity has reached 97.5% of adult size in gorillas (100% in males, 97.8% in females), 99.8% in chimpanzees (99.5% in males, 99.8% in females), and 94.0–99.3% in modern humans (94.4–99.0% in males, 95.2–99.0% in females). Because there is so little growth remaining between 3.73–3.93 years of age and adulthood (Figures 1–2), juvenile estimates increase by only 0–26 cm^3 (Table 3) – a maximum increase of ~6.5%. There is a 3–26 cm^3 (0.75–6.48%) increase according to modern human growth curves, but chimpanzees and gorillas have, respectively, <1% and <3% growth remaining (Table 3, Figure 2, Supplementary table 1). Adult cranial capacity estimates range between 404 cm^3 and 415 cm^3 according to the gorilla curve, 405 cm^3 and 406 cm^3 according to the chimpanzee curve, and 407 cm^3 and 430 cm^3 according to the two human curves (Table 3). The two modern human growth curves differ from each other in terms of percentage-growth change and estimated adult cranial capacities, with a more prolonged growth trajectory in brains derived from 19th-century German autopsy material[46] than in a 20th-century sample of cranial capacities from an Australian research hospital[47] (Table 3, Figures 1–2). Adult cranial capacity estimates for different starting cranial capacities and ages are presented in Supplementary tables 1 and 2.

Regardless of which growth curve is used, all the adult estimates noted above place Taung at the low end of the range of adult *A. africanus* variation. Original estimates for *A. africanus* cranial capacities were fairly large (>500 cm^3 in several cases), whereas revised estimates tend to be smaller. One cranium in particular, Sts 25 (350–375 cm^3), does not appear in many comparative analyses of *A. africanus* cranial variation because it is fragmentary and still partially embedded in matrix, so its cranial capacity has been estimated by regression[30] (Wolpoff M 2018, personal communication). We tested the effect of including Sts 25 in the *A. africanus* hypodigm on CV values by running analyses with and without this specimen. Table 4 shows CV values for the *A. africanus* hypodigm. Including new adult estimates for Taung in the *A. africanus* sample produces CV values that range between 11.8% and 13.6%, which are similar to values for the sample including Holloway's[22] 440-cm³ adult estimate but lower than CV values for samples that included larger adult estimates (Supplementary table 3). Coefficient of variation values for 404–440 cm³ estimates in this study and others[22,28,32] can be accommodated within the range of variation for all extant samples, but larger adult estimates (>550 cm³ and >600 cm³ for samples that include and exclude Sts 25, respectively) fall outside the 95% confidence intervals for modern humans (Figure 3).

Table 3: Adult cranial capacity estimates for Taung produced using narrowly defined starting ages

Species	Sex	Adult estimates for starting age (years)			
		3.83		3.73–3.93	
		% Increase	Adult estimate	% Increase	Adult estimate(s)
Gorilla[42]	M	0.000	404	0.000	404
	F	2.248	413	2.087–2.421	412–414
	Comb	2.492	414	2.314–2.685	413–415
Chimp[43-44]	M	0.442	406	0.388–0.503	406
	F	0.178	405	0.153–0.207	405
	Comb	0.250	405	0.217–0.289	405
Human1[47]	M	0.978	408	0.874–1.094	408
	F	1.111	408	0.997–1.238	408–409
	Comb	0.845	407	0.753–0.949	407–408
Human2[46]	M	6.038	428	5.625–6.483	427–430
	F	4.848	424	4.498–5.227	422–425
	Comb	5.971	428	5.563–6.411	426–430

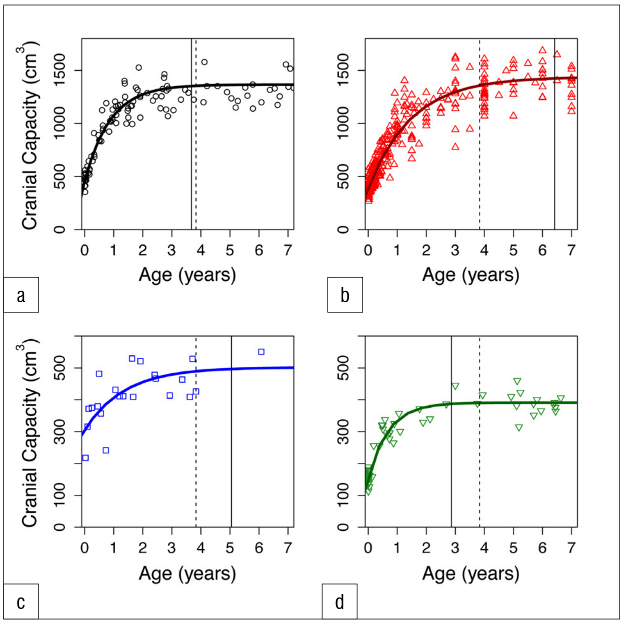


Figure 2: Close-up views showing the first 7 years of cranial capacity growth in (a,b) modern humans, (c) mountain gorillas and (d) chimpanzees. In each panel, the dashed vertical line indicates the 'starting age' for Taung and the solid vertical line indicates the point of 99% growth completion.

Table 4: Coefficient of variation (CV) estimates for the *Australopithecus africanus* hypodigm computed using different values for Taung's adult cranial capacity

Adult value	Reference/model	w/Sts 25	w/o Sts 25
		CV (%)	CV (%)
405–406	C[this study]	13.5–13.6	12.3
404–415	G[this study]; C[28]	13.4–13.6	12.1–12.4
423–430	H[this study]	13.2–13.3	11.8–11.9
440	C[22,27]	13.1	11.6
520	G,C[24]	13.7	12.1
540	C[19]	14.2	12.6
550	C,G[19]	14.5	12.9
562	H,C,G,O	14.8	13.3
570	C[Table 1]	15.1	13.6
600	G[25]	16.2	14.8
625	NR[2]	17.2	15.9
650	NR[26]	18.3	17.1
728	C (min–max size/age)[19]	22.1	21.1

C, chimpanzee; G, gorilla; H, human; O, orangutan; NR, not reported

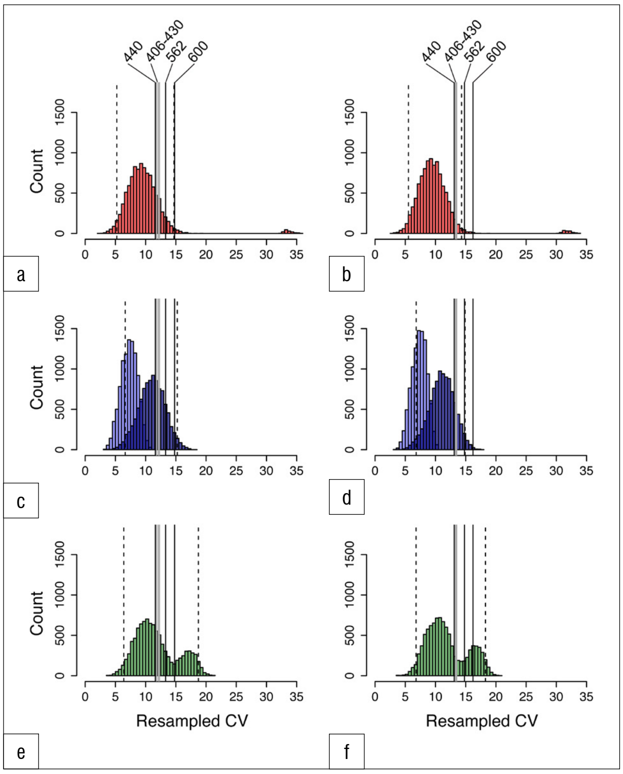


Figure 3: Resampled coefficients of variation (CVs) for (a,b) modern humans (red), (c,d) lowland gorillas (dark blue), mountain gorillas (light blue) and (e,f), chimpanzees (green) for sample sizes $n=9$ (left, excluding Sts 25) and $n=10$ (right, including Sts 25). Solid vertical lines indicate CVs for the *Australopithecus africanus* sample with different adult estimates for Taung – from left to right in each panel: 440 cm³,[22] estimates from this study (404–430 cm³), 562 cm³,[34] and 600 cm³.[25] Dashed lines indicate 95% confidence intervals for each distribution.

Discussion

Adult cranial capacities estimated for Taung with starting ages between 3.73 and 3.93 years (Table 3) ranged between 404 cm³ (no increase) and

430 cm^3 (6.4% increase). These values are at the lower end of the adult *A. africanus* range of variation (Table 2). As expected, the largest of these adult estimates are produced by modern human growth curves. Estimates calculated using chimpanzee and gorilla growth curves ranged between 404 cm^3 and 415 cm^3 (Table 3). If *A. africanus* brain growth followed a chimpanzee trajectory as previously suggested[12,13,15,16,18,19,24,25,31,32,39], then Taung had <1% growth remaining, so 405–406 cm^3 may be the most reasonable estimate of adult cranial capacity.

Five modern estimates of Taung's adult cranial capacity are available for comparison. Based on an average value of 92% for brain growth completion prior to eruption of the first molar in hominoids[19,33,34], Holloway[22] added 35 cm^3 to Taung's juvenile cranial capacity to predict an adult value of 440 cm^3. Following the same logic we followed here, Falk[31] increased Taung's juvenile estimate using a chimpanzee brain development curve, producing estimates ranging between 404 cm^3 and 420 cm^3. Somewhat presciently, Falk[31(p.19)] wrote: ''If Taung was as old as 3.7 years, … (the) curve suggests that its adult cranial capacity had already been achieved!'' Falk and Clarke[28] estimated a juvenile cranial capacity of 382 cm^3 based on a different reconstruction of Taung's endocast, then increased this value to 406 cm^3 based on a chimpanzee growth curve. Holloway and Broadfield[29] updated Falk and Clarke's[28] adult estimate (to 390 cm^3) by setting their juvenile value to 98% growth-complete. Conroy et al.[32] replaced the averaged value for hominoid brain growth completion prior to eruption of the first molar[33] used by Holloway[22] with a different set of values based on growth in chimpanzees, increasing Holloway's 404/405-cm^3 juvenile estimate to 431 cm^3 based on a mixed-sex sample, 422 cm^3 based on a female sample, and 455 cm^3 based on a male sample. Our new estimates for Taung based on chimpanzee growth curves – 405–406 cm^3 – corroborate Falk's[31] estimates using a different data set and are similar to estimates from Holloway and Broadfield[29], but are lower than estimates from Holloway[22], Conroy et al.[32] and Falk and Clarke[28].

The range of *A. africanus* cranial capacity variation in this study, 205 cm^3 (363–568 cm^3), is fairly large compared to older studies (Supplementary table 3) because of the inclusion of two small crania: Sts 25[30] (350–375 cm^3), a small specimen which preserves the left half of the cranial base and a partial vault still covered with breccia; and a new, smaller estimate of 391 cm^3 for Sts 60[23]. If Sts 25 is excluded from the sample, the range drops to 177 cm^3, with end-points of the range defined by crania that have been digitally reconstructed with a high degree of confidence.[23] Coefficients of variation range between 11.8% and 13.6% for different iterations of the sample (Table 4), which are on par with recent studies but higher than in earlier work (Supplementary table 3).[22,27,31,34] It has been recognised previously that variability between specimens can be underestimated when regression formulae are used to estimate cranial capacities as well as when complete fossil crania like Sts 5 are used as 'templates' to reconstruct less complete crania.[34,55] Both of these conditions apply historically (Supplementary table 3). Coefficient of variation values for *A. africanus* cranial capacity (Table 4) generated in this study fall within the ranges of lowland gorilla, chimpanzee and modern human variation, but fall outside the range of mountain gorilla variation derived from limited samples (Figure 3). Adult estimates for Taung >550–600 cm^3 yield CV values outside the range of modern human variation and estimates >565 cm^3 yield CV values outside the range of gorilla variation (Figure 3). High CV values (>10%) support results based on craniofacial linear measurements.[56]

As noted above, another way to approach this problem is to estimate how much the brain would have had to grow to reach target adult *A. africanus* cranial capacities and to compare these values to growth data for hominoid comparative samples. Starting at 3.83 years of age with a 404 cm^3 cranial capacity, Taung would have needed to grow ~2% to match Sts 71's cranial capacity (Figure 4, Table 5) – a value that falls within the range of variation of two of the comparative samples (Figure 4). Setting aside for a moment Sts 25 and Sts 60 (two small specimens with adult cranial capacities smaller than Taung's juvenile cranial capacity), Taung's cranial capacity would otherwise have had to increase 7.9–40.6% to match any of the other *A. africanus* specimens and 17.6% or 26.2–40.6% to match growth in male specimens. This amount of brain growth remaining at 3.83 years of age is unlikely. In fact, to match these values, Taung would have had to have been between 1.19 and 2.63 years old based on human growth curves and 0.51 and 1.45 years old based on chimpanzee curves (Table 5) – values outside the range of ages previously suggested for Taung. If Taung grew more like a chimpanzee[12,13,15,16,18,19,24,25,31,39], then it would not have been possible to reach the upper echelons of *A. africanus* adult variation. Another line of evidence supports a small adult cranial capacity estimate for Taung. According to developmental simulations of craniofacial growth[11], Taung would have grown up to resemble Sts 71, a small-brained putative female, more closely than Sts 5 (which is either a large-brained female[35] or a small-brained male[36]) and other early hominin specimens. It is reassuring that different types of data from the brain and craniofacial region point to the same specimen as an adult target. If Taung grew according to an ape-like brain growth trajectory, then its estimated adult size and similarity to Sts 71, both in the face and neurocranium, support interpretations that Taung is a small female.[24,25,57,58]

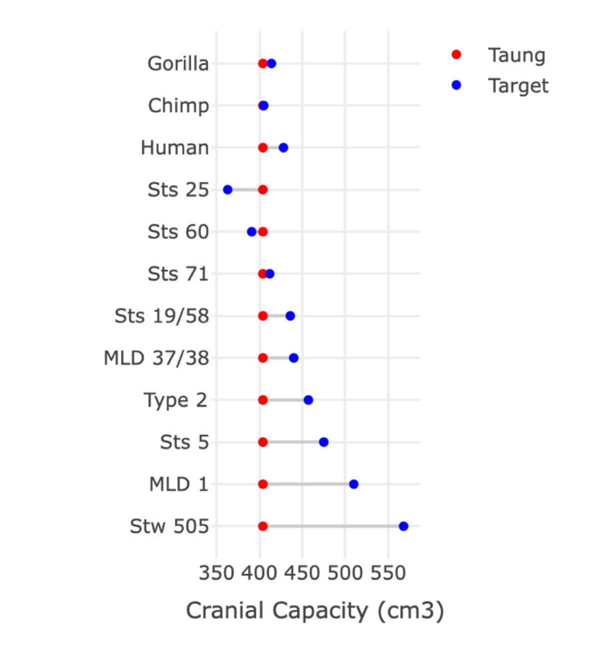


Figure 4: Dumbbell plot showing developmental trajectories necessary to grow a 404-cm^3 Taung cranial capacity to target *Australopithecus africanus* adult specimens starting at 3.83 years of age. Comparative data for cranial capacity growth in gorillas, chimpanzees and modern humans are presented at the top of the graph.

Table 5: Results for growth simulations

Specimen	Cranial capacity (cm^3)		% Change	% Target	Predicted starting age (years)	
	Target adult	Change			Chimpanzee model	Human model
Sts 25	363	-41	-10.148	111.3	–	–
Sts 60	391	-13	-3.218	103.3	–	–
Sts 71	412	+8	+1.980	98.1	2.32	3.41
Sts 19/58	436	+32	+7.921	92.7	1.45	2.63
MLD 37/38	440	+36	+8.911	91.8	1.37	2.54
Type 2	457	+53	+13.119	88.4	1.14	2.22
Sts 5	475	+71	+17.574	85.1	0.96	1.96
MLD 1	510	+106	+26.238	79.2	0.74	1.59
Stw 505	568	+164	+40.594	71.1	0.51	1.19

In this study, we followed the logic of previous studies by using modern human and African ape growth curves to estimate Taung's adult brain size. However, Taung is one of only a few *Australopithecus* juveniles with a fairly secure developmental age that can be used to test hypotheses about how the pattern and rate of hominin growth compare to growth in African apes and modern humans.[45,59,60] The results presented here rely on the assumption that brain growth in *Australopithecus* can be modelled accurately with reference to these comparative samples. We acknowledge that this does not necessarily have to be the case. It will take a combination of new fossil discoveries of juvenile specimens, new comparative data[43-47], innovative analyses[44,45,59,60], and reinterpretations of previous data and analyses[48,49,51] to test this assumption about growth and development in the genus *Australopithecus*.

Conclusions

Our results support the hypothesis that Taung was female and help to clarify the lower end of the *A. africanus* range of variation. This study focused on brain ontogeny in Taung, which is important not only in an historical sense but also because Taung is one of only a few juvenile specimens that can shed light on brain growth in *Australopithecus*. It is possible to extend this type of developmental simulation to other juvenile fossil endocasts and crania, including specimens of *Australopithecus afarensis*, *Paranthropus boisei*, *P. robustus*, *Australopithecus sediba*, *Homo habilis*, *Homo erectus* and *Homo neanderthalensis*. Evidence for early cessation of brain growth in hominoid comparative samples published here and elsewhere[42,48,49,51,59,60] brings to light the intriguing possibility that previous adult cranial capacity predictions from juvenile specimens might be overestimates.

