South African Journal *of* Science

volume 117
number 3/4

COVID-19 amplifies urban inequalities

SAWS weather forecasts and early warnings of severe weather

2018 bolide and meteorite fall over Benenitra, southwestern Madagascar

Review of *Moringa oleifera* production and consumption in South Africa

Dental caries in South African fossil hominins

ASSAf
ACADEMY OF SCIENCE OF SOUTH AFRICA

Volume 117
Number 3/4
March/April 2021

South African Journal of Science

eISSN: 1996-7489

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

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

Correspondence and enquiries
sajs@assaf.org.za

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

Research Article

Research Letter

Corrigendum

Cover caption
A meteor fireball (or bolide) approaching the Earth's surface. During the early evening of 27 July 2018, residents in southwestern Madagascar observed a rapidly moving bright light in the sky which exploded into multiple smaller glowing fragments. In their article on page 91, Gibson and colleagues confirm that this event was a meteorite fall – only the second known meteorite fall over Madagascar and the first that can be linked to a bolide.

Behind the scenes

This is the first issue of the *South African Journal of Science* with my name on the masthead as Editor-in-Chief. Though I have been with the Journal for a few weeks, I was involved in almost none of the work that went into producing this issue. In this regard, I want to pay credit to the remarkable work undertaken for the Journal, and for the science community, by my predecessor, Prof. Jane Carruthers, and, indeed by previous editors-in-chief, some of whom have been kind enough to reach out to me and to wish me well in my new role. Given these COVID times, I have never met Jane Carruthers face to face, but she has facilitated an exemplary and very supportive handover, for which I am very grateful indeed. This handover has occurred in the context of an atmosphere of support and encouragement at the Journal and more broadly at ASSAf, with many people behind the scenes working very hard to ensure that the Journal continues as smoothly as possible in our science community.

There is a convention in scientific and other journals which run regular lead articles, for the writers of those leaders to attempt to knit together some, or all, of the content in each issue to create a thread or theme in that issue, even when the issue was not designed to focus on one set of concerns. In what we call the 'front section' of our Journal, in this issue, we have a discussion piece on POPIA (*Protection of Personal Information Act*) – an important concern for all researchers in South Africa and a topic on which we will be running more comment in forthcoming issues, and a piece on Plan S and questions of open science – also very important to the science community. We have two interesting pieces on COVID, another issue which inevitably you will be hearing more about in the Journal, as you have before. One of the pieces deals with artificial intelligence, and the other with pharmacists as vaccinators. We celebrate 200 years of the South African Astronomical Observatory, an ongoing resource to the science community. There are book reviews on diverse topics, and obituaries of people who gave much to the academic and scientific community in our country.

There may be many ways of finding links in the content of these and in the review and research articles in this issue, but a common factor I want to discuss here is that all the contributions to the front section of this issue – those I have mentioned above – were written by people who chose to help our Journal by providing relevant, up-to-date, material for us. In his piece, and in other pieces he has written for this Journal, Keyan Tomaselli notes the ongoing thorny problem of incentivising research outputs in the South African higher education system, and in particular the issue of perverse incentives. The fact that Tomaselli himself, and all our other contributors to the front section of the Journal, choose to write for us pieces which do not 'count' in the current subsidy system – that is, they do not generate financial rewards for universities – tells us a story, and I think it is an important one.

It is almost inevitable that anyone entering a well-functioning operation, like our Journal, will ask themself what the operation (in this case our Journal) is for. If you look at our 'About the Journal' web page, you will see our formal statement on our mission and vision. Implicit in what is written there is something about the role of the Journal in helping to sustain and develop the scientific community in South Africa and beyond. I believe that this wish to sustain and develop our community may be what motivates people to send us copy which is not obviously rewarded. Different people who help our Journal have different personal motivations for doing so, but what unites them, I think, is this central commitment.

In an early contribution to the field that is now generally known as 'ethics of care', in 1990, Berenice Fisher and Joan Tronto[1] provided the following definition of care:

> *On the most general level, we suggest that caring be viewed as a species activity that includes everything that we do to maintain, continue, and repair our 'world' so that we can live in it as well as possible. That world includes our bodies, our selves, and our environment, all of which we seek to interweave in a complex, life-sustaining web.*

As someone who does research in the field of care, and who has seen this definition often, I am struck, returning to it in my new role, of how appropriate the definition is for much of the work, in a wide range of fields, which is published in the Journal. Science, appropriately done and regardless of discipline, can and often should be part of the work of care. But one of the ways in which we 'maintain, continue, and repair' the world of science itself is through having journals like ours. And a key feature of research on ethics of care more broadly is its focus on the work, often hidden work, that people do to maintain our world.

Part of the work of care for a Journal like ours is the writing our contributors do for us. We are grateful for the support we receive from authors from a wide range of disciplines. But there is also a lot of hidden work of care. As Editor-in-Chief, I am given space to write leaders like this one, but my work would be impossible without the labour of many people, some of whose names you will find on our website, who keep the journal going – amongst them, the Managing Editor, Dr Linda Fick, and the Online Publishing Systems Administrator, Nadia Grobler, the Associate Editors, the Associate Editor Mentees, our Editorial Advisory Board, and others in ASSAf and beyond. When you read anything in our Journal, between the lines you are reading their care work. We are also wholly dependent on people who review for us. There is generally speaking little or no overt, tangible incentive to be a reviewer, but without peer review, journals like ours would collapse. There are others who support us, often in roles considered menial and not worthy of attention, but without cleaners, administrators and technicians, for example, the science world, and our Journal, would struggle to continue.

In our competitive world, with its emphasis on outputs and prestige, hidden care work is often not thought about. Worse, it may be disparaged and looked down upon, especially when done by women. But we are interconnected and need one another. At the Journal, I have entered a new, exciting and sustaining web of care. Without the ongoing support of our readers, writers, reviewers and others, we could not fulfil our care functions to the science world and the broader community. Care work is important. Thank you for your role in the web of care.

Reference

1. Fisher B, Tronto J. Toward a feminist theory of caring. In: Abel EK, Nelson MK, editors. Circles of care. Albany, NY: SUNY Press; 1990. p. 36.

HOW TO CITE:
Swartz L. Behind the scenes. S Afr J Sci. 2021;117(3/4), Art. #10437. https://doi.org/10.17159/sajs.2021/10437

AUTHOR:
Michael Cardo[1]

AFFILIATION:
[1]Member of Parliament, Democratic Alliance, South Africa

CORRESPONDENCE TO:
Michael Cardo

EMAIL:
mcardo@parliament.gov.za

HOW TO CITE:
Cardo M. Belinda Bozzoli (1945–2020). S Afr J Sci. 2021;117(3/4), Art. #10385. https://doi.org/10.17159/sajs.2021/10385

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

PUBLISHED:
29 March 2021

Belinda Bozzoli (1945–2020): Historical sociologist, academic and politician

Professor Belinda Bozzoli, one of South Africa's leading historical sociologists who entered politics after a distinguished career in academia, died on 5 December 2020.

Her death, after a struggle with cancer, came as a great blow to her colleagues in Parliament. We valued Belinda for her incisive intellect, the original insights into our national condition that she delivered in strikingly wrought prose or speeches that made the House sit up and listen, and her complete mastery of the complex higher education portfolio. She combined warmth and compassion with an independent-mindedness that is all too rare in politics. Belinda had a keen sense of humour, and she could detect the ridiculous in the sublime and the sublime in the ridiculous. These were qualities that endeared her to generations of students, whose understanding of South African history and social dynamics she shaped through her infectious enthusiasm and rigorous scholarship.

Belinda Bozzoli was born on 17 December 1945 in Johannesburg. After matriculating from Parktown Girls' High in 1962, she was encouraged by her father Guerino – then Dean of Engineering at the University of the Witwatersrand (Wits) – to enrol for a BSc degree with science and engineering subjects. The course failed to capture her interest. She channelled her energies into the choral society and politics instead. After a year, Belinda switched to a BA degree in political science and geography, and graduated at the end of 1966.

A spell of teaching at a high school in London followed. During this time Belinda met her future husband, the historian Charles van Onselen, and a circle of Africanist scholars with similar interests. The idea of a career in African studies began to take shape. Belinda returned to Wits in 1969, where her father had recently been installed as Vice Chancellor, and took a first-class honours degree. A brief stint as a journalist on the *Rand Daily Mail* ensued, but the pull of higher education was strong. Belinda proceeded to read for her master's degree and doctorate at the University of Sussex. For much of the time she was based in Oxford where Van Onselen was pursuing his DPhil, and where the ideas of the History Workshop movement – with its emphasis on 'history from below' – exerted a powerful influence. The fortnightly seminar on African studies convened by Shula Marks at the University of London was also formative.

Belinda's doctoral thesis focused on the rise to dominance of South Africa's capitalist ruling class, and the role of intellectuals in expanding the cultural purchase of its worldview. The study, which drew on Gramsci's concept of hegemony, formed the basis of her first book, *The Political Nature of a Ruling Class: Capital and Ideology in South Africa, 1890-1933*, published in 1981.

It was natural that Belinda should return to Wits to build her academic career. The university was personally and intellectually part of her heritage. At the prompting of Dunbar Moodie, she successfully applied for a post in the Sociology Department, which she filled in January 1977. Over the next three and a half decades, Belinda served in a variety of leadership roles at Wits, including Head of Sociology (1996–1998); Head of the School of Social Sciences (2001–2003) and, ultimately, Deputy Vice Chancellor for Research.

Belinda's own research focused on the intersection between culture and politics. She was interested in the cultural tapestries of working-class people's lives. Although she was among the vanguard of Marxist scholars who charted a new course for South African historiography in the 1970s and 1980s, Belinda never reduced culture and politics to a mechanistic function of economics. She questioned the simplistic conflation of apartheid with capitalism. She resisted dogma and orthodoxy. And she was an early adopter of interdisciplinarity. Belinda's research output included numerous articles, among them the seminal 'Marxism, Feminism and Southern African Studies' (1983). She wrote a further two single-author books, *Women of Phokeng: Consciousness, Life Strategy, and Migrancy in South Africa* (1991) and *Theatres of Struggle and the End of Apartheid* (2004). *And she* edited several books that emerged from the Wits History Workshop which she helped establish in the late 1970s.[1-3] Modelled on the Oxford History Workshop, it weaved historical, sociological and political research and forged links with creative arts practitioners as well as trade unions. In 2006, Belinda was the first sociologist to receive an A-rating from the National Research Foundation (NRF).[4] She later served on the NRF Board as Chairperson. Belinda's scholarship was internationally recognised, and she held fellowships at Yale, Oxford, Cambridge and the *École des hautes études en sciences sociales* in France.

Having reached retirement age at Wits, but with much still to offer, Belinda successfully stood for Parliament in the 2014 election under the Democratic Alliance (DA)'s banner. She was appointed as the DA's Shadow Minister for Higher Education and Training (the portfolio was expanded to include Science and Technology after the 2019 election), a position for which she was equipped with a wealth of experience. Belinda held successive cabinet ministers' feet to the fire with her in-depth subject knowledge, dogged questioning and assiduous oversight. She used the parliamentary mechanisms at her disposal to maximum effect, but was horrified by the thin veneer of executive accountability.[5]

For all that, she took her own responsibilities seriously and mined her portfolio. She realised the core problem with South Africa's higher education system is its unsustainable funding model. In Belinda's own words, the country was 'trying to operate the most generous…system in the world in one of the most desperate and failing economies in the world'.[6] She worried that the Department of Higher Education and Training (DHET) had become so consumed with the practicalities of student funding and welfare that it had lost sight of teaching, learning and research excellence. DHET, she maintained, was 'stagnant and bureaucratically overburdened', with several dysfunctional entities.[7]

By contrast, the Department of Science and Innovation (DSI) – within whose ambit the NRF and *Council for Scientific and Industrial Research* (*CSIR*) fall – was, in Belinda's view, a 'well-run gem'. She marvelled at the research being done locally in robotics, nanotechnology, advanced astronomy, climate science, bio-innovation, and the hydrogen economy, among others, and she celebrated our 'many top researchers, some world class, who do this work'. But Belinda acknowledged that the DSI was 'one of the most neglected Departments in government' with a paltry (and shrinking) budget that barely equalled the budget of just one of our major universities. She worried how the DSI would fare in a merger with DHET, predicting it would be a 'litmus test' of the seriousness with which the government takes its commitment to the Fourth Industrial Revolution.

Belinda Bozzoli was as enthused by science and technology as she was by her own disciplinary interests. As her eldest son, Gareth, said of her in a moving eulogy: Belinda 'marveled at architecture and engineering and medicine. She was in awe of nature and biology, of both the universe and the atom. And she immersed herself in theory and debate, in sociology and history, in politics, science and philosophy'.[8]

Having touched the lives of so many personally, Belinda leaves a great void in our communal intellectual and political life, too. Heartfelt condolences to her husband Charles, and children, Gareth, Jessica and Matthew.

References

1. Bozzoli B, editor. Labour, township and protest. Johannesburg: Ravan Press; 1979.
2. Bozzoli B, editor. Town and countryside in the Transvaal. Johannesburg: Ravan Press; 1983.
3. Bozzoli B, editor. Class, community, and conflict. Johannesburg: Ravan Press; 1987.
4. Vaughan CL. On the shoulders of Oldenburg: A biography of the academic rating system in South Africa. Pretoria: NRF; 2015. p. 186–193.
5. Bozzoli B. The engine room of parliament is nothing more than a merry go round. Business Day. 2019 July 16. Available from: https://www.businesslive.co.za/bd/opinion/2019-07-16-belinda-bozzoli-the-engine-room-of-parliament-is-nothing-more-than-a-merry-go-round/
6. Bozzoli B. The crisis in NSFAS: A response [speech in Parliament]. 2018 September 11. Available from: https://www.da.org.za/2018/09/the-crisis-in-nsfas-a-response
7. Bozzoli B. The ANC vs the Fourth Industrial Revolution [speech in Parliament]. 2019 July 09. Available from: https://www.politicsweb.co.za/documents/the-anc-vs-the-fourth-industrial-revolution
8. Van Onselen G. Belinda Bozzoli: A tribute [webpage on the Internet]. 2020 December 11. Available from: https://www.politicsweb.co.za/opinion/belinda-bozzoli-a-tribute

AUTHOR:
Michèle Ramsay[1]

AFFILIATION:
[1]Sydney Brenner Institute for Molecular Bioscience, University of the Witwatersrand, Johannesburg, South Africa

CORRESPONDENCE TO:
Michèle Ramsay

EMAIL:
Michele.ramsay@wits.ac.za

HOW TO CITE:
Ramsay M. Marianne Alberts (1928–2020): Caring, compassionate and humble biochemist and pioneer. S Afr J Sci. 2021;117(3/4), Art. #9458. https://doi.org/10.17159/sajs.2021/9458

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

PUBLISHED:
29 March 2021

Marianne Alberts (1928–2020): Caring, compassionate and humble biochemist and pioneer

In this tribute, we celebrate the remarkable life and work of Marianne Alberts (née Andersson) – biochemist by training, pioneer, teacher, researcher, and caring human being – who was born on 6 August 1928 in Ostersund, Sweden, and who developed her academic and research career, spanning almost seven decades, in South Africa. After graduating from the University of Lund (Sweden), she travelled to South Africa in 1954, where she met and married a South African, Dr H.W. Alberts, in 1955. After briefly working at the CSIR, she joined the South African Institute for Medical Research as a biochemist and research assistant until 1974, during which time she obtained her doctorate from the University of Pretoria, graduating in 1963. She first joined the University of the North (now the University of Limpopo) as a lecturer in the Department of Physiology in 1975, and soon became involved in major academic developments. There she was integral to the establishment of the Department of Medical Laboratory Sciences (1979), School of Anatomy (1980) and Department of Nutrition (1984), and played various roles throughout her career at the university. Although Marianne first retired from the university in 1986, she continued to be involved in part-time lecturing, only to be called back to the Department of Medical Laboratory Sciences and remaining in full-time service until 2016. She contributed to the education and careers of generations of young, mostly black, South African scientists, graduating 18 MSc and 10 PhD students and mentoring many more. She published over 60 research papers. In 2019, at the time of her retirement, she received the University of Limpopo's Lifetime Excellence Award and was deeply touched by this honour.

Her legacy project was to establish the Dikgale Health and Demographic Surveillance System (DHDSS) and Research Centre in 1995. The DHDSS was a founder member of the International Network for the Demographic Evaluation of Populations and their Health (INDEPTH) established in 1998. Professor Osman Sankoh, who stepped down as Executive Director of INDEPTH in 2018, said that Marianne never missed an Annual General Meeting or International Scientific Meeting, whether in Africa, Asia or elsewhere. Professor Steve Tollman, Scientific Director of INDEPTH, described her as 'a grande dame of INDEPTH and a testament to what determination can achieve; and that, if anything, age is not a barrier but a blessing'.

Through her work she gave the communities from the villages around the university a voice, and created awareness about the rising levels of non-communicable and infectious diseases and their devastating toll on health and well-being. The DHDSS collected longitudinal population data on vital events, health and socio-economic factors to monitor lifestyle changes, to study prevalent diseases and their causes, with a view to applying these insights to health programmes. In 2018, the target surveyed population was expanded to reach 100 000 individuals, enabled through an ambitious national programme, funded by the Department of Science and Innovation, with the DHDSS becoming a nodal centre of the South African Population Research Infrastructure Network, and being renamed DIMAMO. Marianne was the Emeritus Director of the DIMAMO Population Health Research Centre that was formally inaugurated on 10 December 2018.

The grand opening of DIMAMO was in a marquee tent in the garden on a stiflingly hot day in Limpopo, and it was a great pleasure for me to spend 2 days with Marianne at that time. The opening event was packed with university and governmental partners, funders, collaborators, students and her family, all of whom came to celebrate this great achievement, spearheaded by such a remarkable woman. She energetically stepped up unassisted onto the podium and in her low-key manner explained the genesis and challenges that had led to that moment. The next day Marianne took me on a long drive to the villages of the region and we talked about the people and their complex lives, and the communities and their hardships. It was clear that she knew them well, had spent much time within the villages and truly cared for the people. She showed me the schools, the clinics and the community areas, and on discovering some beautiful flowers on a cluster of succulent plants, we stopped for a while to have a closer look.

Marianne's research and her collaborative work highlighted the rise in non-communicable diseases such as hypertension and diabetes, which are serious public health problems in the country. She lamented the fact that patients had little knowledge of the risks and that compliance with medication remained low. Referring to DIMAMO, she told audiences that the centre's work would boost South Africa's research into inequality, poverty and population health, including non-communicable diseases, and that the work was expected to inform interventions to significantly improve the health and socio-economic well-being of the whole population. The DIMAMO HDSS will continue to collect data that will document the changing exposure to risk factors to various non-communicable diseases in rural South Africa, hopefully saving lives through better knowledge and targeted interventions.

I started working closely with Marianne in 2012. Our team, spanning four African countries, met for the first time in Addis Ababa, Ethiopia, at the inauguration of the Human Heredity and Health in Africa (H3Africa) Consortium. I discovered Marianne's love for walking and for a glass of red wine with her evening meal – activities we were to pursue on many occasions. Our H3Africa study is referred to as AWI-Gen, the Africa Wits-INDEPTH Partnership for Genomic Research, and Marianne was the principal investigator of one of its six study centres. She was always responsive and involved her team in activities, personally attending many workshops and meetings of H3Africa in different African countries. As recently as October 2019, despite formally retiring on 31 July 2019, she attended the H3Africa meeting in Tunisia and remained a valued member of the team to the end of her life. She passed on 1 August 2020 in Polokwane, South Africa.

Described by her granddaughter as a 'rebellious trailblazer', and by her students, colleagues and friends as extraordinary, humble, knowledgeable, dignified, dependable, wise and compassionate, thoughtful and generous of mind and heart, the world is a better place for Marianne's life, well lived.

AUTHOR:
Wieland Gevers[1]

AFFILIATION:
[1]Emeritus Professor of Medical Biochemistry, University of Cape Town, Cape Town, South Africa

CORRESPONDENCE TO:
Wieland Gevers

EMAIL:
wieland.gevers@uct.ac.za

HOW TO CITE:
Gevers W. Stuart J. Saunders (1931–2021): Mover and shaker, mostly behind the scenes, and key South African vice-chancellor of his time. S Afr J Sci. 2021;117(3/4), Art. #9739. https://doi.org/10.17159/sajs.2021/9739

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

PUBLISHED:
29 March 2021

Stuart J. Saunders (1931–2021): Mover and shaker, mostly behind the scenes, and key South African vice-chancellor of his time

One of the unresolved questions of a free society is how to remunerate citizens in proportion to the contribution that they individually make. Top business executives, bankers, footballers and pop stars are very well rewarded because one can supposedly clearly quantify the financial benefits they generate by simply doing their jobs, although the so-called externalities are ignored in publicly making the calculation. Top teachers, researchers, thinkers and reformers are at the opposite end of this spectrum – the visible and direct financial difference they make is minimal, and everything lies in the hidden externalities.

Thus it is that the contributions made to South Africa's national competitiveness in science and technology by Stuart John Saunders, a canny doctor of humble origin who became a productive researcher in liver disease and then vice-chancellor of a university and still-active sage of a higher education system, have not been widely enough appreciated. This is because they were intrinsic and systemic rather than overtly financial, and because his were focused and well-timed interventions that took time to bear fruit. (It is interesting in our present context that many top businesspeople seemed to appreciate quite readily that he was their equal, which made him a master fundraiser and influencer.)

Saunders was born in Cape Town on 28 August 1931 of English immigrant parents, both of whom had known financial hardship and career disappointment. His schooling was in the same city and he studied medicine, graduating with distinction, at the University of Cape Town (UCT). His subsequent work towards specialisation in internal medicine led to a fellowship that allowed him to obtain the specialist qualification of MRCP at the first attempt, and to establish himself as a researcher in London at the Royal Postgraduate Medical School at Hammersmith, and at Harvard University. He returned to Cape Town, completed the research-based MD degree at UCT in 1965, and became professor and chair of Medicine in 1971. Together with surgical colleagues, he established a combined service and research centre for basic and translational research on liver disease. Over his relatively short major-time research career he became a well-recognised and well-connected scholar in his field, with about 200 publications to his credit as author or co-author. Together with the parallel focus of eminent researchers at the University of the Witwatersrand, these activities contributed significantly to South Africa's well-deserved international reputation for clinical research on the liver.

Saunders became the vice-chancellor of UCT at the beginning of 1981, the third medical scholar to achieve this distinction. (Significantly, no fewer than five of the ten vice-chancellors of the university so far have been medically qualified. In his case, it was reflected in the fact that he had an intense and principled interest in people and their welfare, without sentimentality but with a practical and realistic approach to doing what, in his view, was needed.) As related in his autobiographical book *Vice-Chancellor on a Tightrope – A Personal Account of Climactic Years in South Africa* (David Philip Publishers, Cape Town, 2000), he fearlessly took on the hostile apartheid government by making moves that were difficult to counter that would open the institution fully to black students, by resisting state defunding and other threats skilfully in the courts, and even privately by personal contacts with powerful people made possible by application of his considerable professional skills. He opened the residences, ended the racialised training of registrars, and began an 'academic development programme' to assist underprepared but talented students. He strove for academic freedom as a necessary prerequisite for excellence in teaching and scholarship.

Internally, Saunders left an indelible legacy on the academic process and research at UCT. The current high ranking of the university in Africa and the world at large can be ascribed in no small part to his insistence on chairing every selection committee for a university chair, making no concessions to special pleading or parochialism. His emphasis on research as the decisive 'edge' of the institution in its quest for academic excellence was a hallmark of his leadership role at UCT. Significant internationalisation took place. He took the lead in garnering support from overseas foundations, and raised significant funding for the university by establishing dedicated trusts for this purpose in the USA and the UK.

Saunders's overall emphasis on academically informed leadership of a university helped to shape the country's entire university system. This assessment does not take away from the achievements of other vice-chancellors who headed institutions in South Africa in the last two decades of the 20th century, but it was a constant inspiration and stimulus for them to know that liberal, values-based leadership was not only viable but could be very successful, and he was steady as a rock on basic principles. Not everything that was needed could be done in his 16 years at the helm, and it was left to his successors to restructure the faculties and try to address the basic pedagogic issues that affected the coherence and appropriateness of the general undergraduate curriculum, to modernise and consolidate the postgraduate model, to reach critical mass in research groups, and to introduce postdoctoral fellowships as an indispensable element of high-level university-based research. (Wise national policy in the introduction of well-supported centres of excellence and research chairs was also enormously helpful, as well as the pouring in of international grants to counter the twin epidemics of HIV infection and tuberculosis.) But Saunders had laid the groundwork for strong research development through his emphasis on quality and his sheer force of character. South Africa's relatively strong presence in most long lists of the world's best universities owes a lot to him. It is now one of the plusses in an otherwise somewhat patchy national performance dashboard.

After Saunders's retirement in 1996 he set out on a path that would turn him from the 'very model of a vice-chancellor' into the 'very model of a retired vice-chancellor'. Apart from continued active roles in UCT's convocation, legacy society and fundraising trusts, and a 9-year period of service on the Council of Higher Education, he was the senior advisor of the New York based Andrew W. Mellon Foundation, where he advocated a major expansion of grant-making in South Africa, particularly in the arts, humanities and the interpretative social sciences. For example, there has to date been support for 250 scholars through faculty development awards, and 1840 scholarships (1700 at graduate level and140 postdoctoral fellowships) to cultivate future generations of arts and humanities scholars. Library resources, digital collections, and information and communication technology facilities have been strengthened at higher education institutions. A major transformative contribution has been made to opera, helping to harvest a rich vein of talent in the population. Digitisation of historically significant 'rock art' has made thousands of rock art paintings accessible online, preserved for future generations. All this came from the personal chemistry of the relationship between Saunders and Dr Bill Bowen, president of the Andrew W. Mellon Foundation at that time.

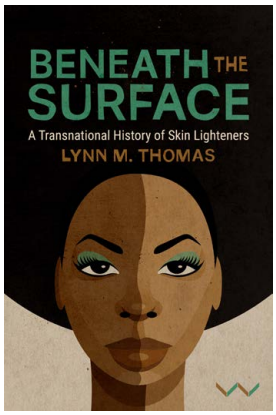
Another major contribution made by Saunders was his role as the essential midwife of TENET (originally the 'Tertiary Education Network' for South Africa). When the unsuccessful government-run Uninet was closed down, the vice-chancellors of the universities and the then technikons decided to form a company which would set up and supply bandwidth to all institutions as well as the research councils. They asked Saunders to take overall responsibility for this project, partly on account of his close involvement with the Andrew W. Mellon Foundation, which had also invested in the computerisation of university libraries and was concerned about the high costs of bandwidth. He established a task team to investigate options, and it was decided to form a non-profit company with a board of directors chaired by Saunders. TENET was incorporated in August 2000 and has continued since then to play a crucial role in ensuring an adequate and uninterrupted supply of bandwidth to the higher education and research council system. The importance of these far-sighted arrangements cannot be over-emphasised in the light of subsequent developments and the COVID-19-induced extension of online learning, teaching and research.

Saunders's other post-retirement activity of significance to the country's knowledge system was his active involvement in the rapid expansion of postdoctoral fellowships in the research landscape. As an influential member of the board of the Claude Leon Foundation he focused on the higher education component of its grant-making in South Africa. Since 1998, over 500 multi-year postdoctoral fellowships were awarded across the system – a very significant fraction of the total awarded in the country as a whole. The critical gap in bursary support for honours students has also been met to some extent.

The total 'value-added' of Stuart Saunders's career cannot be calculated in rands and cents, or even in buildings, new institutions or donations. Yet I venture to assess it as systemically immense within our country, and wholly inadequately reflected in the personal memories of those of us who knew him as a friend and colleague, a companionable and funny yet deadly serious and purposeful individual. Many millions of citizens who have never heard of him are the real beneficiaries of his remarkable life.

Photo: UCT News

BOOK TITLE:
Beneath the surface: A transnational history of skin lighteners



AUTHOR:
Lynn M. Thomas

ISBN:
9781776146154 (softcover, 368 pp)

PUBLISHER:
Wits University Press, Johannesburg; ZAR347

PUBLISHED:
2020

REVIEWER:
Caradee Y. Wright[1,2]

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

EMAIL:
Caradee.Wright@mrc.ac.za

HOW TO CITE:
Wright CY. Unravelling historical and intertwined complexities of light and dark skin. S Afr J Sci. 2021;117(3/4), Art. #9092. https://doi.org/10.17159/sajs.2021/9092

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

PUBLISHED:
29 March 2021

Unravelling historical and intertwined complexities of light and dark skin

Diving 'Beneath the Surface', Lynn Thomas, Professor of History, uncovers the secrets, stories and complexities of one of humankind's greatest 'obsessions' – the 'self', particularly how skin colour contributes to that image of self. Thomas' involvement in a project called 'The Modern Girl Around the World' spurred her immersion into cosmetic advertising and her realisation of the widespread use of skin lightening products among both black and white women during the early 20th century.[1]

Through multiple lenses, Thomas expresses the different human perspectives related to the skin and how these perspectives came about. By producing a critical effect through understanding the interconnectedness of human beings, emotion and consumption, over a long period of history, Thomas identifies and describes the causal changes that create consuming individuals.

In six chapters, plus an introduction and a conclusion, the reader is taken on a chronological and historical journey that unravels the practices of skin lightening, how it occurred in different geographic locations, between cultures and across cultural frontiers, including very diverse groups of people. In the chapter entitled 'Modern Girls and Racial Responsibility', female beauty and black women's application of skin lightening products are considered in the light of beauty ideals and embracing consumer culture while taking into account disrupting factors related to patriarchy, gender and culture. Another chapter is dedicated to reconstructing how commercial manufacturing and marketing of skin lighteners took place over time. Through social, political and racial interactions and triggers, the skin lightening industry transformed and reinvented itself according to the current context of human need and other issues that were significant at the time.

Thomas explains how skin lighteners operated as a 'technology of visibility' and people used them to attract favourable attention, enhance their appearance and render themselves 'legible' (even acceptable) in multiple scenarios. As a 'technology', skin lightening was deemed necessary in confronting racism, for example, in attempts to obtain official racial reclassification, or to enhance prospects in social and work settings where people with so-called 'privileged light skin' were included more readily. However, resistance to these skin lightening products, as well as to changes in skin colour to alter one's racial group, was rife from individuals, political parties, community leaders and even nations. As Thomas describes: 'The anti-racist activists framed skin lighteners as overdetermined evidence of the psychological effects of structural racism and bearing ground for what was called the modern goal, versus the term loving blackness and ethnic that rejected white beauty standards.'

The book gives a transnational account of skin lightening, but the South African content is especially prominent with examples and occurrences in our country referred to in detail. One such example relates to non-white medical school students founding the South African Student Organisation (SASO) – the all-black group from which Black Consciousness thought emerged. The leader of SASO, Steve Biko, crafted a political ideology that addressed the demands of life under apartheid with new ways of thinking. 'Black is Beautiful' was embraced by activists and skin lighteners were condemned in order to promote racial self-respect. When hydroquinone, the ingredient in skin lighteners, became a medical concern, the political and health issues became complexly intertwined. While skin lightening creams containing hydroquinone were banned in 1990 in South Africa (the first country globally to do so), products with hydroquinone (and mercury, also used in skin lightening products) are still available and used in the country today.[2] These products are used, for example, by individuals who have skin conditions that lead to scarring and pigmentation side-effects, as a result of ingesting certain medications prescribed to them for other health outcomes.

Beneath the Surface is the most comprehensive book regarding skin lighteners available to date and it is both interesting and innovative. It is indeed an international description of skin lightening (and brightening) presented in an historical account with a smart balance between text and original advertisements, photographs, newspaper clippings and personal accounts that bring the complicated narrative to life. Thomas evidently spent considerable time gathering material and literature from a variety of sources. With 56 pages of notes, a 47-page bibliography and a useful 12-page index, this book is nothing short of an encyclopaedic production.

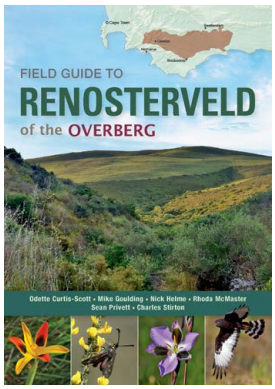
During my review of this intriguing book, I dog-eared dozens of pages to return to and re-read because they contained statements that were revelations to me. For anyone wishing to read the book as a 'novel', it is recommended that personal notes are made to help navigate the content when searching to revisit items of interest.

The book has value as a postgraduate textbook relevant to the fields of history, social science, geopolitics, gender studies, geography, psychology, dermatology, and others. The layered, integrated history presented by Thomas in *Beneath the Surface* is indeed 'a landmark study' of skin colour and skin lighteners that interrogates every influencing factor from slavery and segregation to consumer capitalism, political protests and reinforced social inequities, and beyond.

References

1. Moosa F. In conversation with Lynn Thomas, author of book on skin lighteners. The Daily Vox. 2020 July 03 [cited 2020 Oct 14]. Available from: https://www.thedailyvox.co.za/in-conversation-with-lynn-thomas/
2. Davids LM, Van Wyk J, Khumalo NP, Jablonski NG. The phenomenon of skin lightening: Is it right to be light? S Afr J Sci. 2016;112(11/12), Art. #2016-0056. http://dx.doi.org/10.17159/sajs.2016/20160056

BOOK TITLE:
Field guide to renosterveld of the Overberg



AUTHORS:
Odette Curtis-Scott, Mike Goulding, Nick Helme, Rhoda McMaster, Sean Privett, Charles Stirton

ISBN:
9781775845805 (epub, 488 pp)

PUBLISHER:
Struik Nature, Cape Town; ZAR290

PUBLISHED:
2020

REVIEWER:
Brian W. van Wilgen[1]

AFFILIATION:
[1]Centre for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, Stellenbosch, South Africa

EMAIL:
bvanwilgen@sun.ac.za

HOW TO CITE:
Van Wilgen BW. A guide to a vanishing flora. S Afr J Sci. 2021;117(3/4), Art. #9169. https://doi.org/10.17159/sajs.2021/9169

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

PUBLISHED:
29 March 2021

A guide to a vanishing flora

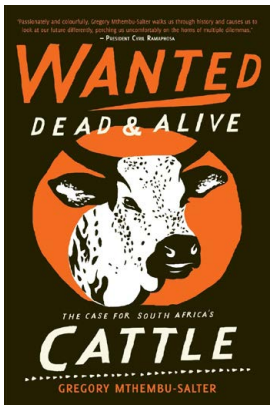
Renosterveld is a highly endangered vegetation type that once covered over 2 million hectares within the Cape Floristic Region. Today, the estimated 18 000 remaining patches are widely scattered across the landscape, collectively covering less than 4% of the original extent of this once-expansive veld type. Renosterveld differs from the neighbouring and better-known fynbos in that it occurs on nutrient-rich soils derived from shale, as opposed to the nutrient-poor sandstone soils that support fynbos. Fynbos is best known for its tall proteas and showy ericas, as well as reed-like restios. Renosterveld, on the other hand, is more grassy and characterised by shrubs of low to medium height. In the past, and unlike fynbos, renosterveld supported an abundance of large grazing mammals and their accompanying predators. This began to change around 2000 years ago when the Khoekhoen arrived with domesticated livestock whose more selective feeding habits would have differed from those of wildlife, thus impacting on the flora. Large wild mammals were virtually exterminated when European settlers arrived with modern firearms some 300 years ago, and domestic livestock became the dominant grazers. The fate of the renosterveld was finally sealed after World War II, when mechanised farming allowed large tracts of land to be converted to crops such as wheat and canola. Whether the original renosterveld was a shrubby grassland or a grassy shrubland is a topic that ecologists debate today, as the changing grazing pressures and human-influenced fire regimes would have affected the proportional contribution of these two important components of the vegetation. Better understanding of this issue would be important for managing this vegetation type correctly, and this book provides some insights into these intriguing questions. Nonetheless, forming an acceptably robust grasp of the functioning and dynamics of such a fractured ecosystem is akin to visualising the image of a 100-piece jigsaw puzzle from four remaining pieces.

This book is, however, not primarily intended as an ecological treatment. Rather, it provides a guide to almost 1000 plant species that occur in the renosterveld of the Overberg, a region that stretches over 300 km between Caledon and Mossel Bay, south of the Riviersonderend and Langeberg mountains. Many of these plant species are found nowhere else on earth, and many are classed as Endangered, or Critically Endangered, in South Africa's Red List of Plants. The concise documentation of such a large, rich and often unique flora is a notable achievement. Each species is described in adequate detail, and where appropriate its IUCN 'Red List' status is noted. A key feature that makes this book so attractive is the remarkably high quality of the photographs of each plant. Modern digital photography has resulted in vast improvements that will allow readers to identify species much more easily. Over 70% of the book (350 pages) is devoted to plant species, but an additional 50 pages also cover some of the more interesting birds, mammals, reptiles, fish and invertebrates of the region. While these sections are not comprehensive, they will certainly be useful for visitors who are unfamiliar with the area's fauna.

The intended readership would presumably include nature lovers and plant enthusiasts, as well as tourists. However, renosterveld is not readily accessible to members of the public, as 99% of the remaining land is in private ownership, presenting a challenge to potential users of this book. The authors devote part of this book to landowners who are taking positive steps towards conserving renosterveld patches on their land, and they describe the initiatives of 12 prominent farmers who are labelled 'conservation heroes'. The book also introduces the 500-ha Haarwegskloof Renosterveld Reserve, which is home to the Overberg Renosterveld Conservation Trust and a first-ever Renosterveld Visitor and Research Centre. The reader is informed that 'all nature lovers are welcome', and although no further details are given, they are available on the Internet.

This book is the first to be devoted to the flora, fauna, ecology and management of the renosterveld, and as such it is a major step forward. It should be of considerable interest to farmers and other landowners of the area, but it will also raise awareness of the conservation importance of renosterveld among town dwellers and tourists and will be a useful resource for students of botany and ecology. Odette Curtis-Scott and her co-authors are to be congratulated on producing such a handy and attractive guide.

BOOK TITLE:
Wanted dead & alive: The case for South Africa's cattle



AUTHOR:
Gregory Mthembu-Salter

ISBN:
9781928466130 (softcover, 184 pp)

PUBLISHER:
Cover2Cover Books, Cape Town; ZAR250

PUBLISHED:
2019

REVIEWER:
Wandile Sihlobo[1,2]

AFFILIATIONS:
[1]Chief Economist, Agricultural Business Chamber of South Africa (Agbiz), Pretoria, South Africa
[2]Visiting Research Fellow, Wits School of Governance, University of the Witwatersrand, Johannesburg, South Africa

EMAIL:
wandile@agbiz.co.za

HOW TO CITE:
Sihlobo W. Why do South Africans appreciate cattle farming so much? S Afr J Sci. 2021;117(3/4), Art. #9168. https://doi.org/10.17159/sajs.2021/9168

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

PUBLISHED:
29 March 2021

Why do South Africans appreciate cattle farming so much?

Agricultural books are not widely available in popular bookstores in South Africa. And even when they are available, they tend to be written in a style that is not accessible to non-experts in the field. Such challenges usually mean that the literature is restricted to academics and practitioners in the agriculture sector alone. While it could be argued that this is not a phenomenon unique to agriculture, the conditions are different, especially if we take a South African perspective. Agriculture is one of the sectors in which the South African government aims to increase the participation of black people, and also to ensure that farming of all kinds will assist to revitalise the rural economy and create jobs.

The dearth of accessible agricultural books is what makes Gregory Mthembu-Salter's new book, *Wanted: Dead & alive: The Case for South Africa's Cattle,* so important. Mthembu-Salter takes the reader on a historical, social, economic and cultural journey around cattle in South Africa. Mthembu-Salter is neither farmer nor animal scientist, but he does a brilliant job in narrating cattle farming in this book. Born in the United Kingdom and a South African immigrant in the 1990s, as a researcher and journalist, his curiosity about the value (not merely in economic terms) of cattle can be traced to his lobola negotiation days, and through various cultural activities that followed after his marriage.

Upon appreciating how valuable cattle were in African culture, Mthembu-Salter took it upon himself to trace the history of cattle in South Africa, going back to the Khoekhoen era in the 1400s. He begins by walking the reader through this historical journey, relating how European explorers in 1488 first encountered the Khoekhoen and their cattle, moving on to colonial times when black South Africans experienced massive losses of cattle through wars and diseases, until our democratic era in which cattle are part of modern farming in South Africa.

The second chapter is particularly important to readers with an interest in understanding the cultural relationship that South African black families have with cattle. Mthembu-Salter captures this vividly through observations of various events he attended in the process of his research. He describes, in the smallest of details, how meat is served at cultural gatherings, and what specific pieces are reserved for the elders.

The third chapter will possibly answer questions that some readers might have about why South Africans are so fond of eating beef or meat in general. The per capita consumption of beef in South Africa is 17.42 per kg as per 2019 data from the Department of Agriculture, Land Reform and Rural Development. This is the highest in Africa, albeit lower than that in several developed economies. Aside from a general narrative of what drives meat consumption – which is boosted by higher income levels to a certain extent – Mthembu-Salter presents the reader with a 'taste' of various meats that South Africans enjoy and provides their history. This includes biltong, whose story is told from the perspective of the San people, moving on to boerewors, to processed meat cuts and off-cuts and also Jewish kosher meat products that were introduced into South Africa. The home braai (barbecue), shisanyama and steakhouse are some of the venues in which South Africans consume their meat. The value of the book is in providing the reader with a clear picture of the setting and history of these meat-eating venues.

Chapter four deals with what may be familiar to some readers: the structure of a modern beef value chain, from farm to fork. This information is provided through the infusion of both hard science research and the narratives from various farmers whom the author interviewed. This will also be an important chapter for anyone interested in starting a beef farm or other business within the value chain. Mthembu-Salter gives a broad view of the industry, although perhaps the author could have given more detail about the economic or commercial business side of the beef industry – which is what the chapter title promised but failed to deliver.

An important component of the livestock industry is dairy farming, which is not often covered in various popular agriculture books in South Africa. Mthembu-Salter gives a brief window into this important sub-sector, but he leaves the reader yearning for more information on the structure of the dairy industry, both its commercial and non-commercial aspects. Perhaps there may be further writing from this author that could include a contribution that captures the evolution of the dairy industry in South Africa through the various centuries, along with government policies and changes in consumer preferences that influence the industry.

The book ends with important discussions that locate cattle farming in the land reform and democratic state agricultural policies. Cattle farming is positioned within the current rural economy and the author emphasises its potential role in uplifting rural economies, especially in the provinces that were former homelands. This is particularly significant because an estimated 34% of cattle in South Africa in 2019 were owned by black farmers according to the National Agricultural Marketing Council. The persistent question is how best to ensure that these farmers' cattle are integrated into the formal beef value chains. On the positive side, however, there are private-public-partnership models such as those spearheaded in the Free State by the Sernick Group, a diversified agricultural organisation, which are doing encouraging work into commercialising black farmers.

The book was published in 2019 but it will increasingly become more relevant as agriculture is one of the key sectors that government hopes will drive economic activity and job creation as the country recovers from the COVID-19 pandemic. The publication carries an important history of cattle farming and explains the structure of the industry which will be useful to any non-specialist joining the sector. Importantly, Mthembu-Salter should be commended for ending his work by positioning it in current agricultural development policies. Overall, this is a welcome contribution, rich with history, culture and insights into the economic development of cattle farming in South Africa.

AUTHORS:
Donrich Thaldar[1,2]
Beverley Townsend[1,2]

AFFILIATIONS:
[1]School of Law, University of KwaZulu-Natal, Durban, South Africa
[2]African Health Research Flagship, University of KwaZulu-Natal, Durban, South Africa

CORRESPONDENCE TO:
Donrich Thaldar

EMAIL:
ThaldarD@ukzn.ac.za

HOW TO CITE:
Thaldar D, Townsend B. Protecting personal information in research: Is a code of conduct the solution? S Afr J Sci. 2021;117(3/4), Art. #9490. https://doi.org/10.17159/sajs.2021/9490

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
exemption, cross-border transfer, functional equivalent, Information Regulator, POPIA, specific consent

PUBLISHED:
29 March 2021

Protecting personal information in research: Is a code of conduct the solution?

In 2009, the South African Law Reform Commission published its report on privacy and data protection. Four years later this culminated in the enactment of the *Protection of Personal Information Act 4 of 2013* (POPIA) by Parliament. To provide society with sufficient time to prepare to be POPIA compliant, POPIA's substantive provisions only entered into force in July 2020. In addition, POPIA itself provides for a one-year grace period before compliance becomes compulsory. During the latter part of 2020, the Academy of Science of South Africa (ASSAf) initiated a project to develop a code of conduct in terms of POPIA for all research activities. In this essay, we explore (1) the purpose of codes of conduct and (2) the concerns of the science community regarding POPIA, and (3) pose the question: is a code of conduct the solution to address these concerns?

The purpose of codes of conduct

POPIA provides that the Information Regulator may issue codes of conduct for particular sectors. Codes of conduct can be useful tools to facilitate compliance, as codes of conduct can explain and apply POPIA's principles to sector-specific activities. What are these principles? Most importantly, POPIA sets out eight 'conditions' for the lawful processing of personal information, namely: accountability, processing limitation, purpose specification, further processing limitation, information quality, openness, security safeguards, and data subject participation. Researchers and research institutions (or 'responsible parties') are required to ensure that all measures are taken to adhere to these conditions, unless one or more particular conditions have been specifically excluded or exempted from operation under specific provisions within POPIA.

POPIA is clear: a code of conduct must incorporate all the conditions for the lawful processing of personal information, or set out obligations that provide a *functional equivalent* of the obligations set out in the conditions. A position or practice of 'functional equivalence' is one that performs the *same function* and provides the *same utility* as is required by the provision. It is therefore not envisaged in POPIA that a code of conduct can fundamentally deviate from the eight conditions for the lawful processing of personal information.

What a code of conduct can do – and in fact, *must* do – is to prescribe how the conditions for the lawful processing of research data are to be applied, and how the provisions within POPIA are to be complied with within a particular sector. In doing this, the code of conduct should focus on the typical activities particular to a given sector, in an attempt to explain and demonstrate how POPIA applies to these activities.

Concerns about POPIA compliance

During an ASSAf workshop on the intended code of conduct, held online on 10 December 2020, the following concerns about POPIA were most prominent among participants (in our observation):

- uncertainty about what exactly POPIA entails in general;
- the issue of cross-border transfers of personal information; and
- POPIA's requirement that consent must be specific, which is a break from the past where broad consent was deemed sufficient, and which poses a challenge especially to biobanks with historical data that were collected without specific consent.

Can a code of conduct offer a solution to these three concerns? We analyse these concerns seriatim.

Uncertainty

A well-drafted code of conduct should clearly explain the principles of POPIA and also provide guidelines on exactly how to protect personal information at every stage of the research process. Such a code of conduct would indeed address the concern about uncertainty.

Cross-border transfers

A code of conduct will offer a partial solution to concerns about cross-border transfers. First, consider a scenario in which health information and biometric information are to be transferred to a research institution in a foreign country that does not provide an adequate level of protection. This scenario would be typical in the context of health research. The default position in this scenario is that the South African research institution must obtain prior authorisation for the intended transfer from the Information Regulator. However, this requirement of having to obtain prior authorisation is obviated if the South African research institution is operating under the auspices of a code of conduct that has been approved by the Information Regulator. As such, in this scenario, a code of conduct does offer a solution.

Next, consider a scenario in which a South African research institution intends to transfer personal information (not limited to health information and biometric information) to a research institution in a foreign country. In the absence of specific consent by research participants to transfer their personal information to this research institution in a foreign country, the South African research institution may still transfer the personal information if there is an adequate level of protection of personal information by either the law in the foreign country or by an agreement between the two research institutions. (This is of course subject to the paragraph above, and also subject to the terms of the informed consent provided by the research participants.) What is needed by the South African science

community is a *standard data transfer agreement* that can be used to ensure that an adequate level of protection of personal information is in place. Ideally, such a standard data transfer agreement should be developed together with the code of conduct as complementary legal instruments.

Specific consent

POPIA's definition of consent as 'voluntary, specific and informed expression of will' (section 1) has been the subject of academic debate: Staunton et al.[1,2] argue that POPIA allows for *broad* consent. We find this argument unconvincing, and have proffered a full critique of Staunton et al.'s position[3,4], to which they have replied[5]. Our position can be encapsulated as follows: POPIA contains exceptions for research; however, these exceptions are subject to certain requirements, including the requirement that the original collection of data must be done for a 'specific, explicitly defined and lawful purpose' (section 13). This poses a particular challenge to researchers who might have gathered data of thousands of data subjects in the past without obtaining consent for a 'specific, explicitly defined and lawful purpose'. Also, going forward, researchers may, for good reasons, wish to collect data for a broad range of possible research projects. Is there – within the interpretation of POPIA requiring specific consent, as we suggest – a solution to these issues?

The appropriate strategy to address issues such as biobanks with historical data (obtained without specific consent and in the absence of any other of the legitimate grounds described in section 11) would be to approach the Information Regulator for an *exemption* from specific consent in specified circumstances in terms of section 37. The Information Regulator may only grant an exemption if she is satisfied that the public interest outweighs any possible interference with the privacy of the research participant *to a substantial degree*. Accordingly, the mere inconvenience of complying with the requirement of specific consent is unlikely to suffice as a reason for granting an exemption. More solid, principled reasons would need to be put forward. For example, a principled reason in the context of *health* research would be that the research links with the right to access to health care and ultimately with the right to life. Furthermore, an exemption application should be supported by evidence, preferably in the form of empirical studies on representative samples of South African research participants to ascertain their opinions on the sufficiency of less specific kinds of consent. Provided that the outcome of such studies is favourable, and provided that principled reasons can be presented for a specific field of scientific endeavour, and under specified circumstances, the combination of evidence and principled reasons would constitute a good argument in support of an exemption.

A code of conduct remains important to explain and clarify how specific consent ought to be obtained in the context of scientific research, and how the requirements for the research exclusions (where no consent for further research is required by POPIA) are to be met, for example, by setting out what exactly 'sufficient guarantees' are (section 27(1)(d)). A code of conduct should also take cognisance of the Department of Health's Ethics Guidelines[6], which under certain circumstances require research participant consent for further research (unlike POPIA). Ideally, a code of conduct would provide a *consolidated guide* for the science community. When engaged in this consolidation exercise, care should be taken to not erode any of the constituent sources' requirements – whether POPIA or the Ethics Guidelines.

Conclusion

A code of conduct is not a panacea. To address concerns about POPIA compliance in the South African science community, we recommend that ASSAf expands its current code-of-conduct-development initiative to include: (1) the development of a standard data transfer agreement that can be incorporated as an annexure to the code of conduct, and (2) an investigation into the need, scope and justification for a possible exemption from specific consent in certain contexts.

Acknowledgements

We acknowledge an African Health Research Flagship Grant from the University of KwaZulu-Natal and a South African National Research Foundation Grant (grant no. 116275).

Competing interests

We declare that there are no competing interests.

References

1. Staunton C, Adams R, Botes M, Dove ES, Horn L, Labuschaigne M, et al. Safeguarding the future of genomic research in South Africa: Broad consent and the Protection of Personal Information Act No 4 of 2013. S Afr Med J. 2019;109(7):468–470. https://doi.org/10.7196/SAMJ.2019.v109i7.14148
2. Staunton C, Adams R, Anderson D, Croxton T, Kamuya D, Munene M, et al. Protection of Personal Information Act 2013 and data protection for health research in South Africa. Int Data Priv Law. 2020;10(2):160–179. https://doi.org/10.1093/idpl/ipz024
3. Thaldar D, Townsend B. Genomic research and privacy: A response to Staunton et al. S Afr Med J. 2020;110(3):172–174. https://doi.org/10.7196/SAMJ.2020.v110i3.14431
4. Townsend B, Thaldar D. Navigating uncharted waters: Biobanks and informational privacy in South Africa. S Afr J Hum Rights. 2019;35(4):329–350. https://doi.org/10.1080/02587203.2020.1717366
5. Staunton C, Adams R, Botes M, Adams R, Botes M, Dove ES, et al. Correspondence. S Afr Med J. 2020;110(3):175–176. https://doi.org/10.7196/SAMJ.2020.v110i3.14450
6. South African Department of Health (DoH). Ethics in health research: Principles, processes and structures. Pretoria: DoH; 2015.

AUTHOR:
Keyan G. Tomaselli[1,2]

AFFILIATIONS:
[1]Department of Communication Studies, University of Johannesburg, Johannesburg, South Africa
[2]Centre for Communication, Media and Society, School of Applied Human Sciences, University of KwaZulu-Natal, Durban, South Africa

CORRESPONDENCE TO:
Keyan Tomaselli

EMAIL:
keyant@uj.ac.za

HOW TO CITE:
Tomaselli KG. The geography of Plan S open science. S Afr J Sci. 2021;117(3/4), Art. #9256. https://doi.org/10.17159/sajs.2021/9256

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
open-access publishing, OA2020, open-access funding

PUBLISHED:
29 March 2021

The geography of Plan S open science

Plan S is a grand plan, but the devil is in the detail.[1]

So says Robin Crewe, Chair of the Academy of Science of South Africa (ASSAf)'s Committee on Scholarly Publishing in South Africa. This European-sponsored plan devised by research funders replaces barriers to reading with barriers to authorship. The 13 European funding agencies (cOAlition S) promoting Plan S require research to be published in compliant open access (OA) journals from 2020.

Universities South Africa (USAf) is working with OA2020, an initiative aligned with Plan S, but devised by actual research producers. The global subscription system is argued to contain sufficient funds to transition to OA. OA2020 assumes that scholars publish for impact, not for money. If the 'devil is in the detail', then this assumption is contradicted by the effects of the well-intentioned Department of Higher Education and Training (DHET) publishing incentive that has enabled South African universities to position their researchers as relentless rent-seekers.

Academic labour occurs within three interlocking spheres: a scholarly community, a bureaucracy and a corporation. Plan S and OA2020 add further overlays on these often contradictory sites of labour, production and consumption. For example, Plan S requires that academic authors (or their institutions) pay to get published. Plan S aims to 'accelerate the transition to a scholarly publishing system that is characterised by immediate, free online access to, and largely unrestricted use and re-use of scholarly publications'[2].

The academic world would divide into different research coalitions. These would be (1) the traditional global subscription model, other than (2) the European Plan S enforced silo; and (3) the Open Knowledge for Latin America and the Global South (AmeliCA) university-based communication infrastructure option. Plan S signatories, however, represent only 5% of global research output.

South Africa at 1% has already committed to developing OA, with ASSAf's SciELO SA providing an electronic OA platform for participating journals. But SciELO is not a funding mechanism. OA2020's argument is that the legacy publishers are hoarding and monetising public knowledge, allegedly preventing access to readers. Legacy publishers are thus accused of commoditising a good that should be 'free'. At the social level, says USAf's Ahmed Bawa, there is a 'growing unaffordability even for our research-focused institutions' (Bawa A 2020, written communication, 20 October). The huge inequalities of access in our system and 'the growing disconnect between society and 'science' need to be addressed.

The questions arising, however, are: Is Plan S another Northern imposition on the South? Will Plan S ringfence EU research only for funding, and restrict permission to publish? For Lyn Kamerlin[3], the Plan S 'embargo requirements and repository technical requirements … are so draconian that paid-for gold becomes the easiest way to fulfil them'. Plan S thus simply flips who pays, entrenching the for-profit publishing firms.

The outcome will impoverish the research community by disadvantaging those unable to afford article-processing charges (APCs). It might also discourage cooperation across publishing systems. The funders' aim is to prevent publishers and authors from gaining income from what they seem to implicitly claim as `their paid-for' research, positioning themselves as the new scholarly oligarchy. The beneficiaries of Plan S will be mainly sponsored researchers able to publish in fully fledged (non-hybrid) OA journals. The DHET incentive would need to be paid immediately on acceptance of an article rather than 18 months after, as is the DHET norm.

The global publishing system could be fractured by Plan S into non-compatible ecosystems of research reporting – what will be legitimate for one system could be declared illegitimate in the other. This is already the case with the way that the DHET incentive, as myopically applied by universities, herds South African authors into particular qualifying lists.

By limiting the legacy publishers with their added values of peer review, plagiarism and libel checks, cross-referencing, copy editing, legal protections, ethical regimes, marketing and so on, further opportunities will be opened to the ever-opportunistic predators. Such kinds of differential access will result in automatic inequality in publishing opportunities based on geographic location and funding availability. This means that the cost of publishing rather than the quality of research will decide where and what research is published.

AmeliCA wants a collaborative, non-commercial, sustainable and non-subordinated system returned to the academy. Under this scenario, DHET would require universities to invest in infrastructure and technology for science communication – i.e. journals – to be located within universities, and to delink from 'legitimation systems' like Scopus and Web of Science. But then the ASSAf-identified problem of 'house journals', and the DHET requirement of 'accreditation', could be exacerbated in an already overprovisioned local environment.

Kamerlin[3] further states that 'openness … needs to be community driven, not funder driven' as OA2020 appears to be. But can a one-size-fits-all model work? In the event of the aligned Plan S and OA2020, the DHET incentive will have to shift from rewarding universities (and their authors) to now awarding APCs to journals for articles approved for publication. South African journals are currently excluded from the DHET funding value chain and *themselves subsidise the publication of research at no or small cost to authors, universities and the state*.

The devil will be in the detail no matter the initiatives adopted: USAf's national approach is to secure broadened access to scientific material and to utilise the current national spend of about ZAR600 million on subscriptions to achieve OA and to eliminate the need for individual researchers to pay APCs.

The question is, in light of the corporatisation of academic labour, production and consumption, who will be supported, and how will under-funded authors be assured of access? That is, which authors, which journals, and which paradigms?

Should not scientists' academic freedom remain at arm's length from state, funders and universities? And, should we not be publishing for impact, not just DHET monetary incentives?

Competing interests

There are no competing interests to declare.

References

1. Crewe R. Plan S is a grand plan, but the devil is in the detail: Robin Crewe on open access in South Africa [webpage on the Internet]. c2019 [cited 2020 Nov 17]. Available from: https://council.science/current/blog/plan-s-is-a-grand-plan-but-the-devil-is-in-the-detail-robin-crewe-on-open-access-in-south-africa/
2. Amsterdam Call for Action on Open Science [document on the Internet]. c2016 [cited 2020 Nov 17]. Available from: https://government.nl/binaries/amsterdamcallforactiononopenscience
3. Reaction of researchers to Plan S; Too far, too risky? [webpage on the Internet]. No date [cited 2020 Nov 17]. Available from: https://sites.google.com/view/plansopenletter/open-letter?authuser=0

AUTHOR:
Daniel Cunnama[1]

AFFILIATION:
[1]South African Astronomical Observatory, Cape Town, South Africa

CORRESPONDENCE TO:
Daniel Cunnama

EMAIL:
daniel@saao.ac.za

HOW TO CITE:
Cunnama D. The South African Astronomical Observatory 200 Virtual Symposium. S Afr J Sci. 2021;117(3/4), Art. #9392. https://doi.org/10.17159/sajs.2021/9392

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
National Heritage Site, African astronomy, SALT, astro-tourism, astronomy for society, National Astrophysics and Space Science Programme

PUBLISHED:
29 March 2021

The South African Astronomical Observatory 200 Virtual Symposium

The South African Astronomical Observatory (SAAO), formerly The Royal Observatory Cape of Good Hope, celebrated 200 years of existence as an astronomical observatory on 20 October 2020 and is the oldest scientific institution in South Africa.

The Royal Observatory Cape of Good Hope was founded on 20 October 1820, and for much of its history, it was the major contributor to positional astronomy in the southern hemisphere. The SAAO is not only known for its rich history and various contributions to science, its buildings are also of special architectural significance. Consequently, on 21 December 2018, the South African Heritage Resources Agency (SAHRA) officially declared the SAAO a National Heritage Site and on 20 October 2020 it was formally unveiled.

History of the Observatory

At the time of its foundation, the Royal Observatory was formally controlled by the British Admiralty and was initially intended for the improvement of navigation. Its main duty was to chart the southern skies and provide a time service for passing ships in Cape Town Harbour.

The work of the Observatory, however, soon moved into scientific inquiry and discovery. Indeed, only a few years after the completion of the Royal Observatory in 1832/1833, Thomas Henderson made the first measurements to find the distance to a star. The Royal Observatory found its forte in the late 19th century by leading the way in astronomical photography and the cataloguing of stars. For this, it gained a global reputation. The 20th century saw research move into the more fundamental physics of stellar dynamics and evolution, including the development of spectroscopy with the donation of the Victoria Telescope by Frank McClean. This led to further breakthroughs such as the discovery of oxygen, silicon, and europium in stars.

In 1951, the Observatory gained access to the 1.9-metre Radcliffe telescope in Pretoria which remained the largest telescope in South Africa until 2004. In 1971, a decision was taken to amalgamate the major facilities for optical astronomical research into one body, which became known as the South African Astronomical Observatory.

There was a new beginning in South African optical astronomy in 1972 with the foundation of the SAAO's Sutherland Observatory, situated far from the Cape Town city lights, wet winters and air pollution. The Sutherland site offered ideal conditions for astronomy being at high altitude with cold, clear air and a steady atmosphere. One of the primary telescopes at the new Sutherland site was the 1.9-metre telescope from the Radcliffe Observatory in Pretoria which was moved to Sutherland following the closure of the Radcliffe Observatory, and it became operational again in January 1976.

The South African astronomy community began considering the idea of acquiring a new telescope towards the late 1980s and had initially envisaged a rather modest 4-metre instrument. However, these dreams were far surpassed when, in 1998, South Africa became the leading partner in the Southern African Large Telescope (SALT), which was to become the largest optical and infrared telescope in the southern hemisphere. It is designed to provide maximum collecting area for minimum cost and is especially suited for spectroscopy of faint objects.

SALT comprises 91 hexagonal mirrors arranged to form a combined diameter of over 10 metres and was additionally backed by a consortium of international partners. The SALT consortium involved institutions from nearly a dozen countries in Europe, the USA, Britain and New Zealand, with South Africa constituting the largest partner. Based on the Hobby-Eberly Telescope in Texas, this facility has provided South African astronomers with outstanding capacity in the fields of optical photometry and high-speed spectroscopy with South African scientists having access to one-third of the observing time.

The South African National Research Foundation stated in 1999 that SALT's 'two primary missions' were (1) to provide a state-of-the-art facility for local and international astronomers and astrophysicists and (2) to overcome and redress past government policies that 'dislocated the majority of South Africans from science, engineering and technology education'.

From 2011, after a period of commissioning and performance verification, SALT started full science operations, coming into its own as Africa's 'giant eye' on the universe, with significant contributions to global astronomy.

One of the most notable and dramatic discoveries in recent years came in August 2017 when four telescopes at Sutherland contributed to a global effort to detect the first visible counterpart of a gravitational wave source. Gravitational waves were predicted by the theory of General Relativity but are extremely difficult to detect as they are ripples in the structure of space–time and are not like electromagnetic radiation. Gravitational waves were observed from a source in a known galaxy 130 million light years away on 17 August 2017. SAAO observed the aftermath of this neutron star merger with four telescopes: SALT, MASTER-SAAO, the 1.0-m and the Infrared Survey Facility, with SALT obtaining the first spectrum of the source.

Today Sutherland is home to more than 20 telescopes of various shapes and sizes with SALT remaining the flagship and the 1.9-m telescope still fully operational.

SAAO 200 celebrations

To celebrate the bicentenary of the Observatory, various events were arranged including the unveiling of the SAAO as a National Heritage Site, the SAAO 200 Astronomy Symposium, and the SAAO 200 Virtual Astronomy Festival.

SAAO Managing Director Prof. Petri Vaisanen stated:

> *This occasion is an opportunity to recall some great scientific achievements. But more than that, it is an opportunity to celebrate our country's and continent's rich heritage in attempting to understand the universe and our place in it. In particular, we want to convey the pure excitement of exploring the amazing universe we are part of, and also highlight the many benefits that science brings to society. The theme of our event is 'Beyond 200 Years of Astronomy', and we see this future full of opportunity, inspiration, and pride in the excellence of decidedly African astronomy at the forefront of a cutting-edge global pursuit.*

Unveiling of the SAAO as a National Heritage Site

The formal 'virtual' unveiling of the National Heritage Site was held on the 20 October 2020 and was attended by a limited number of dignitaries including Director-General of the Department of Science and Innovation (DSI), Dr Phil Mjwara, and CEO of SAHRA, Adv. Lungisa Malgas. The unveiling included pre-recorded addresses by the Minister of Sports, Arts and Culture, the Honourable Nkosinathi Mthethwa, and the Minister of Higher Education and Training, Science and Innovation, the Honourable Dr Blade Nzimande.

The programme included interviews with SAAO staff, and the premiere of a new animation of indigenous Khoesan starlore, 'Moons Message', which was very well received by dignitaries, the community, and the media. The livestream of the unveiling was viewed by over 1000 people.

SAAO 200 Virtual Symposium

The SAAO 200 Symposium was planned as an in-person symposium due to take place from 20 to 22 October 2020. Owing to the COVID-19 pandemic, the decision was taken to shift from an in-person symposium to a virtual one. The symposium was then extended to become a 4-day event, starting with the unveiling of the SAAO as a National Heritage Site, followed by the SAAO 200 Symposium.

The original target was set at 300 attendees for an in-person symposium. Shifting to virtual allowed for more participants and complimentary registration fees. The symposium kicked off on 20 October with the number of registrations at 575. In total, 626 attendees registered and, at any given stage, an average of 127 viewers were engaged.