Acknowledgements

R.C.M. acknowledges support from the Dr. Scholl Foundation and E.Z. acknowledges the support of the Benedictine University Natural Science Summer Research Program. We thank Philip Novack-Gottshall for help with R programming, Milford Wolpoff and Ralph Holloway for providing information about Sts 25 and Sts 60, and two anonymous reviewers for suggestions that helped improve the manuscript.

Authors' contributions

R.C.M. conceptualised the study, designed the methodology, validated data, performed some of the data analyses, and wrote and revised all drafts of the paper. E.Z. collated and curated data, performed data analysis, and helped revise early drafts of the paper.

Data availability

Data in this article are available as an open data set.[61]

References

1. Dart RA. *Australopithecus africanus*: The man-ape of South Africa. Nature. 1925;115(2884):195–199. https://doi.org/10.1038/115195a0
2. Dart RA. Taungs and its significance. Nat Hist. 1926;26:315–327.
3. Dart RA. The relationships of brain size and brain pattern to human status. S Afr J Med Sci. 1956;21:23–45.
4. Falk D. A reanalysis of the South African australopithecine natural endocasts. Am J Phys Anthropol. 1980;53:525–539. https://doi.org/10.1002/ajpa.1330530409
5. Holloway RL. Some additional morphological and metrical observations on *Pan* brain casts and their relevance to the Taung endocast. Am J Phys Anthropol. 1988;77:27–33. https://doi.org/10.1002/ajpa.1330770106
6. Falk D, Redmond JC, Guyer J, Conroy GC, Recheis W, Weber GW, et al. Early hominid brain evolution: A new look at old endocasts. J Hum Evol. 2000;38:695–717. https://doi.org/10.1006/jhev.1999.0378
7. Holloway RL, Clarke RJ, Tobias PV. Posterior lunate sulcus in *Australopithecus africanus*: Was Dart right? Comptes Rendus Palevol. 2004;3:287–293. https://doi.org/10.1016/j.crpv.2003.09.030
8. Bromage T. Taung facial remodeling: A growth and development study. In: Tobias PV, editor. Hominid evolution: Past, present, and future. New York: Alan R. Liss; 1985. p. 239–245.
9. Bromage T. Ontogeny of the early hominid face. J Hum Evol. 1989;18:751–773. https://doi.org/10.1016/0047-2484(89)90088-2
10. Ackermann RR, Krovitz GE. Common patterns of facial ontogeny in the hominid lineage. Anat Rec. 2002;269:142–147. https://doi.org/10.1002/ar.10119
11. McNulty KP, Frost SR, Strait DS. Examining affinities of the Taung child by developmental simulation. J Hum Evol. 2006;51:274–296. https://doi.org/10.1016/j.jhevol.2006.04.005
12. Bromage TG, Dean MC. Re-evaluation of the age at death of immature fossil hominids. Nature. 1985;317:525–527. https://doi.org/10.1038/317525a0
13. Conroy GC, Vannier MW. Dental development of the Taung skull from computerized tomography. Nature. 1987;329:625–627. https://doi.org/10.1038/329625a0
14. Mann AE. The nature of Taung dental maturation. Nature. 1988;333:123. https://doi.org/10.1038/333123a0
15. Conroy GC, Vannier MW. Dental development in South African australopithecines. Part II: Dental stage assessments. Am J Phys Anthropol. 1991;86:37–156. https://doi.org/10.1002/ajpa.1330860205
16. Smith BH. Life history and the evolution of human maturation. Evol Anthropol. 1992;1:134–142. https://doi.org/10.1002/evan.1360010406
17. Lacruz RS, Ramirez Rossi F, Bromage TG. Dental enamel hypoplasia, age at death, and weaning in the Taung child. S Afr J Sci. 2005;101:567–569.
18. Dean MC. Measures of maturation in early fossil hominins: Events at the first transition from australopiths to early *Homo*. Phil Trans Roy Soc B. 2016;371, Art. #20150234. https://doi.org/10.1098/rstb.2015.0234
19. Zuckerman S. Age changes in the chimpanzee, with special reference to growth of brain, eruption of teeth, and estimation of age; with a note on the Taungs ape. Proc Zool Soc Lond. 1928;1:1–142. https://doi.org/10.1111/j.1469-7998.1928.tb07138.x
20. Falk D, Zollifkofer CPE, Morimoto N, Ponce de Leon MS. Metopic suture of Taung (*Australopithecus africanus*) and its implications for hominin brain evolution. Proc Natl Acad Sci USA. 2012;109:8467–8470. https://doi.org/10.1073/pnas.1119752109
21. Holloway RL, Broadfield DC, Carlson KJ. New high-resolution computed tomography data of the Taung partial cranium and endocast and their bearing on metopism and hominin brain evolution. Proc Natl Acad Sci USA. 2014;111:13022–13027. https://doi.org/10.1073/pnas.1402905111
22. Holloway RL. Australopithecine endocast (Taung specimen, 1924): A new volume determination. Science. 1970;168:966–968. https://doi.org/10.1126/science.168.3934.966
23. Neubauer S, Gunz P, Weber GH, Hublin J-J. Endocranial volume of *Australopithecus africanus*: New CT-based estimates and the effects of missing data and small sample size. J Hum Evol. 2012;62:498–510. https://doi.org/10.1016/j.jhevol.2012.01.005
24. Keith A. The fossil anthropoid ape from Taungs. Nature. 1925;115:234–235. https://doi.org/10.1038/115234a0
25. Keith A. New discoveries relating to the antiquity of man. New York: W.W. Norton and Company; 1931.
26. Elliot Smith G. The fossil anthropoid ape from Taungs. Nature. 1925;115:235. https://doi.org/10.1038/115234a0
27. Holloway RL, Broadfield DC, Yuan MS. The human fossil record. Volume 3: Brain endocasts – the paleoneurological evidence. New York: John Wiley and Sons; 2004. https://doi.org/10.1002/0471663573
28. Falk D, Clarke R. Brief communication: New reconstruction of the Taung endocast. Am J Phys Anthropol. 2007;134:529–534. https://doi.org/10.1002/ajpa.20697
29. Holloway RL, Broadfield DC. Technical note: The midline and endocranial volume of the Taung endocast. Am J Phys Anthropol. 2011;146:319–322. https://doi.org/10.1002/ajpa.21570
30. De Miguel C, Henneberg M. Variation in hominid brain size: How much is due to method? Homo. 2001;52:3–58. https://doi.org/10.1078/0018-442X-00019
31. Falk D. Hominid paleoneurology. Ann Rev Anthropol. 1987;16:13–30. https://doi.org/10.1146/annurev.an.16.100187.000305

32. Conroy GC, Falk D, Guyer J, Weber GW, Seidler H, Recheis W. Endocranial capacity in Sts 71 (*Australopithecus africanus*) by three-dimensional computed tomography. Anat Rec. 2000;258:391–396. https://doi.org/10.1002/(SICI)1097-0185(20000401)258:4%3C391::AID-AR7%3E3.0.CO;2-R

33. Ashton EH, Spence TF. Age changes in the cranial capacity and foramen magnum of hominoids. Proc Zool Soc Lond. 1958;130:169–181. https://doi.org/10.1111/j.1096-3642.1958.tb00567.x

34. Tobias PV. The brain in hominid evolution. New York: Columbia University Press; 1971. https://doi.org/10.5962/bhl.title.15880

35. Grine FE. The alpha taxonomy of *Australopithecus africanus*. In: Reed KE, Fleagle JG, Leakey RE, editors. The paleobiology of *Australopithecus*. Vertebrate paleobiology and paleoanthropology. Dordecht: Springer; 2013. p. 73–104. https://doi.org/10.1007/978-94-007-5919-0_6

36. Tawane GM, Thackeray JF. The cranium of Sts 5 ('Mrs Ples') in relation to sexual dimorphism of *Australopithecus africanus*. S Afr J Sci. 2018;114, Art. #a0249, 4 pages. http://dx.doi.org/10.17159/sajs.2018/a0249

37. Pffister H. Die kapazität des schädels (der kopfhöhle) beim säugling und älteren kinde [The capacity of the skull (the head cavity) in infants and older children]. Monatsschr Psychiatr Neurol. 1903;13(6):577–589. German. https://doi.org/10.1159/000219469

38. Berry RJA. Brain and mind: Or, the nervous system of man. New York: The MacMillan Company; 1928.

39. Biggerstaff RH. Time-trimmers for the Taungs child, or How old is '*Australopithecus africanus*'? Am Anthropol. 1967;69:217–220. https://doi.org/10.1525/aa.1967.69.2.02a00110

40. Wolpoff MH, Monge JM, Lampl M. Was Taung human or an ape? Nature. 1988;335:501. https://doi.org/10.1038/335501a0

41. Mann AE. Paleodemographic aspects of the South African australopithecines. Philadelphia, PA: University of Pennsylvania; 1975.

42. McFarlin SC, Barks SK, Tocheri MW, Massey JS, Eriksen AB, Fawcett KA, et al. Early brain growth cessation in wild Virunga mountain gorillas (*Gorilla beringei beringei*). Am J Primatol. 2013;75:450–463. https://doi.org/10.1002/ajp.22100

43. Herndon JG, Tigges J, Anderson DC, Klumpp SC, McClure HM. Brain weights throughout the lifespan of the chimpanzee. J Comp Neurol. 1999;409:567–579. https://doi.org/10.1002/(SICI)1096-9861(19990712)409:4%3C567::AID-CNE4%3E3.0.CO;2-J

44. DeSilva JM, Lesnick JM. Chimpanzee neonatal brain size: Implications for brain growth in *Homo erectus*. J Hum Evol. 2006;51:207–212. https://doi.org/10.1016/j.jhevol.2006.05.006

45. DeSilva JM, Lesnick JJ. Brain size at birth throughout human evolution: A new method for estimating neonatal brain size in hominins. J Hum Evol. 2008;55:1064–1074. https://doi.org/10.1016/j.jhevol.2008.07.008

46. Marchand F. Ueber das Hirngewicht des Menschen [About human brain weight]. Leipzig: B.G. Teubner; 1902. German.

47. Abbott AH, Netherway DJ, Niemann DB, Clark B, Yamamoto M, Cole J, et al. CT-determined intracranial volume for a normal population. J Craniofac Surg. 2000;11:211–223. https://doi.10.1097/00001665-200011030-00002

48. Jolicoeur P, Baron G, Cabana T. Cross-sectional growth and decline of human stature and brain weight in 19th-century Germany. Growth Dev Aging. 1988;52:201–206.

49. Cofran Z. Brain size growth in wild and captive chimpanzees. Am J Primatol. 2018;80(7), e22876. https://doi.org/10.1002/ajp.22876

50. Rohatgi A. WebPlotDigitizer v. 4.1. 2018. Available from: https://automeris.io/WebPlotDigitizer

51. Cofran Z, DeSilva JM. A neonatal perspective on *Homo erectus* brain growth. J Hum Evol. 2015;81:41–47. https://doi.org/10.1016/j.jhevol.2015.02.011

52. Kretschmann H-J, Schleicher A, Wingert F, Zilles K, Löblich H-J. Human brain growth in the 19th and 20th century. J Neurol Sci. 1979;40:169–188. https://doi.org/10.1016/0022-510X(79)90202-8

53. Sokal RR, Rohlf FJ. Biometry. 4th ed. New York: W.H. Freeman and Company; 2012.

54. Beaudet A, Clarke RJ, De Jager EJ, Bruxelles L, Carlson KJ, Crompton R, et al. The endocast of StW 573 ('Little Foot') and hominin brain evolution. J Hum Evol. 2019;126:112–123. https://doi.org/10.1016/j.jhevol.2018.11.009

55. Holloway RL. Australopithecine endocasts, brain evolution in the Hominoidea, and a model of hominid evolution. In: Tuttle R, editor. The functional and evolutionary biology of primates. Chicago, IL: Aldine Atherton; 1972. p. 185–203. http://dx.doi.org/10.4324/9781315132129-8

56. Lockwood CA. Sexual dimorphism in the face of *Australopithecus africanus*. Am J Phys Anthropol. 1999;108:97–127. https://doi.org/10.1002/(SICI)-1096-8644(199901)108:1<97::AID-AJPA6>3.0.CO;2-O

57. Hrdlička A. The Taungs ape. Am J Phys Anthropol. 1925;8:379–392. https://doi.org/10.1002/ajpa.1330080402

58. Robinson A. The Taungs skull. Brit Med J. 1925;March 21.

59. Cofran Z. Brain size growth in *Australopithecus*. J Hum Evol. 2019;130:72–82. https://doi.org/10.1016/j.jhevol.2019.02.006

60. Zollikofer CPE, Ponce de Leon MS. Pandora's growing box: Inferring the evolution and development of hominin brains from endocasts. Evol Anthropol. 2013;22:20–33. https://doi.org/10.1002/evan.21333

61. McCarthy R. Revised estimates of Taung's brain size growth [data set]. Mendeley Data v1 2020. http://dx.doi.org/10.17632/wyfvwd4s22

A systematic analysis of doctoral publication trends in South Africa

AUTHORS:
Susan van Schalkwyk[1]
Johann Mouton[2]
Herman Redelinghuys[2]
Sioux McKenna[3]

AFFILIATIONS:
[1]Centre for Health Professions Education, Faculty of Medicine and Health Sciences, Stellenbosch University, Stellenbosch, South Africa
[2]Centre for Research on Evaluation, Science and Technology, Stellenbosch University, Stellenbosch, South Africa
[3]Centre for Postgraduate Studies, Rhodes University, Makhanda, South Africa

CORRESPONDENCE TO:
Susan van Schalkwyk

EMAIL:
scvs@sun.ac.za

DATES:
Received: 03 Feb. 2020
Revised: 23 Mar. 2020
Accepted: 26 Mar. 2020
Published: 29 July 2020

HOW TO CITE:
Van Schalkwyk S, Mouton J, Redelinghuys H, McKenna S. A systematic analysis of doctoral publication trends in South Africa. S Afr J Sci. 2020;116(7/8), Art. #7926, 9 pages. https://doi.org/10.17159/sajs.2020/7926

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Jenni Case

KEYWORDS:
publication, doctoral graduates, doctoral education, PhD, dissertations

FUNDING:
None

It is incumbent upon doctoral students that their work makes a substantive contribution to the field within which it is conducted. Dissemination of this work beyond the dissertation, whether whilst studying or after graduation, is necessary to ensure that the contribution does not remain largely dormant. While dissemination can take many forms, peer-reviewed journal articles are the key medium by which knowledge is shared. We aimed to establish the proportion of doctoral theses that results in journal publications by linking South African doctoral thesis metadata to journal articles authored by doctoral candidates. To effect this matching, a customised data set was created that comprised two large databases: the South African Theses Database (SATD), which documented all doctoral degrees awarded in South Africa (2005–2014), and the South African Knowledgebase (SAK), which listed all publications submitted for subsidy to the South African Department of Higher Education and Training (2005–2017). The process followed several iterations of matching and verification, including manual inspection of the data, in order to isolate only those records for which the link was established beyond doubt. Over the period under review, 47.6% of graduates, representing 22 of the 26 higher education institutions, published at least one journal article. Results further indicate increasingly higher publication rates over time. To explore whether the journal article identified was a direct product of the study, a similarity index was developed. Over 75% of records demonstrated high similarity. While the trend towards increasing publications by graduates is promising, work in this area should be ongoing.

Significance:

- In spite of increasing trends in publications by graduates, many are not disseminating their work, suggesting that significant bodies of research are potentially not being shared with the academic community and are therefore not contributing to the relevant discipline or field.
- This study provides baseline data from which a number of further investigations can be launched, such as exploring the extent to which doctoral candidates who are also academics are publishing their work; the factors that enable or constrain publication; the other avenues of dissemination used; and whether publishing or not publishing can serve as a proxy for the quality of the doctoral work.

Introduction

The South African Higher Education Qualifications Sub-Framework describes doctoral studies as needing to 'make a significant and original academic contribution especially at the frontiers of a discipline or field'[1]. Knowledge creation is an unconditional expectation of doctoral work.[2,3] However, this knowledge contribution will remain largely dormant and 'invisible' if the work is not disseminated beyond the student, the supervisor and the examiners, limiting the opportunity for sharing the academic contribution with others.[4] While institutional e-repositories have made access to dissertations easier than before, most of the work documented in the thesis will require additional forms of dissemination if it is indeed to contribute to the frontiers of the relevant discipline or field. One could even argue that there is a moral obligation on doctoral graduates to ensure that their research is readily accessible to other researchers. Doctoral study is the most highly subsidised higher education qualification and taxpayers have a right to benefit from the potential value and contributions of such research. Implicit in conducting research at postgraduate level, therefore, is the notion of dissemination, typically through publication in an academic journal or similar artefact.[5] While dissemination may take many forms – such as patents, community workshops, news articles and so on – accredited conference proceedings and journal articles remain the key means by which knowledge is shared and thereby cumulatively built.

The South African Higher Education Qualifications Sub-Framework further indicates that the doctorate should 'satisfy peer review and merit publication'[1]. Thus, the expectation of publishing some aspect of one's research more widely, is not only for ensuring dissemination of the knowledge, but also could be used as an additional measure of the quality of the thesis. Having a reliable estimate of the number of publications emanating from the doctorate over an extended time frame provides a proxy for understanding the extent to which the doctoral graduate translated their work into more accessible publication outlets as well as continued a specific line of research or not.