In terms of participants, 75.6% were from South Africa. The Western Cape represented 52.5%, Gauteng 12.8%, Eastern Cape 3.2%, KwaZulu-Natal 2.2%, Mpumalanga 1.5%, Free State 1.4%, North West 1.3%, and Limpopo 0.9%. Participants from other countries included those based in Britain, Mauritius, Nigeria, India, USA, Turkey, Japan, Uganda and Zimbabwe.

The virtual symposium provided an ideal opportunity to incorporate some leading international speakers without incurring exorbitant travel costs. This helped to develop a very strong programme.

Astronomy for society

The SAAO 200 Virtual Symposium saw presentations covering a wide range of topics, including current and future science, the history of astronomy on the continent, as well as the cultural and sociological aspects of astronomy.

The opening session of the Symposium included keynote addresses focusing on a wide range of astronomy-related topics. The President of the International Astronomical Union (IAU), Prof. Ewine van Dishoeck, addressed the IAU's role in development, outreach and education including its Office Astronomy for Development (OAD) at SAAO in Cape Town. Prof. Vanessa McBride provided further details of the excellent work being done by the OAD to utilise astronomy to achieve the United Nation's Sustainable Development Goals.

One highlight of the Symposium was the address by the SAAO Manager of Collateral Benefits in Sutherland, Mr Anthony Mietas, who delivered a stirring account of the remarkable achievements of the programme over the past decade. Through this work, Sutherland has become an astro-tourism hub and has created 302 jobs directly and, indirectly, a number more. SAAO/SALT remains the single largest employer in the town of Sutherland, and SALT and SAAO continue to utilise the Sutherland-based local companies for various projects.

The SAAO and the National Research Foundation have refurbished both the primary and high schools' laboratories and SALT purchased school desks for the intermediate learners in the Roggeveld Primary School.

Additionally, the Sutherland Community Development Centre is an initiative from the SAAO, with support from SALT, the local community, and various partners, in particular the DSI as the main sponsor. The centre provides connectivity, childcare and several cultural, artistic, sports and social events throughout the year, enhancing the life of the community.

Various talks at the symposium addressed critical societal issues within astronomy such as transformation, human capital development and successful outreach initiatives across South Africa and the rest of Africa. Teams such as the SAAO Outreach department, the Astronomical Society of Southern African and the African Astronomical Society presented on their efforts to engage with the community.

A rich history

The Symposium featured a variety of presentations on the diverse history of the Observatory and of astronomy in Africa, including presentations on indigenous knowledge and ethno-astronomy and efforts to better communicate this valuable knowledge and the efforts of astronomers at the Cape before 1820.

Dr Ian Glass delivered his keynote address on 'The Cape Observatory from 1820 to 1972', highlighting the myriad scientific achievements during those years and this was complemented by more detailed discussions of some of the key figures involved, such as Sir David Gill, as well as the role of the Royal Astronomical Society in South African astronomy.

More recent developments in African astronomy, such as the advent of SALT and the genesis of the Square Kilometre Array, were presented by Dr David Buckley and Dr Bernie Fanaroff, respectively, illustrating the tremendous advances in astronomy facilities on the continent in recent decades. The SKA Director-General, Prof. Phil Diamond, addressed some of the challenges and opportunities arising from one of the largest science experiments in his talk entitled 'SKA: Building an Observatory to Study the Dawn of Time and the Origins of Life'.

A bright future

The bright future for African astronomy would not be possible without the development of African astronomers. Prof. Patricia Whitelock gave an overview of the hugely successful National Astrophysics and Space Science Programme (NASSP) founded in 2003. Hosted at the University of Cape Town, the University of KwaZulu-Natal and North-West University, the programme has seen over 300 students graduate with honours degrees and more than 140 with master's degrees.

In addition, since 2011, 87 students who were trained at SAAO and SALT have graduated – approximately 40% with honours, 30% with MSc and 25% with PhD degrees. Of these students, 63 (72%) were people of colour. Students at NASSP and SAAO represent the rich diversity and culture of our continent, coming from a range of countries including

Rwanda, Uganda, Ethiopia, Mauritius, Sudan, Burkina Faso, Lebanon and Madagascar.

Many of these students were represented in the Symposium schedule, both in oral and poster presentations, with talks covering observational and theoretical aspects of stellar and extragalactic astronomy, astronomical instrumentation, and computer modelling. Many projects involve combining and analysing data from different wavelengths, e.g. optical data from SAAO/SALT with radio data from MeerKAT and/or X-ray measurements from spacecraft.

Astronomy is expanding across the African continent, with the newly revitalised African Astronomical Society. On the horizon is the largest meeting in the astronomy calendar, the General Assembly of the International Astronomical Union. This meeting in Cape Town in 2024 will be the first General Assembly held on the African continent since the establishment of the Union over 100 years ago – securing the recognition of Africa's contribution to global science.

Competing interests

There are no competing interests to declare.

AUTHOR:
Ntombizodwa Mathe[1]

AFFILIATION:
[1]National Laser Centre, Council for Scientific and Industrial Research, Pretoria, South Africa

CORRESPONDENCE TO:
Ntombi Mathe

EMAIL:
NMathe@csir.co.za

HOW TO CITE:
Mathe N. Insights from 'Unlocking COVID-19 current realities, future opportunities: Artificial intelligence in the time of COVID-19'. S Afr J Sci. 2021;117(3/4), Art. #9311. https://doi.org/10.17159/sajs.2021/9311

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
4IR technologies, healthcare technologies, autonomous systems, machine learning

PUBLISHED:
29 March 2021

Insights from 'Unlocking COVID-19 current realities, future opportunities: Artificial intelligence in the time of COVID-19'

The University of Cape Town, in partnership with Standard Bank, hosted a webinar entitled 'Unlocking COVID-19 current realities, future opportunities: Artificial intelligence in the time of COVID-19' on 19 August 2020. The webinar was facilitated by Professor Tommie Meyer from the University of Cape Town's Centre of Artificial Intelligence Research at the Department of Computer Science. The two speakers for the event were Professor Tshilidzi Marwala, Vice Chancellor and Principal of the University of Johannesburg and a pioneer in the field of artificial intelligence (AI) in South Africa and the Deputy Chair of the Presidential Commission on the Fourth Industrial Revolution (4IR). The second speaker was Mr Nanda Padayachee, Head AI, Automation and APIs at the Standard Bank Group. He has extensive experience in digitisation of capabilities and established the Inaugural Data Science capabilities within the Standard Bank Group.

Prof. Marwala started the session with an overview of the COVID statistics globally, including South Africa, with the number then at 22 million cases. Economies have been devastated, with many companies closing down just 5 weeks into the hard lockdown, with a negative impact on families and on the whole of South Africa. He observed that, coincidentally, the pandemic has occurred alongside the acceleration of 4IR technologies. This is a time when technologies such as AI are permeating all aspects of our lives and merging the identities of people and identities of machines into one. AI has been used extensively in all areas of our lives and health care is one of the main industries using this technology. For instance, a study by Lancet Digital Diagnostic AI uses deep learning models with the aim of being equivalent to that of healthcare professionals in diagnosing illnesses correctly. However, the use of AI is not intended to make doctors obsolete, but rather to take over many of the time-consuming and tedious aspects of patient care while allowing for faster diagnoses. Prof. Marwala also used an example of a video circulated in China at the beginning of the pandemic in February 2020 that showed a drone communicating with an elderly woman instructing her to wear a mask and wash her hands. However, this kind of application can also generate the fear that, by using 4IR technologies, state surveillance agencies are able to track people and their movements in instances that are not benign. Nonetheless, AI has been an important tool in diagnosing and preventing illness while also assisting many people to deal with the impact of isolation through the availability of online platforms that allow video calls for families to stay connected and robot assistants that assist the elderly with household chores.

Prior to lockdown, the Beijing subway system used AI to scan body temperatures and identify any COVID-19 symptoms in the passengers using the trains. Moreover, algorithms are being used by healthcare professionals to scan chest X-rays of patients in order to differentiate between pneumonia and COVID-19. An AI program has also been developed to determine which coronavirus patients will develop serious complications, thus allowing doctors to determine which patients are more likely to need ventilators. There are, of course, limitations to current AI technologies which rely on reading pictorial data of ultrasounds and CT scans, yet these have been proven to be useful tools in managing the disease and patient care. An AI healthcare start-up company in Daejeon, South Korea, used deep learning models to determine how strongly a molecule would bind to a protein, for determination of the ideal medication for patients with COVID-19.

Similarly, Benevolent AI in the United Kingdom used AI and existing data to sort through existing drugs that could be used to treat patients with COVID-19 until a vaccine became available. VIR Biotechnologies in San Francisco, USA, used algorithms to identify molecules that could be targeted towards treating COVID-19. In March 2020, the US government launched a project in collaboration with Tech Giants and academics to make coronavirus research accessible to AI researchers and their algorithms in an effort to aid health experts. Such measures are needed because AI cannot be used unless data are available.

It was explained in the webinar how the use of AI in combating the pandemic has ranged from robotic cleaners that spray disinfectants in hospital wards, to AI voice assistants calling people to gather information on home quarantine. In China, autonomous robots have replaced human cleaners, which has not only reduced infection rates but also prolonged working time. Voice-assisted AI has contributed to monitoring and control of the spread of the virus by collecting personal information and symptoms to alert officials to possible hotspots. An AI search tool called COVID-Scholar was developed by a researcher at the Lawrence Berkeley National Laboratory, University of California Berkeley, to assist researchers and clinicians to sift through the literature on the virus – much needed because there have been over 28 000 scholarly publications related to coronavirus since the beginning of the pandemic.

In South Africa, the robot 'Pepper', piloted at Nedbank, has been deployed as a robot companion to counteract the surge of loneliness during the pandemic, particularly among the older population. Pepper can make phone calls and identify missing items in the kitchen. AI has also been used as a government services digital platform called GovChat in South Africa, which allows for the easy access of information as well as for tracking and tracing to assist authorities in identification of COVID-19 hotspots so as to manage resource allocation. At the University of Johannesburg, Prof. Marwala and his student have developed an AI tool that they describe in their publication in *PLoS ONE*[1] entitled 'Bayesian inference of COVID-19 spreading rates in South Africa' and which is able to estimate vital statistics such as infection and incubation rates. The CSIR and the Department of Science and Innovation have developed a dashboard that can streamline data that will help the government COVID-19 National

Command Council to better manage the pandemic. By teaming up with Telkom and Samsung, the South African government is able to track and trace people who may have contracted the coronavirus using data such as geographical information systems (GIS) and mapping from their cellphones. While if a person has already contracted the coronavirus, they are able to personally assist by doing a daily check-in on the Impilo app developed by the Gauteng Department of Health in 2019.

Despite the successes of AI, there is a growing debate for regulation. Prof. Marwala is a member of the World Health Organization's committee that is developing guidelines for the application of AI in medicine so that the technology cannot be misused. AI technologies are often not guided by policy but rather used to pilot new ideas. The consequence is that there are great unknowns as well as potential dangers given the speed of adoption of these technologies and the frantic nature of trying to find speedy solutions to the pandemic. We have not experienced a pandemic like this since the 'Spanish' influenza in 1918 that killed over 50 million people, 350 000 of whom in South Africa, and the isolation, social distancing, total lockdown, and working and schooling from home are collectively a new experience. Prof. Marwala concluded his presentation with a quote from the 2014 book *The 1918 Spanish Flu Pandemic: The History and Legacy of the World's Deadliest Influenza Outbreak* by Charles River Editors:

> *In many ways, it is hard for modern people living in First World countries to conceive of a pandemic sweeping around the world, and killing millions of people. And it is even harder to believe that something as common as influenza could cause such widespread illness and death.*

This is now our reality, and yet in the 4IR we are far better equipped than we have previously been, with AI being the most powerful tool in our arsenal.

Prof. Marwala was followed by Nanda Padayachee who pointed out that any notion of AI being in the realm of science fiction is no longer the case and AI has impacted our lives and society as a whole. He noted that even before COVID-19 there were technological advances that made our lives 'not normal'. For instance, we need to consider the global inequalities and remember that there are over 600 million people untouched by the First Industrial Revolution, without access to electricity or any form of mechanisation, while half the current population lacks access to Internet connectivity. We are reaching a point at which our impact as a species is becoming detrimental to the planet. The question is: 'How do we use AI to create a better normal?' AI must have a far better reach in creating the 'new normal' to address societal needs in a far more equitable manner. In terms of what we want to achieve with AI, Padayachee broke it down into three specific categories:

1. Using AI to improve existing processes, services and products with the opportunity for incremental growth such as autonomous systems.
2. Using AI to create new ideas, for instance in medicine, agriculture, and in productivity generally.
3. Considering how AI can lead to breakthrough advances, e.g. personalised medicine, agriculture, and mining for using resources more efficiently. For example, one mining company has utilised machine learning to optimise productivity and has achieved a 20% increase simply by streamlining their processes, without having to drill new mines or expend additional resources.

Often when we think of new technologies, we consider them in a private sector context or consumer context. However, there is a substantial opportunity for these technologies to be used in the public sector, i.e. from government to consumer and the transition to a digital government. Countries like Estonia are making substantial leaps towards a digital government. COVID-19 has shown that as a country, South Africa can move quickly and make bold decisions, and more of such thinking and action will allow us to serve our people, entrepreneurs, and organisations to a far greater extent. One of the creative features of AI is the ability to recommend actions with human oversight.

In conclusion to this fascinating and information-packed webinar, Prof. Marwala emphasised that the three areas mentioned by Nanda Padayachee needed to be assessed urgently and reminded the audience that South Africa can gain advantage by collaborating with other countries to take full advantage of the technologies on offer.

Competing interests

There are no competing interests to declare.

Reference

1. Mbuvha R, Marwala T. Bayesian inference of COVID-19 spreading rates in South Africa. PLoS ONE. 2020;15(8), e0237126. https://doi.org/10.1371/journal.pone.0237126

AUTHOR:
Velisha A. Perumal-Pillay[1]

AFFILIATION:
[1]Discipline of Pharmaceutical Sciences, University of KwaZulu-Natal, Durban, South Africa

CORRESPONDENCE TO:
Velisha Perumal-Pillay

EMAIL:
perumalv@ukzn.ac.za

HOW TO CITE:
Perumal-Pillay VA. Pharmacists as vaccinators in South Africa – addressing COVID-19 and beyond. S Afr J Sci. 2021;117(3/4), Art. #8761. https://doi.org/10.17159/sajs.2021/8761

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
vaccines, immunisation, health promotion, health education, pharmacy

PUBLISHED:
29 March 2021

Pharmacists as vaccinators in South Africa – addressing COVID-19 and beyond

As the world grapples with subsequent waves of COVID-19, the only hope for ending this pandemic is successful vaccination. As of March 2021, over 2.7 million people have succumbed to the virus globally and the number is still rising.[1] Hence, the news that some pharmaceutical companies such as Pfizer and Moderna had managed to successfully develop a vaccine was uplifting and welcomed. South Africa joined the COVAX Facility, thereby ensuring the country would have access to these vaccines via the financing mechanism.[2] Professor Salim Abdool Karim, Chair of the Ministerial Advisory Committee on COVID-19 for South Africa, advised that the appropriate vaccine option should be a single dose vaccine to reduce the burden on medical infrastructure and cost, and the selected candidate vaccine should also be compatible with existing platforms of vaccine technology currently used in the country for ease of roll-out.[3] Once vaccines are approved as safe and efficacious by our regulatory authority, the South African Health Products Regulatory Authority (SAHPRA), they are registered and roll-out will begin.

The South African national roll-out strategy for COVID-19 vaccines was outlined by the Ministry of Health in a national briefing on 3 January 2021.[4] The strategy is a three-phased approach: Phase 1 – vaccination of all frontline healthcare workers; Phase 2 – vaccination of essential workers, persons in congregate settings, persons >60 years, persons >18 years with comorbidities; and Phase 3 – other persons >18 years. The roll-out strategy makes provision for individuals to receive the vaccine through both public healthcare settings (primary healthcare clinics, hospitals, outreach-based vaccination programmes, mobile clinics) and private healthcare settings (pharmacies, doctors, work-based vaccination programmes). To support the successful roll-out of these vaccines, a trained workforce is required. Currently, those healthcare professionals allowed to vaccinate are predominantly doctors and nurses trained to do so in line with the Vaccinators Manual and Expanded Programme on Immunisation for South Africa (EPI-SA).[5] South Africa should be looking to expand the workforce that is able to provide extended immunisation services to meet the immediate demand for vaccinators to fulfil the needs of the COVID-19 vaccine roll-out. There should be widespread availability of vaccinators to reach the proportion of the population, in all locations including remote areas, identified for initial vaccination according to the roll-out strategy and then the rest of the population, as more vaccines are procured and areas of acute COVID-19 infection are identified. With a population of over 59 million[6], South Africa will need a suitably sized workforce to meet the demand of the COVID-19 vaccine roll-out. According to 2020 statistics of the South African Pharmacy Council (SAPC) (the statutory professional body for pharmacists), there are 17 842 registered pharmacists (including interns and community service pharmacists) and 3481 community pharmacies, 306 institutional private and 613 institutional public pharmacies that can contribute to this workforce.[7]

Pharmacists in South Africa form part of the healthcare workforce but are limited in their scope of practise as the *Pharmacy Act (Act 53 of 1974)*[8] does not explicitly make provision for pharmacists to practise as vaccinators in the country, although it is not prohibited if adequate training is provided. The required skills can be acquired through an additional qualification offered on a part-time basis. Although vaccines may be administered in a pharmacy, they cannot be administered by the pharmacist without the required training, and are therefore predominantly administered by a nurse at the pharmacy clinic offering pharmacy-based immunisation services. In comparison, in other countries such as the USA, UK and Australia, pharmacists have been providing vaccination services for years.[9] South Africa is lagging behind in advocating for, and establishing, all pharmacists as vaccinators in the healthcare system as pharmacists are available both at public sector institutions such as primary healthcare clinics and hospital settings as well as at private healthcare institutions and community pharmacies in rural and urban locations.

The International Pharmaceutical Federation, a non-governmental organisation and global body representing pharmacy, strongly promotes pharmacists as vaccinators and is advocating for countries to update regulations on vaccine administration by pharmacists to improve vaccine coverage globally.[10] The COVID-19 pandemic has highlighted the importance and need for this update to be seriously considered and instituted in countries that have not developed pharmacists in vaccine administration in their basic training. The pharmacist has an important role to play in immunisation. Under the current scope of practise in South Africa, pharmacists are involved in the provision of healthcare education, patient counselling, and health promotion related to vaccines, immunisation and vaccine-preventable diseases. This role can be further expanded to include management of patients' immunisation schedules and adverse events reporting following immunisation. This aspect of pharmacovigilance is vital to monitoring the potential successes and adverse effects of vaccines and positive health outcomes for patients post-immunisation.

The current Bachelor of Pharmacy (BPharm) degree curriculum in South Africa is designed to train pharmacists as generalists and content includes learning about vaccine-preventable diseases, vaccines for routine and travel immunisation, immunisation concepts, and logistics of vaccine transport and cold chain management. Hence, pharmacists have theoretical knowledge on vaccines. The Good Pharmacy Practice Guidelines and related SAPC rules[11] state that pharmacists performing immunisation services must be familiar with handling of syringes and administration devices and pharmacists must ensure that they have adequate training, knowledge and skills to provide the service. To administer vaccines, pharmacists require supplementary training such as the Primary Care Drug Therapy certification. There are also private organisations that offer training opportunities for those healthcare professionals, including pharmacists, who wish to train as vaccinators such as the Online Higher Certificate in Vaccinology [HCert (Vacc)] Programme offered by the South African Vaccination and Immunisation Centre at the Sefako Makgatho Health Sciences University in Pretoria[12] and the Health Science Academy[13] (a subsidiary of the

Foundation for Professional Development) that offers an Immunisation and injection technique course. There may be other training providers but there is no list of providers accredited by the SAPC. However, this is supplementary training post-registration as a pharmacist and does not uniformly form part of the undergraduate BPharm training programme in the country, which is what is being advocated.

The benefit of training and including all pharmacists as vaccinators in South Africa would be both immediate and long term. Immediate benefit would be the ability to better respond to the immediate increased demand for immunisation services during the COVID-19 vaccine roll-out by increasing the number of trained vaccinators available in the healthcare sectors. This may also ensure that healthcare services currently offered at hospitals, clinics, and other healthcare facilities are not adversely affected by diverting doctors and nurses to immunisation services. The burden on other healthcare professionals would be decreased if pharmacists could also vaccinate. The long-term effect would be the sustainability of the current EPI-SA which saw interruptions during the COVID-19 pandemic and would ensure continued immunisation services to sustain and maybe even improve vaccination coverage. Currently, vaccination coverage is poor, as evidenced by the South Africa Demographic and Health Survey for 2016 which reported 61% of children (12–23 months of age) received all basic vaccinations, and only 53% received all age-appropriate vaccinations.[14] Community pharmacists have always been within easy reach and easily accessible to the public without consultation appointments. Vaccine accessibility and vaccination coverage could be improved in the long term given this easy access to pharmacists and pharmacy-based immunisation services. Pharmacists, if adequately trained, could also ensure continuity of immunisation services in the absence of nursing staff at their clinic.

If South Africa heeds the call to train all pharmacists as vaccinators, it will involve amending legislation to include training on vaccine administration techniques in the core competencies for pharmacists so that they develop adequate skills as vaccinators during their undergraduate BPharm degree which would also require expanding the scope of practise for pharmacists. These legal changes to the *Pharmacy Act*, which governs the regulations related to the practise of pharmacy, would need to be advocated by pharmacists with the support of professional bodies such as the Pharmaceutical Society of South Africa (PSSA) and the aforementioned SAPC. The PSSA is a professional society which supports and promotes pharmacists by improving medication use and advancing patient care.[15] The PSSA is also well positioned to mobilise an advocacy campaign as a member organisation to the International Pharmaceutical Federation, the very organisation globally advocating for pharmacists to be vaccinators.

The greater the number of pharmacists who are trained and able to administer vaccines, the stronger the vaccinator workforce will be and the greater the vaccine coverage will be due to the ability of pharmacists to identify at-risk individuals, such as children and ageing adults, who are often overlooked where immunisation is concerned, from regular close contact; to provide trusted health information sources to people across all age groups and from several diversities; and to increase accessibility to remote populations.[10] Thus, there is much value in recognising pharmacists' ability to join the global imperative of improving vaccination coverage. Pharmacists of South Africa with the support of their professional bodies should lead the advocacy campaign to train all pharmacists as vaccinators. If the government were receptive, it would provide evidence of South Africa embracing a multidisciplinary approach to health care, as the country moves towards universal health coverage under the national health insurance.

Acknowledgements

Research reported in this publication was supported by the Fogarty International Center (FIC), NIH Common Fund, Office of Strategic Coordination, Office of the Director (OD/OSC/CF/NIH), Office of AIDS Research, Office of the Director (OAR/NIH), National Institute of Mental Health (NIMH/NIH) of the National Institutes of Health under Award Number D43TW010131. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health. Research reported in this publication was also supported by the South African Medical Research Council (MRC) under a Self-Initiated Research Grant. The views and opinions expressed are those of the author and do not necessarily represent the official views of the MRC.

Competing interests

There are no competing interests to declare.

References

1. Covid-19 coronavirus pandemic [webpage on the Internet]. No date [updated 2021 Jan 04; cited 2021 Jan 04]. Available from: https://www.worldometers.info/coronavirus/
2. SA expecting first Covid-19 vaccines from Covax scheme in first half of next year. News24. 2020 December 04. Available from: https://www.news24.com/news24/southafrica/news/sa-expecting-first-covid-19-vaccines-from-covax-scheme-in-first-half-of-next-year-20201204
3. Daniel L. SA will have vaccine by mid-2021, but must be one dose only, says top Covid advisor prof Karim. Business Insider South Africa. 2020 December 04. Available from: https://www.businessinsider.co.za/sa-will-have-vaccine-by-mid-2021-but-must-be-one-dose-only-says-top-covid-advisor-prof-karim-2020-12
4. South African National Department of Health. Covid-19 vaccine rollout strategy [document on the Internet]. c2021 [cited 2021 Jan 04]. Available from: https://bhekisisa.org/wp-content/uploads/2021/01/COVID-Vaccine-Plan-03012021.pdf
5. South African National Department of Health (DoH). Vaccinators manual 2015. 4th ed. Pretoria: DoH; 2015.
6. Stats SA. Mid-year population estimates [document on the Internet]. c2020 [cited 2020 Dec 04]. Available from: http://www.statssa.gov.za/publications/P0302/P03022020.pdf
7. South African Pharmacy Council. Statistics [webpage on the Internet]. No date [cited 2020 Dec 06]. Available from: https://interns.pharma.mm3.co.za/Statistics/PersonsByPharmacistRoleProvince?class=btn%20btn-primary
8. Republic of South Africa. Pharmacy Act 53 of 1974. Government Gazette no.4442, Vol. 112. Available from: https://www.gov.za/sites/default/files/gcis_document/201505/act-53-1974.pdf
9. International Pharmaceutical Federation. An overview of current pharmacy impact on immunization – A global report. The Hague: FIP; 2016. Available from: https://www.fip.org/files/fip/publications/FIP_report_on_Immunisation.pdf
10. International Pharmaceutical Federation. News: FIP releases resource for pharmacists to expand their roles in immunisation [webpage on the Internet]. c2020 [cited 2020 Jul 06]. Available from: https://www.fip.org/news?news=newsitem&newsitem=321
11. Juta Statutes and Regulations of South Africa. The good pharmacy practice manual and associated SAPC rules [document on the Internet]. [updated 2008 May 30; cited 2020 Dec 11]. Available from: https://www.pharmcouncil.co.za/Media/Default/Documents/Rules%20published%20in%20terms%20of%20section%2035A%20of%20the%20Pharmacy%20Act%2053%20of%201974.pdf
12. South African Vaccination and Immunisation Centre. Online Higher Certificate in Vaccinology [HCert (Vacc)] Programme [webpage on the Internet]. No date [cited 2020 Dec 06]. Available from: https://savic.ac.za/vaccinology-courses/
13. Health Science Academy. Immunisation and injection technique [webpage on the Internet]. c2015 [cited 2020 Dec 06]. Available from: https://www.hsa.co.za/course/5
14. South African National Department of Health (DoH). South Africa demographic and health survey 2016 [document on the Internet]. c2019 [cited 2020 Dec 06]. Available from: https://dhsprogram.com/pubs/pdf/FR337/FR337.pdf
15. Pharmaceutical Society of South Africa. About us [webpage on the Internet]. No date [cited 2020 Dec 06]. Available from: https://www.pssa.org.za/aboutus.html

AUTHORS:
Chuene V. Mashamaite[1]
Petrus J. Pieterse[1]
Palesa N. Mothapo[2]
Ethel E. Phiri[1]

AFFILIATIONS:
[1]Department of Agronomy, Stellenbosch University, Stellenbosch, South Africa
[2]Division for Research Development, Stellenbosch University, Stellenbosch, South Africa

CORRESPONDENCE TO:
Chuene V. Mashamaite

EMAIL:
vickymashamaite@gmail.com

DATES:
Received: 22 July 2020
Revised: 18 Nov. 2020
Accepted: 19 Nov. 2020
Published: 29 Mar. 2021

HOW TO CITE:
Mashamaite CV, Pieterse PJ, Mothapo PN, Phiri EE. *Moringa oleifera* in South Africa: A review on its production, growing conditions and consumption as a food source. S Afr J Sci. 2021;117(3/4), Art. #8689. https://doi.org/10.17159/sajs.2021/8689

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITORS:
Teresa Coutinho
Salmina Mokgehle

KEYWORDS:
acceptability, cultivation practices, moringa, nutrition, production areas

FUNDING:
Southern African Systems Analysis Centre, National Research Foundation (South Africa), South African Department of Higher Education, Science and Technology

Moringa oleifera in South Africa: A review on its production, growing conditions and consumption as a food source

Moringa oleifera (moringa) trees excel mainly in tropical and sub-tropical regions and are known to thrive in a wide range of soil types. The recent rise in moringa production in various agro-ecological zones of South Africa could be attributed to its multiple benefits, including nutritional and medicinal properties. Since its introduction as a cultivated crop, there has been a growing interest from farmers, researchers and government on various aspects of the tree such as its morphology, chemistry, growing conditions, production, processing and utilisation. We reviewed the work done on moringa within the South African context in terms of production, growing conditions and cultivation practices. The involvement of government departments on moringa-oriented activities and its consumption as food were also reviewed. In addition, gaps were outlined on its utilisation that need to be addressed, and recommendations provided on what could be done to ensure successful production of moringa in South Africa.

Significance:

- This review highlights moringa research that has been done on growing conditions, production and human consumption in South Africa.
- The review further addresses the potential commercialisation of moringa and existing knowledge gaps.

Introduction

Increased demand for food to alleviate hunger and malnutrition has been pertinent over the last few decades among emerging countries across the globe.[1] In mid-2019, the world population reached 7.7 billion and is projected to reach 9.7 billion in 2050.[2] Moreover, the African population was estimated at 1.3 billion[2] and the South African population reached 58.8 million in mid-2019[3]. Thus, ensuring food security for a growing population has become the leading global challenge. For centuries, humans have been dependent on the diversity of plants for food, nutrition, medicine and shelter, as well as for an energy source and for their overall well-being.[4] Of about 390 000 known plant species, only 7000 species are cultivated or collected for nourishment.[5] Furthermore, less than 150 plant species are farmed commercially for food, with 95% of human food being provided by only 30 species.[4,6] Although these cultivated species could fulfil energy requirements, they could not solely meet dietary nutrient requirements.[6] Therefore, it is imperative to diversify and improve the production of less studied, locally adapted plant species that are used in rural communities as food or as a raw material.[7]

Moringa oleifera (hereafter referred to as moringa) is native to the sub-Himalayan parts of Northern India, and is widely utilised and known among 13 documented species of the Moringaceae family.[8-10] The plant is propagated by either seeds or cuttings, with seeds planted directly in the field or seedlings raised in nurseries.[11] It is well known for its multitude of uses, including nutrition, medicine, livestock feed, plant growth enhancer, cosmetics, water purification and biofuel production.[12-16] Almost all its parts contain nutrients such as proteins, carbohydrates, vitamins (A, B1, B2, B3, C and E) and minerals.[17-19] Additionally, the leaves and roots are an excellent source of natural antioxidants and contain significant amounts of total phenols, tannins and flavonols.[13,20] According to Marcu[21], no adverse consequences from the daily intake of moringa leaves have been documented. The leaves and seeds are often eaten raw, cooked or added into food in a powder form for nutritional and medicinal purpose.[12,17,22] As such, moringa has been used in various developmental projects across the globe for malnutrition and poverty alleviation as well as climate change mitigation.[15,23]

Moringa was introduced to rural communities of the Limpopo Province (South Africa) as a cultivated crop in 2006 by the Lammangata moringa project which is based in Tooseng village.[22] Since then, its production and utilisation have been on the rise in various agro-ecological zones of the country.[24-26] Consequently, various stakeholders such as the South African government, farmers and higher education institutions have started flagships with moringa-oriented projects due to growing national interest in the tree. For example, the Agricultural Research Council's Vegetable and Ornamental Plants group in Roodeplaat have done research on moringa propagation, cultivation practices, processing, storage (shelf-life) and analysis on biological activities, safety and phytochemistry. Nationally, moringa has the potential to improve the nutrition, income and livelihood of marginal communities.[18,26-28] Therefore, we reviewed the work done on moringa within South Africa in terms of its production, growing conditions and brief involvement of government departments on moringa-oriented projects. Moreover, we report on its current and prospective consumption by rural communities and potential uses, as well as other aspects that have not yet been explored in the country. To our knowledge, no research work of this nature has been carried out within the South African context.

Overview of moringa production areas

In South Africa, moringa is produced in six of the nine provinces: Limpopo, Gauteng, Mpumalanga, KwaZulu-Natal, Free State and North West.[24,25] Among these provinces, it is mainly grown in the Limpopo Province by farmers

and at household level.[22] In 2013, the Moringa Development Association of South Africa (MDASA) was formed with a mandate to promote the production, use and commercialisation of moringa. MDASA served as a hub for moringa farmers, product developers and consumers through collaboration with research institutes and organisations for new knowledge and developments. Since its establishment, the number of farmers growing moringa has increased in some parts of South Africa. According to a survey conducted by Mabapa et al.[26], moringa farmers were found present in all five districts of the Limpopo Province (Capricorn, Mopani, Sekhukhune, Vhembe and Waterberg Districts). Moringa was produced on an area of more than 0.25 ha, with seed yields of 50–100 kg/ha. Furthermore, annual enterprise income was estimated at USD13 000 and a gross margin of USD 6000 through selling moringa leaves.[26]

Tshabalala et al.[29] forecasted that, of South Africa's land area, about 17% (200 837 km^2) had optimum growing conditions for moringa cultivation, 18% (216 758 km^2) had suitable conditions, 46% (560 794 km^2) was less suitable and only 19% (240 699 km^2) was not suitable. Approximately 80.3% of Limpopo Province's total area had ideal conditions for growing moringa. Furthermore, the ideal conditions were predicted on the eastern coast of the country from KwaZulu-Natal to Eastern Cape as well as some parts of Northern Cape and Western Cape Provinces[29], thus, suggesting that it could be produced in all nine South African provinces. The production of moringa in South Africa is still at developmental and infant stage[26], thus, making it difficult to quantify the areas under production and number of hectares dedicated for its cultivation. Therefore, it is necessary to encourage farmers in provinces with suitable cultivation areas to participate in its production and to make them realise its potential benefits. However, cultivating moringa beyond its preferred areas is also possible through manipulation of the growing environment; for instance, growing it under controlled conditions such as greenhouses.

Government bodies on moringa-oriented projects

In sub-Saharan Africa, moringa is considered as a 'developmental tree' of choice by governments to combat several socio-economic challenges such as poverty, malnutrition and food insecurity.[10,15] Likewise, the government of South Africa has been supporting the production of moringa in some parts of the country with the mission of alleviating malnutrition in disadvantaged communities (Table 1). The Department of Science and Technology (now Department of Science and Innovation; DSI) has made available a substantial amount of funds to support moringa-oriented activities. Moreover, DSI continues to build and support rural community-based infrastructures for growing and processing moringa.[30,31] Since 2010, DSI has funded research projects aimed at value additions on the quality and product development of moringa to encourage the interface of science and indigenous knowledge systems.[29] Similarly, the Department of Agriculture, Forestry and Fisheries (now known as the Department of Agriculture, Land Reform and Rural Development (DALRRD)) has dedicated substantial financial and technical support to emerging moringa farmers.[24] It has listed moringa as the most cultivated and used medicinal plant.[25] The Department of Rural Development and Land Reform (now part of DALRRD) established a moringa agro-processing business at Temotua farm in Mopani District, Limpopo Province.[32] Consequently, DALRRD continues to encourage several communities to consider moringa as an alternative solution to reduce unemployment. Recently, several studies were carried out on promoting the incorporation of moringa leaf powder in the daily diets of children in school yards.[33,34] These may suggest that the Department of Basic Education supports and recognises the potential use of moringa in schools for alleviating malnutrition among children due to its nutritional benefits.

Growing conditions and cultivation practices

Work done on climatic conditions

Amongst all climatic factors that affect plant growth, temperature is one of the most important factors governing natural geographical plant distribution, tree performance, physiology and productivity.[12] Studies conducted at the University of Pretoria's experimental farm in Gauteng Province revealed that among three evaluated night/day temperature regimes (10/20 °C, 15/25 °C, 20/30 °C), 20/30 °C was the most suitable for germination and plant growth of moringa seedlings (Table 2). Furthermore, it was confirmed that tropical and sub-tropical conditions with hot summers and mild winters are ideal for cultivation.[13,35] According to Manduwa et al.[36], the performance of the reproductive phase of moringa depends on temperature. Temperatures of between 30 °C and 35 °C encouraged fruit set whereas low temperatures (≤15 °C) reduced fruit set during bloom. Although high temperatures are well suited for moringa growth, decreased but satisfactory growth and yields could still be attained in below-optimal climates.[36] This is attributed to the ability of moringa to withstand lower temperatures through physiological adaptation by thickening of leaves when under temperature stress.[37] Successful moringa cultivation in cooler climates would greatly increase its production; however, the effect of growing it under such conditions should be well understood before planting.

These sun- and heat-loving plant species grow best with an annual rainfall of between 250 mm and 1500 mm.[11] Photosynthesis, growth and respiration are processes controlled by metabolic pathways and are affected by temperature and seasonal changes.[39] Gaseous exchange was affected by seasonal change in a study conducted at NTL Baraka Eco-farming, Limpopo Province.[40] In summer, an increase in sub-stomatal carbon dioxide (CO_2) was observed while stomatal conductance transpiration and photosynthetic rate were all reduced. Moringa plants reduced these parameters as an adaptation mechanism and increased water use efficiency under low rainfall and extreme temperature. This is due to the capability of the plant to store carbon in its succulent parts during growing periods.[40] According to Gandji et al.[41], the rate at which moringa trees absorb CO_2 is about 55 times higher when compared to Japanese cedar trees and 20 times higher than that of typical vegetation. Subsequently, cultivation of moringa in parts of South Africa could contribute to national climate change mitigation and adaptation plans.[42] Moringa trees are highly susceptible to frost and cold conditions.[11] This suggests that the southwestern parts of South Africa, which are characterised by cold winters with rainfall[43], may not be suitable for its cultivation during the winter season. Alternatively, moringa seedlings are usually grown inside greenhouses for overwintering and transplanted shortly after winter in temperate zones of South Africa.[44,45]

Table 1: Participation of various government departments in moringa production

Government department	Programmes/activities	References
Department of Agriculture, Forestry and Fisheries (DAFF)	Included moringa on the list of medicinal plants in South Africa	DAFF[24,25]
Department of Science and Technology (DST)	Offered financial support for moringa-oriented activities	DST[30,31]
Department of Rural Development and Land Reform (DRDLR)	Established agro-processing business at Temotua farm	DRDLR[32]
Department of Basic Education	Allowed research on the incorporation of moringa in the daily diets of school children	Ntila et al.[33]; Zungu et al.[34]

Table 2: Temperature requirements proposed by various studies conducted within South Africa for growth and development stages and storage of moringa

	Required temperature	References
Germination and seedling growth	20/30 °C (night/day)	Muhl et al.[13,35]
Fruit set	30–35 °C	Manduwa et al.[36]
Seed storage	Between -19 °C and 4 °C	Du Toit et al.[38]

Soil, nutrient and water requirements

Studies elsewhere report that moringa is well adapted to a wide range of soil types and thrives in poor soils with little or no fertilisation.[46] According to Price[11], moringa trees prefer well-drained sandy loam soils due to their susceptibility to waterlogged conditions. Within South Africa, Mashela[47] assessed its response to loam, sandy, clayey and calcareous soils collected from Madisha-Ditoro and Moletlane villages (Limpopo Province). Moringa seedlings grown on clay and sandy soils produced slightly similar above-ground biomasses that were higher than those produced in calcareous and loam soils. Comparatively, plants grown in clay soils had both higher above-ground biomass and root length, whilst those in sandy soils also had higher above-ground biomass but with reduced below-ground biomass. Additionally, it was demonstrated that calcareous soils with their high pH reduced growth of seedlings.[47]

Reducing or withholding fertilisers tends to slow the overall growth of moringa.[48] Thus, it is advisable to apply fertilisers for improving its growth and yield in areas with low rainfall and extreme temperatures.[49] Moringa is capable of surviving dry seasons and tolerates drought conditions mainly due to its long tuberous taproot that grows very deep into the soil to absorb water and minerals from the sub-soils.[50] The effects of three irrigation intervals (300, 600 and 900 mm per annum) were evaluated on flowering and fruit development of moringa in Gauteng Province. Low irrigation treatment (300 mm/annum) resulted in high bud initiation; however, fruit set was delayed. Therefore, it is necessary to limit water supply to enhance flower initiation prior to floral initiation, but this should be followed by sufficient irrigation to ensure pollination, fruit set and yield.[51]

Planting density, biomass and storage conditions

Mabapa et al.[49] investigated the effect of four plant densities (c. 435 000, c. 300 000, c. 200 000 and c. 100 000 plants/ha) on above-ground biomass of moringa at Syferkuil and Ofcolaco farms (Limpopo Province). Their results showed that the highest planting density of c. 435 000 plants/ha led to the highest biomass of 527–2867 kg/ha on both sites with different agro-climatic conditions.[49] The study carried out at NTL Baraka Eco-farming indicated that various planting densities (c. 1250, c. 1667, c. 2500 and c. 5000 plants/ha) did not influence gaseous exchange parameters such as stomatal conductance, and photosynthetic and transpiration rates. It was recommended that a planting density of c. 5000 plants/ha should be used for moringa cultivation in many parts of the province with favourable conditions, for enhancing the livelihoods and well-being of farmers.[40] Therefore, lower planting density may be used for seed production while high density is ideal for optimum leaf production. Flower and fruit development stages of moringa were affected by pruning intensity in a study carried out at the University of Pretoria. Moderate (2 m) pruning above the soil surface (pollarding) was found to be an ideal practice for moringa production rather than severe (1 m), light (3 m) or no pruning under sub-optimal conditions.[52] In order to ease mechanical harvesting, moringa can be harvested at a height of 0.5 m above the soil surface when the stem is relatively soft (Table 3).[49]

Storage, temperature, moisture content and seed aging are the main factors affecting seed viability and quality of moringa.[38,53] According to Fotouo-M[53], moringa seeds can be stored inside paper or aluminium bags for up to 6 months; beyond this period, they should be stored in sealed containers at temperatures of between -19 °C and 4 °C. Farmers are advised to store their seeds at temperatures below 20 °C in paper bags for up to 12 months, given that the seeds' moisture content remains below 8%.[38] Another method of seed storage includes storing them in their fruits under cool and dry conditions for up to 12 months. Alternatively, one can dry the seeds for 30 days before storing them in airtight containers to ensure high-quality seeds.[54] Moringa leaves may be transported and consumed throughout the year without losing nutrient quality.[55]

Moringa consumption for nutrition

Moringa has been included in some diets to combat malnutrition, especially among infants and breastfeeding women in developing countries.[19,56] Interest in moringa consumption among South Africans is expected to rise. Farmers in the Limpopo Province use moringa as a nutritional source, an income source and for health purposes.[26] The dried moringa leaves of South African ecotypes were found to have the following nutrients: crude protein (30.3%), calcium (3.7%), phosphorus (0.3%), magnesium (0.5%), iron (490.0 mg/kg), sodium (0.6%), zinc (13.0 mg/kg), potassium (1.5%), copper (8.3%), manganese (86.8 mg/kg), sulfur (0.6%) and selenium (363 mg/kg). Additionally, the leaves contained 19 amino acids, 17 fatty acids and high concentrations of vitamin E and beta-carotene.[18] The findings of Pakade et al.[57] revealed that moringa leaves collected from Limpopo (Tooseng village) and Gauteng (Atteridgeville) provinces did not contain metals and were thus safe and suitable for human consumption. Their results also showed that moringa leaves had higher concentrations of calcium and magnesium than locally sourced vegetables such as spinach (*Spinacia oleracea*), cabbage (*Brassica oleracea* var. *capitata*), peas (*Pisum sativum*), cauliflower (*Brassica oleracea* var. *botrytis*) and broccoli (*Brassica oleracea* var. *italica*). 'The concentrations of other major nutrients in moringa leaves were similar to those of the vegetables'[57].

Table 3: Cultivation practices and storage conditions as suggested by studies conducted within South Africa

Production practices	Description	References
Planting density	Minimum = 5000 plants/ha; maximum = 450 000 plants/ha	Mabapa et al.[40,49]
Harvesting/cuttings	50 cm above the soil surface to facilitate mechanical harvesting	Mabapa et al.[49]
Irrigation	Bud initiation = 300 mm/annum; fruit set = 600–900 mm/annum	Muhl et al.[51]
Pruning	Moderate (2 m) pruning above soil surface is ideal for moringa production	Du Toit et al.[52]

Incorporation of moringa should be encouraged in meals that contain these vegetables for enhancing their nutritional value.[58] Furthermore, incorporating moringa into common dishes could promote awareness of moringa as a food source nationally rather than being limited only to medicinal uses.[25] Table 4 shows various consumption patterns of moringa by communities that were reported within the South African context.

Over a decade ago, the acceptability test outcomes of dishes prepared with moringa revealed that moringa was well accepted and recommended as an additional food ingredient to enhance the nutritional value of diets among Bapedi communities in Limpopo Province.[27] Subsequently, few studies have since been carried out to evaluate its acceptability in complementary foods. The acceptability and perceptions of soft white maize porridge modified with different levels of moringa leaf powder (MLP) were evaluated in Hammanskraal (Gauteng Province) and Lebowakgomo (Limpopo Province) townships.[33] Increased levels of MLP added in the porridge led to a decrease in its acceptability among caregivers, and this was due to the bitter taste, making it undesirable to children.[33] Therefore, it is imperative to consider low quantities of MLP in foods prepared for children to increase its acceptance. According to Ntila et al.[59], increasing levels of MLP (0–3%) in maize soft porridge contained higher amounts of nutrients, phenols and antioxidant activities. Moreover, it was emphasised that foods mixed with MLP are likely to be unacceptable to consumers due to the unfamiliar taste and colour. The addition of MLP in *mahewu* (non-alcoholic cereal grain beverage) led to a substantial increase in the nutrient content of the beverage.[60] However, its acceptability by the Ntambanana community (KwaZulu-Natal Province) declined with an increasing percentage of MLP. This was ascribed to sensory attributes of unfamiliar change in colour and aroma. Subsequently, 2% MLP supplemented *mahewu* was more acceptable than 4% and 6%.[60] Another study in KwaZulu-Natal Province assessed nutritional composition and acceptability of MLP-based chip snacks and it was found that the snack containing 1% MLP was almost as acceptable as the untreated snacks. Moreover, 1% MLP snacks had a higher concentration of crude protein and minerals with less fat compared to untreated snacks.[34] Incorporation of MLP in snacks could significantly improve nutrition among children and could lead to a reduction in the intake of unhealthy snacks which pose a health threat. Indeed, these recent studies show that the use of moringa-based products could have a positive impact on the nutritional status of South Africans, particularly malnourished children.

Research should be conducted to fully document and assess local knowledge on consumption patterns of moringa as a food source among individuals living in populations in which its use is established. Despite their nutritional benefits, moringa foodstuffs in southern Africa are at times categorised as 'famine food' eaten during periods of food shortage and are often associated with a lower social class.[61] As such, moringa products are still underutilised and yet to be fully exploited. There is, therefore, a need to promote awareness on the nutritional importance of moringa among marginal communities of South Africa. Moreover, guidance and training on its processing and incorporation in foods is necessary.[33]

Prospects of moringa within South Africa

Promotion of production and commercialisation

Studies should be conducted to record the total number of major moringa production areas in provinces across South Africa, as well as the number of hectares dedicated to its cultivation. Such surveys could easily be conducted with the help of the government departments involved in moringa production and the MDASA. Although moringa food products are available in local markets in a few areas of South Africa, its production is still insufficient and unstable.[26] Furthermore, its commercialisation is still in its infancy, which makes it challenging to quantify the hectares under production, the volume and product value nationally.[22,29] Hence, it is necessary to shift from small-scale, backyard production into commercial economies of scale. This would necessitate the formation of an inclusive value chain starting from cultivation to postharvest. However, access to adequate and reliable markets is the main constraint.[26] Subsequently, the selling price of leaf powder and other moringa products varies widely due to the lack of formal markets and price control. Lots of work and funding are required to support moringa production. Government bodies involved in moringa activities should assist farmers in attaining reliable markets for their produce. One practical way to publicise its production and consumption is by organising workshops and conferences that aim at discussing its nutritional, medicinal, and other various benefits. For example, the Second International Symposium on Moringa held in Pretoria (10–13 November 2019) brought together moringa experts, researchers, farmers, product producers and entrepreneurs from all over the world. The symposium offered an opportunity for sharing, networking and increasing knowledge. The display of various moringa products in those workshops and conferences could also increase its interest and acceptance among the public. In order to extend the audience, information about the potential benefits of moringa should be broadcast on television, radio and influential social platforms such as Twitter, Instagram and Facebook. The distribution of free or low-cost seeds/seedlings to potential farmers would also assist in popularising its production. Currently, there are varieties of foodstuffs such as moringa spice, *mahewu*, tea leaves, ice tea, peanut butter, energy drinks and yoghurts (Table 5).[22,26]

Nutritional analysis of moringa food products could be prerequisites to marketing as these are important in offering quality assurance and integrity to consumers. Despite limited documentation on safety, there are currently no reported adverse effects of moringa products and extracts in human use.[62] However, quality assurance should be addressed because Dar[63] stated that moringa leaf powder could lead to gastrointestinal disorders in consumers due to poor sanitation standards during production and postharvest handling. Indeed, this could improve the market value of moringa food products as it would increase the confidence of traders as well as consumers and this would be a big step in its commercialisation in the country.

Table 4: Consumption patterns of moringa for nutritional purposes in South Africa

Moringa part	Consumption pattern	References/source
Leaves	Grounded moringa leaf (leaf powder) used as a spice in common dishes; fresh leaves added in salads or cooked as a vegetable; dry leaf powder added to meals for children; moringa tea prepared by adding one spoon of leaf powder to boiled water	Pakade et al.[58]; Lekgau[22] Mabapa et al.[26]; personal observation
Seeds	Seeds are eaten like groundnut; grounded seeds used as a spice in food	Personal observation and communication
Roots and other parts	Uses of moringa roots and other parts as a food source are currently unknown and undocumented	

Table 5: Moringa-based food products in markets (formal and informal) within South Africa

Food products	Description	References
Capsules	The capsules are manufactured using moringa dry leaf powder and used as a nutritional supplement	Pakade et al.[57,58]; Mabapa et al.[26]
Energy drinks	Made from moringa leaf extracts or powder	Tshabalala et al.[29]
Flavoured peanut butter	Addition of moringa in peanut butter to increase its nutritional value	Tshabalala et al.[29]
Ice tea	Made from moringa leaf extracts or powder	Tshabalala et al.[29]
Moringa juice	Extracts from either fresh or dry leaves	Mabapa et al.[26]
Moringa leaf powder	Dry leaves are grounded into powder	Mabapa et al.[26]
Soft porridge	Incorporation of moringa leaf powder in soft maize porridge	Ntila et al.[33]
Tea leaves/bags	Made from dry moringa leaf powder	Lekgau[22]; Mabapa et al.[26]
Yoghurt	Made from either dry or fresh leaf extracts	Tshabalala et al.[29]
Mahewu	Incorporation of moringa leaf powder in a non-alcoholic beverage commonly known as *mahewu*	Olusanya et al.[60]
Chip snacks	Incorporation of moringa leaf powder in chip snacks	Zungu et al.[34]

Research gaps on utilising other moringa parts

All the various parts of moringa are known across the world for their multiple uses, including consumption of flowers, seeds and immature pods, fresh leaves as a relish and water purification (seeds).[17,64] However, not all of these uses are recognised in South Africa (i.e. consumption of immature fruits and seeds). Moringa seeds are known to be an important source of essential nutrients and, in other countries, the immature pods are cooked and consumed as food.[65] This practice ought to be adopted in South Africa. Another example is the use of crushed moringa seeds and/or a seedcake for water purification. Many households in rural villages use river water which can be full of suspended matter such as microorganisms, objects and silt particles and such material necessitates elimination prior to water use. This use should be implemented in South Africa as it will be beneficial to rural communities that rely on river water; however, mechanisms involving coagulation by moringa need to be well understood.

In South Africa, moringa leaves are the most researched and utilised part of the tree, whereas little work has been done on the application of other parts of the tree such as seeds, pods and roots. Almost all moringa parts have been researched for use and product development in many countries including India, the Philippines and Nigeria.[15,66] While there are several publications on morphology and chemical composition, there is scant documentation on familiarity and perceptions about moringa in rural communities that make use of the tree. Current studies on moringa consumption and acceptability have been narrowly conducted on caregivers (and/or mothers) and children, and to our knowledge, insufficient research has been done on perceptions and acceptability of moringa in elders, men and youths. Hence, research seems to be gender and age biased. Therefore, there is a need to encourage more intensive and comprehensive utilisation of the plant. In addition, higher education institutions should conduct studies to develop processing techniques and establish a value chain of moringa in South Africa.

Other knowledge gaps on moringa

Studies conducted in Limpopo Province showed that moringa seedlings were infested by spider mites (*Tetranychus urticae*), which contributed to seedling mortality.[67] There is a scarcity of information on pests and diseases that affect moringa production and yields. Knowledge on this aspect would allow the development of proper management strategies which would be beneficial to the moringa industry. Although research has been done on the application of moringa leaf extracts as growth enhancers on certain crops[68], its effectiveness has not yet been tested on a larger commercial scale. Also, the use of seedcakes as biofertiliser and/or organic manure is yet to be researched in South Africa and this could reduce overreliance on chemical fertilisers, which are known to have adverse effects on the environment.

Conclusion

Over the last decade, moringa has risen from a minor crop to one of the most auspicious multi-use crops in South Africa. Its cultivation has expanded in recent years and a diversity of moringa-based food products is available in rural communities and local markets. Stakeholders such as government departments, farmers and higher education institutions have initiated flagship projects with moringa-oriented activities to ensure successful cultivation, production and utilisation of this miracle tree. Moringa is a perennial, drought-tolerant and resilient crop that is capable of surviving a wide range of environmental and climatic conditions. Considering the predicted climatic changes which are likely to have adverse consequences for farming, moringa can serve as a feasible alternative crop for rural inhabitants. Therefore, promoting its production, food products and consumption would be beneficial for the well-being of people residing in marginal communities. This review highlights research that has been done on growing conditions, production and human consumption, and addresses existing knowledge gaps on these topics within South Africa.

Acknowledgements

We are grateful to Southern African Systems Analysis Centre (South Africa), the National Research Foundation (South Africa), and the Department of Higher Education, Science and Technology (South Africa) for financial assistance.

Competing interests

We declare that there are no competing interests.

Authors' contributions

C.V.M.: Conceptualisation; writing – the initial draft; writing – revisions; project management. P.N.M.: Conceptualisation; student supervision; writing – review and editing; validation; project leadership. P.J.P.: Conceptualisation; student supervision; writing – review and editing; validation; project leadership. E.E.P.: Conceptualisation; student supervision; funding acquisition; validation; review and editing; project leadership. All authors discussed, read and approved the final manuscript.

References

1. Saini RK, Sivanesam I, Keum Y. Phytochemicals of *Moringa oleifera*: A review of their nutritional, therapeutic and industrial significance. 3 Biotech. 2016;6:203–217. https://doi.org/10.1007/s13205-016-0526-3
2. United Nations Department of Economic and Social Affairs, Population Division. World population prospects 2019: Data booklet (ST/ESA/SER.A/424). New York: United Nations Department of Economic and Social Affairs, Population Division; 2019.
3. Statistics South Africa (StatsSA). Statistical release P0302: Mid-year population estimates 2019. Pretoria: StatsSA; 2019. Available from: https://www.statssa.gov.za/publications/P0302/P03022019.pdf
4. Chang Y, Liu H, Liu M, Liao X, Sahu SK, Fu Y, et al. The draft genomes of five agriculturally important African orphan crops. GigaScience. 2018;8:1–16. https://doi.org/10.1093/gigascience/giy152
5. Development Initiatives. Global nutrition report 2017: Nourishing the SDGs. Bristol, UK: Development Initiatives; 2017. Available from: https://globalnutritionreport.org/reports/2017-global-nutrition-report/
6. Wang SM, Zhang ZW. The state of the world's plant genetic resources, food and agriculture. J Plant Genet Resour. 2011;47:574–574.
7. Foyer CH, Lam HM, Nguyen HT, Siddique KHM, Varshney RK, Colmer TD, et al. Neglecting legumes has compromised human health and sustainable food production. Nat Plants. 2016;2:16112. https://doi.org/10.1038/nplants.2016.112
8. Murro JK, Muhikambele VRM, Sarwatt SV. *Moringa oleifera* leaf meal can replace cottonseed cake in the concentrate mix feed with Rhodes grass (*Choris gayana*) hay for growing sheep livestock. Res Rural Dev. 2003;15:11–17.
9. Brilhante RSN, Sales JA, Pereira VS, Castelo-Branco DSCM, Cordeiro RA, Sampaio CSM, et al. Research advances on the multiple uses of *Moringa oleifera*: A sustainable alternative for socially neglected population. Asian Pac J Trop. Med. 2017;10:621–630. http://doi.org/10.1016/j.apjtm.2017.07.002
10. Hassan FG, Ibrahim MA. *Moringa oleifera*: Nature's most nutritious and multi-purpose tree. Sudan Int J Sci Res Publ. 2013;3:2250−3153.
11. Price ML. The moringa tree: Technical notes. New York: ECHO Publishers; 2007.
12. Anwar F, Latif S, Ashraf M, Gilani AH. *Moringa oleifera*: A food plant with multiple medicinal uses. Phytother Res. 2007;21:17−25. https://doi.org/10.1002/ptr.2023
13. Muhl QE, Du Toit ES, Robbertse PJ. Temperature effect on seed germination and seedling growth of *Moringa oleifera* Lam. Seed Sci Technol. 2011;39:208–213.
14. Nouman W, Siddique MT, Basra SMA, Farooq H, Zubair M, Gull T. Biomass production and nutritional quality of *Moringa oleifera* as a field crop. Turk J Agric For. 2013;37:410–419.
15. Leone A, Spada A, Battezzati A, Schiraldi A, Aristil J, Bertoli S. Cultivation, genetic, ethnopharmacology, phytochemistry and pharmacology of *Moringa oleifera* leaves: An overview. Int J Mol Sci. 2015;16:12791–12835.
16. Coles ZS, Du Toit ES. Open air-layering of *Moringa oleifera* utilizing seedling plug containers. S Afr J Bot. 2020;129:225–228. https://doi.org/10.1016/j.sajb.2019.07.016
17. Fahey JW. *Moringa oleifera*: A review of the medical evidence for its nutritional, therapeutic, and prophylactic properties. Part 1. Trees Life J. 2005;1:1–15.
18. Moyo B, Masika PJ, Hugo A, Muchenje V. Nutritional characterization of moringa (*Moringa oleifera* Lam.) leaves. Afr J Biotechnol. 2011;10:12925–12933.
19. Srikanth VS, Mangala S, Subrahmanyam G. Improvement of protein energy malnutrition by nutritional intervention with *Moringa oleifera* among Anganwadi children in rural area in Bangalore, India. Int J Sci. 2014;2:32–35.
20. Tshabalala T, Ndhlala AR, Ncube B, Abdelgadir HA, Van Staden J. Potential substitution of the root with the leaf in the use of *Moringa oleifera* for antimicrobial, antidiabetic and antioxidant properties. S Afr J Bot. 2020;129:106–112. https://doi.org/10.1016/j.sajb.2019.01.029
21. Marcu MG. Miracle tree. La Canada, CA: KOS Health Publications; 2005.
22. Lekgau J. *Moringa oleifera*: A tree giving life to rural communities. Pretoria: National Agricultural Marketing Council; 2011.
23. Ekesa BN. Selected superfoods and their derived super diets. Kampala: InTech Biodiversity International; 2017. https://doi.org/10.5772/67239
24. South African Department of Agriculture, Forestry and Fisheries (DAFF). Medicinal plants of South Africa. Pretoria: Plant Production, DAFF; 2013.
25. South African Department of Agriculture, Forestry and Fisheries (DAFF). A profile of the South African plants market value chain. Pretoria: Marketing, DAFF; 2016.
26. Mabapa MP, Ayisi KK, Mariga IK, Mohlabi RM, Chuene RS. Production and utilization of moringa by farmers in Limpopo Province, South Africa. Int J Agric. Res. 2017;12:160−171. https://doi.org/10.3923/ijar.2017.160.171
27. Agyepong AO. The possible contribution of *Moringa oleifera* Lam. leaves to dietary quality in two Bapedi communities in Mokopane, Limpopo Province [master's dissertation]. Pretoria: University of South Africa; 2009.
28. Mashamaite CV, Mothapo PN, Albien AJ, Pieterse PJ, Phiri EE. A SUSPECT under the National Environmental Management Biodiversity Act (NEM:BA) – *Moringa oleifera*'s ecological and social costs and benefits. S Afr J Bot. 2020;129:249–254. https://doi.org/10.1016/j.sajb.2019.07.019
29. Tshabalala T, Ncube B, Moyo HP, Abdel-Rahman EM, Mutanga O, Ndhlala AR. Predicting the spatial suitability distribution of *Moringa oleifera* cultivation using analytical hierarchical process modelling. S Afr J Bot. 2020;129:161–168. https://doi.org/10.1016/j.sajb.2019.04.010
30. South African Department of Science and Technology (DST). FY2017/18 Annual performance plan: Third draft submission: 10 March 2017. Pretoria: Technology Innovation Agency; 2017.
31. South African Department of Science and Technology (DST). DST Second quarter financial and performance report (July to September 2017). Pretoria: Portfolio Committee on Science and Technology; 2018.
32. South African Department of Rural Development and Land Reform (DRDLR). Annual report 2013/2014 Vote 33: Part E: Financial information. Pretoria: DRDLR; 2014.
33. Ntila S, Ndhlala AR, Kolanisi U, Abdelgadir H, Siwela M. Acceptability of moringa-added complementary soft porridge to caregivers in Hammaskraal, Gauteng Province and Lebowakgomo, Limpopo Province, South Africa. Afr J Clin Nutr. 2018;1:7. https://doi.org/10.1080/16070658.2018.1449377S
34. Zungu N, Van Onselen A, Kolanisi U, Siwela M. Assessing the nutritional composition and consumer acceptability of *Moringa oleifera* leaf powder (MOLP)-based snacks for improving food and nutrition security of children. S Afr J Bot. 2020;129:283–290. https://doi.org/10.1016/j.sajb.2019.07.048
35. Muhl QE, Du Toit ES, Robbertse PJ. Adaptability of *Moringa oleifera* Lam. (horseradish) tree seedlings to three temperature regimes. Am J Plant Sci. 2011;2:776–780.
36. Manduwa DM, Du Toit ES, Robbertse PJ. Study on the effect of temperature and flower age on pollen performance, stigma receptivity and fruit-set of *Moringa oleifera* Lam. Acta Hortic. 2018;1204:57–64. https://doi.org/10.17660/ActaHortic.2018.1204.8
37. Muhl QE, Du Toit ES, Robbertse PJ. *Moringa oleifera* (horseradish tree) leaf adaptation to temperature regimes. Int J Agric Biol. 2011;13:1021–1024.
38. Du Toit ES, Fotouo H, Robbertse PJ. Seed storage conditions influence germination of *Moringa oleifera* Lam. seed. Acta Hortic. 2017;1158:441–446. https://doi.org/10.17660/ActaHortic.2017.1158.51
39. Ouma GB. Growth responses of rough lemon (*Citrus limon* L.) rootstock seedlings to different container sizes and nitrogen levels. Agric Tropica ET Subtrop. 2006;39:183−189.
40. Mabapa MP, Ayisi KK, Mariga IK. Seasonal effect on *Moringa oleifera* gaseous exchange and water use efficiency under diverse planting densities. J Appl Bot Food Qual. 2018;91:219–225. https://doi.org/10.5073/JABFQ.2018.091.029
41. Gandji K, Chadare FJ, Idohou R, Salako VK, Assogbadjo AE, Glèlè-Kakaï RL. *Moringa oleifera*: A review on its importance and research avenues. Afr Crop Sci J. 2018;26:137–156.
42. Mabapa PM, Ayisi KK, Mariga IK. Comparison of gas exchange in *Moringa oleifera* and other drought tolerant tree species for climate change mitigation under semi-arid condition of Northern South Africa. Int J Agric Biol. 2018;20:2669–2676. https://doi.org/10.17957/IJAB/15.0808

43. Daron JD. Regional climate messages: Southern Africa. Scientific report from the CARIAA Adaptation at Scale in Semi-Arid Regions (ASSAR) Project. Unpublished report; December 2014.

44. Muhl QE. Seed germination, tree growth and flowering responses of *Moringa oleifera* Lam. (horseradish tree) to temperature [MSc thesis]. Pretoria: University of Pretoria; 2009.

45. Mashela PW. Responses of *Moringa oleifera* root growth to container size during overwintering in temperate regions. Afr J Agric Res. 2019;14:65–68. https://doi.org/10.5897/AJAR2016.11667

46. Mridha MAU. Prospects of moringa cultivation in Saudi Arabia. J Appl Environ Biol Sci. 2015;5:39–46.

47. Mashela PW. Growth of moringa (*Moringa oleifera*) seedlings in calcareous, clayey and sandy soils relative to loamy soil. Afr J Agric Res. 2017;12:3508–3512. https://doi.org/10.5897/AJAR2016.11617

48. Mahn LH, Dung NNX, Ngoi PT. Introduction and evaluation of *Moringa oleifera* for biomass production and as feed for goats in the Mekong Delta. Livest Res Rural Dev. 2005;17, Art. #104. Available from: http://www.lrrd.org/lrrd17/9/manh17104.htm

49. Mabapa MP, Ayisi KK, Mariga IK. Effect of plant density and harvest interval on leaf yield and quality of moringa (*Moringa oleifera*) under diverse agro-ecological conditions of Northern South Africa. Int J Agron. 2017: Art. #2941432, 9 pages. https://doi.org/10.1155/2017/2941432

50. Mohammed FA. Antioxidants composition of moringa (*Moringa oleifera* Lam.) in different plant organs [master's dissertation]. Pietermaritzburg: University of KwaZulu-Natal; 2015.