It has been argued that the African continent needs 'tens of thousands more PhDs'[6]. In South Africa, there is a national mandate to significantly increase the number of PhD graduates to 100 per million of the population by 2030.[7] In various policy documents – the National Research and Development Strategy of 2002[8] and the Ten-year Innovation Plan of 2008[9] – the target is 5000 PhDs per year by 2030. In a recent report by SciSTIP[10] on *The State of the South African Research Enterprise*, it is shown that this target is in fact achievable under certain conditions, and given the current upward trajectory. Data show that slightly over 3300 students graduated with a doctoral degree in 2018, a sharp increase from the 1420 graduates in 2010 and 973 in 2000.[10]

Debates about the quality of a doctorate play out quite differently from one country to the next. In many Global North countries, the PhD agenda has become one of employability and relevance for industry as the number of PhDs produced far exceeds the number required by academia.[10] But, in South Africa, a significant proportion of doctoral candidates (around 35%) are in fact academics[6] who pursue doctoral studies as part of their 'training for an academic career'[1]. The demand for academics to obtain a doctoral degree is emphasised when one keeps in mind that

2018–2019 Department of Higher Education and Training data indicate that only 45% of permanent academic staff in South Africa have PhDs[11] while the National Development Plan goal is for 75% of academics to have doctorates by 2030[7].

As alluded to above, South Africa has witnessed a steady and, since 2008, rapid growth in the number of both doctoral enrolments and doctoral graduates. There probably are a number of different drivers for this growth, including the significant subsidy paid to institutions by the DHET for doctoral study both in terms of teaching input (enrolment) and research output (graduates). The extent of the increase in the numbers of doctoral graduates and the commensurate increase in the subsidy amounts paid since the promulgation of the 2005 funding formula are evident in Table 1.

Table 1: Subsidies generated through graduation of doctorates in South Africa (2005–2017)

Academic year	Number of doctoral graduates	Subsidy amount (ZAR)
2005	1189	303 287 742
2006	1100	291 779 400
2007	1274	392 148 666
2008	1182	415 392 624
2009	1380	528 421 320
2010	1421	508 708 053
2011	1576	562 759 656
2012	1878	648 202 968
2013	2051	696 421 152
2014	2258	736 286 382
2015	2530	813 814 980
2016	2782	923 610 090
2017	3040	1 111 463 520

Source: Table compiled from data provided by the South African Department of Higher Education and Training (DHET)

Other drivers for increasing doctoral numbers could include the setting of national targets, and the restructuring of higher education institutions in South Africa that has seen, for example, universities of technology mount multiple initiatives to encourage academic staff to obtain their doctorates. Either way, the rapid increase since 2005 (compound average growth rate of 7.7%) has been accompanied by growing concerns about the quality of our doctoral graduates and their theses. The increase in the quantum of doctoral enrolments has placed strain on the capacity of the system to supervise such students as the ratio of enrolled students to staff with PhDs increased to 2.5 to 1 (22 572 students that could potentially be supervised by the 9032 staff with doctoral degrees). It is worth pointing out that this ratio is an average across all fields and universities. Because doctoral enrolments are not evenly distributed, but concentrated in certain fields and especially concentrated in the more research-intensive universities, this ratio in some cases is much higher. This means that the 'burden' of doctoral supervision[6] has increased significantly over the past decade. Together with a commensurate increase in the burden of master's supervision, it is not surprising that concerns over quality have arisen.

While measures such as external examination are meant to ensure that the doctoral graduates the country produces are of the appropriate standard, anecdotal evidence suggests that quality in doctoral education is already being compromised in some instances. In the first round of institutional audits undertaken by the Council on Higher Education between 2005 and 2012, for example, concerns were raised about supervisors being used as examiners and about the incidence of repeated use of a very small pool of examiners for numerous theses. This has led to many universities improving their quality assurance procedures in order to ensure that good practice in the appointment of examiners is followed. The current national review of the doctorate by the Council on Higher Education suggests that the quality of the doctorate remains an area of concern and requires scrutiny.[12]

The obvious approach to assess the quality of doctoral education in a system would be to review examiners' reports, the names and reputations of examiners as well as the records of the decisions of higher degree committees when awarding doctoral degrees. However, as far as we know, no such study has been done in South Africa and access to examiner reports is extremely difficult to obtain. Instead we analysed the links between doctoral graduation and publication. We report on the first part of this investigation here. Our aims were modest – namely to attempt to estimate what proportion of doctoral theses result in journal publications. Specifically, we focused on doctoral students who graduated between 2005 and 2014 and who published the results of their studies in any of the publications appearing on the lists of journals as accredited by the Department of Higher Education and Training (DHET) between 2005 and 2017. As far as we are aware, this is the first systematic and comprehensive study aimed at establishing what proportion of 'materials' contained in South African doctoral theses eventually found their way into peer-reviewed journals. Our research questions address whether this proportion has increased over time, and whether there are significant differences in these proportions by university and gender of graduate. In the next phase of the investigation we intend to do further analyses around the different types of dissemination strategies related to such journal outputs as well as the quality of the journals in which articles appear.

Methodology

The single biggest methodological challenge of our study was to link South African doctoral thesis metadata to journal articles authored by the doctoral candidates. As there is no national database that contains these data, we had to create a customised data set for the analysis by linking two databases housed at the Centre for Research on Evaluation, Science and Technology (CREST) at Stellenbosch University: the South African Theses Database (SATD) and South African Knowledgebase (SAK).

Over the past 5 years, CREST has been building the SATD. It is a dedicated bibliographic database of doctoral dissertations submitted at South African universities since 2000. The data have been extracted from institutional repositories and the South African National Research Foundation (NRF) database of theses and dissertations. Fields captured within the SATD are:

- Thesis title
- Doctoral candidate: surname, initials and first name
- Granting university and year of thesis
- Supervisor(s): surname, initials and first name
- Field of study (not always possible/often only department)
- Abstract (full abstract where available)
- URL (handle) to the actual depository address where the thesis is stored

At the time of conducting our analysis, the SATD included metadata on 23 547 doctoral theses for the period 2000 to 2017. This figure represented a 92.7% coverage of the 25 390 doctoral degrees awarded by South African universities over this period.[13]

The second database, the SAK, is CREST's proprietary database of scientific publications authored by South African academics and scholars. SAK is unique in several respects. Firstly, SAK contains metadata on all scientific articles that earned subsidy for South African universities under the DHET Funding Framework. This means that it includes published articles that appear not only in the Web of Science (WoS), but also in other indexes and lists, including the DHET list of South African journals (not indexed in the WoS), the Proquest International Bibliography of the Social Sciences (IBSS) list, the Norwegian list and Scopus. As a result, we are able to augment the bibliometric analyses with more detailed analyses of output in different indexes as well as to compare the numbers of outputs in national versus international journals.

Specifically, SAK includes:

- Title of the document (article/proceedings/book/chapter)
- Author(s) (surname, initial and first name)
- Source information (journal/publisher/publication year)

- Author demographics (race, gender, nationality, year of birth, rank)
- Indexing (journal list: WoS, Scopus, DHET, IBSS, SciELO)
- Scientific field (four levels of increasing detail)

Although the coverage for each of these variables is not equally good, we deem it to be sufficient to present general trends for each of the variables. Our coverage of these variables in all cases varies between 80% and 90% depending on the variable, allowing us to draw relatively robust conclusions from these analyses. The current version of SAK comprises 426 496 authorships made up of journal articles (82%), conference proceedings (13%), and books and book chapters (5%) that were submitted for subsidy to the DHET for the period 2005 to 2017.

Preparing the two databases for analysis required intensive work over a significant period of time as researchers at CREST spent many months 'cleaning' the data, and seeking to fill in missing information – a process that continued into the matching period as ongoing engagement with the data identified anomalies and gaps that needed to be addressed.

Ultimately, we decided to confine our analysis to doctoral theses granted by South African universities between 2005 and 2014 and these were then compared to all journal articles in SAK for the period 2005 to 2017. This allowed for the counting of publications for a period of between 12 years (for the 2005 graduates) and 3 years (for the 2014 graduates) after the date of graduation.

Matching records in the SATD and the SAK

Our process of matching the data in the two databases was as follows. The first step was to match the records of doctoral theses to the journal articles based on the surnames of graduates and first authors, and on the similarity between the graduating institution (in SATD) and the institution that submitted the request for publication subsidy (in SAK). Once this process had been completed, we needed to establish whether the thesis author (in SATD) and the article author (in SAK) were indeed the same person (this verification was required because of the many instances of the same surnames – especially very common surnames – in both databases). In order to establish a more verifiable link between the records in the two databases, we wrote an algorithm that assigned a similarity score between the thesis titles and the article titles (see Table 2 and Figure 1). This programmatic process was followed by a manual inspection of the results in order to isolate only those records for which the identity of the thesis author and article author had been established beyond doubt.

Despite this systematic and rigorous process, we should note two limitations. First, the likelihood of matching the records beyond reasonable doubt was reduced where graduates changed surnames (as a result, for example, of either marriage or divorce) or used a different surname, or even order of names, in their publications versus the full names provided on the cover of the theses. Second, using the university (graduating and submitting) as one identifier in the matching process could have excluded graduates who soon after graduation moved to another institution or were indeed already working at another institution at the time of graduation, and then published articles under the name of a university other than that from which they graduated. We do not view either of these issues as major weaknesses in our methodology. Given the size of the sample of theses and publications analysed, the impact of these limitations would be minimal and would simply mean that our estimates are (to a small degree) lower than the actual figures.

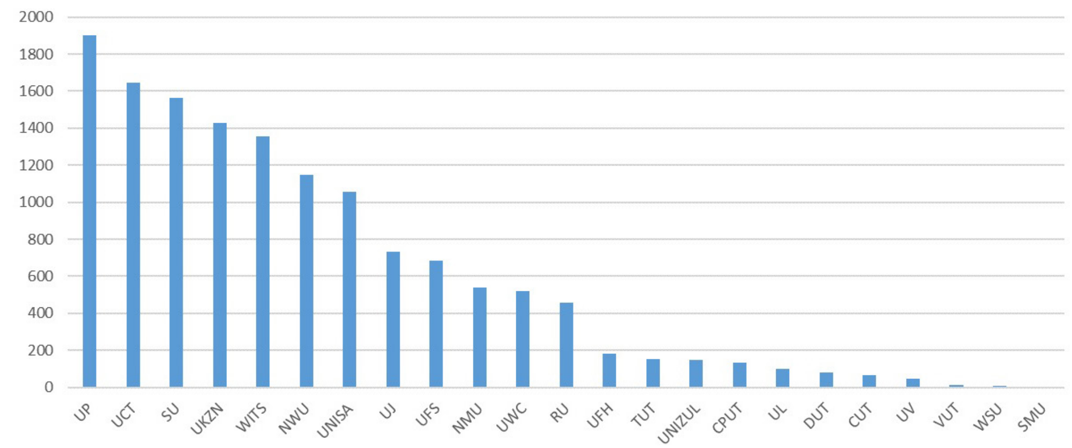
Our focus in this study is on doctoral graduates who published the findings or results of their doctoral studies in journals. We are well aware of the different publication practices across different fields. In the humanities and social sciences, for example, publications in books are deemed to be important. Fields such as the computer sciences, as well as certain subfields of engineering sciences, mathematics and economics, view conference proceedings as essential publication outlets. In the next phase of our investigation, we intend to investigate some of these other modes of publication as well. In some contexts, the dissemination of the knowledge takes place outside of academic fields altogether, through creative outputs, community workshops, policy briefs, patents and more. This is a limitation of the extent to which we can claim that publications in journals represent dissemination of knowledge developed through the doctorate, but the dominance of journal articles as the means of knowledge communication and the study's large sample size mitigate this limitation.

Results

At the time of conducting the analyses for this study, SATD contained information on 13 962 doctoral theses of students who graduated between 2005 and 2014, representing 22 of the 26 higher education institutions in South Africa (Figure 1).

The matching algorithm (Step 1) identified 51 864 *possible* authorships in SAK that could be linked to 7069 of these theses. The distribution of the similarity scores of these are summarised in Table 2.

The next step was to undertake a visual inspection of the 51 864 records and check whether the author of each publication (in SAK) was in fact the same person as the author of the doctoral thesis (the graduate listed in SAT). This process resulted in a reduced number of 44 073 records that could be definitively matched to a thesis author. The 44 073



UP, University of Pretoria; UCT, University of Cape Town; SU, Stellenbosch University; UKZN, University of KwaZulu-Natal; Wits, University of the Witwatersrand; NWU, North-West University; Unisa, University of South Africa; UJ, University of Johannesburg; UFS, University of the Free State; NMU, Nelson Mandela University; UWC, University of the Western Cape; RU, Rhodes University; UFH, University of Fort Hare; TUT, Tshwane University of Technology; Unizul, University of Zululand; CPUT, Cape Peninsula University of Technology; UL, University of Limpopo; DUT, Durban University of Technology; CUT, Central University of Technology; UV, University of Venda; VUT, Vaal University of Technology; WSU, Walter Sisulu University; SMU, Sefako Makgatho University

Figure 1: Number of graduates listed in the South African Theses Database per institution.

records thus identified corresponded to 6650 doctoral theses (Table 3). This produced the first main finding of our study, namely that over this time period, 47.6% of doctoral graduates published at least one journal article between 2005 and 2017.

It is important that we emphasise that the results presented in Table 3 refer to any publication that we could accurately link to a specific doctoral graduate (thesis). Whether the publication (journal article) was a direct *product* of the doctoral study and clearly based on the doctoral thesis is a second question which needs to be addressed.

Table 2: Breakdown of similarity index

Range of similarity score	Count of records	Share
1	66	0.13%
0.00 to 0.1	24 658	47.54%
0.1 to 0.2	14 211	27.40%
0.2 to 0.3	6747	13.01%
0.3 to 0.4	3248	6.26%
0.4 to 0.5	1510	2.91%
0.5 to 0.6	757	1.46%
0.6 to 0.7	331	0.64%
0.7 to 0.8	185	0.36%
0.8 to 0.9	110	0.21%
0.9 to 1	41	0.08%
Grand total	**51 864**	100.00%

Table 3: Proportion of doctoral graduates linked to journal articles by year (2005–2014)

Year	Number of graduates (total)	Number of published graduates	Published graduates as a percentage of all graduates
2005	1118	389	34.8%
2006	1088	433	39.8%
2007	1313	512	39.0%
2008	1289	614	47.6%
2009	1317	600	45.6%
2010	1382	701	50.7%
2011	1425	752	52.8%
2012	1570	836	53.2%
2013	1683	927	55.1%
2014	1777	886	49.9%
Total	**13 962**	**6650**	**47.6%**

The journal articles identified through the matching procedure could in fact have been produced before, during or after the completion of the doctoral degree. In order to establish the proportions of articles that were produced before, during or after the student's thesis was completed, we subsequently recoded the publication dates of the journal articles into three time periods. In our recoding we decided to use the following decision-rules:

- Articles with a publication date of between 3 years and 7 years before the thesis completion date were coded as 'before'.
- Articles with a publication date of 3 years or less before the thesis completion date were classified as 'during'.
- Articles with a publication date after the thesis publication date were coded as 'after'.

This classification was not an entirely arbitrary decision as the average time to degree of doctoral studies in South Africa over the past two decades has been about 4 years. This average duration of time to doctoral degree varies by field, for which we have not corrected in these analyses as our focus in this paper is on general trends in doctoral publication. The decision to define 'during' as thesis year minus 3 years also meant that we only included theses with a publication date of 2008 and later in Table 4.

In addition, we set one further parameter to all our queries: we decided to present the results for 4-year constant rolling windows according to the publication date of the thesis. Setting the time window of the thesis publication year at a constant 4-year rolling window period allowed us to use the thesis publication year as the reference year for all other comparisons over time.

The results presented in Table 4 confirm a central hypothesis of our study as far as linked theses are concerned, namely that doctoral students are publishing from their theses at increasingly higher rates. This trend is clearly illustrated in the steady increase in the proportion of articles that appeared *during* the production of the thesis: from 28% in the earliest period to 39% for the most recent period (2011 to 2014). The decline in the proportion of articles after thesis publication (from 69% for the earliest period to 52% for the most recent period) is the result of the longer time period since the thesis publication date in the early years.

The increased funding by the DHET of journal articles – especially since 2005 – and the subsequent push at higher education institutions for increased publications[10] are the most plausible drivers of these increases. Another explanation for the increasing number of published studies might be the growing use of the 'PhD by publication' model whereby journal articles comprise part of the doctorate itself.[14] Recent work in this area conducted at one institution (2008–2014) suggested that this model has been adopted for approximately 26% of all theses.[15]

Table 4: Linked journal articles according to publication date

Thesis publication year	Number of linked theses	Number of published articles				% Before	% During	% After
		Before	During	After	Grand total			
2008–2011	2667	545	5128	12 575	18 248	3%	28%	69%
2009–2012	2889	987	5763	11 918	18 668	5%	31%	64%
2010–2013	3216	1440	6781	11 232	19 453	7%	35%	58%
2011–2014	3401	1771	7489	10 006	19 266	9%	39%	52%

Table 5: Ratio of articles to theses (2008–2014)

Year	Published graduates	Articles published (during)	Ratio of article to theses (during)	Articles published (during and after)	Ratio of articles to theses (during and after)
2008	614	1242	2.02	2624	4.27
2009	600	1088	1.81	2424	4.04
2010	701	1277	1.82	2676	3.82
2011	752	1521	2.02	3212	4.27
2012	836	1877	2.25	3916	4.68
2013	927	2106	2.27	4179	4.51
2014	886	1985	2.24	3667	4.14
Total	**6650**	11 096	2.09	**26 905**	**4.05**

Although Table 4 shows that an increasing number of doctoral graduates publish from their thesis (the 'during' category), we were also interested to establish whether this finding meant that the average number of articles per thesis had increased over this time period. Table 5 presents the results of these analyses. It is interesting that the results show no significant increase in the average number of articles per thesis over this period: whether one focuses on only those articles that appeared during the generation of thesis (average number of articles of just over 2.1) or on those articles which were published both 'during' and 'after' (up to 3 years after the graduate date, average of 4.1). This finding is interesting because it shows that although there has been an increase in the recent past of doctoral students publishing from their theses (Table 3), this increase does not mean that the average student who publishes from their thesis is generating more articles from their thesis than before. Stated differently: the increase in the number of articles that we have recorded linked to doctoral theses is due not to doctoral students on average becoming more productive, but simply that a larger proportion of doctoral students is in fact publishing from their theses.

The general trends observed above hide large differences in the publication practices of our sample. We give two examples below to illustrate this point. In both examples we have authors who have published before, during and after the completion of their theses.

The first example (Figure 2) is the profile of a student who published at least five journal articles in the period preceding the study years (2005 to 2009); five during the thesis years (2010–2013) and three subsequent to the completion of the thesis. The three articles highlighted in blue are seemingly not directly related to the thesis topic (although this is subject to further validation). The remaining articles – before, during and after – form a clearly cohesive collection of papers with a similar specialisation.

The second example (Figure 3) is the profile of a person whose thesis was published in 2012. This is also an example (similar to the example above) in which the candidate had published articles in a specific field before completion of the thesis. In fact, inspection of the actual thesis and the publication profile as illustrated below, shows the typical profile of a doctoral student who conducted a PhD by publication. All the titles highlighted in green appear as chapters in the thesis. This is clearly the profile of a scholar who had published in their field of specialisation for some time before writing up the thesis as well as continued to disseminate the results of the thesis after the degree was awarded. This is very often the profile of scientists in the natural and life sciences where the knowledge and insights produced over an extended period is cumulative in nature.

These two examples show how the profiles of publishing doctoral graduates can differ. Our very basic classification of articles ('before', 'during' and 'after') clearly hides deeper issues around differences in publication practices. We currently are investigating a more comprehensive typology of these variants.

In the remainder of the article we disaggregate our general findings further. In all of these analyses we revert to the total sample of articles linked to theses irrespective of whether these appeared before, during or after the thesis publication year. Including the 'before' category (which is the smallest category of publications) is justified for our focus now shifts away from the question of whether we have seen a clear increase in doctoral publication productivity (which has been sufficiently demonstrated above) to questions related to different publication practices at the institutional level (breakdown by university below) and at the individual level (breakdown by productivity and gender). We have excluded Walter Sisulu University (5 of their 7 doctoral graduates published articles), Sefako Makgatho University (3 of their 5 graduates

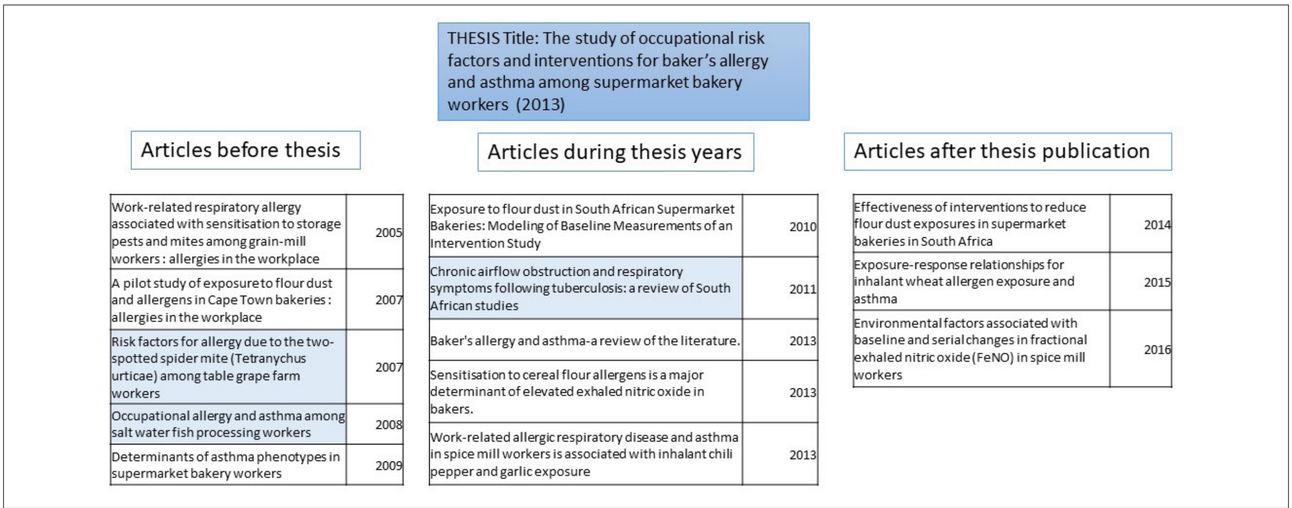


Figure 2: Illustrative example 1 of study and publication trajectory.

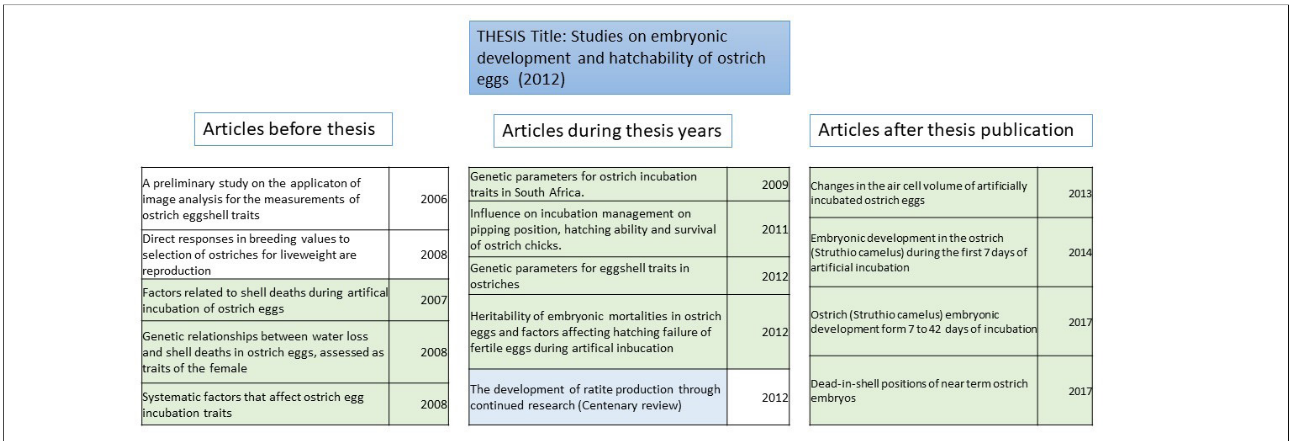


Figure 3: Illustrative example 2 of study and publication trajectory.

published) and Vaal University of Technology (6 of their 11 graduates published) from these tables as the numbers are too small for sensible percentages to be generated.

We first present a breakdown of the percentage of doctoral graduates who published before, during or after their doctorate by university (Table 6). Given the large range in the distributions by universities, we have split the table into two to distinguish between those universities with larger samples of publishing graduates (more than 200) and those with much smaller samples of graduates (between 30 and 200).

Table 6: Percentage of doctoral graduates per institution who have published (2005–2017)

University (2005–2014)	Number of graduates	Number of graduates who have published	Percentage of graduates who have published (in descending order)
n = 200+			
Rhodes University	455	266	58.46%
Stellenbosch University	1565	876	55.97%
University of Cape Town	1645	892	54.22%
University of KwaZulu-Natal	1426	762	53.44%
University of the Witwatersrand	1355	702	51.81%
North-West University	1146	579	50.52%
University of Pretoria	1900	864	45.47%
University of the Free State	682	295	43.26%
University of the Western Cape	518	217	41.89%
University of Johannesburg	731	299	40.90%
Nelson Mandela University	539	207	38.40%
University of South Africa	1056	295	27.94%
n < 200			
Tshwane University of Technology	151	93	61.59%
Cape Peninsula University of Technology	132	72	54.55%
University of Fort Hare	183	78	42.62%
University of Venda	45	19	42.22%
Central University of Technology	68	25	36.76%
Durban University of Technology	82	30	36.59%
University of Limpopo	99	34	34.34%
University of Zululand	150	31	20.67%

The results (excluding the three universities with too small an output), reveal two interesting trends:

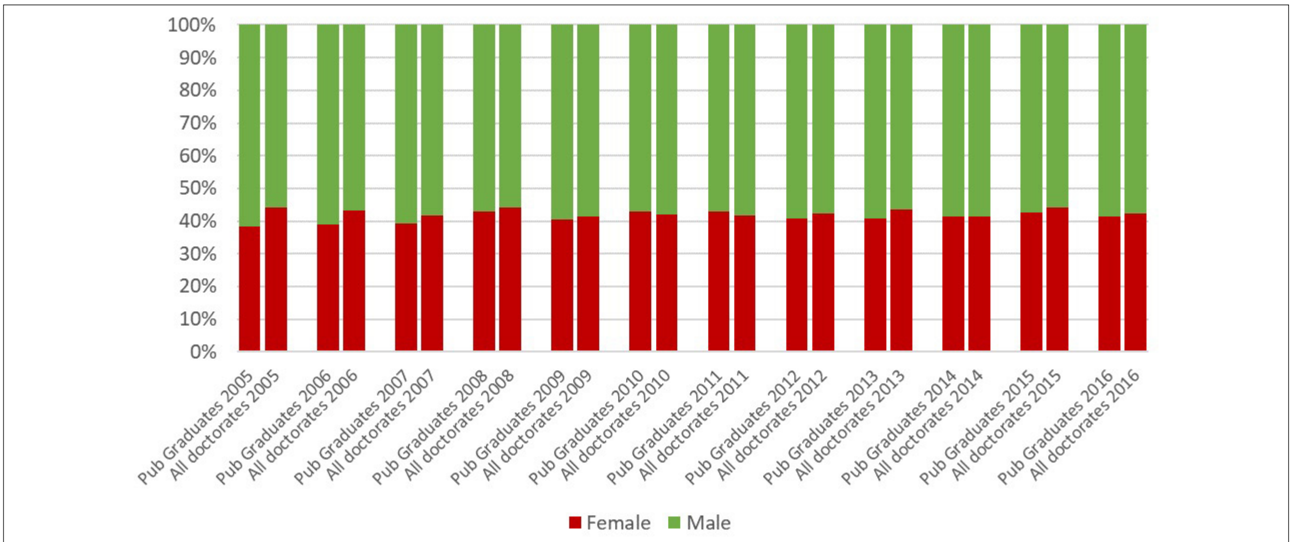
- The proportion of publishing graduates is typically highest at the traditional universities (higher than the average of 46% for the system).
- It is generally the case that doctoral graduates at the universities of technology publish from their PhDs at much lower rates (Tshwane and Cape Peninsula Universities of Technology being the exceptions).

Publication rates thus vary significantly by institution. Analysis at the level of the individual graduate also shows variance (Table 7), with 25% of graduates publishing only one article, and a significant group (41.56%) publishing from two to five articles. At the other end of the scale is a small group of 16 significant outliers who published more than 100 articles each. All of these 16 were, at the time of conducting this analysis, full-time academics who were affiliated with nine different universities; five were women, seven were black researchers, ten worked in medicine and the health sciences, only one came from the social sciences and none came from the humanities. It needs to be emphasised that the numbers in Table 7 relate only to the 6650 graduates to whom we could link publications.

Table 7: Publication rates per published graduate (before, during or after graduation within the 10-year period)

Number of articles	Number of published graduates	Percentage of published graduates
100+	18	0.27%
51–100	47	0.71%
31–50	127	1.91%
21–30	190	2.86%
11–20	676	10.18%
6–10	1112	16.74%
2–5	2779	41.84%
1	1692	25.47%

Further analyses of those doctoral graduates who have highlights the trends in terms of gender (Figure 4) and journal index (Figure 5). The disaggregation by gender shows that there is no significant difference between the proportion of articles authored by men and that authored by women in comparison to the gender distribution of all doctoral graduates for the same year.



Source: Higher Education and Training Management Information System (HEMIS)

Figure 4: Published doctoral graduates disaggregated by gender and year compared to total doctoral graduates disaggregated by gender and year.

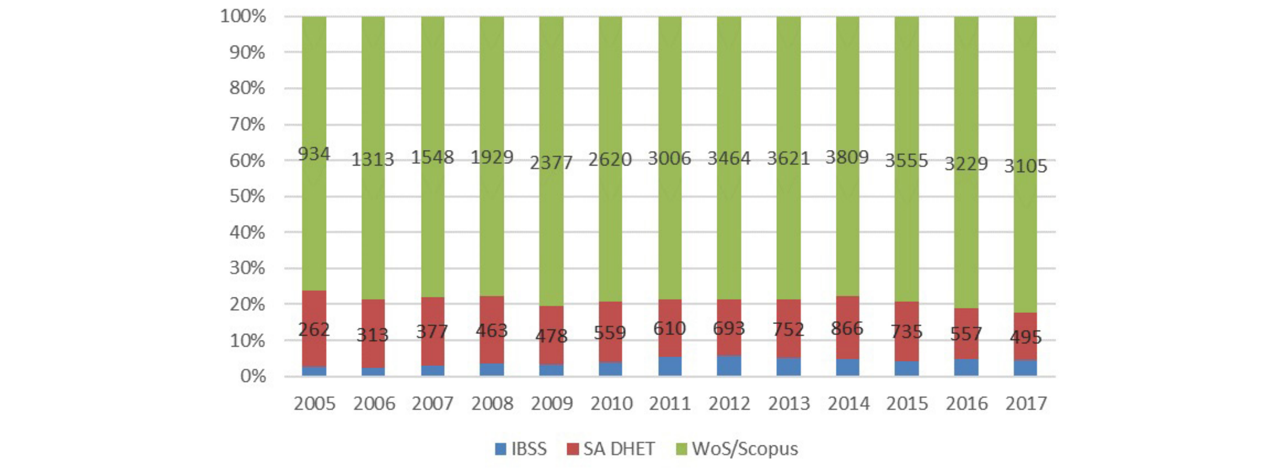


Figure 5: Articles disaggregated by journal index.

Finally, Figure 5 provides a breakdown of the publications by doctoral graduates according to the journal indexes and lists in which the relevant journals feature. For the purposes of this analysis, all articles were coded into three categories:

- Articles that appeared only in local DHET-accredited journals (and not in any other journal databases) were coded as 'SA DHET'.
- Articles that appeared only in the ProQuest IBSS list of journals (and not in any other journal databases) were coded as 'IBSS'.
- Articles that appeared in either the Web of Science and/or Scopus citation databases (and not in any other journal databases) were coded as 'WOS/Scopus'.

There has been some speculation that the pressure to publish that characterises many higher education institutions, both in South Africa and beyond[16,17], would lead academics and, by implication, also our doctoral students or graduates to publish more in local journals (DHET) or IBSS journals rather than in journals that are indexed in the two citation databases ([CA]Web of Science or Scopus). However, the evidence (Figure 5) shows a small and steady increase in the proportion of articles published in journals in the WOS or Scopus citation databases: from 76% in 2005 to 82% in 2017. This finding is important as it suggests that the increase in the rate of publications from PhD theses did not occur at the expense of publishing in more visible international journals.

Discussion

We present an analysis of journal article publications by doctoral graduates in South Africa across a specific time frame to provide a picture of the publication rates of these graduates and, in so doing, consider what this picture might indicate about the quality of doctoral work. While we found that approximately 52.2% of doctoral graduates could not be linked to any DHET-accredited publication, there is evidence of a trend toward increasing publication by graduates over the period under scrutiny, which is promising, particularly given the increase in WOS/Scopus-indexed outputs. It is important, however, to consider what lies behind these findings, the context within which they occurred, and to reflect on what they mean for doctoral education going forward. As mentioned earlier, this paper presents the first level investigation and it is clear that work in this area should be ongoing, with many aspects requiring further investigation.

We have already alluded to the issue of subsidy generation for universities, and the concomitant pressure that institutions then place on supervisors and students to 'produce'. We have also emphasised the role that the 'PhD by publication' model may have played in increasing doctoral outputs. There could, of course, be several other factors that our data cannot expose. In their work on enhancing postgraduate supervision, for example, Nulty et al.[18] identified strengthening supervisor capacity, particularly at research-intensive universities, as a mechanism to grow postgraduate publication outputs. Over the past 10 years in South Africa, initiatives such as the Strengthening Postgraduate Supervision (www.postgraduatesupervision.com) and the Enhancing Postgraduate Environments (www.postgradenvironments.com) have intentionally focused on postgraduate supervision, possibly also contributing to the increase in outputs that we now observe. On the other hand, there may be supervisors who discourage publication during doctoral studies as it may distract the student from doctoral work[4], or alternatively coerce students into publication for their own gain[19]. Clearly silent in this study are the individual student voices which could inform us about what motivates towards publication, what might enable or constrain it.