51. Muhl QE, Du Toit ES, Steyn JM, Apostolides Z. Bud development, flowering and fruit set of *Moringa oleifera* Lam. (horseradish tree) as affected by various irrigation levels. J Agric Rural Dev Trop. 2013;114:79–87.

52. Du Toit ES, Sithole J, Vorster J. Pruning intensity influences growth, flower and fruit development of *Moringa oleifera* Lam. under sub-optimal growing conditions in Gauteng, South Africa. S Afr J Bot. 2020;129:448–456. https://doi.org/10.1016/j.sajb.2019.11.033

53. Fotouo-M H, Du Toit ES, Robbertse PJ. Effect of storage conditions on *Moringa oleifera* Lam. seed oil: Biodiesel feedstock quality. Ind Crop Prod. 2016;84:80–86. https://doi.org/10.1016/j.indcrop.2016.01.032

54. Fotouo-M H, Vorster J, Du Toit ES, Robbertse PJ. The effect of natural long-term packaging methods on antioxidant components and malondialdehyde content and seed viability *Moringa oleifera* oilseed. S Afr J Bot. 2020;129:17–24. https://doi.org/10.1016/j.sajb.2018.10.017

55. Du Toit ES, Sithole J, Vorster J. Leaf harvesting severity affects total phenolic and tannin content of fresh and dry leaves of *Moringa oleifera* Lam. trees growing in Gauteng, South Africa. S Afr J Bot. 2020;129:336–340. https://doi.org/10.1016/j.sajb.2019.08.035

56. Arun PR, Sarita S. Comparative analysis of preservation techniques on *Moringa oleifera*. Asian J Food Agro-Indust. 2011;4:65–80.

57. Pakade V, Cukrowska E, Chimuka L. Metal and flavonol contents of *Moringa oleifera* grown in South Africa. S Afr J Sci. 2013;109(3–4), Art. #835. https://doi.org/10.1590/sajs.2013/835

58. Pakade V, Cukrowska E, Chimuka L. Comparison of antioxidant activity of *Moringa oleifera* and selected vegetables in South Africa. S Afr J Sci. 2013;109(3–4), Art. #1154. https://doi.org/10.1590/sajs.2013/1154

59. Ntila SL, Ndhlala AR, Mashela PW, Kolanisi U, Siwela M. Supplementation of a complementary white maize soft porridge with *Moringa oleifera* powder as a promising strategy to increase nutritional and phytochemical values: A research note. S Afr J Bot. 2020;129:238–242. https://doi.org/10.1016/j.sajb.2019.07.021

60. Olusanya RN, Kolanisi U, Van Onselen A, Ngobese NZ, Siwela M. Nutritional composition and consumer acceptability of *Moringa oleifera* leaf powder (MOLP)-supplemented mahewu. S Afr J Bot. 2020;129:175–180. https://doi.org/10.1016/j.sajb.2019.04.022

61. Seifu E, Teketay D. Introduction and expansion of *Moringa oleifera* Lam. in Botswana: Current status and potential for commercialization. S Afr J Bot. 2020;129:471–479. https://doi.org/10.1016/j.sajb.2020.01.020

62. Stohs SJ, Hartman MJ. Review of the safety and efficacy of *Moringa oleifera.* Phytother Res. 2015;29:796–804. https://doi.org/10.1002/ptr.5325

63. Dar WD. Challenges in the industrialisation of moringa in the Philippines. Acta Hortic. 2017;1158:15–18. https://doi.org/10.17660/ActaHortic.2017.1158.3

64. Coppin J. A study of the nutritional and medicinal values of *Moringa oleifera* leaves from sub-Saharan Africa: Ghana, Rwanda, Senegal and Zambia [MSc thesis]. New Brunswick, NJ: State University of New Jersey; 2008.

65. Ananias NK, Kandawa-Schulz M, Hedimbi M, Kwaambwa HM, Tutu H, Makita C, et al. Comparison of metal content in seeds of *Moringa ovalifolia* and *Moringa oleifera*. Afr J Food Sci. 2016;10:172–177. https://doi.org/10.5897/AJFS2016.1421

66. Palada MC. The moringa industry in the Philippines: Status, challenges and opportunities. Acta Hortic. 2017;1158:447–454. https://doi.org/10.17660/ActaHortic.2017.1158.52

67. Dube ZP, Mashela PW, Abdelgadir AH. Identification of spider mites from *Moringa oleifera* using molecular techniques. Acta Agric Scand B Soil Plant Sci. 2015;65:479–482. https://doi.org/10.1080/09064710.2015.1019555

68. Mvumi C, Tagwira F, Chiteka AZ. Effect of moringa extract on growth and yield of maize and common beans. Greener J Agric Sci. 2013;3:59–62.

AUTHORS:
Brigitta Tóth[1]
Makoena J. Moloi[2]
Lóránt Szöke[1]
Maryke Labuschagne[2]

AFFILIATIONS:
[1]Institute of Food Science, University of Debrecen, Debrecen, Hungary
[2]Department of Plant Sciences, University of the Free State, Bloemfontein, South Africa

CORRESPONDENCE TO:
Brigitta Tóth

EMAIL:
btoth@agr.unideb.hu

DATES:
Received: 01 June 2020
Revised: 15 Sep. 2020
Accepted: 17 Sep. 2020
Published: 29 Mar. 2021

HOW TO CITE:
Tóth B, Moloi MJ, Szöke L, Labuschagne M. Low nitrogen and phosphorus effects on wheat Fe, Zn, phytic acid and phenotypic traits. S Afr J Sci. 2021;117(3/4), Art. #8414. https://doi.org/10.17159/sajs.2021/8414

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITORS:
Teresa Coutinho
Salmina Mokgehle

KEYWORDS:
nutrient deficiency, phenotypic traits, plant height, dry weight, microelements

FUNDING:
National Research Foundation (South Africa)

Low nitrogen and phosphorus effects on wheat Fe, Zn, phytic acid and phenotypic traits

In sub-Saharan Africa, crops are often grown under low nitrogen (N) and low phosphorus (P) conditions, which may impact on the nutritional components of the grains. The aim of this study was to investigate the effect of low N and low P and a combination of the two on iron (Fe), zinc (Zn) and phytic acid content in two commercial South African spring wheat cultivars (PAN3497 and SST806). Phenotypic traits were also investigated. Although cultivar effects were not significant, treatment effects were highly significant for the phenotypic and nutritional traits. Low P stress increased Fe and Zn levels, whereas low N stress had the opposite effect. In addition, low P stress inhibited phytic acid accumulation the most, suggesting that under this treatment, Fe and Zn were more available because of less interaction with phytic acid. Compared to the low N treatment, the low P treatment led to lower reductions in the number of tillers, plant height, stem thickness, number of seeds, weight of seeds and dry weight for both cultivars. While low P had positive effects on the nutritional value of wheat, the combination of low N and P treatment had a negative impact on most of the measured characteristics. Low N conditions had more negative effects on all measured characteristics than low P conditions and was very detrimental to wheat nutritional value and yield.

Significance:

- Results from this study emphasise the impact of fertilisation and the impact of insufficient nitrogen and phosphorus fertiliser on wheat productivity.
- Low nitrogen and phosphorus fertilisation impact grain microelement content and bioavailability which impact nutritional value.

Introduction

Nitrogen deficiency is one of the major crop production constraints in the world.[1] Plants require nitrogen (N) in large quantities to attain normal growth and development because N concentration is strongly related to photosynthetic rate and other photosynthetic parameters such as the electron transport rate and carboxylation capacity.[2] The estimated world supply of nitrogen as ammonia will be 170 761 thousand tons by 2020, of which only 5.5% will be used in Africa.[3] Statistics indicate that the sub-Saharan region utilises very low levels of N for grain crop production, at an average of 11 kg/ha/year despite the 90 to 120 kg/ha/year recommended rates.[4]

Phosphorus (P) is the most widely used fertiliser after N.[5] Its deficiency affects about 40% of the cultivated land of the world and causes loss of productivity and quality.[6] As most of the P is stored in the grain, harvesting grain crops leads to continuous removal of the P from the soil. Consequently, P fertiliser application is required to address soil P deficiencies. Both N and P are essential macronutrients required for vegetative and reproductive plant growth.[7-9] Small-scale farmers in sub-Saharan Africa often do not have access to fertiliser, due mainly to the cost, leading to poor N and P status of soils.[10] Artificial fertilisers mainly contain N, P and potassium (K) while microelements are present in natural organic fertilisers such as compost.[11] Some industrial by-products and waste materials can be used as micro-fertilisers in case of iron (Fe) deficiency.[12]

Sub-optimal concentrations of Fe and zinc (Zn) in crops as well as in wheat grain cause micronutrient deficiencies in humans. Deficiency of Fe is a problem in most developing countries as a result of, amongst other things, inadequate intake, reduced absorption, and deficiency in the soil.[13] Fe is a key component for infection resistance in humans.[13] The World Health Organization (WHO) reported that 30% of the world population, specifically women and children, suffers from anaemia as a result of Fe deficiency.[14] Furthermore, the WHO issued a statement that Zn deficiency ranks 11th among the 20 most important risk factors contributing to the burden of disease in the world.[15] Zn deficiency is 5th among the 10 most important factors in developing countries, while Fe deficiency ranks 6th.[16] Approximately 70–80% of the total P contained in cereal grains is in the form of phytate.[17] The bioavailability of micronutrients for human uptake is limited by phytic acid concentration because it can make complexes with cations such as Zn^{2+}, forming insoluble phytates (such as zinc phytate), which influence the bioavailability of Zn in grains.[18] It has been reported that phytate content affects Fe bioavailability more than the total Fe content[19], although this finding was contradicted by another study[20].

Although studies on the effect of low N and P content in the soil have been conducted extensively, the main focus of these studies was on crop yields. Information on how these macronutrients affect micronutrients in wheat is not evident. Therefore, the aim of this study was to investigate the effect of low N and low P as well as a combination of the two on Fe, Zn and phytic acid content in two commercial South African spring wheat cultivars with excellent baking quality. The results of this study will shed light on the quality of these cultivars under low N, P and a combination of the two. The effects of these treatments on the phenotypic characteristics of the two wheat cultivars further elucidated how the cultivars responded under these conditions.

Materials and methods

Greenhouse trials

Two commercial, South African spring wheat cultivars with excellent baking quality, PAN3497 and SST806 (the commercial standard cultivar for spring wheat baking quality in South Africa), were sown in 2-L pots filled with 2 kg soil in a greenhouse. A randomised complete block design was used with two factors: treatment and cultivar. Soil from a depth of 1.5 m was obtained from Bainsvlei, Bloemfontein, South Africa (29.05° S, 26.11667° E). The soil had very low nutrient content as indicated in Table 1.

Four treatments were applied to the two cultivars, with three replications: 15 pots per replication in 2016 and 20 pots per replication in 2017. Each pot contained three plants. The trials were carried out from June to the end of October 2016 (winter time) and during the same time in 2017. Greenhouse temperatures were set to 18 °C at night (21:00–06:00) and 22 °C during the day. Low N and low P stress and a combination of the two were induced according to the protocol given in Table 2. These treatments were tested against an optimal control. The treatments were initiated at three-leaf stage. Before this, plants were irrigated with deionised water. Once a week, pots were flushed with deionised water to prevent salt build-up. Treatments were applied twice a week (250 mL nutrient solution per pot). The electrical conductivity was maintained at 1.5 mS/cm^2 to tillering stage and at 1.80 mS/cm^2 after tillering.

All treatments received the same micronutrient fertilisation as follows: 3.45 mg/L $C_{10}H_{13}FeN_2O_8$, 0.30 mg/L $MnSO_4$, 0.13 mg/L $ZnSO_4$, 0.62 mg/L H_3BO_3, 0.05 mg/L $CuSO_4$, 0.02 mg/L Na_2MoO_4. After ripening, the seeds were harvested and milled into whole flour using a laboratory mill (IKA A10 Yellowline analysis grinder, Merck Chemicals Pty Ltd. Mountainview, CA, USA). These whole flour wheat samples were used for the determination of Fe, Zn and phytic acid.

Total iron and zinc analysis

Total Fe and Zn were extracted according to the dry-ashing method.[21] Wheat flour (1 g) was placed in glazed, high-form porcelain crucibles and ashed in a furnace at 550 °C for 3 h. A few drops of HNO_3 (55 %, v/v) were added to the samples for digestion. The samples were then placed on a hot sand-bath to completely dry, after which they were returned to the oven for 1 h at 550 °C for further ashing. After cooling, 10 mL of 1:2 HNO_3 was added to the samples for further digestion. The samples were returned to the hot sand-bath until they became warm (100 °C). The samples were then transferred to 100-mL volumetric flasks and filled to the mark with distilled water. Mineral concentrations were measured in triplicate using an atomic absorption spectrophotometer (Agilent Technologies 300 Series AA).

Phytic acid determination

Phytic acid concentration was determined using a rapid colourimetric procedure based on the reaction between ferric acid and sulfosalicylic acid according to the method described by Dragičević et al.[22] with modifications. Ground flour samples (0.25 g) were placed in glass tubes containing 10 mL of trichloroacetic acid (TCA) (5 %, v/v) and placed on a mechanical shaker for 1 h, vortexed at 10-min intervals. The extract (5 mL) was transferred into 15-mL tubes and centrifuged at 12 000 *g* for 20 min. The supernatant (0.5 mL) was transferred into a clean glass tube and 1.5 mL WADE reagent (0.3%, w/v, $FeCl_3$ + $6H_2O$; 3%, v/v, 5"-sulfosalicylic acid) was added. The samples were then centrifuged at 12 000 *g* for 10 min. Absorbance was read at 500 nm with a Helios gamma spectrophotometer (Erlangen, Germany). When phytate is present, the Fe ion present in the WADE reagent binds to the phosphate ester instead of reacting with sulfosalicylic acid, resulting in a decrease in pink colour intensity. The phytic acid concentration was calculated from the phytic acid standard.[23]

The phytic acid standard solution was made from phytic acid sodium salt hydrate from rice (Sigma-Aldrich, P-8810, molecular weight: 660.04 g/mol). A series of standard phytic acid solutions was made from the stock standard solution by appropriate dilutions, with the addition of extraction solutions to simulate conditions similar to the ones in the samples. The concentrations of phytic acid in this series were as follows: 10, 50, 100, 150, 200, 250, 300, 350 and 400 μmol/100 mL.

Phytic acid:iron and phytic acid:zinc molar ratios

The contents of phytic acid, Fe and Zn were converted into moles by division through their molar mass or atomic weight (phytic acid: 660.04

Table 1: Characteristics of the applied soil and the measured nutrient content

pH	N (%)	P (mg/kg)	K (mg/kg)	Ca (mg/kg)	Mg (mg/kg)	Na (mg/kg)	S (mg/kg)	Zn (mg/kg)	C (%)	Sand (%)	Clay (%)	Silt (%)
4.1	0.004	26.5	103.4	136	38.9	1.8	1.62	0.71	0.04	96	4	0

Table 2: Fertiliser applied over two years to two wheat cultivars in four treatments in a greenhouse experiment

Chemical (mg/L)	Optimal		Low N		Low P		Low N and P	
	BT	AT	BT	AT	BT	AT	BT	AT
KNO_3	261	313	0	0	228	273	0	0
K_2SO_4	210	252	210	252	196	235	196	235
KCl	0	0	193	231	56	67	223	268
$NH_4H_2PO_4$	87	104	87	104	0	0	0	0
$Ca(NO_3)_2$	758	909	0	0	797	956	0	0
$CaCl_2$	0	0	353	424	0	0	446	446
$MgSO_4$	348	418	348	418	369	443	443	443

BT, before tillering; AT, after tillering

g/mol, Fe: 55.85 g/mol, Zn: 65.4 g/mol). The molar ratios of phytic acid:Fe and phytic acid:Zn were calculated.

Phenotypic traits

The measured traits were recorded for each plant in a pot at maturity, and values were averaged. These traits included number of tillers per plant, plant height, length of the main ear, stem thickness (main tiller, measured by a ruler), seeds per plant, seeds per main ear, seed weight of main ear, aboveground dry weight and dry leaf area per plant (main stem).

Statistical analysis

Analyses of variance (ANOVA) were performed on the data for both genotypes, four treatments and two years as a three-factor analysis.[24] ANOVA was also done for the two cultivars separately as well as for the two years combined, in order to determine the effects of treatments on the measured parameters within each cultivar. Differences were tested at a $p<0.05$ level of significance. The Tukey test and least significant difference test were used for means separation.

Results

The ANOVA showed that cultivar effect was not significant for nutritional traits (Fe, Zn and phytic acid) (Table 3). The effect of treatment and year was highly significant for Fe, Zn and phytic acid contents. There was a highly significant interaction between cultivar and treatment, and cultivar and year for Zn, and between treatment and year for phytic acid.

Although SST806 recorded higher values for all nutritional traits than did PAN3497, the two cultivars were not significantly different (Table 4).

Under low N conditions, Fe concentration was significantly reduced, but under low P, as well as low N and P treatments combined, Fe content was not significantly higher than under control conditions (Table 5). Zn content increased significantly with low P inducing the highest (33.79 mg/kg) and low N inducing the lowest (20.54 mg/kg) increase. Phytic acid concentrations under both treatments were not significantly different from that of the control. The low P and combination of low N and P treatments produced significantly lower phytic acid content compared to the control, and low N with low P produced the highest reduction (5.03 and 5.31 mg/kg). The molar ratio of phytic acid:Fe was increased under low N conditions, but was significantly decreased under low P (7.13) and a combination of low N and P (4.43). The phytic acid:Zn molar ratio was decreased under low N, low P and a combination of the two but the effect was by far the highest under low P conditions with a 23.15 reduction, followed by that under a combination of low N and P (15.06).

For the two cultivars separately (Figures 1 and 2), PAN3497 Fe content was reduced in all treatments compared to the control. Low N treatment caused the most substantial reduction (21.13 mg/kg) compared to the low P and low N with P combination treatments. Similarly, for SST806, Fe content was reduced (18.08 mg/kg) under low N conditions. However, under low P and a combination of low N and low P stress, there were slight increases in Fe content compared to the control. The Zn content for PAN3497 increased in all treatments compared to the control, with the low P treatment inducing the most significant increase (35.37 mg kg).

Table 3: Analysis of variance for Fe, Zn and phytic acid concentrations in a trial of two wheat cultivars with four treatments over two years

	Cultivar (C)	Treatment (T)	Year (Y)	CxT	CxY	TxY	CxTxY
Fe	11.15	197.44**	382.45**	27.39	55.19	6.38	14.86
Zn	0.26	597.08**	0.01	45.27**	173.09**	96.61	0.05
Phytic acid	0.21	10.26**	5.35**	0.24	0.24	2.92**	0.06

**p≤*0.01*

Table 4: Average values of four treatments for the measured characteristics over two years separated by two cultivars

	PAN3497	SST806	Least significant difference (0.05)
Fe	24.86	25.82	1.85
Zn	23.58	23.70	1.09
Phytic acid	5.89	6.02	0.18
Phytic acid:Fe molar ratio	20.26	21.12	
Phytic acid:Zn molar ratio	23.36	26.06	

Table 5: Average values of two cultivars for the measured nutritional traits over two years separated by treatments

	Control	Low N	Low P	Low N and P	Least significant difference (0.05)
Fe (mg/kg)	26.25	19.60	28.61	26.90	2.61
Zn (mg/kg)	17.68	20.54	33.79	22.56	1.54
Phytic acid (mg/kg)	6.92	6.56	5.03	5.31	0.25
Phytic acid:Fe molar ratio	22.28	27.77	15.15	17.84	
Phytic acid:Zn molar ratio	38.08	26.62	14.93	23.02	

Although a similar pattern was observed for SST806, the low N treatment induced slightly higher Zn than that of PAN3497. Contrary to Fe and Zn, phytic acid content was similar for the two cultivars under the different treatments. In both cultivars, low P and a combination of low N and P caused a significant decrease in phytic acid.

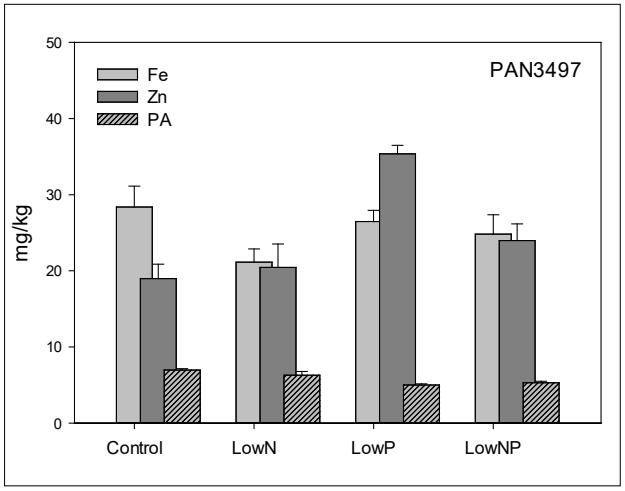


Figure 1: The Fe, Zn and phytic acid contents of PAN3497 under four treatments over two years. Values are means ± s.d. (n=6).

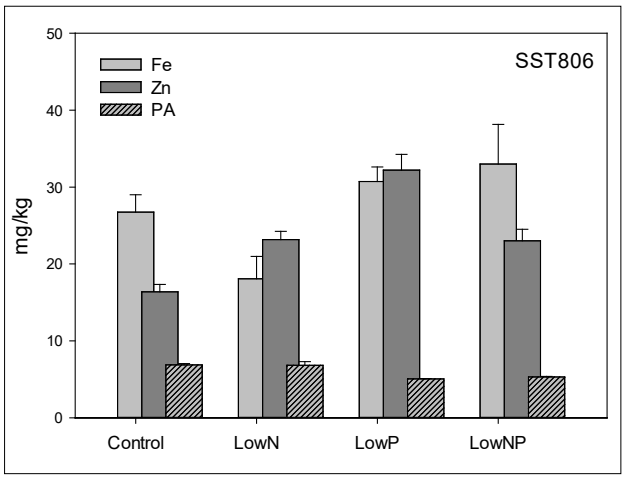


Figure 2: The Fe, Zn and phytic acid contents of SST806 under four treatments over two years. Values are means ± s.d. (n=6).

The results of the ANOVA showed that cultivar effect (as seen from significance of F ratios of mean squares) was significant ($p \leq 0.05$) for main ear, seeds number per main ear, seeds per plant, above ground dry weight and dry leaf area (Table 6). Furthermore, there was a large treatment contribution to variation in the case of plant height, main ear length, main stem thickness, seeds per main ear, seeds per plants and dry leaf area. Year contributed significantly to variation in seeds per main ear and number of tillers. Dry leaf area showed large cultivar x treatment and cultivar x treatment x year interactions.

To better understand the effect of treatment on morphological parameters without genotype playing a role, values were averaged for two cultivars (Tables 7 and 8). Cultivars did not differ significantly for measured phenotypic characteristics, but differences between treatments were significant. For both cultivars, the controls produced the highest values for all phenotypic characteristics and characteristics varied substantially between different treatments. Plants under the low P treatment had significantly higher numbers of tillers than those under low N and low N and P combined treatments for both cultivars. For both cultivars, plant height was reduced by all the treatments, where low N treatment led to the highest reduction in length. The main ear length, stem thickness and seeds per main ear were reduced the most under low N treatment, with low P having the least effect. Total seeds per plant were the least affected by the low P treatment for both cultivars although the weight was not significantly different from that under the low N and P combination for SST806. For the aboveground dry weight, both cultivars showed the largest reduction under the low N treatment. The low N treatment led to the lowest reduction in dry leaf area, but a combination of low N and P stress caused the highest reduction in this parameter for both cultivars, which was unexpected.

Discussion

Wide variation in Fe and Zn concentrations has been recorded in wheat grain.[25-27] The average Fe concentration was reported to be between 30 mg/kg and 73 mg/kg [28], in contrast to this study where it was low – ranging between 19.6 mg/kg and 28.61 mg/kg. The range of Zn concentration was reported to be between 20.4 mg/kg and 30.5 mg/kg in wheat grains.[29] In this study, Zn concentration varied between 17.68 mg/kg and 33.79 mg/kg across the different treatments. This variation in the Fe and Zn contents could be affected by the treatments and the different cultivars used because their concentrations are determined by genetic and environmental factors.[30]

In this study, the effect of cultivar on the measured characteristics was negligible, although SST806 had slightly higher values for all the traits. The effects of the treatments were highly significant for all measured nutritional characteristics. Zn content showed a significant cultivar with treatment interaction, indicating that the two cultivars did not react the same to treatments in terms of Zn content.

Table 6: Analysis of variance for morphological traits in a trial of two wheat cultivars with four treatments over two years

	Cultivar (C)	Treatment (T)	Year (Y)	CxT	CxY	TxY	CxTxY
Tillers	88.74	250.61	66.51*	30.52	89.90*	45.60	52.26
Plant height (cm)	13.89	235.11*	45.67	33.52	45.69	22.11	19.82
Main ear length (cm)	68.20*	513.68*	28.65	35.11	61.12*	48.36	16.52
Main stem thickness (mm)	165.44	495.133*	15.48	29.35	87.11	66.24	24.51
Seeds/main ear	85.35*	134.34*	58.04*	22.11	29.12	37.58	31.89
Seeds/plant	58.22*	363.05*	34.58	17.46	53.64*	87.11*	55.19
Main ear seed weight (g)	35.11	225.18	12.15	65.15	72.32	46.52	28.16
Aboveground dry weight (g)	156.73*	201.07	47.18	26.345	113.58*	61.98	34.58
Dry leaf area/main plant	125.11*	112.835*	288.92	87.11*	126.12	25.69	64.11*

*p<*0.01*

Table 7: Average values for measured characteristics in PAN3479 with four treatments over two years

	Control	Low N	Low P	Low NP
Tillers	2.67±0.80c	1.03±0.16a	1.37±0.52b	1.06±0.24a
Plant height (cm)	63.89±5.06c	57.31±6.75a	59.74±6.67b	58.18±8.15b
Main ear length (cm)	8.60±0.87c	5.72±1.29a	6.01±1.11b	5.91±1.52ab
Main stem thickness (mm)	5.78±0.62c	3.81±0.84a	4.26±0.64b	3.81±0.93a
Seeds/main ear	49.55±9.96c	22.12±9.18a	29.05±8.83b	26.43±11.80b
Seeds/plant	93.80±39.49c	22.45±9.75a	35.76±19.91b	32.11±12.09b
Main ear seed weight (g)	2.02±0.43c	0.84±0.40a	0.98±0.29b	0.98±0.43b
Aboveground dry weight (g)	3.94±0.74c	2.09±1.00a	2.26±0.68b	2.20±1.01ab
Dry leaf area/main plant	56.47±24.02c	13.62±6.18a	17.06±8.39b	12.51±6.21a

*Values in columns are means ± standard deviation (*n*=225; except for dry leaf area* n*=30)*
Letters indicate a significant difference among the treatments; p≤*0.05*

Table 8: Average values for measured characteristics in SST806 with four treatments over two years

	Control	Low N	Low P	Low NP
Tillers	2.57±0.87c	1.01±0.11a	1.32±0.54b	1.05±0.22a
Plant height (cm)	60.09±6.61c	57.55±7.07a	57.77±6.75a	58.68±9.58b
Main ear length (cm)	8.81±0.97d	5.52±1.41a	6.32±1.26c	5.96±1.56b
Main stem thickness/plant (mm)	5.47±0.68d	3.58±0.79a	4.21±0.65c	3.70±0.96b
Seeds/main ear	50.85±13.37c	25.05±12.05a	33.13±10.86b	27.78±13.44a
Seeds/plant	86.91±39.56c	25.21±12.31a	37.93±18.01b	28.38±15.19a
Main ear seed weight (g)	2.02±0.50c	0.90±0.61a	1.03±0.35b	0.91±0.47a
Aboveground dry weight (g)	3.98±0.83d	2.02±0.97a	2.38±0.77c	2.21±1.10b
Dry leaf area/main plant	62.08±27.96d	15.38±6.73b	17.91±7.95c	12.65±6.14a

*Values in columns are means ± standard deviation (*n*=225; except for dry leaf area* n*=30)*
Letters indicate a significant difference among the treatments; p≤*0.05*

There was also a cultivar by year interaction, showing that the trend for Zn content for the two years differed for the two cultivars. There was a treatment by year interaction for phytic acid, showing that the ranking for phytic acid for the different treatments differed for the two years. Only the low N treatment reduced Fe content significantly, and low P stress increased the concentrations of Fe and Zn. With P deficiency or limitation, roots secrete high amounts of organic acids, resulting in the mobilisation of and more intensive uptake of microelements by plants. The hyphae of arbuscular mycorrhizal fungi help to absorb more nutrients from the soil under P-deficient soil conditions.[31] Fe and Zn were found to be negatively correlated in some studies.[32]

Phytic acid, on the other hand, was less reduced under low N stress than under low P stress (Table 5), indicating that availability should increase under low P stress conditions. The highly reduced phytic acid concentration under low P stress was probably due to the fact that most phosphorus in the plant is in the form of phytate[19,33], meaning that reduced availability of P would lead to reduced phytic acid. It has been reported that high N treatment reduced phytic acid and increased Fe content in rice.[34] Similarly in pearl millet, cultivars with low phytic acid had high Fe and Zn bioacess.[32] This finding is in contrast with that of our study, however, where the low P treatment reduced phytic acid and enhanced Fe content, although it was cultivar dependent. Significant correlations between N fertilisation and Fe and Zn concentration in wheat grains has been previously reported.[35] Phytic acid was reduced by increasing N concentration, with the effect of varieties being significant in barley. In contrast to the current study, under low P conditions, one cultivar of barley had the highest phytic acid content.[36] Although phytic acid is the main reservoir of P in plants, it possesses anti-nutritional characteristics[17] which are associated with decreased bioavailability of Fe and Zn[18,19]. Therefore, the current results indicate that in areas experiencing low P stress, Fe and Zn will be more available, which can combat malnutrition and improve human health. This statement is further supported by the lowest phytic acid:Fe and phytic acid:Zn molar ratios under low P stress (Table 5), emphasising higher bioaccessibility of Fe and Zn under low P stress. In agreement, more Zn was available under such low P supply.[32] It has been reported that 55% of Zn was absorbed when the phytic acid:Zn ratio was less than 5, while 35% of Zn was absorbed when the ratio was between 5 and 15, and only 15% was absorbed when the ratio was higher than 15.[16] Low N stress benefitted cultivars selectively because it led to an increase in Zn content for SST806 only (Figure 2), showing that cultivars responded differently

under this treatment compared with under low P stress, which increased Zn content for both cultivars. For all nutritional traits, under combined low N and P stress, it appears that low N effect dominated effects of low P.

PAN3497 and SST806 were not significantly different for the measured phenotypic traits, irrespective of the treatment, which is in agreement with the nutritional quality data. Reduced tillering caused by these treatments, especially under low N stress, shows that the earlier a tiller is formed, the more kernels it will produce.[37] Reduction of other parameters, such as plant height and dry mass, under N and P deficiencies substantiates the fact that they correlate with grain yield positively.[38] As expected, low N, P and a combination of low N with P reduced the phenotypic performance of both cultivars substantially because both macroelements are essential for vegetative and reproductive growth.[7] Similar to nutritional quality results, the effect of low N masked that of low P and low N and low P combined stress, even though low P alone caused the lowest reductions in phenotypic parameters, pointing out the importance of N in these cultivars.

To conclude, our study has revealed that the cultivars were not significantly different for most nutritional and phenotypic traits under various treatments. What is interesting is that, under low P stress, micronutrients (Fe, Zn) increased to levels higher than those of the control while phytic acid was substantially reduced, leading to higher bioavailability of Fe and Zn in these cultivars. Although growing wheat under low P will reduce phenotypic traits and yield, there may be an unexpected benefit for combatting Zn and Fe deficiencies in the areas where the soil is deficient in these macronutrients.

Acknowledgements

This research was funded by the National Research Foundation of South Africa, through the South African Research Chairs Initiative.

Competing interests

We declare that there are no competing interests.

Authors' contributions

B.T. and M.J.M. performed the experiments and collected data. M.L. supervised the experiment. M.L. and L.S. completed the data statistical analysis. B.T., M.J.M. and M.L. wrote the manuscript. All authors read and approved the final manuscript.

References

1. Glass ADM. Nitrogen use efficiency of crop plants, physiological constraints upon nitrogen absorption. Crit Rev Plant Sci. 2003;22:453–470. https://doi.org/10.1080/07352680390243512
2. Hikosaka K. Interspecific difference in the photosynthesis–nitrogen relationship: Patterns, physiological causes, and ecological importance. J Plant Res. 2004;117:481–494. https://link.springer.com/article/10.1007/s10265-004-0174-2
3. UN Food and Agriculture Organization (FAO). World fertilizer trends and outlook to 2020. Summary report. Rome: FAO; 2017. Available from: http://www.fao.org/3/a-i6895e.pdf
4. Chianu J, Mairura JJ. Mineral fertilizers in the farming systems of sub-Saharan Africa. A review. Agron Sustain Dev. 2012;32(2):545–566. https://doi.org/10.1007/s13593-011-0050-0f
5. UN Food and Agriculture Organization (FAO). World fertilizer trends and outlook to 2019. Rome: FAO; 2016.
6. Kochian LV, Hoekenga OA, Piñeros MA. How do crops tolerate acid soils? Mechanisms of aluminium tolerance and phosphorus efficiency. Annu Rev Plant Biol. 2004;55:459–493. https://doi.org/10.1146/annurev.arplant.55.031903.141655
7. Hajabbasi MA, Schumacher TE. Phosphorus effects on root growth and development in two maize genotype. Plant Soil. 1994;158:39–46.
8. Kim HJ, Li X. Effects of phosphorus on shoot and root growth, partitioning, and phosphorus utilization efficiency in Lantana. HortScience. 2016;51(8):1001–1009.
9. Leghari SJ, Wahocho NA, Laghari GM, Laghari AH, Bhabhan GM, Talpur KH, et al. Role of nitrogen for plant growth and development: A review. Adv Environ Biol. 2016;10(9):209–219.
10. Heisy PW, Mwangi WM. Fertiliser use and maize production in sub Saharan Africa. CIMMYT Economics Working Paper. Mexico D.F.: CIMMYT; 1996.
11. Jakubus M, Graczyk M. Microelement variability in plants as an effect of sewage sludge compost application assessed by different statistical methods. Agronomy. 2020;10:642.
12. Tóth B, Moloi MJ. The use of industrial waste materials for alleviation of iron deficiency in sunflower and maize. Int J Recycl Org Waste Agric. 2019;8:145–151.
13. Nag A, Ray O, Rakshit R. Reflection on iron: Strategy for nutritional security in human. Natl Acad Sci Lett. 2019;1–4.
14. World Health Organization (WHO). World health reports. Geneva: WHO; 2014.
15. World Health Organization (WHO). Ten leading causes of illness and disease in low-income countries. The world health report 2002. Geneva: WHO; 2002. Available from: http://www.who.int/mediacentre/factsheets/fs310_2008.pdf
16. Magallanes-López AM, Hernandez-Espinosa N, Velu G, Posadas-Romano G, Ordonez-Villegas VMG, Crossa J, et al. Variability in iron, zinc, phytic acid content in a worldwide collection of commercial durum wheat cultivars and the effect of reduced irrigation on these traits. Food Chem. 2017;237:499–505. https://doi.org/10.1016/j.foodchem.2017.05.110
17. Vanishth A, Rams A, Beniwal V. Cereal phytases and their importance in improvement of micronutrient bioavailability. Biotech. 2017;7:42. https://doi.org/10.1007/s13205-017-0698-5
18. Zhao HJ, Liu LQ, Ren XL, Wu DX, Shu QY. Gene identification and allele-specific marker development for two allelic low phytic acid mutations in rice (*Oryza sativa* L.) Molecular Breeding. 2008;22:603–612.
19. Eagling T, Wawer AA, Shewry PR, Zhao F, Fairweather-Tait SJ. Iron bioavailability in two commercial cultivars of wheat: Comparison between wholegrain and white four and the effects of nicotianamine and 2'-deoxy mugineic acid on iron uptake into Caco–2 cells. J Agric Food Chem. 2014;62:10320–10325. https://doi.org/10.1021/jf5026295
20. Wang Z, Liu Q, Pan F, Yuan L, Yin X. Effects of increasing rates of zinc fertilization on phytic acid and phytic acid/zinc ratio in zinc bio-fortified wheat. Field Crops Res. 2015;184:58–64. https://doi.org/10.1016/j.fcr.2015.09.007
21. AOAC International. Official method 999.11. Determination of lead, cadmium, copper, iron and zinc in foods. J AOAC Int. 2000;73:1204.
22. Dragičević VD, Peric VA, Nisavic AR, Srebric MB. Validation study of a rapid colorimetric method for the determination of phytic acid and inorganic phosphorus from seeds. APTEFF. 2011;42:11–21.
23. Latta M, Eskin M. A simple and rapid colorimetric method for phytate determination. J Agric Food Chem. 1980;28:1313–1315.
24. Agrobase. Agrobase Generation II. Winnipeg: Agronomix; 2014.
25. Cakmak I, Ozkan H, Braun HJ, Welch RM, Römheld V. Zinc and iron concentrations in seeds of wild, primitive, and modern wheats. Food Nutr Bull. 2000;21:401–403. https://doi.org/10.1177/156482650002100411
26. Lestienne I, Icard-Verniére C, Mouquet C, Picq C, Tréche S. Effects of soaking whole cereal and legume seeds on iron, zinc and phytate content. Food Chem. 2005;89:421–425. https://doi.org/10.1016/j.foodchem.2004.03.040
27. Hamnér K, With M, Eriksson J, Kirchmann H. Influence of nitrogen supply on macro- and micronutrient accumulation during growth of winter wheat. Field Crops Res. 2017;213:118–119. https://doi.org/10.1016/j.fcr.2017.08.002
28. Cakmak I, Torun A, Millet E, Feldman M, Fahima T, Korol A, et al. *Triticum dicoccoides*: An important genetic resource for increasing zinc and iron concentration in modern cultivated wheat. Soil Sci Plant Nutr. 2004;50:1047–1054. https://doi.org/10.1080/00380768.2004.10408573
29. Chen XP, Zhang YQ, Tong YP, Xuel YF, Liu DY, Zhang W, et al. Harvesting more grain zinc of wheat for human health. Sci Rep. 2017;1:1–8. https://doi.org/10.1038/s41598-017-07484-2
30. White PJ, Broadley MR. Biofortifying crops with essential mineral elements. Trends Plant Sci. 2006;10:586–593. https://doi.org/10.1016/j.tplants.2005.10.001

31. Xie X, Hu W, Fan X, Chen H, Tang M. Interactions between phosphorus, zinc, and iron homeostasis in nonmycorrhizal and mycorrhizal plants. Front Plant Sci. 2019;10:1172.

32. Krishnan R, Meera MS. Assessment of inhibitory factors on bioaccessibility of iron and zinc in pearl millet (*Pennisetum glaucum* (L.) R. Br.) cultivars. J Food Sci Technol. 2017;54:4378–4386. https://www.readcube.com/articles/10.1007/s13197-017-2911-2

33. Su D, Sultan F, Zhao NC, Lei BT, Wang FB, Pan G, et al. Positional variation in grain mineral nutrients within a rice panicle and its relation to phytic acid concentration. J Zhejiang Univ Sci B. 2014;15:986–996.

34. Zhao N, Zhang Q, Cheng F, Zhou W. Effects of nitrogen, phosphorus and zinc supply levels on grain phytic acid content and its correlation with several mineral nutrients in rice grains. Chin J Rice Sci. 2007;21(2):185–190.

35. Bouis HE, Welch RM. Biofortification – A sustainable agricultural strategy for reducing micronutrient malnutrition in the Global South. Crop Sci. 2010;50:S20–S32. http://dx.doi.org/10.2135/cropsci2009.09.0531

36. Mao Y, Chen X, Qiu BY, Wu FB, Zhang GP. Effect of nitrogen and phosphorus fertilizer application rates on phytic acid content in barley grains. Journal of Zhejiang University (Agriculture and Life Sciences). 2009;35:285–291.

37. Tilley MS, Heigniger RW, Crozier CR. Tiller initiation and its effects on yield and yield components in winter wheat. Agron J. 2019;111:1323–1332.

38. Baye A, Berihun B, Bantayehu M, Derebe B. Genotypic and phenotypic correlation and path coefficient analysis for yield and yield-related traits in advanced bread wheat (*Triticum aestivum* L.) lines. Cogent Food Agric. 2020;6:1752603.

AUTHORS:
Martin Kemler[1,2,3]
Michael J. Wingfield[1,2]
Don A. Cowan[1,4]
Bernard Slippers[1,2]

AFFILIATIONS:
[1]Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa
[2]Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa
[3]AG Geobotanik, Ruhr University Bochum, Bochum, Germany
[4]Centre for Microbial Ecology and Genomics, University of Pretoria, Pretoria, South Africa

CORRESPONDENCE TO:
Martin Kemler

EMAIL:
martin.kemler@rub.de

DATES:
Received: 25 Nov. 2019
Revised: 28 May 2020
Accepted: 17 Aug. 2020
Published: 29 Mar. 2021

HOW TO CITE:
Kemler M, Wingfield MJ, Cowan DA, Slippers B. Foliar fungi of the enigmatic desert plant *Welwitschia mirabilis* show little adaptation to their unique host plant. S Afr J Sci. 2021;117(3/4), Art. #7666. https://doi.org/10.17159/sajs.2021/7666

ARTICLE INCLUDES:
☒ Peer review
☒ Supplementary material

DATA AVAILABILITY:
☒ Open data sets*
☐ All data included
☐ On request from author(s)
☐ Not available
☐ Not applicable
*see accession numbers in text

EDITORS:
Teresa Coutinho
Salmina Mokgehle

KEYWORDS:
Welwitschia mirabilis, Namib Desert, foliar fungi, *Alternaria*, *Aureobasidium*

FUNDING:
Department of Science and Innovation (DSI)–National Research Foundation (NRF) Centre of Excellence in Tree Health Biotechnology (CTHB), University of Pretoria

Foliar fungi of the enigmatic desert plant *Welwitschia mirabilis* show little adaptation to their unique host plant

Foliar fungi, especially endophytic fungi, constitute an important part of the microbiome of plants. Yet little is known about the composition of these communities. In this study, we isolated fungi from leaf tissues of the desert plant *Welwitschia mirabilis* to determine the culturable diversity of the foliar fungal community. The isolated fungal taxa, which grouped into 17 distinct lineages, were identified by sequencing elongation factor 1 alpha, beta-tubulin 1, beta-tubulin 2 and the internal transcribed spacer region. The culturable community was mainly composed of cosmopolitan fungal genera despite the unique taxonomic position of the plant and its geographic isolation. To test for endemism in two of the common fungal genera, *Alternaria* and *Aureobasidium*, we built haplotype networks using a global data set. Even this broad data set showed little evidence for specialisation within this unique host or its geographical location. The data suggest that the culturable members of communities of leaf-associated fungi in habitats with little plant coverage, such as the Namib Desert, are mainly established by long-distance aerially distributed fungal inocula and few of these taxa co-evolve with the host within the habitat.

Significance:

- The culturable members of fungal communities associated with an ecological and evolutionary isolated plant have not co-speciated with their hosts, but to a large extent are composed of globally distributed fungal species.
- Harsh environmental conditions and the geographic isolation of host plants seem to favour ubiquitous fungal species over more specialist fungal species.

Introduction

Fungi and plants have a long history of co-evolution and plant–fungal interactions are thought to have been essential in the establishment of plants in terrestrial environments. For example, fossil records of interaction structures of arbuscular mycorrhizal fungi and Embryophyta have been recovered from the Rhynie cherts and date back to more than 400 Ma[1], possibly as old as 475 Ma[2]. Structures of endophytic fungi (i.e. microorganisms that in part or during their whole life cycle colonise plant tissues without visible symptoms[3]) in prostrate axes have been described from as early as 400 Ma.[4] Similarly, leaf-associated endophytes are known from the 300 Ma Carboniferous era.[5] These studies show that, despite the unknown interaction type between host and fungus, aboveground organs of plants, the so-called phyllosphere, have served as a suitable habitat for fungi for a long time.

Extant phyllosphere-associated fungi including endophytes and epiphytes (i.e. fungi that grow on the surface of plants) mostly show no visible impact on their host under favourable conditions or even enhance plant performance. Some studies have provided evidence that these fungi affect plant physiology. Some epiphytic and endophytic yeasts have been shown to promote plant growth[6,7], whereas some endophytes become pathogenic when the host plant experiences abiotic or biotic stress[8,9]. Despite our limited knowledge of their function, culture-dependent and culture-independent methods have identified species-rich fungal communities in the phyllosphere that add substantially to the hyperdiversity of fungi.[10-14] How these communities become established and the factors that influence them are understood only for a limited number of plant–fungus systems. In the case of endophytes, the greatest impact on the structure of communities results from geographic location and host plant.[15] It is also thought that the evolutionary history of the host and the climate[16] play additional roles. Most likely similar mechanisms apply for epiphytes.[17,18]

Plant species in deserts harbour a greater proportion of fungal endophytes with cosmopolitan distribution than the same plant species under less extreme conditions.[15] This observation is interesting in the light of two aspects. Firstly, many plant–fungus interactions are characterised by local adaptation to the host population[19,20], which is reflected in the genetic sub-structuring of the fungal population[21]. Secondly, many fungi show clear biogeographic patterns when an appropriate species concept that incorporates molecular data is applied.[22] These findings imply that geographic structure should exist in the distribution of cosmopolitan phyllosphere fungi. However, especially culture-dependent studies have recovered many fungal species that seem not necessarily restricted to plant interactions and that have been found in other habitats as well.[17,23,24] Such low host association could counterbalance restrictions to gene flow and lead to low genetic population structure. However, to the best of our knowledge, no studies of phyllosphere-associated fungi have focused on the possibility of biogeographic patterns of cosmopolitan species.

The Namib Desert is one of the oldest and driest deserts on the planet.[25] Like all desert ecosystems, it imposes severe constraints on living organisms and is characterised by the stochasticity of nutrient and water supply, as well as continuous varying UV radiation and extreme temperature changes.[26] The coastal (western) Namib Desert is unusual in that it experiences frequent fog events with a decreasing gradient from the coast towards

the Great Escarpment. It also experiences infrequent and sparse rainfall events with a decreasing gradient from the Great Escarpment towards the coast.[25]

Welwitschia mirabilis Hook, the only living member of the family Welwitschiaceae, is a unique dioecious plant that occurs natively only in the Namib Desert and adjacent savannah ecosystems in Namibia and Angola (Figure 1).[27] It is part of an old lineage of land plants, the Gnetophyta, that diverged from the Pinaceae between 121.8 Ma and 309.5 Ma ago.[28] Welwitschiaceae fossils have been found from the Lower Cretaceous, 112–114 Ma ago.[29] Although these long-lived plants (estimates reach over 2000 years[27]) have been studied extensively[30], little is known regarding their associated mycobiome. A few studies have concentrated on the effect of *Aspergillus* on seed mortality[31,32] and the level of mycorrhizal interactions[33], but to date the diversity of phyllosphere fungi associated with *Welwitschia* is unknown.

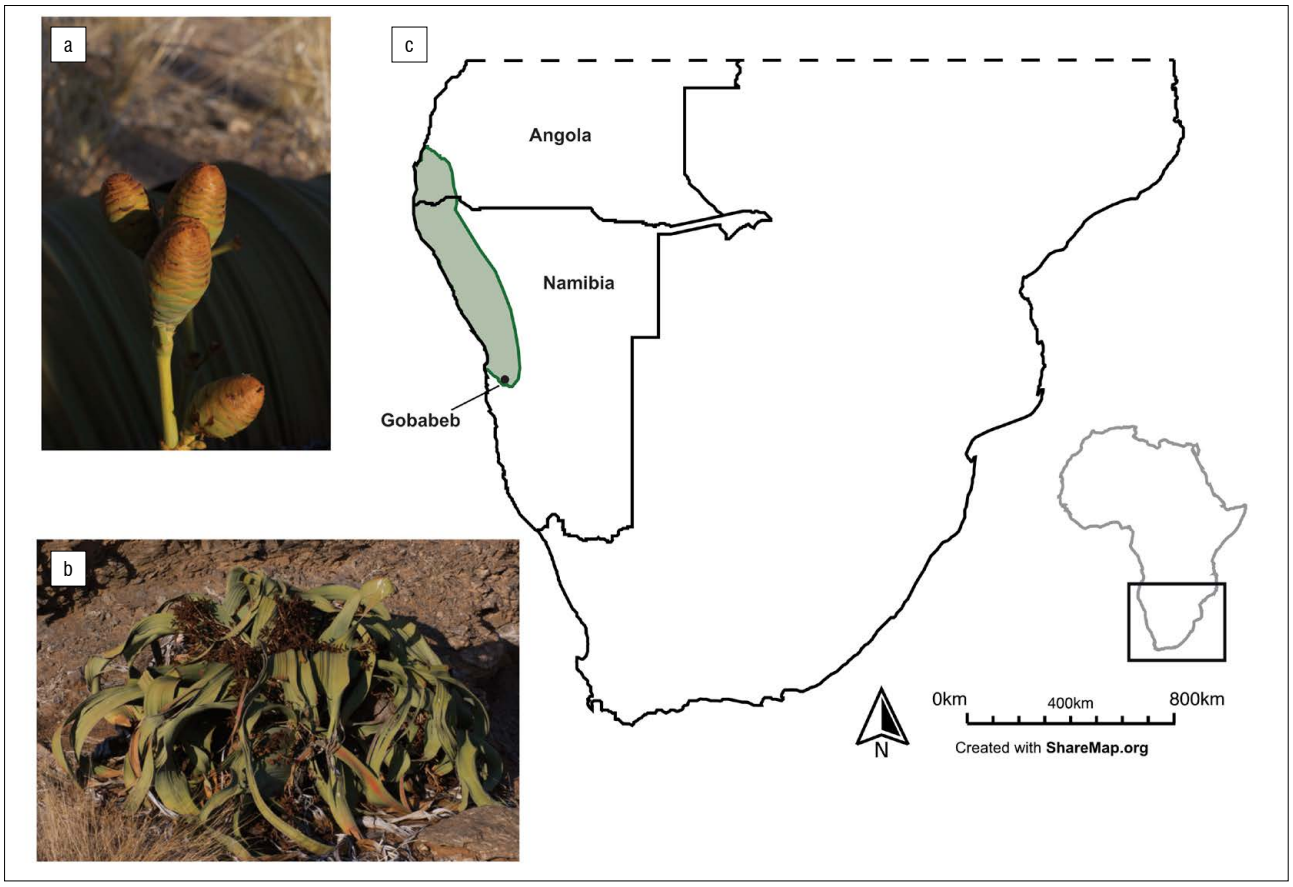
The uniqueness of the phylogenetic position of *Welwitschia*, as well as its remote habitat in one of the oldest and hottest deserts, provides an ideal opportunity to compare factors that might influence the composition of its fungal communities. As a first goal we isolated fungi to assess the diversity and uniqueness of potential endophytes colonising *Welwitschia* plants at a species level. The second goal was to consider the biogeographic structure of intraspecific diversity for some of the common cosmopolitan fungi that were isolated.

Materials and methods

Samples from *Welwitischia mirabilis* were recovered under the auspices of the Namibian Ministry of Environment and Tourism permit, number 1522/2011. Sections of leaves (about 20 cm^2) were obtained from two individuals in the Welwitschia Wash near Gobabeb (23°36.54' S; 15°10.07' E) in the Central Namib Desert in May 2011. Only leaf tissue from young sections near the stem could be obtained, as the plants showed symptoms of intense grazing and lacked any old growth leaf parts. Leaves were washed with ddH_2O and subsequently shaken for 1 h in phosphate-buffered saline[34] to remove superficially attached spores. Leaves were not surface sterilised, as our interest was in all fungi associated with this desert plant, including endophytic and epiphytic fungi. The leaves were cut into pieces of approximately 4 x 4 mm and plated on malt extract agar at room temperature (~22 °C). Emerging mycelia were transferred to fresh malt extract agar media to obtain pure cultures.

DNA from pure fungal cultures was isolated using either a salt extraction method described by Aljanabi and Martinez[35] or the Qiagen DNeasy® Plant Mini Kit (Qiagen, Hilden, Germany). The fungal barcode region, namely the rDNA *ITS*, as well as parts of elongation factor 1 alpha (*EF1a*), beta-tubulin 1 (*BTUB1*) and beta-tubulin 2 (*BTUB2*) genes for selected samples were amplified (Supplementary table 1). The *ITS* rRNA gene region was amplified using ITS1F[36] and ITS4[37]; *EF1a* was amplified using EF1f and EF1r; *BTUB1* was amplified using BT1a and BT1b[38], and *BTUB2* was amplified using BT2a and BT2b[38]. Amplicons were sequenced using the Big Dye™ Terminator Cycle Sequencing Kit V3.1 (Applied Biosystems, Foster City, CA, USA) on an ABI PRISM® 3100 Genetic Analyser (Applied Biosystems, Foster City, CA, USA).

Sequences were checked for quality, trimmed and, where applicable, a consensus sequence was built from forward and reverse strands using CLC Genomics Workbench v4.0.3 (CLC bio, Aarhus, Denmark). Phylogenetic analyses were used to assess the species composition. Phylogenetic trees were inferred for *ITS*, *EF1a* and *BTUB1* individually, as well as by applying a super matrix approach of a concatenated data set for these DNA regions. For the individual gene trees, sequences were compared against the NCBI nucleotide database using BLASTn[39]



Images: ©Martin Kemler

Figure 1: (a) Female seed cones of *Welwitschia mirabilis*. (b) Male *Welwitschia mirabilis* with pollen cones. (c) Geographic distribution of *Welwitschia mirabilis* in Angola and Namibia. The sampling site near Gobabeb is indicated.

and the five hits with greatest similarity were downloaded. Alignments were generated using MAFFT v7.154b[40,41] under the auto option, which chooses the most appropriate alignment strategy for a given data set automatically. The aligned sequences were pasted into the GBlocks web interface[42] to remove ambiguous sites, thus allowing smaller final blocks and gaps in the alignment. Species names for the specimens isolated in this study were assigned based on the single tree. For the super matrix approach, selected NCBI sequences of the single gene trees that clustered close to the sequences retrieved in this study were chosen. Sequences were aligned as described for the individual genes and then concatenated using SequenceMatrix[43], whereby missing sequences were coded as missing data. Maximum likelihood phylogenies and bootstrap with 1000 replicates were inferred using RAxML 8.0.25[44] applied via the raxmlGUI[45] using the GAMMAGTR and the rapid bootstrap option.

To consider the possible effects of geographic distribution and adaptation to *W. mirabilis,* haplotype networks were reconstructed for two of the dominant fungal taxa (*Aureobasidium* and *Alternaria*) with a broader phylogenetic sampling from the NCBI database. In the case of *Aureobasidium*, individual haplotype networks were constructed for ITS and BTUB2, whereas for Alternaria they were constructed for ITS and BTUB1. After aligning the sequences in MAFFT v7.154b, the sequences were collapsed into haplotypes using Map in the SNAP workbench[46,47], thus excluding indels and infinite site violations. The pegas package[48] in R v3.0.0[49] was then used to construct minimum spanning trees from these haplotypes.

Results

The foliar fungi associated with the leaves of *W. mirabilis* that could be isolated and grown in culture belonged to either the Ascomycota or Basidiomycota (Supplementary table 1). The latter group was represented by two isolates (*Rhodotorula* and *Cryptococcus*) and these were not considered in further analyses. Within the Ascomycota, specimens belonging to the Capnodiales, Dothideales, Eurotiales, Hypocreales, Pleosporales or Xylariales were recovered and the most common taxa were *Alternaria*, *Aspergillus*, *Aureobasidium*, Didymellaceae and *Pleospora welwitschiae* s.l. Taxa with few isolates recovered belonged to *Bionectria*, *Cladosporium*, Montagnulaceae, *Penicillium*, *Pestalotiopsis*, *Stigmina* and *Talaromyces* (Figure 2). In total, 17 species or clades that might represent species were identified in the Ascomycota. The majority of taxa clustered together with fungal sequences obtained by other studies from other geographic localities. We inferred only three lineages that did not contain sequences from other studies. However, even these lineages clustered closely to sequences from fungi deposited at NCBI (Figure 2). In the Didymellaceae, we recovered one lineage that showed some genetic divergence from its sister taxon *Didymella pinodella* and another lineage that showed phylogenetic affinity to an uncultured *Phoma* sequence. One lineage clustered together with *Apioplagiostoma aceriferum* and a sequence from *Oryza sativa* (a potential fungal contamination), but with a larger genetic distance and low statistical support. Fungal specimens isolated from *Welwitschia* that fell within the genera *Alternaria* and *Aureobasidium* did not cluster according to their origin but were intermingled with sequences from other geographic origins. The sequences obtained in this study were deposited in GenBank under the accession numbers KT150524–KT150716 (Supplementary table 1).

Sequences downloaded from NCBI for the more thorough analyses of geographic clustering of the two most dominant taxa included: for *Alternaria,* 493 *ITS* and 122 *BTUB1* sequences, and for *Aureobasidium*, 376 *ITS* and 50 *BTUB2* sequences. The analysis of the *ITS* sequences for both genera inferred that the dominant haplotypes contained most of the sequences from many different geographic locations (Figure 3a and 4a). The analysis of the *BTUB1* region in *Alternaria* and the analysis of the *BTUB2* region in *Aureobasidium* showed more haplotype diversity. However, there was no clear correlation between geographic location and the occurrence of fungal groupings (Figure 3b and 4b).

Discussion

This study, although restricted to a few leaf samples because of the high protective state of *W. mirabilis*, revealed a rich fungal diversity of foliar fungi associated with the leaves of this plant. Of the 17 taxa identified, only *Pleospora welwitschiae* is a species known to be specifically associated with *W. mirabilis*. Only three other lineages, which according to our analyses could belong to as-yet unknown taxa, might potentially represent endemic or even host-specific lineages. The other taxa clustered closely with sequences of fungi from different geographic origins and appear to have widespread geographic distributions. To understand geographic sub-structuring and potential adaptation we explored haplotype patterns in more detail for the dominant cosmopolitan genera *Aureobasidium*[50] and *Alternaria*[51].

In the phylogenetic analysis, only the best matching sequences from NCBI were used for identification purposes, as is common for studies of this nature. However, low taxon sampling in phylogenetic analyses has a limited capacity to show any geographic clustering within a species. Many plant-associated fungi show geographic patterns due to local adaptations to a host population and biogeographic patterns in fungi are more prevalent than previously assumed.[22,52,53] We thus hypothesised that for adaptation to the locally restricted host plant to occur, some geographic sub-structure must exist in the common fungal taxa associated with *W. mirabilis*.

In order to identify the existence of geographic sub-structure in the two most frequently isolated fungal taxa isolated from *W. mirabilis*, haplotype networks using larger sequence data sets with publicly available data were reconstructed. Using these larger data sets, we observed no geographic clustering of our samples of *Alternaria* and *Aureobasidium* amongst global collections of closely related sequences (Figures 3 and 4). These results indicate common gene flow between these fungi on *W. mirabilis* and populations represented by downloaded sequences. They most likely also reflect frequent introduction of these fungi into the Namib Desert. Although grazers (e.g. springbok, donkeys and horses[30]) could also distribute phyllosphere fungi, given the general isolation of the environments in which *W. mirabilis* grows, with usually limited human traffic, spores of the commonly isolated and globally distributed fungi are expected to be mostly aerially dispersed. Supporting this view, studies of the biological content of aerosols have shown that fungal biomass can be up to 45% of the total weight of coarse aerosol particles and contain some of the fungi recovered in the present study.[54]

Isolation of cosmopolitan foliar fungal genera has previously been reported in other plant species[15,17,23,24], but seems even more pronounced in individuals in extreme climatic conditions[15]. Studies of endophytes from desert plants, including other gymnosperms, have shown more diverse communities than in plants of the same species from less extreme environments, but cosmopolitan taxa such as *Alternaria*, *Aspergillus*, *Aureobasidium*, *Cladosporium* and *Phoma* constituted the majority of isolated fungi.[15,55,56] In *Deschampsia antarctica*, one of the two angiosperms that grow on the Antarctic continent, endophyte diversity was generally low and the most common endophyte isolated was an *Alternaria* sp.[57] When our results are considered together with these previous studies, it appears that the pattern of high diversity but low uniqueness of plant-associated microbes might be consistent across diverse and isolated extremely dry environments (both cold and hot).

The reason for the high level of occurrence of cosmopolitan and mostly saprobic fungi in harsh environments is not known. It has, however, been hypothesised that infecting any given host in such an environment, even a less optimal host, is preferable to prolonged exposure to extreme temperatures, desiccation or UV radiation.[15] An alternative hypothesis, substantiated by analyses in our study, could be that these fungi, due to their cosmopolitan distribution, are not host specific. In this case, they would be outcompeted in environments where specialised fungi occur, thereby decreasing their abundance. In harsh environments where lack of plant cover reduces overall propagule number, as well as opportunities to infect plants, the adaptations of these opportunistic colonists for survival in aerosols and in harsh environments might provide them with a competitive advantage. Some of the common taxa isolated in this study are well adapted to live under desert conditions.

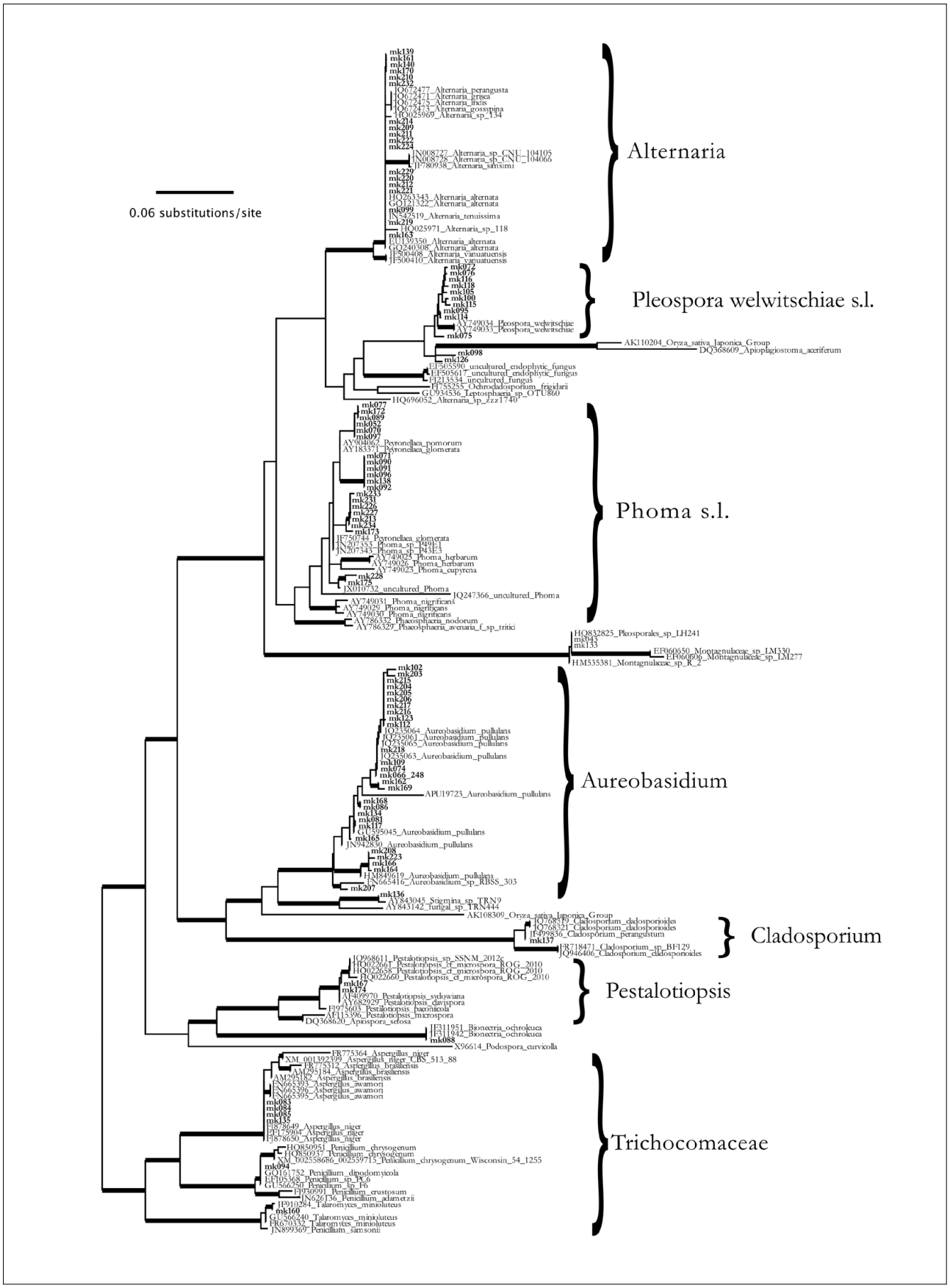


Figure 2: Maximum likelihood reconstruction of Ascomycota groupings isolated from *Welwitschia mirabilis* based on ITS, elongation factor 1 alpha and beta-tubulin 1. A bootstrap of 1000 was conducted. The phylogeny was rooted with members of Eurotiomycetes. Taxa in bold are from this study. Bold branches indicate bootstrap values ≥70. Major fungal lineages are indicated. See also Supplementary data.