There are also a number of caveats that need to be considered. For example, what percentage of doctoral graduates leaves the academic environment after graduation and in so doing moves away from any work-related expectation that they should publish? Graduates may be publishing in journals that are not DHET accredited, but that are specific to their areas of interest. Differentiation also plays out across the representation of fields in the publishing arena. Some fields are more difficult to publish in than others,[20] and some prefer different outputs such as music scores, books and portfolios – our study has not accounted for these.

At a more detailed level, there is evidence of unevenness across the system as the issue of institutional histories that characterise the South African higher education system manifests here. While some universities have such low numbers of doctoral graduates that using publications as a proxy for the quality of the doctorate is highly problematic, there are universities with more than 100 doctoral graduations over this period where very few manage to publish. Two institutions stand out in Table 6. At University of Zululand, only 20.7% of their 150 PhD graduates could be matched to any publications. At Unisa, only 27.8% of their 1056 PhD graduates could be matched to any publications. Overall, the institutional data need to be read with care and interpreted through the lens of our country's complex past.

There is also the issue of academics pursuing doctoral studies. For example, we know that approximately 35% of doctoral candidates are academics. For many academics, the doctorate represents a logical next step in their academic careers[3], undertaken against the backdrop of the 'publish or perish' narrative that has become pervasive in academic circles[16,17]. Further analysis is needed to establish how many of the country's doctoral graduates who have published are indeed academics, what factors have enabled or constrained their participation in publishing, and how some of these statistics might be shifted further upwards. Knowledge production and dissemination are central to the identity of an academic, as well as being key for promotion and for successfully securing grants that will fund future work.[3,17] Focusing specifically on the outputs of doctoral graduates who are in an academic role is also important given that these graduates are expected to take on the supervision of future doctoral students. How does one take on the mantle of research mentor if one is not actively involved in the publication of research?

As far as gender of the publishing doctorates is concerned, no significant difference was found between this cohort and the gender of all doctoral graduates. To put it differently: those doctoral graduates who published were in the same proportion of the gender distribution of all doctoral graduates over the same period. Further analysis by age of the graduate and scientific discipline may reveal deeper differences.

This analysis of doctoral publication outputs has provided a stratified overview of an extremely complex issue. We acknowledge that analysing 'publications' as a collective downplays the significant variance across the publication industry in terms of quality, reach, focus, and the like. Our data do not identify those graduates who could have in fact published in non-accredited or even predatory journals. Recent work in the area of predatory publishing has identified this as a significant area of concern, with 3.4% of articles published by South African authors in the period 2005–2014 being identified as having been published in such journals.[21] It can be assumed that doctoral graduates are included in this number.

At a practical level, this study highlights the need for better tracking of publications given that graduates change their names and institutional affiliations. Many journals now require authors to include their ORCiD. It is clear that encouraging the use of ORCiD on all theses would greatly facilitate future work in this area, and indeed many universities are already implementing this as a requirement for the master's and doctoral graduates as of 2019.

Ultimately, the study provides evidence of Lotka's Law whereby most of the research publications in a system or institution are typically produced by a relatively small group of highly active academics (see also Kamler[4]). Given that 75% of those doctoral graduates who do publish manage to publish two or more articles (Table 7), this study also suggests that if graduates can 'crack the code' and obtain one publication, they are likely to move on to more. Supporting doctoral candidates and new graduates to disseminate their work through publication seems to be an important endeavour, and yet it appears that mentorship to support doctoral publication is not common practice, particularly in the social sciences.[4] Given the enormity of responsibilities associated with supervision and the 'burden of supervision' in a context of scarce resources,[6] some may argue that it is unfair to expect doctoral supervisors to take on the work of inducting graduates into the processes of writing for publication. However, as we have shown, the description of the doctorate in the Higher Education Qualifications Sub-Framework[1] makes it clear that this is indeed central to doctoral education as 'training for an academic career' through the production of knowledge that should 'merit publication'.

The study has opened up several avenues for further research, including more qualitative work, as suggested above. Future analysis of the data will also allow us to interrogate collaborations between doctoral scholars and their supervisors through co-authorships – an approach that has been identified as being enabling and generative for doctoral publication.[4] Such work could expose the 'back story' that further qualitative studies could illuminate. For example, there may be cases where the supervisor undertook extensive mentoring with regard to publication, but chose not to take co-authorship; or where the supervisor wrote the article on their own without any such mentorship and included the student as co-author in recognition for their generation and/or analysis of the data; and so forth. These issues all influence the publication landscape, directly impact on the lives of many academics and students, and therefore warrant our attention.

Finally, it should be noted that the findings from this study are not necessarily out of kilter with work conducted in other countries. In 2008, Lee and Kamler[22] observed similar trends in countries such as the UK, the USA and Australia. At the time, however, they called for doctoral programmes to be more intentionally structured to enable the dissemination of doctoral work, emphasising pedagogic practices that facilitate the graduate's entry into the disciplinary community such as co-authorship, and other writing related initiatives. As South Africa continues to strive towards strengthening research work at doctoral level, institutions should be encouraged to build support for publication throughout the doctoral journey across the initial years and beyond. Much can be learned from those fields, particularly in the natural sciences, where there is a tradition of such collaborative endeavours.

Acknowledgements

We acknowledge Annemarie Visagie and Kyle Ford at CREST who undertook a substantial part of the manual cleaning of the data in SATD and SAK.

Authors' contributions

S.v.S., S.M. and J.M. contributed to the conceptualisation of the study. S.v.S. led the writing process, and S.M. and J.M. contributed to the development of the manuscript from draft to finalisation. H.R. was responsible for data analysis and curation.

References

1. Government Gazette vol. 578 no. 36797. 2013 August 30. Available from: http://www.saqa.org.za/docs/pol/2013/gfetqs_heqsf.pdf
2. Frick BL, Albertyn R, Brodin E, McKenna S, Claesson S, Fourie-Malherbe M, et al. The role of doctoral education in early career academic development. In: Fourie-Malherbe M, Albertyn R, Aitchison C, Bitzer E, editors. Postgraduate supervision: Future foci for the knowledge society. Stellenbosch: Sun Media; 2016. p. 203–219.
3. Jackson D. Completing a PhD by publication: A review of Australian policy and implications for practice. High Educ Res Dev. 2013;32(3):355–368. https://doi.org/10.1080/07294360.2012.692666
4. Kamler B. Rethinking doctoral publication practices: Writing from and beyond the thesis. Stud High Educ. 2008;33(3):283–294. https://doi.org/10.1080/03075070802049236
5. Dinham S, Scott C. The experience of disseminating the results of doctoral research. J Further High Educ. 2001;25(1):45–55. https://doi.org/10.1080/03098770020030498
6. Cloete N, Mouton J, Sheppard C. Doctoral education in South Africa. Cape Town: African Minds; 2015.
7. South African National Planning Commission. National Development Plan: Vision 2030 [document on the Internet]. c2011 [cited 2019 Nov 12]. Available from: http://policyresearch.limpopo.gov.za/bitstream/handle/123456789/941/NDP%20Vision%202030.pdf?sequence=1
8. South African Department of Science and Technology (DST). South Africa's National Research and Development Strategy. Pretoria: DST; 2002. Available from: http://www.dst.gov.za/images/pdfs/National%20research%20%20development%20strategy%202002.pdf
9. South African Department of Science and Technology (DST). Innovation towards a knowledge-based economy: Ten-year innovation plan 2008–2018. Pretoria: DST; 2008. Available from: http://www.foresightfordevelopment.org/library/55/1373-innovation-towards-a-knowledge-based-economy-ten-year-plan-for-south-africa-2008-2018
10. Mouton J, Basson I, Treptow R, Blanckenberg J, Van Lill M, Boshoff N, et al. The state of the South African research enterprise. Stellenbosch: SciSTIP & CREST, Stellenbosch University; 2019. Available from: http://www0.sun.ac.za/scistip/wp-content/uploads/2019/08/state-of-the-South-African-research-enterprise.pdf
11. South African Department of Higher Education and Training. Annual report 2018–2019 [document on the Internet]. c2019 [cited 2020 Jan 15]. Available from: http://www.dhet.gov.za/Commissions%20Reports/DHET_Annual_Report_201819_WEB.pdf
12. McKenna S. South Africa takes steps to assure the quality of its doctorates. The Conversation. 2019 November 06. Available from: http://theconversation.com/south-africa-takes-steps-to-assure-the-quality-of-its-doctorates-125774
13. South African Department of Higher Education and Training. University education [webpage on the Internet]. No date [cited 2019 Nov 12]. Available from: http://www.dhet.gov.za/SitePages/UniversityEducation.aspx
14. Frick L. PhD by publication – panacea or paralysis? Afr Educ Rev. 2019;16(5):47–59. https://doi.org/10.1080/18146627.2017.1340802
15. Odendaal A, Frick L. Research dissemination and PhD thesis format at a South African university: The impact of policy on practice. Innov Educ Teach Int. 2018;55(5):594–601. https://doi.org/10.1080/14703297.2017.1284604
16. Paré A. Slow the presses: Concerns about premature publication. In: Aitchison C, Kamler B, Lee A, editors. Publishing pedagogies for the doctorate and beyond. London: Routledge; 2010. p. 30–46.

17. Carr N, Hayes S. An analysis of tourism PhD students' publication records against the background of 'publish or perish'. Anatolia. 2017;28(2):276–278. https://doi.org/10.1080/13032917.2017.1306715

18. Nulty D, Kiley M, Meyers N. Promoting and recognising excellence in the supervision of research students: An evidence-based framework. Assess Eval High Educ. 2009;34(6):693–707.

19. Li Y. 'Publish SCI papers or no degree': Practices of Chinese doctoral supervisors in response to the publication pressure on science students. Asia Pac J Educ. 2016;36(4):545–558. https://doi.org/10.1080/02188791.2015.1005050

20. Muller J. Forms of knowledge and curriculum coherence. J Educ Work. 2009;22(3):205–226. https://doi.org/10.1080/13639080902957905

21. Mouton J, Valentine A. The extent of South African authored articles in predatory journals. S Afr J Sci. 2017;113(7–8), Art. #2017-0010, 9 pages. https://doi.org/10.17159/sajs.2017/20170010

22. Lee A, Kamler B. Bringing pedagogy to doctoral publishing. Teach High Educ. 2008;13(5):511–523. https://doi.org/10.1080/13562510802334723

AUTHOR:
Jasper A. Slingsby[1,2]

AFFILIATIONS:
[1]Fynbos Node, South African Environmental Observation Network, Centre for Biodiversity Conservation, 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

CORRESPONDENCE TO:
Jasper Slingsby

EMAIL:
jasper@saeon.ac.za

DATES:
Received: 28 Nov. 2019
Revised: 14 Feb. 2020
Accepted: 28 Apr. 2020
Published: 29 July 2020

HOW TO CITE:
Slingsby JA. Forest restoration or propaganda? The need for Transparency and Openness Promotion (TOP) scores to uphold research integrity. S Afr J Sci. 2020;116(7/8), Art. #7684, 4 pages. https://doi.org/10.17159/sajs.2020/7684

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☐ All data included
☐ On request from authors
☐ Not available
☒ Not applicable

EDITOR:
Jane Carruthers

KEYWORDS:
objectivity, advocacy, scholarly publishing, public trust

FUNDING:
National Research Foundation (South Africa)

Forest restoration or propaganda? The need for Transparency and Openness Promotion (TOP) scores to uphold research integrity

In a time of environmental crisis and 'fake news', there are calls for scientists to engage in public debate or advocacy. Some are wary, fearing that revealing subjective views poses a risk to scientific credibility or erodes trust in scholarly publishing. Others are less concerned, seeing it as their duty to society or an opportunity to boost their profile. Ideally, we need better checks and balances that allow scientists to contribute to public discourse without fear of compromising the credibility of their science, while avoiding subjective views influencing the outcomes of peer-reviewed research. For better or worse, scientists have personal views. The question is not whether they should be condoned or condemned, but how they should be managed in the context of scholarly publishing to maximise benefits and minimise negative outcomes. Using the recent contention around global tree 'restoration' potential as an example, I propose we score journals and articles based on the Transparency and Openness Promotion (TOP) guidelines and associated criteria. A high TOP score means readers have sufficient access to information to assess the objectivity and credibility of scientific publications and their authors. I show that current practice provides very little access to information, and readers are essentially being asked to have faith in the scholarly publication system. We must do better.

Significance:

- Science is predicated upon objectivity, yet readers are rarely given enough information to assess the objectivity, and thus integrity, of peer-reviewed research.
- To address this issue, a scoring system is proposed, which is based on the principles of transparency and openness.
- Improving transparency and openness in scholarly publishing is essential for allowing readers to assess the objectivity of published research and researchers, growing public trust, and allowing researchers to engage in public debates without fear of loss of scientific credibility.

A recent publication with a simple message 'The global tree restoration potential'[1] has caused controversy and discomfort in the scientific community. Controversy, because commentaries by leaders in the field highlighted several assumptions or omissions, which they viewed as critical flaws[2-8], but these were largely disregarded by the authors and journal[9,10] (Table 1). Discomfort, because the authors are strongly advocating for the implementation of their research and aim 'to start a global movement'[11] – planting trees on a massive scale to mitigate CO_2 emissions. This despite the perceived flaws in their analysis, many known negative outcomes of afforestation,[12] and a perceived conflict of interest in being the beneficiaries of a USD17 million research grant from a foundation with a stated interest in forest restoration. High-profile publications with potential conflicts of interest are becoming increasingly common, and are challenging scientists to critically assess our role in advocacy and how to balance this against, or integrate it with, the way we do science. Here I use Bastin et al.[1] as an example to argue that we need greater transparency and openness in scholarly publishing to strike the balance between protecting the public from flawed science and protecting scientists from being ostracised for engaging with public issues.

The lack of systemic change in the use of fossil fuels and management of natural resources has increased calls for scientists to communicate their research, become advocates, or even activists, around the global climate and extinction crises.[13,14] Engaging with advocacy raises fears among scientists that their work will lose credibility, because revealing personal views may undermine the scientific objectivity of their research. These fears are unfounded and counterproductive. Scientific objectivity is a noble, but largely unattainable, ideal that is best approached by disclosing all assumptions and biases for others to assess.[15] While publicly airing personal views may incur costs to the individual researcher, there are also many potential gains and the opportunity to improve science in general. Few, if any, scientists do not hold personal views on their subject matter, and denial in any form rarely has positive outcomes. Acceptance and acknowledgement of subjective views can be positive for science as it allows reviewers, editors and readers to assess whether researchers' beliefs may have biased their analyses or findings. Unfortunately, this raises practical drawbacks in that it relies on the honesty of the researcher, and puts the burden on the journal and editors to call out any undue subjectivity. The system fails when sources of bias are not revealed, or where the checks and balances to detect and remedy undue subjectivity are insufficient.

The danger to society is when the facts are misrepresented or concealed to further an agenda – i.e. when what appears to be science or advocacy is actually propaganda. A case in point is the infamous Tobacco Wars, where tobacco companies used marketing, influence and undisclosed funding of scientists to obscure the truth and influence scientific and public debate around the health risks associated with cigarette smoke.[16] Similar approaches have been used to sow doubt about a range of important issues, including global warming.[17] There are a number of pathways by which research can be abused for propaganda, some of which involve dishonesty by various parties, while others rely on poor checks and balances. The production and communication of science includes linkages between backers (i.e. funders and other influences) and researchers, the transfer of manuscripts from researchers to journals for vetting

Table 1: Flaws, errors or omissions of Bastin et al.'s[1] analysis highlighted in technical comments and letters

Issue raised	Response by Bastin et al.
Confounding afforestation with restoration[3]	Argued that afforestation is restoration in all the areas where their model predicts there should be trees[9]
Grossly underestimating historical carbon emissions and overestimating tree sequestration potential by accounting for atmospheric CO_2 only[2,4]	Claimed that this was inconsequential to their argument[9]
Grossly overestimating carbon storage per forest area[2-4,6]	Claimed that this was due to differences in the definition of 'forest' used by the author and the commentaries[9]
Ignoring carbon in existing land cover[2-4,6]	Provided new methods explaining how existing carbon was taken into account. But these methods revealed that they used a limited set of values for different ecosystem types, which include assuming that tundra has 80% and savanna has 100% tree cover[9]
Ignoring feedbacks with albedo, atmospheric CO_2[2-4] and the water cycle[8]	Argued that exploring the effects of albedo was beyond the scope of their study and did not meaningfully comment on the water cycle[9,10]
Ignoring fire and herbivory[3]	Argued that these are included in the training data as they are from protected areas, although they are not explicitly included in the model[9]
Ignoring that their proposed 'solution' treats the symptom (atmospheric CO_2) not the cause (CO_2 emissions)[2-4]	Altered their abstract from reading 'global tree restoration as our most effective climate change solution to date' to 'global tree restoration as one of the most effective carbon drawdown solutions to date'[9]
Ignoring operational feasibility[5] or negative externalities relating to social fairness, water, biodiversity and other opportunity costs[3,7]	Argued that they are merely highlighting possibilities and not proposing actions and that balancing these trade-offs are not for them to decide[9,10]

and publication, and the communication of the findings to the public (including scientists). Perhaps the most common source of propaganda is the hijacking of communication to the public by self-interested parties and the misquoting or other abuse of honest, largely objective research in marketing or social media campaigns. Another pathway is corruption: when researchers and/or journals are dishonest and publish bias or fake science that furthers their own interests, views and agendas or those of their backers. A third pathway is when researchers and/or journals are manipulated by their backers or coerced into nefarious actions in the belief that they are being objective or contributing to a greater cause. This scenario is currently a real fear with the recent offer of USD1 billion in research funding from a tobacco company.[18] Embracing, rather than denying, the subjective views of researchers, editors and backers may actually provide the opportunity to formally improve research integrity and strengthen the checks and balances needed to identify sources of bias and potential propaganda.