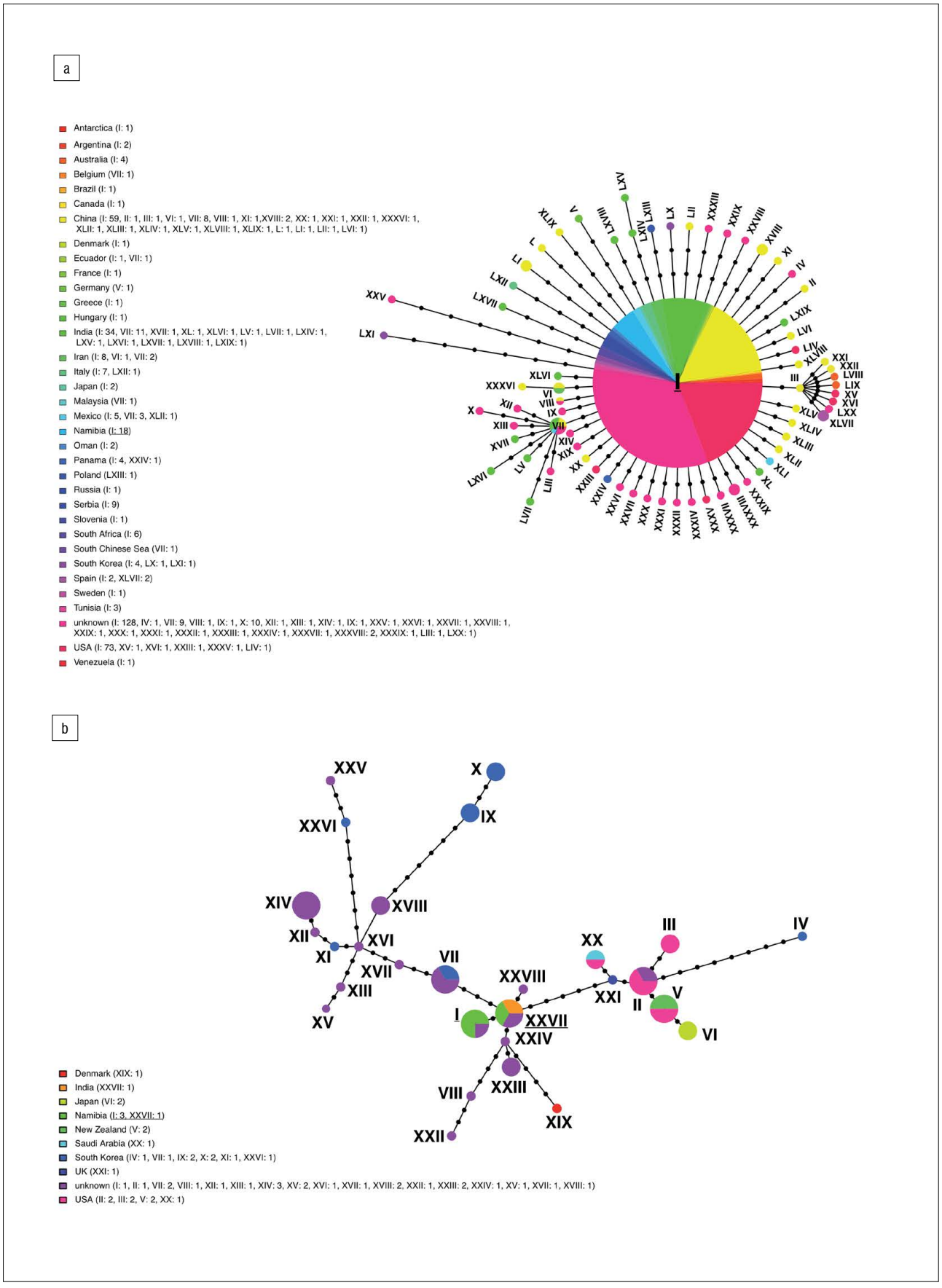


Figure 3: Minimum spanning tree of (a) ITS and (b) beta-tubulin 1 for *Alternaria* sequences. Roman numerals indicate individual haplotypes. Country of origin is colour-coded and text in parentheses indicates the number of sequences per haplotype per country of origin. Haplotypes containing taxa from this study are underlined.

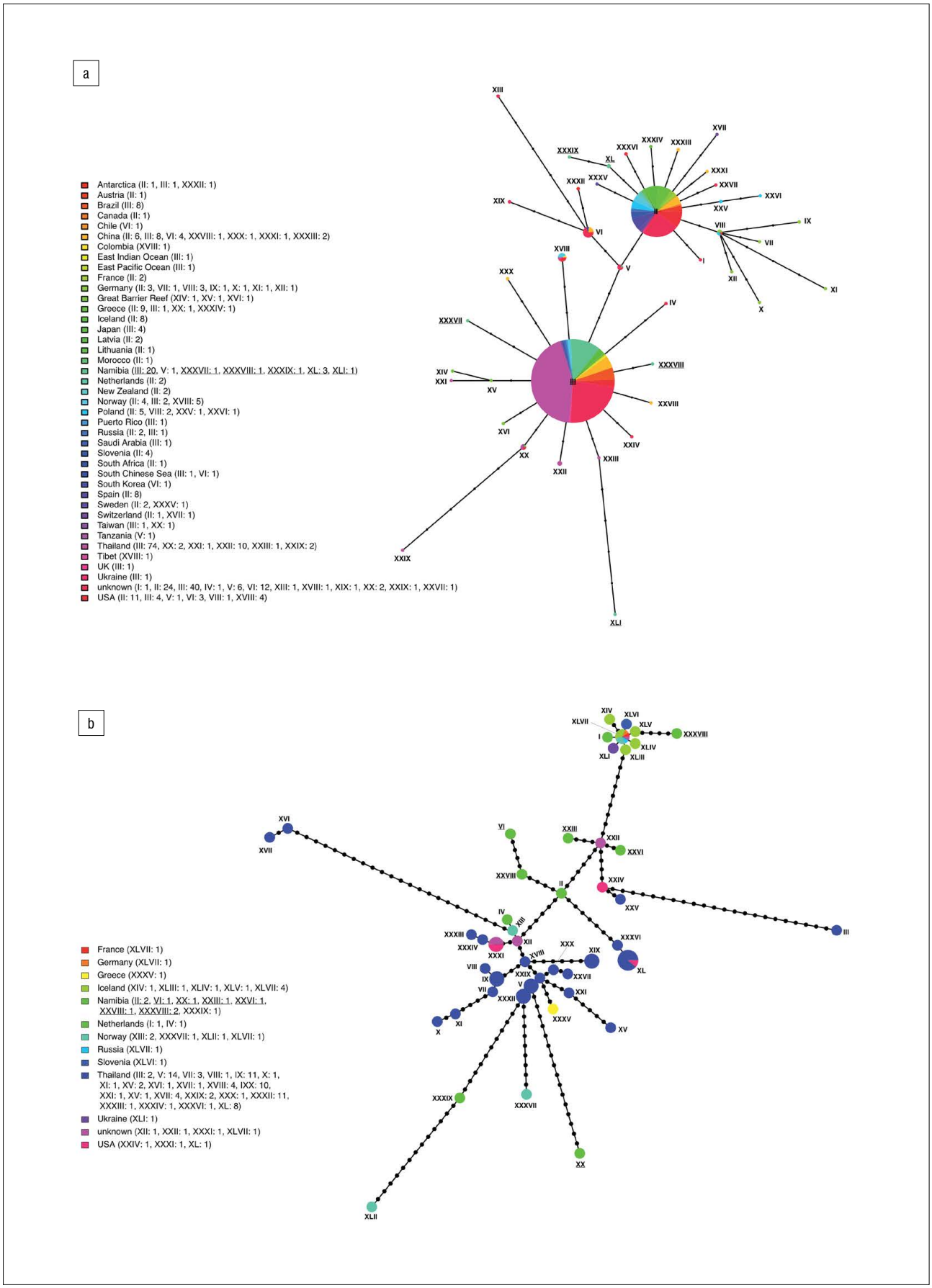


Figure 4: Minimum spanning tree of (a) ITS and (b) beta-tubulin 2 for *Aureobasidium* sequences. Roman numerals indicate individual haplotypes. Country of origin is colour-coded and the text in parentheses indicates the number of sequences per haplotype per country of origin. Haplotypes containing taxa from this study are underlined.

Many of these taxa (e.g. *Alternaria*[58,59], *Aspergillus*[60], *Aureobasidium*[61,62]) have been shown to produce melanins, a group of polymers that fungi produce as protection against harsh environmental conditions.[63] In cases where the fungi enter an endophytic life stage, the production of melanins might be less relevant. However, during long-distance dispersal, arrival, epiphytic growth, and infection in an extreme environment, melanins could be the difference between success and failure of establishment.

In this study, we used a culture-dependent approach to identify fungi. Studies comparing culture-dependent vs. culture-independent next-generation sequencing studies point to the fact that there could be a significant difference in fungi between these approaches.[64] Nevertheless, this culture-based approach remains relevant for the categories of taxa which have been isolated. Future studies, possibly using culture-independent tools and next-generation sequencing, would therefore be especially interesting, particularly with an increased sample size to include gradients from extremely dry to more humid environments. This would make it possible to characterise changes in the patterns common to geographically unique fungal phyllosphere communities. Such studies should be complemented by local air sampling to assess the propagule hypothesis raised in this study.

Acknowledgements

We thank the Department of Science and Technology (DSI)–National Research Foundation (NRF) Centre of Excellence in Tree Health Biotechnology (CTHB) for funding to conduct this study. M.K. thanks the University of Pretoria for postdoctoral funding. We also acknowledge the important improvements suggested by the anonymous reviewers.

Competing interests

We declare that there are no competing interests.

Authors' contributions

D.A.C., B.S. and M.J.W. conceived the study and provided funding. M.K. and D.A.C. conducted the sampling of material. M.K. conducted sample analysis, data analysis and wrote the initial draft. B.S., D.A.C., M.K. and M.J.W. revised the manuscript and wrote the final draft. All authors approved the final draft.

References

1. Remy W, Taylor TN, Hass H, Kerp H. Four hundred-million-year-old vesicular arbuscular mycorrhizae. Proc Natl Acad Sci USA. 1994;91:11841–11843. https://doi.org/10.1073/pnas.91.25.11841
2. Bidartondo MI, Read DJ, Trappe JM, Merckx V, Ligrone R, Duckett JG. The dawn of symbiosis between plants and fungi. Biol Lett. 2011;7:574–577. https://doi.org/10.1098/rsbl.2010.1203
3. Wilson D. Endophyte: The evolution of a term, and clarification of its use and definition. Oikos. 1995;73:274–276. http://dx.doi.org/10.2307/3545919
4. Krings M, Taylor TN, Hass H, Kerp H, Dotzler N, Hermsen EJ. Fungal endophytes in a 400-million-yr-old land plant: Infection pathways, spatial distribution, and host responses. New Phytol. 2007;174:648–657. https://doi.org/10.1111/j.1469-8137.2007.02008.x
5. Krings M, Taylor TN, Dotzler N, Galtier J. Fungal remains in cordaite (Cordaitales) leaves from the Upper Pennsylvanian of central France. Bull Geosci. 2011;86:777–784. https://doi.org/10.3140/bull.geosci.1278
6. Fu SF, Sun PF, Lu HY, Wei JY, Xiao HS, Fang WT, et al. Plant growth-promoting traits of yeasts isolated from the phyllosphere and rhizosphere of *Drosera spatulat*a Lab. Fungal Biol. 2016;120:433–448. https://doi.org/10.1016/j.funbio.2015.12.006
7. Nassar AH, El-Tarabily KA, Sivasithamparam K. Promotion of plant growth by an auxin-producing isolate of the yeast *Williopsis saturnus* endophytic in maize (*Zea mays* L.) roots. Biol Fertil Soils. 2005;42:97–108. https://doi.org/10.1007/s00374-005-0008-y
8. Sieber TN. Endophytic fungi in forest trees: Are they mutualists? Fungal Biol Rev. 2007;21:75–89. https://doi.org/10.1016/j.fbr.2007.05.004
9. Slippers B, Wingfield MJ. Botryosphaeriaceae as endophytes and latent pathogens of woody plants: Diversity, ecology and impact. Fungal Biol Rev. 2007;21:90–106. https://doi.org/10.1016/j.fbr.2007.06.002
10. Arnold AE, Maynard Z, Gilbert GS, Coley PD, Kursar TA. Are tropical fungal endophytes hyperdiverse? Ecol Lett. 2000;3:267–274. https://doi.org/10.1046/j.1461-0248.2000.00159.x
11. Fonseca Á, Inácio J. Phylloplane yeasts. In: Rosa C, Gábor P, editors. Biodiversity and ecophysiology of yeasts. Berlin/Heidelberg: Springer; 2006. p. 263–301. https://doi.org/10.1007/3-540-30985-3_13
12. Jumpponen A, Jones KL. Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. New Phytol. 2009;184:438–448. https://doi.org/10.1111/j.1469-8137.2009.02990.x
13. Kemler M, Garnas J, Wingfield MJ, Gryzenhout M, Pillay K-A, Slippers B. Ion torrent PGM as tool for fungal community analysis: A case study of endophytes in *Eucalyptus grandis* reveals high taxonomic diversity. PLoS ONE. 2013;8, e81718. https://doi.org/10.1371/journal.pone.0081718
14. Zimmerman NB, Vitousek PM. Fungal endophyte communities reflect environmental structuring across a Hawaiian landscape. Proc Natl Acad Sci USA. 2012;109:13022–13027. https://doi.org/10.1073/pnas.1209872109
15. Hoffman MT, Arnold AE. Geographic locality and host identity shape fungal endophyte communities in cupressaceous trees. Mycol Res. 2008;112:331–344. https://doi.org/10.1016/j.mycres.2007.10.014
16. U'Ren JM, Lutzoni F, Miadlikowska J, Laetsch AD, Arnold AE. Host and geographic structure of endophytic and endolichenic fungi at a continental scale. Am J Bot. 2012;99:898–914. https://doi.org/10.3732/ajb.1100459
17. Guimaraes JB, Pereira P, Chambel L, Tenreiro R. Assessment of filamentous fungal diversity using classic and molecular approaches: Case study – Mediterranean ecosystem. Fungal Ecol. 2011;4:309–321. https://doi.org/10.1016/j.funeco.2011.01.006
18. Kemler M, Witfeld F, Begerow D, Yurkov A. Phylloplane yeasts in temperate climates. In: Buzzini P, Lachance MA, Yurkov A, editors. Yeasts in natural ecosystems: Diversity. Cham: Springer Nature; 2017. p. 171–197. https://doi.org/10.1007/978-3-319-62683-3_6
19. Greischar MA, Koskella B. A synthesis of experimental work on parasite local adaptation. Ecol Lett. 2007;10:418–434. https://doi.org/10.1111/j.1461-0248.2007.01028.x
20. Kaltz O, Shykoff JA. Local adaptation in host–parasite systems. Heredity. 1998;81:361–370. https://doi.org/10.1046/j.1365-2540.1998.00435.x
21. Enjalbert J, Duan X, Leconte M, Hovmøller MS, De Vallavieille-Pope C. Genetic evidence of local adaptation of wheat yellow rust (*Puccinia striiformis* f. sp. *tritici*) within France. Mol Ecol. 2005;14:2065–2073. https://doi.org/10.1111/j.1365-294X.2005.02566.x
22. Taylor JW, Turner E, Townsend JP, Dettman JR, Jacobson D. Eukaryotic microbes, species recognition and the geographic limits of species: Examples from the kingdom Fungi. Philos Trans R Soc Lond B Biol Sci. 2006;361:1947–1963. https://doi.org/10.1098/rstb.2006.1923
23. Smith H, Wingfield MJ, Petrini O. *Botryosphaeria dothidea* endophytic in *Eucalyptus grandis* and *Eucalyptus nitens* in South Africa. For Ecol Manag. 1996;89:189–195. https://doi.org/10.1016/S0378-1127(96)03847-9
24. Unterseher M, Schnittler M. Species richness analysis and ITS rDNA phylogeny revealed the majority of cultivable foliar endophytes from beech *Fagus sylvatica*. Fungal Ecol. 2010;3:366–378. https://doi.org/10.1016/j.funeco.2010.03.001
25. Viles HA, Goudie AS. Weathering in the central Namib Desert, Namibia: Controls, processes and implications. J Arid Environ. 2011;93:20–29. https://doi.org/10.1016/j.jaridenv.2011.09.012
26. Noy-Meir I. Desert ecosystems: Environment and producers. Annu Rev Ecol Syst. 1973;4:25–51. https://doi.org/10.1146/annurev.es.04.110173.000325
27. Von Willert DJ. *Welwitschia mirabilis* Hook. fil. – das Überlebenswunder der Namibwüste [*Welwitschia mirabilis* Hook. fil. – the survival wonder of the Namib Desert]. Naturwissenschaften. 1994;81:430–442. German. https://doi.org/10.1007/BF01136642
28. Clarke JT, Warnock RC, Donoghue PC. Establishing a time-scale for plant evolution. New Phytol. 2011;192:266–301. https://doi.org/10.1111/j.1469-8137.2011.03794.x

29. Dilcher DL, Bernardes-De-Oliveira ME, Pons D, Lott TA. Welwitschiaceae from the Lower Cretaceous of northeastern Brazil. Am J Bot. 2005;92:1294–1310. https://doi.org/10.3732/ajb.92.8.1294

30. Henschel JR, Seely MK. Long-term growth patterns of *Welwitschia mirabilis*, a long-lived plant of the Namib Desert (including a bibliography). Plant Ecol. 2000;150:7–26. https://doi.org/10.1023/A:1026512608982

31. Pekarek E, Jacobson K, Donovan A. High levels of genetic variation exist in *Aspergillus niger* populations infecting *Welwitschia mirabilis* Hook. J Hered. 2006;97:270–278. https://doi.org/10.1093/jhered/esj031

32. Whitaker C, Pammenter NW, Berjak P. Infection of the cones and seeds of *Welwitschia mirabilis* by *Aspergillus niger* var. *phoenicis* in the Namib-Naukluft Park. S Afr J Bot. 2008;74:41–50. https://doi.org/10.1016/j.sajb.2007.08.008

33. Jacobson KM, Jacobson PJ, Miller Jr OK. The mycorrhizal status of *Welwitschia mirabilis*. Mycorrhiza. 1993;3:13–17. https://doi.org/10.1007/BF00213462

34. Mendes R, Pizzirani-Kleiner AA, Araujo WL, Raaijmakers JM. Diversity of cultivated endophytic bacteria from sugarcane: Genetic and biochemical characterization of *Burkholderia cepacia* complex isolates. Appl Environ Microbiol. 2007;73:7259–7267. https://doi.org/10.1128/AEM.01222-07

35. Aljanabi SM, Martinez I. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Res. 1997;25:4692–4693. https://doi.org/10.1093/nar/25.22.4692

36. Gardes M, Bruns TD. ITS primers with enhanced specificity for basidiomycetes – application to the identification of mycorrhizae and rusts. Mol Ecol. 1993;2:113–118. https://doi.org/10.1111/j.1365-294X.1993.tb00005.x

37. White TJ, Bruns T, Lee S, Taylor J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M, Gelfand D, Shinsky J, White T, editors. PCR protocols: A guide to methods and applications. San Diego, CA: Academic Press; 1990. p. 315–322. https://doi.org/10.1016/B978-0-12-372180-8.50042-1

38. Glass NL, Donaldson GC. Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. Appl Environ Microbiol. 1995;61:1323–1330.

39. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

40. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30:3059–3066. https://doi.org/10.1093/nar/gkf436

41. Katoh K, Kuma K, Toh H, Miyata T. MAFFT version 5: Improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 2005;33:511–518. https://doi.org/10.1093/nar/gki198

42. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 2000;17:540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334

43. Vaidya G, Lohman DJ, Meier R. SequenceMatrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics. 2011;27:171–180. https://doi.org/10.1111/j.1096-0031.2010.00329.x

44. Stamatakis A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033

45. Silvestro D, Michalak I. raxmlGUI: A graphical front-end for RAxML. Org Divers Evol. 2012;12:335–337. https://doi.org/10.1007/s13127-011-0056-0

46. Aylor DL, Price EW, Carbone I. SNAP: Combine and Map modules for multilocus population genetic analysis. Bioinformatics. 2006;22:1399–1401. https://doi.org/10.1093/bioinformatics/btl136

47. Price EW, Carbone I. SNAP: Workbench management tool for evolutionary population genetic analysis. Bioinformatics. 2005;21:402–404. https://doi.org/10.1093/bioinformatics/bti003

48. Paradis E. pegas: An R package for population genetics with an integrated-modular approach. Bioinformatics. 2010;26:419–420. https://doi.org/10.1093/bioinformatics/btp696

49. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2016.

50. Zalar P, Gostinčar C, De Hoog GS, Uršič V, Sudhadham M, Gunde-Cimerman N. Redefinition of *Aureobasidium pullulans* and its varieties. Stud Mycol. 2008;61:21–38. https://doi.org/10.3114/sim.2008.61.02

51. Woudenberg JHC, Groenewald JZ, Binder M, Crous PW. *Alternaria* redefined. Stud Mycol. 2013;75:171–212. https://doi.org/10.3114/sim0015

52. Ellison CE, Hall C, Kowbel D, Welch J, Brem RB, Glass NL, et al. Population genomics and local adaptation in wild isolates of a model microbial eukaryote. Proc Natl Acad Sci USA. 2011;108:2831–2836. https://doi.org/10.1073/pnas.1014971108

53. Geml J, Tulloss RE, Laursen GA, Sazanova NA, Taylor DL. Evidence for strong inter- and intracontinental phylogeographic structure in *Amanita muscaria*, a wind-dispersed ectomycorrhizal basidiomycete. Mol Phylogenet Evol. 2008;48:694–701. https://doi.org/10.1016/j.ympev.2008.04.029

54. Elbert W, Taylor PE, Andreae MO, Pöschl U. Contribution of fungi to primary biogenic aerosols in the atmosphere: Wet and dry discharged spores, carbohydrates, and inorganic ions. Atmos Chem Phys Discuss. 2007;7:4569–4588. https://doi.org/10.5194/acp-7-4569-2007

55. Sun Y, Wang Q, Lu X, Okane I, Kakishima M. Endophytic fungal community in stems and leaves of plants from desert areas in China. Mycol Prog. 2012;11:781–790. https://doi.org/10.1007/s11557-011-0790-x

56. Suryanarayanan TS, Wittlinger SK, Faeth SH. Endophytic fungi associated with cacti in Arizona. Mycol Res. 2005;109:635–639. https://doi.org/10.1017/S0953756205002753

57. Rosa LH, Vaz ABM, Caligiorne RB, Campolina S, Rosa CA. Endophytic fungi associated with the Antarctic grass *Deschampsia antarctica* Desv. (Poaceae). Polar Biol. 2009;32:161–167. https://doi.org/10.1007/s00300-008-0515-z

58. Kimura N, Tsuge T. Gene cluster involved in melanin biosynthesis of the filamentous fungus *Alternaria alternata*. J Bacteriol. 1993;175:4427–4435. https://doi.org/10.1128/jb.175.14.4427-4435.1993

59. Pridham JB, Woodhead S. The biosynthesis of melanin in *Alternaria*. Phytochemistry. 1977;16:903–906. https://doi.org/10.1016/S0031-9422(00)86689-3

60. Youngchim S, Morris-Jones R, Hay RJ, Hamilton AJ. Production of melanin by *Aspergillus fumigatus*. J Med Microbiol. 2004;53:175–181. https://doi.org/10.1099/jmm.0.05421-0

61. Durrell LW. Studies of *Aureobasidium pullulans* (de Bary) Arnaud. Mycopathologia et Mycologia Applicata. 1968;35:113–120. https://doi.org/10.1007/BF02049574

62. Gadd GM. Melanin production and differentiation in batch cultures of the polymorphic fungus *Aureobasidium pullulans*. FEMS Microbiol Lett. 1980;9:237–240. https://doi.org/10.1111/j.1574-6968.1980.tb05644.x

63. Gessler NN, Egorova AS, Belozerskaia TA. Melanin pigments of fungi under extreme environmental conditions (Review). Appl Biochem Microbiol. 2014;50:105–113. https://doi.org/10.1134/S0003683814020094

64. Siddique AB, Khokon AM, Unterseher M. What do we learn from cultures in the omics age? High-throughput sequencing and cultivation of leaf-inhabiting endophytes from beech (*Fagus sylvatica* L.) revealed complementary community composition but similar correlations with local habitat conditions. MycoKeys. 2017;20:1–16. https://doi.org/10.3897/mycokeys.20.11265

AUTHORS:
Hélen Prinsloo[1]
Brian Reilly[1]
Willem Myburgh[1]

AFFILIATION:
[1]Department of Nature Conservation, Tshwane University of Technology, Pretoria, South Africa

CORRESPONDENCE TO:
Hélen Prinsloo

EMAIL:
helen@bucandi.co.za

DATES:
Received: 21 Oct. 2019
Revised: 22 June 2020
Accepted: 17 Aug. 2020
Published: 29 Mar. 2021

HOW TO CITE:
Prinsloo H, Reilly B, Myburgh W. Identifying potential protected areas in the Grassland Biome of South Africa. S Afr J Sci. 2021;117(3/4), Art. #7507. https://doi.org/10.17159/sajs.2021/7507

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITORS:
Teresa Coutinho
Salmina Mokgehle

KEYWORDS:
biodiversity, conservation, ecosystem services, focus areas, Heilbron–Petrus Steyn Target Area

FUNDING:
None

Identifying potential protected areas in the Grassland Biome of South Africa

South Africa is considered to be one of the most biologically diverse countries in the world. The conversion, degradation, and fragmentation of natural habitats have caused a loss of biodiversity in many areas. Grasslands have been widely recognised as important for both biodiversity and economic development. Many conservation efforts have in the past been theory driven, without actionable, feasible results. We hypothesised that correct implementation of the available data will indicate where conservation efforts should be focused to move closer to achieving targets for biodiversity conservation in the Grassland Biome in South Africa. We identified an area (near Heilbron and Petrus Steyn in the Free State) that is representative of the biodiversity in the region and is suitable for modern conservation efforts in the 'real world'. This approach provides a practical look at conservation in the modern era and a feasible result for conservation efforts.

Significance:

- An area in the Free State Province was identified that has a high level of biodiversity representative of the Grassland Biome in South Africa. A subjective process was implemented to ensure that the area represents the factors that are considered important for conservation, land use and appropriate location in order to make the area a feasible target for conservation. The result therefore represents not only a theoretical finding that is based on data analysis, but also a practical approach to ensure that the findings can be implemented. This study is an effort to provide a baseline for a more modern approach to conservation, in which current land use, economic value and conservation value are all considered to achieve a sustainable result in terms of the conservation of biodiversity.

Introduction

The need for conservation

Grasslands have been widely recognised as important for both biodiversity and economic development. They occur within a range of climates from semi-arid (400 mm mean annual precipitation) to moist (1000 mm mean annual precipitation).[1]

The Grassland Biome is the second largest biome in South Africa and covers 29% of South Africa's land area (339 237 km^2). It harbours many species of birds, mammals, reptiles and butterflies and it is the preferred habitat for South Africa's national bird, the blue crane (*Anthropoides paradiseus*), the critically endangered blue swallow (*Hirundo atrocaerulea*), oribi (*Ourebia ourebi*) and the vulnerable sungazer lizard (*Smaug giganteus*). Nearly half of South Africa's Red Listed butterfly species (28/60 = 46.7%) occur in the Grassland Biome.[2] The Grassland Biome is considered to have extremely high species diversity, second only to the well-known Fynbos Biome, and includes many rare and threatened species, but is one of the most critically threatened southern African ecosystems.[3]

Grasslands provide essential ecosystem services that are necessary for economic development, but are in turn threatened by some of these developments.[4] The major land uses of the Grassland Biome include urban development (South Africa's largest urban and industrial area of Gauteng is located entirely within the Grassland Biome), coal mining (mainly in Mpumalanga), plantation forestry (mainly in KwaZulu-Natal and Mpumalanga), as well as agriculture, including cultivation and livestock production. If these activities are located in the correct areas and managed appropriately, they offer opportunities to align development and conservation objectives. The intention of biodiversity or grassland conservation should therefore not be to prevent development, but to ensure that it is appropriately located and managed.[5]

Biodiversity conservation planning in South Africa

Since the 2000s, South Africa has attempted an outcome-based, modern approach to biodiversity conservation that has been fine-tuned and measured consistently and from different angles. This approach included the compilation and publication of various documents and plans such as the *National Spatial Biodiversity Assessment – Terrestrial Component*[4], the *National Protected Areas Expansion Strategy*[6], *South Africa's Fourth National Report to the Convention on Biological Diversity*[5] and *South Africa's Fifth Annual Report to the Convention of Biological Diversity*[7].

Threats to biodiversity

The conversion, degradation, and fragmentation of native habitats are widely recognised among ecologists as the principal causes of biotic impoverishment.[8,9] The loss and fragmentation of native habitats caused by agricultural development and conversion of agricultural lands into urban sprawl are generally considered the most serious modern threats to the conservation of biodiversity.[10] Habitat loss, habitat degradation and invasive alien species are considered the greatest threat specifically, but not exclusively, to plant species.[5] Activities such as agriculture, infrastructure, housing and industrial development, amongst others, require land clearance, which causes habitat destruction, fragmentation, disturbance of ecosystems and species loss.[11-14]

Biodiversity targets

The Fifth World Parks Congress in Durban, South Africa, announced in September 2003 that the global network of protected areas then covered 11.5% of the planet's land surface. This surpassed the 10% target proposed a decade earlier, at the Caracas Congress, for 9 of 14 major terrestrial biomes.[15] Rodrigues et al.[15] recognised the challenges with setting such uniform targets based on percentage of area that, at that stage, had become deeply embedded into national and international conservation planning. Although politically expedient, the scientific basis and conservation value of these targets were being questioned.[15]

Modern conservation efforts have started setting conservation targets by incorporating biodiversity priority areas, rather than a 'blind' percentage of available land. When biodiversity priority areas are identified, biodiversity has to be measured, biodiversity goals have to be set and methods for implementing those goals have to be applied.[16]

Even though vast amounts of data are available on the processes needed to achieve biodiversity conservation targets, the knowledge is not translating into action and providing the necessary benefits to biodiversity conservation. This led us to hypothesise that correct implementation of the available data will indicate where conservation efforts should be focused to move closer to achieving targets for biodiversity conservation in the Grassland Biome. This should be done using a 'real-world' approach incorporating reality into the data. To investigate this hypothesis, we formulated three key objectives:

1. Identify focus areas for conservation in the Grassland Biome.
2. Apply filters and assumptions to identify a case study area.
3. Identify the most suitable polygon within the study area to be the target for conservation efforts.

Methods

Over the last 25 years, South Africa has gradually changed its approach to conservation planning, moving away from the ad hoc method and towards the aim of establishing a functional network of biodiversity protection. While the initial focus was on getting the policies and legislation in place, South Africa has also developed strategies and plans that support the policies and legislation for biodiversity management. The *National Spatial Biodiversity Assessment*, completed in 2004, provides a spatial picture of the location of threatened and under-protected ecosystems. This largely informed the National Biodiversity Strategy and Action Plan, which was finalised after comprehensive stakeholder participation, and sets out a comprehensive long-term strategy for the conservation and sustainable use of South Africa's biodiversity. The National Biodiversity Framework distils the thematic and spatial priority actions from the National Biodiversity Strategy and Action Plan and *National Spatial Biodiversity Assessment* for the 5-year period from 2008 to 2013.[5]

The methodology of this study was based on that recommended by Bourgeron et al.[17] It required:

1. Delineating a set of land units from which candidate areas were selected. In this case the Grassland Biome was chosen and all 'land units' within it were considered candidate areas.
2. Determining the suitability of these land units for conservation based on their ecological conditions. This was done by relying on the data that were collected and represented in the various available regional, provincial and municipal biodiversity conservation plans.
3. Selecting land units for inclusion in a conservation network. Many different iterations of selections were run according to the selection criteria described below to reach one primary target area.

Identify focus areas for conservation in the Grassland Biome

The available data sets on biodiversity at national, provincial and municipal scales were acquired from the South African National Biodiversity Institute (SANBI) Biodiversity Geographic Information Systems (BGIS) database or the provincial governments directly in the form of shapefiles. A full list of all the data sets incorporated can be obtained on request.

All the shapefiles were incorporated into ArcGIS 10.1 (hereafter referred to as GIS (Geographic Information Systems)) and a stepwise procedure was followed to run iterations to identify potential focus areas. The steps are described in order below.

The 'Grassland Biome' shapefile was isolated by importing the 'vegm2006 Biomes' shapefile as issued by Mucina and Rutherford[18] and extracting all the polygons that had the 'Grassland Biome' attribute. All the provincial data sets were then clipped (using the Geoprocessing/Clip tool in GIS) to the 'Grassland Biome' shapefile to select those polygons from all the shapefiles that fell within the Grassland Biome. The Grassland Biome outline was then dissolved to provide contiguous areas where applicable (using the Geoprocessing/Dissolve tool in GIS).