Fear of subjectivity and propaganda in science is not new, and there have been several mechanisms put in place and refined over the decades to help reduce their prevalence and improve public trust in science. Perhaps the longest standing and best known are scholarly peer review and 'conflict of interest' statements, but these have deficiencies and are not applied in a consistent manner across journals. Peer review is predominantly performed behind closed doors, with no accountability, while conflict of interest statement requirements are highly varied, poorly reported and predominantly apply to financial interests only. Moves towards open peer review[19] and the expansion, standardisation and public registration of researcher conflict of interest statements[20] are positive moves in this regard. Additional refinements or additions could include establishing a code of ethics or peer review for press releases associated with the publication of articles, and conflict of interest statements for journals and funders that are lodged with discoverable registries, disclosing their funding sources and ideologies.

A recent move to improve our ability to assess the credibility of scientific contributions and their authors is the development of standards to promote a culture of transparency and open science.[21] These Transparency and Openness Promotion (TOP) guidelines are aimed at journal procedures and policies for publication and are increasingly and incrementally being adopted by journals. While many of these principles and standards are not yet implemented or enforced by journals, they can easily be voluntarily adopted and implemented by researchers engaged in science communication or advocacy to defend their credibility. Table 2 presents an approach for scoring the transparency of an article or journal based on applying the TOP guidelines and others based on peer review and the declaration of conflicts of interest. I have indicated scores for each criterion achieved by Bastin et al.[1] and *Science*, based on information available from the article and the journal website. The system allows scores to range from 0 (no transparency or openness) to 1 (maximum transparency and openness). This scoring system could be extended to authors by averaging the TOP scores of all their research outputs over a particular time window such as 2 or 5 years, as is done for the h-index. TOP scores align closely with, and provide a method to quantify, many of the principles proposed in the draft *Hong Kong Manifesto for Assessing Researchers: Fostering Research Integrity* presented at the recent 6th World Conference on Research Integrity in Hong Kong.[22]

While some criteria (e.g. preregistration) are often less feasible in ecology, the scores are generally low (Bastin et al. = 8/36 = 0.222; *Science* = 9/39 = 0.23). These scores are of concern because they are likely to be among the highest scores in ecology. Bastin et al.[1] went out of their way to make their analyses repeatable, while *Science* is one of the leading TOP journals. Together, this highlights that there is great room for improvement in the transparency and openness enforced by scientific journals in general. Until this happens, it is up to authors to go the extra mile to improve the TOP scores of their articles. While the TOP guidelines improve openness and repeatability, they do little to counter any subjectivity in the presentation or interpretation of results. This is where open peer review and improved disclosure of interests could make a telling contribution.

The TOP scoring exercise presented drives home that readers are really being asked to have faith in scientists and publishers, and are not given enough information to assess the objectivity, and thus credibility, of scientific publications, editors and authors. This is highly problematic, because in an era of fake news there is an increasing need for scientists to engage with public debate without threat to their credibility. There is also an increasing risk of scientific propaganda.

Whether you trust the science put forward by Bastin et al.[1] or agree with the approach they have adopted or not, get used to it – it is a model that is likely to become increasingly prevalent. The onus is on the scientific community to adopt and enforce principles and standards that ensure openness and transparency, allowing scientists to contribute to public discourse without fear of losing their credibility, but also rooting out and debunking propaganda. Finally, an additional advantage of greater transparency and openness is that as our philosophy of science evolves, such as becoming more inclusive of methods of knowledge generation and verification beyond the Western paradigm, we should have the materials available to assess and validate the record of research through a new lens.

Table 2: A transparency scoring system for articles and journals based on the standards and levels of the Transparency and Openness Promotion (TOP) guidelines, type of peer review, and disclosure of financial and ideological conflicts of interest. The minimum standards implemented by *Science* are indicated with an ***S***, while the score for the Bastin et al.[1] article is indicated with a ***B***.

TOP guidelines	Score			
	0	1	2	3
Citation standards	None or encouraged	Journal describes standards and rules in author guidelines	Article adheres to guidelines, but this is not a requirement for publication	Appropriate citation for data and materials provided ***B, S***
Data transparency	None or encouraged	Article states whether data are available	Data posted to a trusted repository ***B, S***	+ analyses reproduced independently prior to publication
Analytic methods (code) transparency	None or encouraged	Article states whether code is available	Code made available ***B, S***	+ analyses reproduced independently prior to publication
Research materials transparency	None or encouraged	Article states whether materials are available, and, if so, where to access them ***B, S***	Materials posted to a trusted repository. Exceptions allowed.	+ analyses reproduced independently prior to publication
Design and analysis transparency	None or encouraged ***B, S***	Journal articulates standards	Adherence to journal standards required	Adherence to journal standards required and enforced
Study preregistration	None or encouraged ***B, S***	Article states whether preregistration exists	+ allows journal access during peer review for verification	Preregistration required and link and badge in article to meeting requirements provided
Analysis plan preregistration	None or encouraged ***B, S***	Article states whether preregistration exists	+ allows journal access during peer review for verification	Preregistration required and link and badge in article to meeting requirements provided
Replication *[Not relevant to individual articles]*	None or discouraged	Encouraged ***S***	Encouraged and conducts results blind review	Uses Registered Reports for replication studies with peer review prior to observing the study outcomes
Additional guidelines				
Openness of peer review	None ***B, S***	Reviews reported	Reviews and reviewers reported	Transparent review: public can see manuscripts, reviews, reviewer names and author responses
Open disclosure of funders and potential financial and ideological conflicts of interest by authors	Funders and financial interests only ***B, S***	+ ideological and non-financial	+ in open online registry	+ in open online registry, time stamped and up to date
Open disclosure of funders and potential financial and ideological conflicts of interest by journal	None ***B, S***	Funders and financial interests	+ ideological and non-financial	+ in open online registry, time stamped and up to date
Open disclosure of funders and potential financial and ideological conflicts of interest by funders	None ***B, S***	Funders and financial interests	+ ideological	+ in open online registry, time stamped and up to date
Peer review of press release	None ***B, S***	Reviewed by journal staff	Peer-reviewed	Open peer review

Acknowledgements

Thanks to Nicky Allsopp, William Bond, two anonymous reviewers and the editor for constructive feedback on an earlier version of the manuscript. This work is based on research supported in part by the National Research Foundation of South Africa.

Competing interests

The author has no competing financial interests, but declares that he is ideologically opposed to the inappropriate afforestation of open ecosystems without due consideration of the trade-offs with water delivery, biodiversity and livelihoods.

References

1. Bastin J-F, Finegold Y, Garcia C, Mollicone D, Rezende M, Routh D, et al. The global tree restoration potential. Science. 2019;365(6448):76–79. http://dx.doi.org/10.1126/science.aax0848
2. Lewis SL, Mitchard ETA, Prentice C, Maslin M, Poulter B. Comment on 'The global tree restoration potential'. Science. 2019;366(6463), eaaz0388. http://dx.doi.org/10.1126/science.aaz0388
3. Veldman JW, Aleman JC, Alvarado ST, Anderson TM, Archibald S, Bond WJ, et al. Comment on 'The global tree restoration potential'. Science. 2019;366(6463), eaay7976. http://dx.doi.org/10.1126/science.aay7976
4. Friedlingstein P, Allen M, Canadell JG, Peters GP, Seneviratne SI. Comment on 'The global tree restoration potential'. Science. 2019;366(6463), eaay8060. http://dx.doi.org/10.1126/science.aay8060
5. Grainger A, Iverson LR, Marland GH, Prasad A. Comment on 'The global tree restoration potential'. Science. 2019;366(6463), eaay8334. http://dx.doi.org/10.1126/science.aay8334
6. Luedeling E, Börner J, Amelung W, Schiffers K, Shepherd K, Rosenstock T. Forest restoration: Overlooked constraints. Science. 2019;366(6463):315. http://dx.doi.org/10.1126/science.aay7988
7. Delzeit R, Pongratz J, Schneider JM, Schuenemann F, Mauser W, Zabel F. Forest restoration: Expanding agriculture. Science. 2019;366(6463):316–317. http://dx.doi.org/10.1126/science.aaz0705
8. Sheil D, Bargués-Tobella A, Ilstedt U, Ibisch PL, Makarieva A, McAlpine C, et al. Forest restoration: Transformative trees. Science. 2019;366(6463):316–317. http://dx.doi.org/10.1126/science.aay7309
9. Bastin J-F, Finegold Y, Garcia C, Gellie N, Lowe A, Mollicone D, et al. Response to Comments on 'The global tree restoration potential'. Science. 2019;366(6463), eaay8108. http://dx.doi.org/10.1126/science.aay8108
10. Bastin J-F, Finegold Y, Garcia C, Mollicone D, Rezende M, Routh D, et al. Forest restoration: Transformative trees – Response. Science. 2019;366(6463):317. http://dx.doi.org/10.1126/science.aaz2148

11. Irwin A. The ecologist who wants to map everything. Nature. 2019;573(7775):478–481. http://dx.doi.org/10.1038/d41586-019-02846-4
12. Bond WJ, Stevens N, Midgley GF, Lehmann CER. The trouble with trees: Afforestation plans for Africa. Trends Ecol Evol. 2019;34(11):963–965. http://dx.doi.org/10.1016/j.tree.2019.08.003
13. Ripple WJ, Wolf C, Newsome TM, Galetti M, Alamgir M, Crist E, et al. World scientists' warning to humanity: A second notice. Bioscience. 2017;67(12):1026–1028. https://doi.org/10.1093/biosci/bix125
14. Gardner CJ, Wordley CFR. Scientists must act on our own warnings to humanity. Nat Ecol Evol. 2019;3(9):1271–1272. http://dx.doi.org/10.1038/s41559-019-0979-y
15. Kuhn TS. The structure of scientific revolutions. Chicago, IL: University of Chicago Press; 1962. http://dx.doi.org/10.7208/chicago/9780226458106.001.0001
16. Ferber D. Tobacco wars. Research on secondhand smoke questioned. Science. 2004;306(5700):1274. http://dx.doi.org/10.1126/science.306.5700.1274
17. Oreskes N, Conway EM. Merchants of doubt: How a handful of scientists obscured the truth on issues from tobacco smoke to global warming. New York: Bloomsbury Publishing; 2011.
18. Enserink M. Big tobacco's offer: $1 billion for research. Should scientists take it? ScienceMag. 2018 February 08 [cited 2019 Oct 28]. Available from: https://www.sciencemag.org/news/2018/02/big-tobacco-s-offer-1-billion-research-should-scientists-take-it
19. Ford E. Defining and characterizing open peer review: A review of the literature. J Scholar Pub. 2013;44(4):311–326. https://doi.org/10.3138/jsp.44-4-001
20. Dunn AG, Coiera E, Mandl KD, Bourgeois FT. Conflict of interest disclosure in biomedical research: A review of current practices, biases, and the role of public registries in improving transparency. Res Integr Peer Rev. 2016;1, Art. #1. http://dx.doi.org/10.1186/s41073-016-0006-7
21. Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, et al. Scientific standards: Promoting an open research culture. Science. 2015;348(6242):1422–1425. http://dx.doi.org/10.1126/science.aab2374
22. Moher D, Bouter L, Kleinert S, Glasziou P, Sham MH, Barbour V, et al. The Hong Kong principles for assessing researchers: Fostering research integrity. OSF Preprints. 2019 September 17 [cited 2019 Nov 28]. https://doi.org/10.31219/osf.io/m9abx

Check for updates

Aerobiology in South Africa: A new hope!

AUTHORS:
Linus Ajikah[1,2]
Frank H. Neumann[1]
Dilys Berman[3,4]
Jonny Peter[3,4]

AFFILIATIONS:
[1]Evolutionary Studies Institute and School of Geosciences, University of the Witwatersrand, Johannesburg, South Africa
[2]Department of Plants and Ecological Studies, University of Calabar, Calabar, Nigeria
[3]Allergy and Immunology Unit, University of Cape Town Lung Institute, University of Cape Town, Cape Town, South Africa
[4]Division of Allergy and Clinical Immunology, University of Cape Town, Cape Town, South Africa

CORRESPONDENCE TO:
Frank Neumann

EMAIL:
frank.neumann@wits.ac.za

DATES:
Received: 31 Mar. 2020
Revised: 18 May 2020
Accepted: 24 May 2020
Published: 27 July 2020

HOW TO CITE:
Ajikah L, Neumann FH, Berman D, Peter J. Aerobiology in South Africa: A new hope! S Afr J Sci. 2020;116(7/8), Art. #8112, 4 pages. https://doi.org/10.17159/sajs.2020/8112

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:
aerospora, pollen monitoring network, pollen allergy, public health

FUNDING:
Wits School of Governance/University of the Witwatersrand, A. Vogel Echinoforce, Clicks, Durand Foundation, Thermo Fisher Scientific, Twinsaver, Zeiss

Pollen and fungal spores (aerospora) are the major atmospheric bioaerosols. Aerospora occurrence and concentration vary by geographical region and blossoming period, and with meteorological factors. Allergic respiratory diseases affect about 20 million South Africans, with pollen and fungal allergens amongst the leading triggers. Asthma triggered by aerospora can be life threatening; allergic rhinitis causes considerable morbidity and carries financial implications for individuals and health systems. Thus, knowledge about geographical variation, seasonal timing and intensity, as well as annual aerospora fluctuations in South Africa, where climate and vegetation are exceptionally diverse, is essential for effective diagnosis and treatment of allergies. Unfortunately, there is a lack of continuous aeropalynological data from South Africa. The longest annual monitoring exists for Cape Town (~20 years), with almost no data out of the Cape since the 1990s, and many parts of South Africa have never been monitored. In this brief review, we highlight the cross-disciplinary need for strengthened and expanded continuous aeropalynological study in South Africa, the history of efforts to date, and the introduction of the interdisciplinary South African Pollen Network (SAPNET). SAPNET was launched in seven major cities in August 2019 in order to monitor weekly variations of aerospora, and provide online data for allergy sufferers and health-care providers (www.pollencount.co.za) with the aim to establish regional pollen calendars.

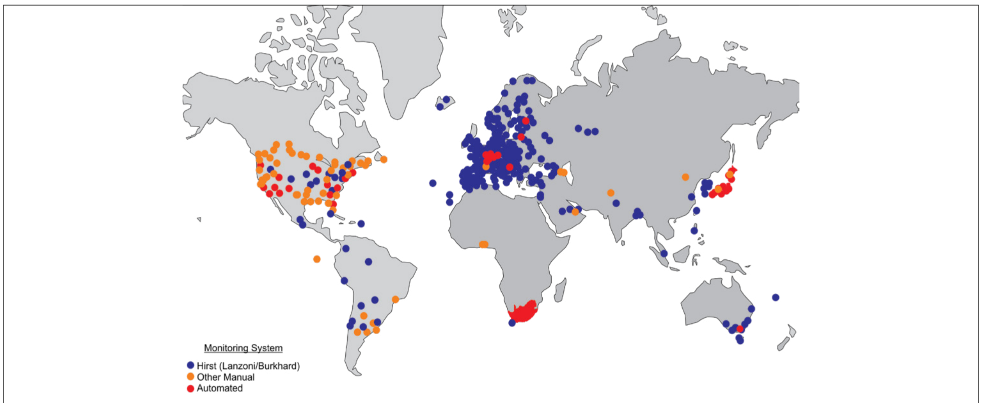
Significance:

- The significance of aerobiology as the study of airborne organic particles which include pollen, spores of fungi and cryptogams, plant hair, insect remains etc. especially as agents of allergic rhinitis (commonly known as hay fever) is outlined. Cases of allergic rhinitis are globally increasing due to current climate change.
- Very limited pollen monitoring data are available for South Africa; such data are a prerequisite for establishing pollen calendars to help medical practitioners and allergy sufferers. SAPNET (South African Pollen Network) addresses this lack of data. Since August 2019, seven 7-day volumetric spore traps have been operating in South Africa in major cities, and are continuously gathering data about the pollen and spore contents in a country which is highly affected by allergic rhinitis.

The problem

Allergic rhinitis affects 20–30% of South Africans, with co-morbid asthma in 20% of cases. Pollen and house dust mite allergens are the commonest triggers of allergic rhinitis and asthma. Most of the common global aeroallergens are found in South Africa, due to a variety of climates and biomes.[1] Non-native trees have been planted in South Africa to produce timber, for example eucalypt (*Eucalyptus* spp.) and pine (*Pinus* spp.) trees, or because of their ornamental value, for example London plane tree (*Platanus acerifolia*). Many non-native trees have undesired effects, such as fire proneness or pollen allergenicity.[2] Global warming is responsible for alarming changes, including increased total pollen amount, earlier shifts of pollen seasons with levels expected to be four times higher by 2050, invasion of allergenic plants with resultant increases in allergic sensitisation, and increased amounts of allergenic protein in pollen.[3] The changing environment, the major health ramifications and the biodiversity of South Africa underline a need for widespread pollen monitoring. In this paper, we follow Galán et al.[4] in regard to aerobiological terminology.

There are more than 879 active pollen monitoring stations globally.[5] Pollen monitoring coverage is biased towards the northern hemisphere; Europe has more than 500 spore traps. Pollen monitoring in Africa was, until August 2019, restricted to four spore traps in northern Africa (Morocco, Algeria, Tunisia), four spore traps in western Africa (Nigeria, Benin) and one spore trap in southern Africa (Cape Town, South Africa) (Figure 1[5]) – the only location of continuous monitoring in South Africa.[6] Monitoring in the Cape has shown large seasonal variability and increasing pollen productivity; changes in other parts of the country were until recently unknown.[6]



Source: after Buters et al.[5], updated in accordance with https://www.eaaci.org/19-activities/task-forces/4342-pollen-monitoring-stations-of-the-world.html, 2020 May 13

Figure 1: Global map of aerospora monitoring stations. Area in red represents South Africa and Lesotho.

Aerospora are globally increasingly inducing allergic respiratory disorders such as asthma, rhinitis, allergic conjunctivitis and atopic dermatitis, especially in arid to semi-arid regions in the world.[7] Atopic individuals have a genetic predisposition to produce IgE antibodies against usually harmless environmental proteins (allergens). Several factors influence the development of new allergic sensitisations, including: (1) allergen exposure (large geographical pattern in aeroallergen profiles); (2) nature of the allergenic proteins, e.g. certain aeroallergens are considerably more allergenic; and, increasingly, (3) the co-factor of air populates such as particulate matters NO_2, ozone, CO and SO_2.[8]

A major challenge for aerobiologists in South Africa is the diverse vegetation and climate across its nine biomes (Figure 2) – ranging from the arid Karoo to the humid coast of KwaZulu-Natal in the East and from the cool Drakensberg grasslands to the wet, windy Cape.[9,10] The Cape Floristic Region, which consists of the Fynbos and Succulent Karoo Biomes, illustrates this diversity, with its approximately 9000 flowering plants[11] (Figure 2) compared to Europe which harbours about 10 600 flowering plants[12]. Other South African biomes are the tropical Indian Ocean Coastal Belt, the temperate Grassland, and the hot Savanna Biome which represents the southern extension of the largest biome of Africa (Figure 2).[10] These different environments have distinct flora and include, next to indigenous vegetation, alien plants from Europe, North and South America and Australasia (548 taxa).[13] Aerospora diversity is associated with distinct allergy patterns, and it is inaccurate to apply aerospora data from one biome to another; multiple monitoring stations are the only option for understanding regional aerospora diversity.

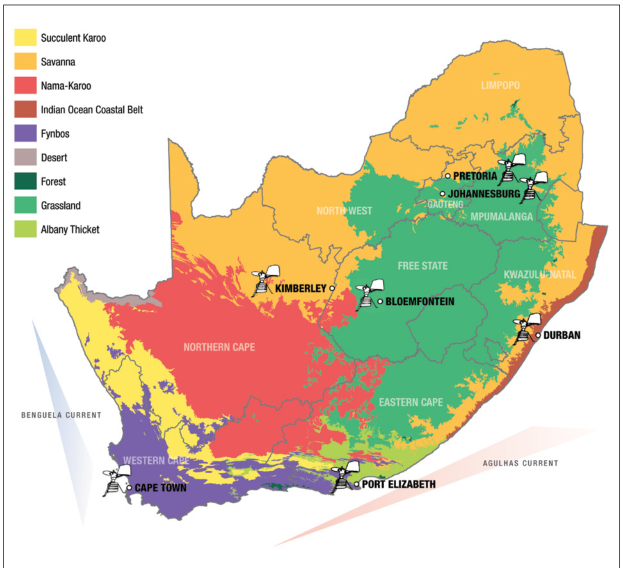


Figure 2: Biomes of South Africa, Lesotho and Swaziland after Mucina and Rutherford[10], with oceanic currents which trigger climate (blue arrow: Benguela Current; red arrow: Agulhas Current). Dots show cities covered by SAPNET since August 2019.

What we do know? History of aerobiology in South Africa

Aerobiological studies in South Africa began with Davidson's studies in 1941 on grasses in Johannesburg. Aerobiological studies in South Africa are summarised in Table 1. Ordman and Etter[14,15] contributed to the understanding of pollinosis in South Africa. Hawke and Meadows[16] studied seasonal allergens and their morphology by observing, measuring and documenting aeroallergens in 1945–1970 in some areas of South Africa, and unravelling some of the unknown elements of the environmental allergens that trigger allergies. Potato dextrose agar settle plates and Durham (gravity) spore traps were used to collect pollen and to culture fungi by exposing a sticky slide at the top of a three-storey building in Johannesburg. Grasses were identified as the most important pollen allergens in South Africa.[17] The late winter peak pollination period for Cupressaceae (cypress) pollen in Cape Town and Johannesburg was first documented by Ordman[14]. He identified *Prosopis* (mesquite) as an important allergenic pollen in South West Africa (Namibia). Ordman[18] hypothesised a link between climate and respiratory allergy, observing that humidity in coastal areas is a key factor in perennial respiratory allergy. This discovery was remarkable considering that house dust mites had yet to be identified.

Early pioneering contributions were followed by a second wave of allergists and aerobiologists. In 1981, Weinberg et al.[19] successively upgraded and operated a 7-day volumetric spore trap in the Cape. By 1985, the Air Pollution Department in Cape Town added pollen sampling to their Air Quality Monitoring Programme, developing a partnership with the Red Cross Allergy Department. This allowed for pollen monitoring in different suburbs of Cape Town and provision of data to the Allergy Clinic, which treated patients from many areas of the Cape Town metropole.[6] Cadman and Dames[20] added insights and observations on pollen seasonality and profiles of Gauteng and KwaZulu-Natal through pollen monitoring from 1987 to 1991. Potter and colleagues[21] and Dames and Cadman[22] concluded that *Alternaria* was amongst the most prominent fungal spore perennially, both in Cape Town and Durban. In 2008, Berman[6] set up a 3-year aerobiological survey for an Environmental Impact Assessment commission in Mpumalanga. Here, differences in seasonality of aerospora in the Lowveld were tracked, and differed from those in the Cape. Grass pollen peaked in January, and tree pollen was low and dominated by non-native trees (*Platanus* spp.), oak (*Quercus* spp.) and eucalypt (*Eucalyptus* spp.). High levels of fungal spores were observed, probably due to increased humidity. Berman[23] provided a profile for all aerospora for respiratory diseases at sites close to a residential area. Berman[24] continued to conduct research in different parts of South Africa to monitor aerospora for clinical studies.

Aerobiological studies outlined above have provided important insights into the most allergenic aerospora in South Africa. Some of these insights are given below.

- At the Cape, the grass pollen season is September–March, with peaks in October. On the Highveld, the grass pollen season is longer (September–May); grass is regarded as a 'perennial' allergen.[1] In Gauteng, a noticeable variation was reported between the models for both grass pollen and spores for Johannesburg and Pretoria due to meteorological factors.[25] Allergenic grass pollens in South Africa are *Eragrostis curvula,* kikuyu (*Pennisetum clandestinum*) and buffalo grass (*Stenotaphrum secundatum*).[1,23] *Zea mays* is problematic in the Free State Province.
- Unlike in Europe, tree pollen is a relatively uncommon cause of seasonal allergy in South Africa, although further studies are needed as allergenic pollen of non-native trees such as oak (*Quercus* spp.), Italian cypress (*Cupressus sempervirens*), plane trees (*Platanus* spp.), and mulberry (*Morus* spp.) are a 'growing problem', especially in the cities where they are planted as ornamentals or fruit trees[1,23] (own unpublished data).
- Weeds are arguably a minor problem, although the invader plantain *Plantago* spp. as well as several, mostly indigenous, members of the daisy family are known to cause allergies.[1,23]
- Fungal spores are globally the most abundant aerospora. In South Africa, *Cladosporium*, *Alternaria*, *Aspergillus* and *Epicoccum* are the most important.[1,23]

These insights provided a basic understanding; however, many of these studies were conducted decades ago and environments might have changed. Furthermore, Table 1 and Figure 2 show that aerospora concentrations in many South African provinces and cities have never been monitored (Northern Cape, North West and Limpopo Provinces), and even for well-monitored areas, uninterrupted data are limited. There is an urgent need for up-to-date monitoring data to address knowledge gaps including: (1) the main allergenic aerospora in many areas/regions/biomes; as well as the (2) flowering times and (3) peak concentrations of important allergens.

Table 1: Historical and chronological order of aeropalynology studies and purpose of each study (updated/modified from Buters et al.[5])

Investigator	Period	Region in South Africa	Purpose	Reference
Davidson	1941	Johannesburg	Pollen research	Ordman[14,17,18]
Ordman	1945–1970	Gauteng: Johannesburg KwaZulu-Natal: Durban, Pietermaritzburg Mpumalanga: Nelspruit	Monitoring for clinical practice	Ordman[14,15,17,18]
Weinberg	1973–2004	Western Cape: Rondebosch, Table View	Clinical trials monitoring	Weinberg et al.[19]
Dames	1987–1991	Gauteng: Johannesburg, Pretoria KwaZulu-Natal: Durban	Fungal spore research / pollen research	Cadman and Dames[20,22,25]
Cadman	1987–1994	Gauteng: Johannesburg, Pretoria KwaZulu-Natal: Durban Western Cape: Kirstenbosch, Parow	Fungal spore research / pollen research	Cadman and Dames[20,22,25]
Berman / Potter	1985–1990 1985–1986 1986–1986 2000–2001 2000–2003 2008–2009 2008–present 2010	Cape Town suburbs Edgemead Bothasig Somerset West Parow Mowbray Observatory, Parow Mowbray Observatory	Monitoring for clinical practice Monitoring for clinical practice Monitoring for clinical practice Monitoring for clinical practice Monitoring for clinical practice Clinical trials monitoring Monitoring for clinical practice	Potter et al.[21] Berman[8] Berman[8] Berman[23] Berman[23] Berman[24] Berman[6] Potter[1]

The way forward: South African Pollen Network and Monitoring Programme (SAPNET)

The need for a pollen monitoring network in South Africa has been outlined above. Already in 2007, the necessity of a national network of spore traps in South Africa as 'a collaborative exercise between allergists and the botany or palynology departments of the local universities' was underlined.[23] A network of this scale requires substantial funding; a strong, dedicated team of collaborators; and widespread public and institutional support. SAPNET was launched in August 2019 at the University of Cape Town Lung Institute and 7-day volumetric Burkard spore traps were placed in seven major South African cities (Bloemfontein, Cape Town, Durban, Johannesburg, Kimberley, Port Elizabeth, Pretoria; Figure 2). This new cycle of monitoring, which covers several biomes and includes major population centres, is a necessity for South Africa, considering the growing and mobile population, climate and environmental changes and the expansion and likely worsening of aeroallergen-triggered allergic diseases. This will provide insight into the high–low risk periods for sensitised individuals through weekly online updates (http://www.pollencount.co.za/), as well as the development of updated pollen calendars for major cities and provinces spanning different biomes. Information about allergenic aerospora will assist health-care providers, including general practitioners, in making clinical diagnoses and decisions. It will help pollen allergy sufferers to identify plants which have the tendency to produce allergenic pollen in their environment and adopt prophylactic measures as advised on the website.

Future goals

1. Expansion of SAPNET into northern provinces (Mpumalanga, Limpopo, North West) and neighbouring countries.
2. Coverage of rural areas in South Africa, including townships in the mining belt of Mpumalanga and other, highly industrialised areas or along major transport routes where – due to high air pollution – the risk of asthma, often triggered by pollen allergy[24-26], is high.
3. Collaboration with above-mentioned northern and western African nations (see Figure 1), as well as southern hemisphere pollen monitoring networks.
4. Training of aerobiologists and knowledge exchange with other countries, where pollen monitoring is established, with the aim to improve aerobiological methods in South Africa.

Acknowledgements

SAPNET acknowledges all investigators (in alphabetical order): Marion Bamford, Prosper Bande, Angela Effiom, Kevwe Eweto, Moteng Moseri and Nikiwe Ndlovu (Johannesburg); Rebecca Garland, Robin Green, Juanette John and Seneca Naidoo (Pretoria); Lize Joubert, Mawethu Ndiki, Louis Scott, Riaz Seedat and Andri van Aardt (Bloemfontein); Jemma Finch, Trevor Hill and Ahmed Ismail Manjra (Durban/Pietermaritzburg); Erin Hilmer and Lynne Quick (Port Elizabeth); Joseph Chikumbirike, Werner Hoek and Tendai Musvuugwa (Kimberley). SAPNET also acknowledges financial support from (in alphabetical order): A. Vogel Echinoforce, Clicks, Durand Foundation, Thermo Fisher Scientific, Twinsaver and Zeiss. We thank the Wits School of Governance for sponsoring the Life in the City project and providing funds for the postdoctoral research of L.A. at the University of the Witwatersrand. We also thank Karen Adamson of Chocolate Cow Design for designing Figure 2 and the anonymous reviewers for their valuable comments.

Authors' contributions

L.A.: Conceptualisation, formulation of the overarching research aims, preparation and creation of the published work, specifically writing the initial draft. F.H.N.: Critical review, commentary, extension and revision of text, tables, figures, corresponding author. D.B.: Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of the review and other research outputs, contributions to text. J.P.: Acquisition of the financial support for the project leading to this publication, management and coordination responsibility for the research activity planning and execution, contributions to text.

References

1. Potter PC. Common indoor and outdoor aeroallergens in South Africa. CME. 2010;28(9):426–432.
2. CastroDíez P, Vaz AS, Silva JS, Vanloo M, Alonso Á, Aponte C, et al. Global effects of non-native tree species on multiple ecosystem services. Biol Rev. 2019;94:1477–1501.
3. Damialis A, Häring F, Gökkaya M, Rauer D, Reiger M, Bezold S, et al. Human exposure to airborne pollen and relationships with symptoms and immune responses: Indoors versus outdoors, circadian patterns and meteorological effects in Alpine and urban environments. Sci Total Environ. 2019;653:190–199. https://doi.org/10.1016/j.scitotenv.2018.10.366
4. Galán C, Ariatti A, Bonini M, Clot B, Crouzy B, Dahl A, et al. Recommended terminology for aerobiological studies. Aerobiologia. 2017;33:293–295.

5. Buters JTM, Antunes C, Galveias A, Bergmann KC, Thibaudon M, Galán C, et al. Pollen and spore monitoring in the world. Clin Transl Allergy. 2018;8:9–24. https://doi.org/10.1186/s13601-018-0197-8

6. Berman D. Aerobiology studies in South Africa. Curr Allergy Clin Im. 2014;27(4):255–258.

7. International Study of Asthma and Allergies in Childhood (ISAAC). ISAAC Steering Committee. Clin Exp Allergy. 1998;28(suppl.5):52–66; discussion 90-1. https://doi.org/10.1046/j.1365-2222.1998.028s5052.x

8. Berman D. Variations in pollen and fungal air spora: An analysis of 30 years of monitoring for the clinical assessment of patients in the Western Cape [PhD thesis]. Cape Town: University of Cape Town; 2018.

9. Rutherford MC, Westfall RH. Biomes of southern Africa – an objective categorization. 2nd ed. Mem Bot Surv S Afr. 1994;63:1–94.

10. Mucina L, Rutherford MC, editors. The vegetation of South Africa, Lesotho and Swaziland. Pretoria: South African National Biodiversity Institute; 2006. p.748–789.

11. Goldblatt P, Manning JC. Plant diversity of the Cape region of southern Africa. Ann Missouri Bot. 2002;89(2):281–302. https://doi.org/10.2307/3298566

12. Mutke S, Sánchez JG, Chambel MR, Prada MA, Álvarez D, Iglesias S, et al. Phenotypic plasticity is stronger than adaptative differentiation among Mediterranean stone pine provenances. Forest Syst. 2010;19(3):354–366. https://doi.org/10.5424/fs/2010193-9097

13. Henderson L. Invasive, naturalized and casual alien plants in southern Africa: A summary based on the Southern African Plant Invaders Atlas (SAPIA). Bothalia. 2007;37:215–248. https://doi.org/10.4102/abc.v37i2.322

14. Ordman D. Pollinosis in South Africa. S Afr Med J. 1947;21:38–48.

15. Ordman D, Etter KG. The air-borne fungi in Johannesburg: A five-year survey as a basis for the study of fungus allergy in South Africa. S Afr Med J. 1956;30(11):1054–1058.

16. Hawke, P, Meadows M. Winter airspora spectra and meteorological conditions in Cape Town, South Africa. Grana. 1989;28(3):187–192. https://doi.org/10.1080/00173138909427430

17. Ordman D. The evolution and present status of the concepts and practice of allergy in South Africa. S Afr Med J. 1963;36:419–424.

18. Ordman D. Respiratory allergy in the coastal areas of South Africa. S Afr Med J. 1955;19(8):173–179.

19. Weinberg EG, Potter PC, Motala C. The Allergy Clinic: 50 years. S Afr Med J. 2006;96:9.

20. Cadman A, Dames JF. Airspora of Durban: A coastal South African city. I: Pollen component. Grana. 1994;32:346–348. https://doi.org/10.1080/00173139409429023

21. Potter PC, Berman D, Toerien A, Malherbe D, Weinberg EG. Clinical significance of aero-allergen identification in the western Cape. S Afr Med J. 1991;79(2):80–84.

22. Dames JF, Cadman A. Airspora of Durban: A subtropical, coastal South African city. II: Fungal spore component. Grana. 1994;33(6):346–348. https://doi.org/10.1080/00173139409429023

23. Berman D. Pollen monitoring in South Africa. Curr Allergy Clin Im. 2007;20:184–187.

24. Berman D. Regional-specific pollen and fungal spore allergens in South Africa. Curr Allergy Clin Im. 2013;26(4):196.

25. Cadman A. Airspora of Johannesburg and Pretoria, South Africa, 1987/88. I: Pollen calendars. Grana. 1991;29(4):311–317. https://doi.org/10.1080/00173139009428942

26. Motta AC, Marliere M, Peltre G, Sterenberg PA, Lacroix G. Traffic-related air pollutants induce the release of allergen-containing cytoplasmic granules from grass pollen. Int Arch Allergy Immunol. 2006;139(4):294–298. https://doi.org/10.1159/000091600

A few hours in the Seychelles with Alex du Toit in 1938

AUTHOR:
Sharad Master[1]

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

CORRESPONDENCE TO:
Sharad Master

EMAIL:
sharad.master@wits.ac.za

DATES:
Received: 19 Dec. 2019
Revised: 28 Apr. 2020
Accepted: 04 May 2020
Published: 29 July 2020

HOW TO CITE:
Master S. A few hours in the Seychelles with Alex du Toit in 1938. S Afr J Sci. 2020;116(7/8), Art. #7747, 3 pages. https://doi.org/10.17159/sajs.2020/7747

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Yali Woyessa

KEYWORDS:
Mahé, Silhouette, Mascarene Plateau, Mauritius nucleus, continental drift

FUNDING:
None

Alexander Logie du Toit (1878–1948) was South Africa's most famous geologist during his lifetime, having authored five books which brought him world renown. In December 1937 to January 1938, accompanied by his wife Evelyn, he visited India in order to attend the Jubilee Indian Science Congress in Calcutta and to do field work in coal and diamond mines. On the return journey to Africa by ship, they stopped for a few hours in Port Victoria on Mahé Island in the Seychelles archipelago. They also passed by Silhouette Island. Du Toit recorded his activities in a diary, and his geological observations in a notebook, where he also drew a sketch of Mahé, and recorded steep structures on the east coast of Silhouette. Although he had not visited the Seychelles before, his deep understanding of the problems of Seychelles geology resulted from his comprehensive research on Indian Ocean geology for his 1937 book *Our Wandering Continents*. He made remarkably accurate observations on the geomorphology and structure, some of which were only confirmed decades later when the Seychelles were mapped in the 1960s to 1990s. His bold and prescient ideas on the breakup of the Gondwana continent, and on the formation of the Indian Ocean, have been amply confirmed by modern studies, especially by those of Lewis D. Ashwal and his collaborators.

Significance:

- South African geologist Alexander Logie du Toit's impressions of the Seychelles in 1938 are recorded for the first time, based on entries in his diaries. His observations of structures on Mahé and Silhouette Islands were prescient. His deep understanding of Seychelles geology was the result of his research for his 1937 book *Our Wandering Continents*. His bold conjecture that the Mascerene Ridge, made of continental crust, was the nucleus of Mauritius, was finally proved in 2017.

Introduction

In 1938, the South African geologist Alexander Logie du Toit (1878–1948) attended the Jubilee Indian Science Congress, from 3 to 9 January, in Calcutta (now Kolkata), India. He was accompanied by his wife Evelyn. They arrived in India by boat on 11 December 1937. Du Toit was then in the employ of the De Beers corporation, as Consulting Geologist, so after the Indian Science Congress, he visited the Panna District of Madhya Pradesh, central India, to report on the diamond mines there, on behalf of his company.

Du Toit was at that time the most famous geologist in South Africa, having published books on the *Geology of Cape Colony* (with Arthur Rogers, 1909)[1], on *Physical Geography for South African Schools* (1912)[2], *Geology of South Africa* (1924)[3], *Geological Comparison of South America and South Africa* (1927)[4], and most recently his passionately argued and hugely controversial book *Our Wandering Continents* (1937)[5]. He was particularly interested in the fossil evidence supporting the concept of continental drift. Prior to the Indian Science Congress he had visited the Jharia coalfield (in the Satpura Gondwana Basin some 200 km northwest of Calcutta) as well as the fossil collections of the nearby Dhanbad School of Mines, and during the congress he went on a field excursion to the Ramgahr coalfield (also in the Satpura Basin), which overlies the Talchir glacial beds (and is thus the equivalent of the Ecca coalfield overlying the Dwyka conglomerate in South Africa). He also visited the palaeontological museum of the Geological Survey of India in Calcutta, with its rich collections of Gondwana *Glossopteris* fossil flora.

Seychelles

On the return sea journey following his lengthy and strenuous time in India, du Toit was able to relax. The du Toits had departed Bombay (now Mumbai) on 19 January on the 60 000-tonne ship S.S. *Karanja* (Figure 1), and after putting into Murmagao (Goa) the next day, they headed across the Indian Ocean for Africa. On 24 January 1938, du Toit recorded in his diary[6] that they had re-entered the southern hemisphere (i.e. had crossed the equator). There was 'an absolutely calm sea', and he spent his time 'reading novels all day'.[6] The weather was 'dull at times in the afternoon, with occasional showers'[6].

Du Toit's time for relaxation was short-lived, for the very next day, the ship sailed into the waters of the Seychelles archipelago. Du Toit recorded in his diary again[6]:

> *25th January 1938.*
>
> *Seychelles. Fine day but cloudy at times and a spot of rain. Arrived at Seychelles at 8 AM and with Eve went ashore. Some 1¾ hours on pier, then round about to Hotel des Palmes for morning tea and back to ship at 12:45 for lunch. More passengers left at 4:30 PM, and passed North Island at 6 PM. Excellent Day.*

The ship had sailed through the night, and by the next day, it was back to reading novels for du Toit. His full entry for 26 January 1938 reads: 'Fine fresh day. Nothing to record – remaining due west.'[6] For the remainder of the journey, du Toit spent his time working on and completing his report on the Panna diamond fields, and writing letters. On 29 January the ship reached Kilindini on the Kenyan Coast. Over the next week the ship put in at Zanzibar, Dar-es-Salaam and Beira, before finally arriving at its destination, Lourenço Marques in Portuguese East Africa (now Maputo, Mozambique), on 8 February. The du Toits then took an overnight train, arriving back home in Johannesburg on 9 February 1938.[6]

If du Toit's diary was all we had to go on, it would seem that he and his wife Evelyn had a pretty pleasant and uneventful few hours in the Seychelles, just enough to enjoy a stroll on shore and have tea at the Hotel des Palmes. But du Toit's mind was furiously engaged with trying to explain the great conundrum of the Seychelles: almost uniquely among coral-fringed oceanic islands, the Seychelles are made of granite, and not basalt – and therefore are made up of rocks characteristic of continental crust. How could that be explained?

Source: ©Tim Webb; reproduced with permission

Figure 1: The passenger liner S.S. *Karanja*, which operated between India and the East Coast of Africa between 1931 and 1939. The du Toits travelled on this ship in January 1938 from Bombay to Zanzibar and Lourenço Marques, with a short stopover in Port Victoria on Mahé Island, Seychelles. This ship was later commandeered by the British Royal Navy in 1940, and was sunk in 1942 off the coast of Algeria during Operation Torch, the invasion of North Africa.

Aside from his calendar diary, in which it had been his habit to record his daily activities, always starting with the weather (more than four decades of his diaries are preserved in his Archives stored at the University of Cape Town), du Toit, when on field excursions, also kept an excursion diary. On this occasion, he noted the following in his excursion diary, concerning the few hours he spent in the Seychelles[7]:

> *Seychelles 25 January 1938.*
>
> *Fine group of islands made by bare crystalline rocks, often black through crust of lichens. At north end of Mahé, Port Victoria, there is a remarkably regular sheeting of the granite dipping SW at about 15° or so* [Figures 2 and 3]. *Granite without any marked gneissic structure or banding. In Silhouette Island to the NW there are indications of some structure standing at a high angle and trending rudely NNW. (Extraordinary is the small and large scale fluting of the solid granite, visible in Mahé Island, visible even in crags high up in the range, due to wetting caused by dripping of vegetation or guidance of rain down incipient fluting)* [Figure 4]. *Noticeable is the long axis of Mahé parallel to this same direction. The group of islands rises from a shelf with relatively shallow water which, at least on NW side (track to Mombasa) has an extremely sharp edge. Its north margin is outlined by shoals and islands.*

Although du Toit had spent only a few hours in the Seychelles, and made a few observations from on board ship concerning the geomorphology and structure of the islands he saw, he already knew more than just about anybody else on earth about the significance of this picturesque group of tropical islands. He was thus particularly concerned about the structure, because he believed that the granitic rocks of the Seychelles must have been part of a continent, from which they had been faulted away.

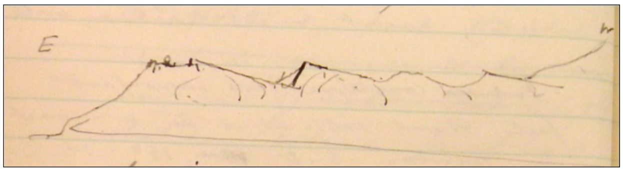


Figure 2: Du Toit's sketch of the outline of Mahé Island, Seychelles, 25 January 1938, looking towards the south.[7] Three shallow westward-dipping planes represent a former peneplain, which has been downfaulted and rotated by steeply eastward dipping faults, one of which is drawn with a prominent thick line (centre).

Source: CC-BY-2.0 Wikimedia Commons

Figure 3: Jagged skyline of Mahé Island, Seychelles, which was interpreted by du Toit[7] as having formed by block faulting of a peneplained continental granitic terrain.

Source: Hotel La Roussette; reproduced with permission

Figure 4: Strongly fluted outcrops of granite, due to wetting caused by dripping of vegetation or guidance of rain down incipient fluting, as observed by du Toit on Mahé Island.

He was quite alert to the possibility that the jagged outline of these rocks (Figures 2 and 3) could represent a formerly peneplained continental region which had been broken up by block faulting. His observations on the steep structures on Silhouette Island were confirmed many years later by Stephens[8] who showed that on the east coast of Silhouette there is a north to north-northwest trending trachytic breccia traversed by steep trachytic, micrograncitic and microsyenitic dykes (which have been dated at between 64 Ma and 63 Ma)[9]. In 1937, in his magnum opus *Our Wandering Continents*[5], du Toit wrote, what we now know, in hindsight more than 80 years later, to be a breathtakingly daring, intricately detailed and presciently accurate account of the geological setting and palaeogeographic positions of Madagascar, the Seychelles, Mauritius and the Mascarene Plateau, and India, in relation to the breakup of the Gondwana continent and the formation of the Indian Ocean. This is what du Toit wrote, in part, about the Indian Ocean, based on his comprehensive and detailed researches[5(p.226–227)]:

> *THE INDIAN OCEAN*
>
> *Floor. – With area slightly less than the area of the Atlantic this basin has a mean depth of about -4000 m. with maximum of -7450 m. in the Sunda or Java foredeep.*

> *It is like the Atlantic through the presence of a medial rise that runs northwards from the Gauss Berg in Antarctica through Heard, Kerguelen, St. Paul, New Amsterdam, Chagos, Maldive, and Laccadive Islands to the Western side of India, which in similar fashion gives off branches to the Cape (via the Crozets and Agulhas Bank), Seychelles, Somaliland (via the Carlsberg Ridge and Socotra) and possibly Tasmania. The floor is thus divisible into an eastern and a western portion, and those again into subsidiary deeps.*

> *Madagascar stands on a deep shelf prolonged somewhat to the south, which is connected to the African mass, and has a crystalline basement. The curious curving Mascarene ridge to the east reaches above the sea in the granite of the Seychelles in the north, and the volcanic islands of Mauritius and Reunion in the south.*

About the Seychelles he had this to say[5(p.125)]:

> *...the Mozambique Channel was developed by the movement of Madagascar to the east of south, while the crystallines of the Seychelles – perhaps the nucleus of Mauritius also – are visualized as fragments left behind in the rear of India.*

Considering how little was known then about the sea floor, and about the ages of rocks, what du Toit said in 1937[5] presages the insights that were to be gained by the geological mapping of the Seychelles by Baker in 1963[10], and the more recent geochronological and palaeomagnetic studies of Tucker et al.[11] and Torsvik et al.[12], summarised most recently in the 35th Alex du Toit Memorial Lecture by Professor Lewis D. Ashwal[13]. Du Toit's remarkable conjecture, made in 1937, in visualising the crystalline rocks of the Seychelles, and 'perhaps the nucleus of Mauritius also' as being fragments left behind by the rifting away of India from Madagascar, was finally proven to be correct 80 years later. Ashwal et al.[14] showed that the island of Mauritius was underlain by crust of a much older continental fragment called 'Mauritia', which extended towards the Seychelles, and was originally situated between Madagascar and India, exactly as du Toit[5] had envisioned it.

Du Toit in the 1930s had been one of the very few crusaders (together with people like Reginald Daly and Arthur Holmes)[15-17] championing Alfred Wegener's ideas[18] of continental drift (the Displacement Hypothesis), while single-handedly opposing the ideas of isthmian and insular links between the continents (continental linking)[19], which had been invoked to explain the remarkable similarities in the fossil fauna and flora of the Gondwana continental fragments through a process of island-hopping. He ended his Chapter X, on 'The Oceans'[5], with this strongly worded, declamatory statement:

> *The Displacement Hypothesis is, on the contrary, competent to explain these and other puzzles of biological distribution, particularly in this great Austral region, in a simple and logical manner and without the violation of isostatic principles.* Current views of continental linking must therefore be firmly rejected.

Acknowledgements

I am grateful to the librarians (especially to Clive Kirkwood) of the Jagger Library, University of Cape Town, for their help in accessing the du Toit archives, and for permission to publish this work. I also thank Lewis Ashwal for discussions, and for a preprint of his du Toit Memorial Lecture. This paper was presented at the conference 'Earth, Moon, Mars and Ashwal: 70 years of fun with Lew' on 22 November 2019, and is dedicated to Lewis D. Ashwal on the occasion of his 70th birthday.

References

1. Rogers AW, Du Toit AL. An introduction to the geology of Cape Colony. 2nd ed. London: Longmans, Green & Co.; 1909.
2. Du Toit AL. Physical geography for South African schools. Cambridge: Cambridge University Press; 1912.
3. Du Toit AL. The geology of South Africa. Edinburgh: Oliver & Boyd; 1926.
4. Du Toit AL. A geological comparison of South America and South Africa. Publication no. 381. Washington DC: Carnegie Institution of Washington; 1927.
5. Du Toit AL. Our wandering continents: An hypothesis of continental drifting. Edinburgh: Oliver & Boyd; 1937.
6. Du Toit AL. Diary for 1938. In: Alex L. du Toit Archives, BC722, A1, Jagger Library, University of Cape Town, Cape Town, South Africa; 1938.
7. Du Toit AL. Field notebook: India 1938. In: Alex L. du Toit Archives, BC722, D. Jagger Library, University of Cape Town, Cape Town, South Africa; 1938.
8. Stephens WE. Geology of Silhouette Island. Phelsuma. 1996;4:11–18.
9. Ganerød M, Torsvik TH, Van Hinsbergen DJJ, Gaina C, Corfu F, Werner S, et al. Palaeoposition of the Seychelles microcontinent in relation to the Deccan Traps and the Plume Generation Zone in Late Cretaceous–Early Palaeogene time. London: Geological Society, Special Publications. 2011;357:229–252. https://doi.org/10.1144/SP357.12
10. Baker BH. Geology and mineral resources of the Seychelles archipelago. Geological Survey of Kenya Memoir. 1963;3:1–140.
11. Tucker RD, Ashwal LD, Torsvik TH. U–Pb geochronology of Seychelles granitoids: A Neoproterozoic continental arc fragment. Earth Planet Sci Lett. 2001;187(1–2):27–38.
12. Torsvik TH, Ashwal LD, Tucker RD, Eide EA. Neoproterozoic geochronology and palaeogeography of the Seychelles microcontinent: The India link. Precambrian Res. 2001;110(1–4):47–59.
13. Ashwal LD. Wandering continents of the Indian Ocean. Alex L. du Toit Memorial Lectures, No. 35, Geological Society of South Africa. S Afr J Geol. 2019;122(4):397–420. https://doi.org/10.25131/sajg.122. 0040
14. Ashwal LD, Wiedenbeck M, Torsvik TH. Archaean zircons in Miocene oceanic hotspot rocks establish ancient continental crust beneath Mauritius. Nat Commun. 2017;8, Art. #14086. https://doi.org/10.1038/ncomms14086
15. Oreskes N. The rejection of continental drift. New York: Oxford University Press; 1999.
16. Frankel HR. The continental drift controversy. Volume 1: Wegener and the early debate. Cambridge, UK: Cambridge University Press; 2012.
17. Master S. The reception of "Our Wandering Continents" (1937), from the correspondence of its author, South African geologist A.L. du Toit. Paper presented at: 35th International Geological Congress; 2016 August 28 – September 05; Cape Town, South Africa; abstract #5230.
18. Wegener A. The origin of continents and oceans. London: Methuen; 1924.
19. Willis B. Isthmian links. Bull Geol Soc Am. 1932;43:917–952.