The individual provinces used different classification systems to attribute a biodiversity conservation value to the provincial polygons. All the categories used can be provided on request. All the categories from the provincial databases that related to protected areas were extracted per province. This extraction was done by selecting, by attribute, the relevant category in each of the provincial shapefiles and exporting these data to create a 'Protected Area' shapefile for each province. All of these shapefiles were then combined (using the Geoprocessing/Merge tool in GIS) with the 'SANBI Protected Areas' shapefile to create a 'Combined Protected Areas' shapefile. The categories in each province that related to Critical Biodiversity Areas (CBA), regardless of the subcategory (excluding protected areas), were then selected (using 'Select by category' in GIS) and exported to create a CBA shapefile for each province. These shapefiles were then combined (using the Geoprocessing/Merge tool in GIS) to form a CBA shapefile for the Grassland Biome. The categories in each province that related to Irreplaceable Critical Biodiversity Areas, were then selected (using 'Select by category' in GIS) and exported to create a 'CBA–Irreplaceable' shapefile for each province. These shapefiles were then combined (using the Geoprocessing/Merge tool in GIS) to form a 'CBA-Irreplaceable' shapefile for the Grassland Biome. The CBA–Irreplaceable shapefiles were dissolved to create contiguous areas where possible.

Apply filters and assumptions to identify a case study area

At this point a subjective method was implemented after investigating perceived biological and socio-economic factors to narrow down the case study area. The assumptions and subjective criteria used to identify a final case study area are described below.

Mpumalanga is the hub of coal mining in South Africa, causing a considerable threat to the grassland in this province. This can certainly be seen as a reason to prioritise conservation of grasslands in Mpumalanga. The grassland in Mpumalanga is as valuable to biodiversity as any other grassland and should not be discounted. However, the aim of the study was not only to identify areas that are important from a biodiversity aspect, but also to look at the easiest and most efficient way of attaining and managing these areas. A never-ending political war between conservation of biodiversity and the economic and social value of coal mining is considered impractical for achieving optimum results for conservation. The Mpumalanga Province was therefore discarded as a potential case study area.

For KwaZulu-Natal, the assumption was that very successful conservation efforts were already undertaken by Ezemvelo KZN Wildlife, and through the creation of the uKhahlamba Drakensberg Park and the Maloti-Drakensberg Park UNESCO World Heritage site (Maloti-Drakensberg Park), making this particular area already well represented in terms of conservation. Even though the formally protected areas still fall short of conservation targets, the percentage area that is protected is high in relation to other provinces and it was decided to exclude KwaZulu-Natal based on its adequate level of conservation success.

We assumed that the high population density in Gauteng Province would result in large, transformed areas, habitat fragmentation and a more complicated approach to management. Taking into consideration the rapid expansion of formal and informal settlements in the region in addition to the high level of habitat transformation, it was decided that, although important areas for conservation exist in the province, management of these areas will be a big challenge politically and practically. For the purpose of this study, that is combining scientific value with practical implication of conservation, it was decided to discard the Gauteng Province as a potential case study area.

Very small portions of the Grassland Biome occur in the Northern Cape and Western Cape Provinces with 412 873 ha in the Northern Cape and only 14 673 ha in the Western Cape. In the case of the Northern Cape, the sections of grassland occurring in the province are on the edge of the Grassland Biome. In the case of the Western Cape, the grassland sections are completely isolated and removed from the rest of the Grassland Biome. Based on these characteristics, the Northern Cape and Western Cape grassland patches were discarded as potential areas.

There was no obvious assumption for discarding any of the remaining three provinces – North-West, Eastern Cape and the Free State. No excessive development threat or habitat degradation occurred in any of the provinces and none of them proved to be well protected. We therefore based our decision on the level of current protection of the grasslands in the remaining provinces. The Free State Province was singled out as it has the lowest percentage of formally protected grasslands (0.91%). The Free State Province also contains a total of 10 914 757 ha of grassland, a larger area of Grassland Biome than any of the other provinces.

An overview of the protected areas occurring in the Free State showed that, even though only 28.83% of the Free State grasslands occurred in the mountainous region including the Drakensberg and Maloti mountain ranges, 55.65% of the protected grassland areas in the Province occurred in this area. Based on this, the mountainous region was discarded and only the non-mountainous area of the Free State Province was further evaluated as a potential case study area.

Identify the most suitable polygon to be the target

The remaining polygons were sorted according to size to determine which size classes would be feasible for further analysis. The 'real-world' idea here was that larger contiguous areas will be more practical for conservation management and it was decided to use the largest areas available for further analysis. Areas were grouped into the following categories: >20 000 ha, 10 001–20 000 ha, 5001–10 000 ha, 3001–5000 ha, 1001–3000 ha and ≤1000 ha. The 10 largest focus areas were singled out for further investigation in a shapefile called 'target_area_step_1'. A buffer zone of 1 km was established around each of the 10 areas that were selected in the previous step. Where this buffer zone overlapped with another target area that had already been identified, the overlapping areas were joined and a new buffer of 1 km was established around the joined areas. This approach continued until all the target areas and focus areas within 1 km of each other were identified and combined into 'Total Target Areas'.

Each target area was then evaluated according to percentage transformed, other vegetation types occurring, percentage road cover, distance to the edge of the Grassland Biome and percentage urban areas. 'Percentage transformed' was calculated using the data from the 2013–2014 National Land Cover map[19] created by the South African Department of Environmental Affairs. All the transformed categories were selected and calculated for each target area. Because only imagery was available, and not shapefiles, the transformed area was estimated by consulting the map. Other vegetation types were estimated through analysis of data from the 2013–2014 National Land Cover[19] map. This map showed that a large percentage of land contains more shrubland than grassland and these areas were also estimated. The percentage area under road coverage was calculated by adding a layer containing all the South African roads (obtained from the South African Department of Agriculture, Forestry and Fisheries). For primary roads (according to the shapefile), a width of 15 m was used. This was calculated using the SANRAL specifications[20] stipulating the tarred width of a two-lane carriage way at 13.4 m. It was rounded up to 15 m to provide for an additional buffer of disturbance. The width of secondary roads was rounded to 10 m. This was assuming that the road width would be approximately the same as a primary road, but the shoulders (2.5 m on each side according to SANRAL[20]) would not be tarred. To incorporate minimisation of edge effects, the distance to the nearest edge of the Grassland Biome was measured for each target area. Percentage urban areas was then estimated using the imagery data from the 2013–2014 National Land Cover map.[19]

Results

Gross results

Table 1 provides the results after each assumption or filter was applied. After discarding Mpumalanga, 13.95% of the Grassland Biome remained. When the polygons from KwaZulu-Natal Province had been discarded, 12.09% of the Grassland Biome remained. Removal of the Gauteng Province polygons resulted in 11.9% of the Grassland Biome remaining. Further removal of the polygons from the North-West Province resulted in 10.34% of the Grassland Biome remaining. The polygons from the Eastern Cape and Western Cape were then removed, resulting in 3.49% and eventually 3.48% of the Grassland Biome remaining. Application of the final assumption resulted in polygons from the Free State highlands being discarded, leaving a total of 2349 polygons covering an area of 410 176.97 ha or 1.16% of the Grassland Biome. Tables 2 and 3 show the remaining areas sorted according to size, and finally the 10 largest areas that were selected for further analysis.

Table 1: Polygons within the Grassland Biome of South Africa that are considered Level 1 Critical Biodiversity Areas, after application of each filter in order

Area discarded	Remaining number of polygons (dissolved)	Total area (ha) remaining	Largest polygon (ha) remaining
Mpumalanga	8892	4 947 775	1 498 496
KwaZulu-Natal	8289	4 287 850	1 498 496
Gauteng	7417	4 221 198	1 498 496
North-West	4228	3 667 642	1 498 496
Eastern Cape	3412	1 238 212	215 579
Western Cape	3401	1 232 564	215 579

Table 2: Polygons identified as potential target areas for conservation within the Grassland Biome of South Africa grouped into size classes

Size class	Number of polygons	Total area of polygons (ha)
>20 000 ha	1	23 000
10 001–20 000 ha	1	12 188
5001–10 000 ha	8	58 096
3001–5000 ha	10	41 282
1001–3000 ha	41	67 290
≤1000 ha	2288	208 318

Table 3: Sizes of the 10 largest areas identified as potential target areas for conservation within the Grassland Biome of South Africa

Area number	Size (ha)
1	23 000
2	12 188
3	8477
4	8146
5	7838
6	7653
7	7411
8	6674
9	6246
10	5647

Implementation of the buffer zone resulted in the identification of five target areas (TA): one located in the northeastern Free State between Heilbron and Petrus Steyn, three located in the central Free State near the towns of Dealesville, Brandfort and Petrusburg, and one located in the southern Free State near Springfontein. The target areas were named for the towns and therefore were identified as Heilbron–Petrus Steyn TA (HP-TA), Hertzogville–Dealesville–Soutpan TA (HDS-TA), Brandfort TA (B-TA), Petrusburg TA (P-TA) and Springfontein TA (S-TA). The locations of these five target areas are shown in Figure 1. The sizes of the identified target areas are as follows:

HP-TA: 39 836 ha

HDS-TA: 171 847 ha

B-TA: 107 629 ha

P-TA: 12 723 ha

S-TA: 10 642 ha

Nett results – the ultimate target area

Determining the final target area for conservation action

After identifying five possible target areas for conservation, it was necessary to apply a 'real-world' approach to target any one of these areas. Even though shapefiles for land use and land cover were not available, these factors were included as estimates based on the 2013–2014 National Land Cover map.[19] This provided estimates for transformed land and urban land. Transformed land refers to land that is no longer covered with indigenous vegetation as a result of cultivation, afforestation, mining, infrastructure development or any other consumptive land use. Urban areas, although according to definition also constituting transformed land, were given a separate category as they are deemed to have no value for conservation, while other transformed areas can still be conserved, depending on the land use and management practices. Table 4 shows the percentage land occurring in each of these categories for each target area.

Due to the small amount of grassland still occurring in P-TA and S-TA and the domination of Karoo-type low shrubland, these target areas were not further considered as options for conservation for the purpose of this study. The percentage road cover for each area and the distance to the nearest edge of the Grassland Biome were incorporated for the remaining three target areas. A summary of the results of this section is given in Table 5 with each target area receiving a ranking of 1 (worst) to 3 (best) for each of the categories. Based on the results of the scoring in Table 5, the HP-TA was identified as the most suitable target area for conservation in the Grassland Biome.

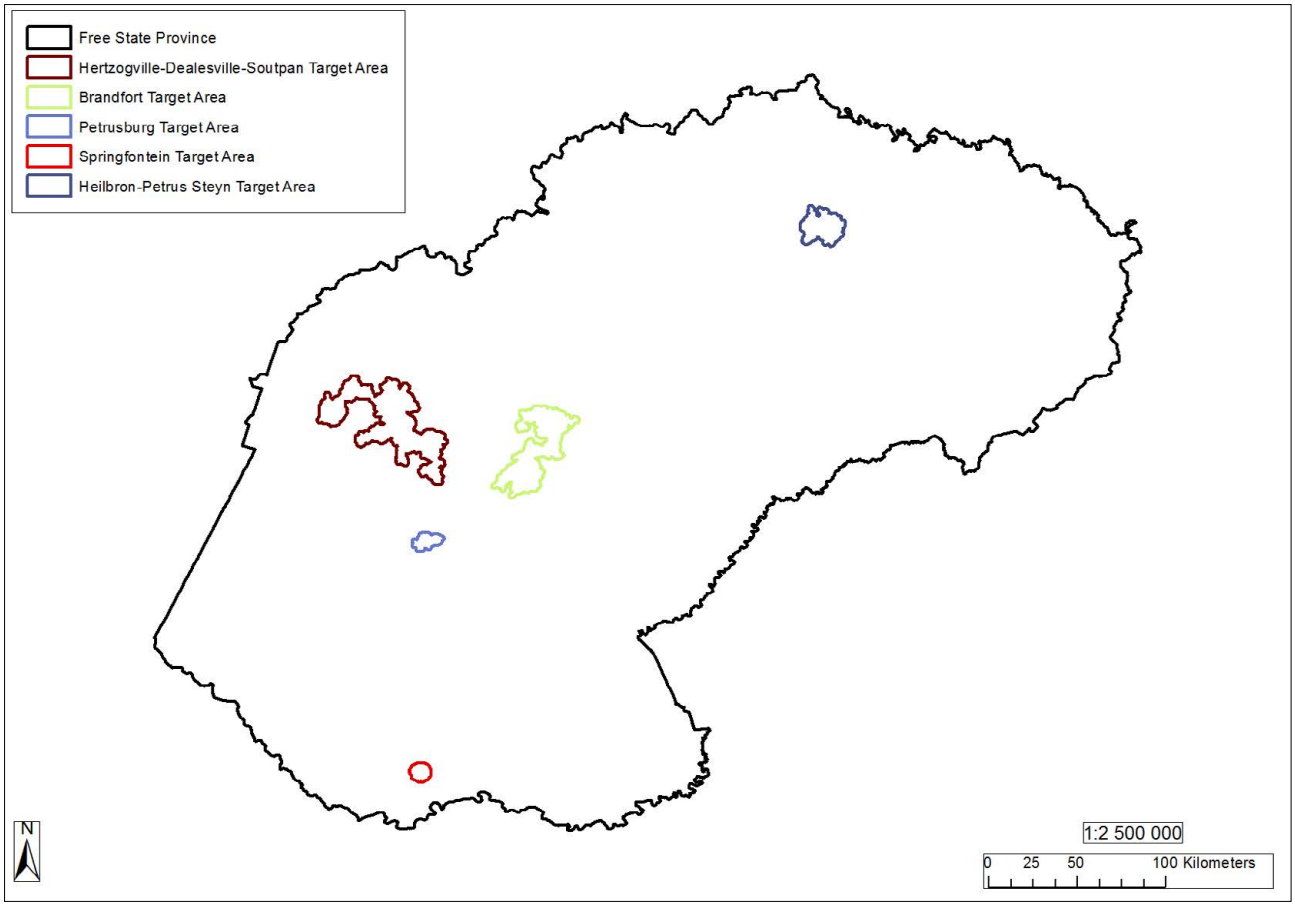


Figure 1: The five identified target areas for conservation action within the Grassland Biome of South Africa.

Table 4: The percentage of land in each target area occurring in the categories identified by the 2013–2014 National Land Cover map[19] created by the South African Department of Environmental Affairs

TA	Untransformed land (%)	Cultivated land (%)	Low shrubland (%)	Urban (%)
HDS-TA	45	20	30	5
B-TA	65	20	10	5
HP-TA	60	40	0	0
P-TA	5	10	80	5
S-TA	10	5	85	0

TA, target area; HDS-TA, Hertzogville–Dealesville–Soutpan target area; B-TA, Brandfort target area; HP-TA, Heilbron–Petrus Steyn target area; P-TA, Petrusburg target area; S-TA, Springfontein target area

Table 5: The scoring system applied to determine the most suitable target area for conservation in the Grassland Biome of South Africa

TA	Size	Urban areas (%)	Untransformed grassland (%)	Road cover (%)	Distance to edge of biome (km)	Total score
HDS-TA	3	1	1	1	1	7
B-TA	2	1	3	2	1	9
HP-TA	1	3	2	3	3	12

TA, target area; HDS-TA, Hertzogville–Dealesville–Soutpan target area; B-TA, Brandfort target area; HP-TA, Heilbron–Petrus Steyn target area

The Heilbron–Petrus Steyn Target Area

Location

The HP-TA is located in the northeastern Free State between the towns of Heilbron (7 km to the north), Frankfort (39 km to the northeast), Tweeling (37 km to the east), Petrus Steyn (10 km to the south) and Edenville (25 km to the west). Its outermost co-ordinates are N: 27º21'14.81"S; 27º57'31.34"E, E: 27º26'20.27"S; 28º09'14.87"E, S: 27º33'56.67"S; 28º04'22.30"E, and W: 27º28'36.13"S; 27º54'13.04"E. The majority of the area lies within the Ngwathe Local Municipality and Fezile Dabi District Municipality, with a small section crossing into the Nketoana Local Municipality and Thabo Mofutsanyane District Municipality.

Climate

The area normally receives about 530 mm of rain per year, with most rainfall occurring during summer. It receives the lowest rainfall (0 mm) in July and the highest (94 mm) in January. The average midday maximum temperatures range from 16.2 °C in June to 27.1 °C in January. The region is the coldest during June, with average night-time temperatures of 0 °C[21]

Land type

The dominant land type for the area is Ea28 with a portion of Dc10 located in the west, Ca7 located in the north and Ca6 located in the south.[22]

Historic vegetation

According to Mucina and Rutherford[18] the area contains mostly Frankfort Highveld Grassland (Vulnerable) with some Central Free State Grassland (Vulnerable) located in the western section, Eastern Free State Clay Grassland (Endangered) in the southeast and two very small portions of Northern Free State Shrubland (Least Threatened).

Discussion and recommendations

This study was not only a scientific exercise, but involved many subjective inputs. This approach constitutes a radical change in relation to previous conservation plans that lack subjective input for streamlining the possibility of implementation. Conservation planning has had very little targeted success, with large gaps between planning and implementing. The vast amount of data available are not being translated into action, rendering the data without purpose in the application of conservation. However, correct implementation of the data should have a sizeable benefit to conservation planning, necessitating the need for a modern approach. The subjective sections of this study were intended to minimise this gap between planning and implementation of conservation action, and to be as realistic as possible in terms of the feasibility of the area right from the onset of designing a conservation area. It is an effort at applying a new and modern approach and the accuracy of this approach can only be tested through concerted conservation efforts in the identified area.

It will now be of cardinal importance to design a modern approach to ownership and management of the identified area. Environmental legislation in South Africa is advancing at a rapid rate and informal protected areas and biodiversity offsets are becoming buzzwords in the conservation sector. For the identified area to be a conservation success, it will be critical that income streams are diverse and, rather than focusing on tourism alone, they should be based on the land use of the area, as well as unique opportunities that can be offered by the specific area. A major benefit of the HP-TA is that it provides habitat for native endemic species such as black wildebeest (*Connochaetes gnou*) and pure-bred blesbok (*Damaliscus pygargus phillipsi*). The area is also well known for its gamebird hunting, which provides an additional income stream that can be utilised.

In addition, conservation in the 21st century has the potential to be incredibly interesting and to redefine the norms of science and historical approaches to conservation. With the availability of funding and ideas through social media and crowd-funded potential, modern ideas for redesigning objectives are limitless. There are well-funded, scientific projects underway for re-establishment of extinct species through genetic engineering and selective breeding. One such project is focusing on reintroducing woolly mammoths (*Mammuthus primigenius*), or at least a mammoth-like hybrid, to the tundra and, through the habitat changes which would be facilitated in the process, drastically contribute to mitigation of climate change.[23] In South Africa, we have the Quagga Project which has already produced at least six individuals that very closely resemble this extinct species in the Western Cape.[24] The original quagga (*Equus quagga quagga*) was a plains species and would theoretically be very suitable for reintroduction in an area such as HP-TA. This sounds completely far-fetched, but in the modern era of connectivity and social media platforms, new and radical ideas are more available than ever before and their impact should be utilised for conservation just as it is utilised in any other industry.

One of the major areas to focus on, and perhaps more immediately achievable than bringing back the quagga from extinction, is the implication of biodiversity offsets. This concept is relatively new and unexplored in South Africa and it is a much-underutilised avenue. If biodiversity offsetting is applied intelligently, it could result in the desirable areas of the Grassland Biome directly benefitting from the destruction of areas where practical implementation of biodiversity conservation would not be feasible. The Grassland Biome is rapidly disappearing due to the pressures of mining, urbanisation and agriculture. If the actions responsible for the destruction of grassland could be utilised, through biodiversity offsetting, to benefit the ultimate conservation of grassland, this approach could be developed and used as a major source of securing land for conservation as part of the HP-TA.

Acknowledgements

We thank SANBI and all the data processors responsible for maintaining accurate and up-to-date information on BGIS – an incredibly helpful and valuable scientific tool. We also thank the provincial governmental departments that were involved in gathering the biodiversity data, including the provincial and/or municipal governments of Limpopo, KwaZulu-Natal, Gauteng, Eastern Cape, North-West, Mpumalanga, Free State, the Western Cape and the following district and local municipalities in the Western Cape: Namaqua District, Drakenstein, Stellenbosch, Knysna, Bitou, Central Karoo District, George, Breede Valley, Bergrivier, Cederberg, Kannaland and Oudtshoorn.

Competing interests

We declare that there are no competing interests.

Authors' contributions

H.P.: Study design, research, writing of manuscript. B.R.: Study design and editing of manuscript. W.M.: Editing of manuscript.

References

1. O'Connor TG. Local extinction in perennial grasslands: A life-history approach. Am Nat. 1991;137(6):753–773. https://doi.org/10.1086/285192
2. Henning GA, Terblanche RF, Ball JB. South African Red Data Book: Butterflies. Pretoria: SANBI; 2009.
3. Rutherford MC, Westfall RH. Biomes of Southern Africa: An objective categorization. 2nd ed. Memoirs of the Botanical Survey of South Africa. 1994;63:1–94.
4. Rouget M, Reyers B, Jonas Z, Desmet P, Driver A, Maze K, et al. National Spatial Biodiversity Assessment 2004: Technical report. Volume 1: Terrestrial component. Pretoria: SANBI; 2004.
5. DEAT, SANBI. South Africa's fourth national report to the Convention on Biological Diversity. Pretoria: DEAT/SANBI; 2009.
6. DEAT, SANBI. National protected areas expansion strategy for South Africa. Draft for Mintech. Pretoria: DEAT/SANBI; 2008.
7. DEA. South Africa's fifth national report to the Convention on Biological Diversity. Pretoria: DEA; 2014.
8. Noss RF, LaRoe ET III, Scott JM. Endangered ecosystems of the United States: A preliminary assessment of loss and degradation. In: Biological report 28. Washington DC: National Biological Service; 1995. https://doi.org/10.2307/4089098
9. Carroll CR, Meffe GK. Management to meet conservation goals: General principles. In: Carroll CR, Meffe GK. Principles of conservation biology. 2nd ed. Sunderland, MA: Sinauer Associates; 1997. p. 347–383. https://doi.org/10.1086/509465
10. Main MB, Roka FM, Noss RF. Evaluating costs of conservation. Conserv Biol. 1999;13(6):1262–1272. https://doi.org/10.1046/j.1523-1739.1999.98006.x
11. Gibbons P, Lindenmayer DB. Offsets for land clearing: No net loss or the tail wagging the dog? Ecol Manag Restor. 2007;8(1):26–31. https://doi.org/10.1111/j.1442-8903.2007.00328.x
12. Maron M, Hobbs RJ, Moilanen A, Matthews JW, Christie K, Gardner TA, et al. Faustian bargains? Restoration realities in the context of biodiversity offset policies. Biol Conserv. 2012;155:141–148. https://doi.org/10.1016/j.biocon.2012.06.003
13. Secretariat of the Convention on Biological Diversity. Global biodiversity outlook 4. Montréal: UNEP; 2014.
14. WWF. Living planet report 2014: Species and spaces, people and places. Gland: WWF; 2014.
15. Rodrigues AS, Andelman SJ, Bakarr MI, Boitani L, Brooks TM, Cowling RM, et al. Effectiveness of the global protected area network in representing species diversity. Nature. 2004;428(6983):640–643. https://doi.org/10.1038/nature02422
16. Margules CR, Pressey RL, Williams PH. Representing biodiversity: Data and procedures for identifying priority areas for conservation. J Biosciences. 2002;27(4):309–326. https://doi.org/10.1007/BF02704962
17. Bourgeron PS, Humphries HC, Reynolds KM. Conducting large-scale conservation evaluation and conservation area selection using a knowledge-based system. Paper presented at: The 4th International Conference on Integrating GIS and Environmental Modelling (GIS/EM4); 2000 September 02–08; Alberta, Canada.
18. Mucina L, Rutherford MC. The vegetation of South Africa, Lesotho and Swaziland. In: Strelitzia19. Pretoria: SANBI; 2006.
19. DEA. 2013-14 National land-cover – 72 classes. Pretoria: DEA; 2015.
20. The South African National Roads Agency Ltd. (SANRAL). Design guidelines for single carriageway national roads. Pretoria: SANRAL; 2009.
21. Cedar Lake Ventures, Inc. Average weather in Heilbron [webpage on the Internet]. No date [cited 2013 May 06]. Available from: https://weatherspark.com/y/94192/Average-Weather-in-Heilbron-South-Africa-Year-Round
22. Land Type Survey Staff. Land types of South Africa: Digital map (1:250 000 scale) and soil inventory datasets. Pretoria: ARC-Institute for Soil, Climate and Water; 1972–2006.
23. Revive and Restore. The woolly mammoth revival [webpage on the Internet]. No date [cited 2016 Aug 23]. Available from: https://reviverestore.org/projects/woolly-mammoth/
24. The Quagga Project. The quagga revival South Africa [homepage on the Internet]. c2008 [cited 2016 Aug 23]. Available from: https://quaggaproject.org/

AUTHORS:
Mary-Jane M. Bopape[1]
Ezekiel Sebego[1]
Thando Ndarana[2]
Bathobile Maseko[1]
Masindi Netshilema[1]
Morne Gijben[1]
Stephanie Landman[1]
Elelwani Phaduli[1]
Gift Rambuwani[1]
Louis van Hemert[1]
Musa Mkhwanazi[1]

AFFILIATIONS:
[1]South African Weather Service, Pretoria, South Africa
[2]Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa

CORRESPONDENCE TO:
Mary-Jane Bopape

EMAIL:
mary-jane.bopape@weathersa.co.za

DATES:
Received: 30 Jan. 2020
Revised: 09 June 2020
Accepted: 15 Oct. 2020
Published: 29 Mar. 2021

HOW TO CITE:
Bopape MM, Sebego E, Ndarana T, Maseko B, Netshilema M, Gijben M, et al. Evaluating South African Weather Service information on Idai tropical cyclone and KwaZulu-Natal flood events. S Afr J Sci. 2021;117(3/4), Art. #7911. https://doi.org/10.17159/sajs.2021/7911

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:
weather extremes, early warnings, floods, Idai, weather forecasting

FUNDING:
Climate Research for Development (CR4D) Postdoctoral Fellowship (CR4D-19-11) of the African Academy of Sciences (AAS) in partnership with the UK Department for International Development (DfID) Weather and Climate Information Services for Africa (WISER) programme and the African Climate Policy Centre (ACPC) of the United Nations Economic Commission for Africa (UNECA)

Evaluating South African Weather Service information on Idai tropical cyclone and KwaZulu-Natal flood events

Severe weather events associated with strong winds and flooding can cause fatalities, injuries and damage to property. Detailed and accurate weather forecasts that are issued and communicated timeously, and actioned upon, can reduce the impact of these events. The responsibility to provide such forecasts usually lies with government departments or state-owned entities; in South Africa that responsibility lies with the South African Weather Service (SAWS). SAWS is also a regional specialised meteorological centre and therefore provides weather information to meteorological services within the Southern African Development Community (SADC). We evaluated SAWS weather information using near real-time observations and models on the nowcasting to short-range forecasting timescales during two extreme events. These are the Idai tropical cyclone in March 2019 which impacted Mozambique, Zimbabwe and Malawi resulting in over 1000 deaths, and the floods over the KwaZulu-Natal (KZN) province in April 2019 that caused over 70 deaths. Our results show that weather models gave an indication of these systems in advance, with warnings issued at least 2 days in advance in the case of Idai and 1 day in advance for the KZN floods. Nowcasting systems were also in place for detailed warnings to be provided as events progressed. Shortcomings in model simulations were shown, in particular on locating the KZN flood event properly and over/underestimation of the event. The impacts experienced during the two events indicate that more needs to be done to increase weather awareness, and build disaster risk management systems, including disaster preparedness and risk reduction.

Significance:

This paper is relevant for all South Africans and the SADC region at large because it provides information on:

- the weather forecasting processes followed at the South African Weather Service,
- available early warning products in South Africa and for the SADC region made possible through the public purse,
- the performance of nowcasting and modelling systems in the case of predicting two extreme weather events that had adverse impacts on southern African society, and
- the dissemination of warnings of future extreme weather events.

Introduction

Weather affects our daily lives from the manner in which we dress, whether we host an event outdoors or indoors, the traffic we experience and our electricity consumption.[1-3] Impacts of weather can also be severe to a point of causing damage to property, injuries and even death. According to the Emergency Events database (EM-DAT), of the total number of people affected by weather-related disasters during the period 1995 to 2015, 56% were affected by floods, 26% by droughts, 16% by storms and 2% by extreme temperatures.[4] When adhered to, timeous, accurate and detailed weather forecasts can reduce impacts associated with adverse weather events.[5] The task to issue weather forecasts and warn the public is usually given to meteorological services, which are mostly state owned and follow the World Meteorological Organization (WMO) guidelines. Weather forecasting requires large investments in observations, as well as computational and human resources primarily financed by the public purse.[6]

The South African Weather Service (SAWS) is one such organisation that has a mandate to provide weather and climate information, products, services and solutions that contribute to the safety of life and property in the air, land and sea over South Africa and adjacent oceans. The organisation is a public entity in terms of the *SAWS Act (No. 8 of 2001)*, as amended in 2013. In addition, SAWS is a recognised WMO regional specialised meteorological centre (RSMC) and therefore provides weather products to meteorological services within the Southern African Development Community (SADC) region. Extreme events that are experienced over the SADC region are as a result of tropical, subtropical as well as mid-latitude systems.[7] Analyses of 30 years (1974 to 2003) of data on natural disasters in Africa, show that droughts and floods are the most frequently occurring natural hazards in southern Africa.[8] The South African Department of Environmental Affairs' Long Term Adaptation Strategy indicated that the most common weather-related disasters in South Africa are floods, droughts, fires and large storms, as reported in the period 1900 to 2014.[9]

A variety of synoptic situations may be conducive to convective storm formation over South Africa. Tropical-temperate troughs[10] associated with northwest–southeast aligned cloud bands were found to be responsible for a high proportion of heavy rainfall days – 30 of 52 extreme rainfall events analysed during the period 1979–1999.[11] Cut-off lows, which are closed low-pressure systems in the upper air, were found to be associated with some

of the extreme rainfall events. During spring and autumn, cut-off lows bring thunderstorms to the southern parts of the subcontinent.[12] Land-falling tropical cyclones can also cause heavy rainfall which can result in flooding. Tropical cyclones contribute to more than 50% of multi-day heavy rainfall events that occur in the Limpopo river basin.[13]

According to EM-DAT data, floods have resulted in the highest economic cost of damages, while droughts have affected a larger proportion of the country's population.[4] The total damage due to the four most common weather-related disasters is approximately USD3.5 billion, with floods amounting to approximately USD1.3 billion (in the period 1900–2014). It may be noted that the events reported by EM-DAT do not account for all disasters because of the strict criteria followed on the determination of which events are listed. Improvements in weather forecasting and greater awareness of risks have reduced impacts associated with adverse weather events in the USA since the 1990s.[5,14]

In March 2019, tropical Cyclone Idai caused floods over Mozambique, Malawi and Zimbabwe, resulting in more than 1000 deaths, a cholera outbreak with over 4000 confirmed cases and an estimated damage of at least USD2 billion.[15,16] In April 2019, the coastal province in the east of South Africa, KwaZulu-Natal (KZN), experienced floods which led to 71 deaths, more than 1400 people displaced and an estimated damage of about USD71 million.[17,18] These impacts were felt in the very recent past despite the availability of weather information disseminated through mobile apps, television and radio stations, and social media. The available weather information is produced using state-of-the-art infrastructure and observations. SAWS uses numerical weather prediction models, satellite, radar and lightning detection data, and ground observations (discussed in the next section) to generate its forecasts and warnings. As an RSMC, SAWS also provides services to forecasters in the SADC region for the production of local warnings and alerts.

A successful weather-related early warning system requires (1) a good observation network used as input to the forecasting process, (2) state-of-the-art numerical models that can provide reliable forecasts with a long enough lead time, (3) skilled forecasters who understand how to interpret observations and model output, (4) relevant dissemination channels, and (5) response. The WMO recommends a move from providing information on what the weather will be to providing impact-based forecasting, which requires information on vulnerability and hazard scenarios.[19-21] Challenges in any of the areas will limit the value and effectiveness of the early warning systems.

In this study, we evaluated the weather information (what the weather will be) issued by SAWS for both the Idai tropical cyclone and the KZN floods. Our main objectives were to determine (1) if models were able to capture the location and intensity of the events; (2) how far in advance models were able to provide guidance on the events; (3) if forecasters issued reliable warnings and (4) if information needed to update the warnings as informed by near real-time observations was available. An understanding of the performance of aspects of the early warning system value chain will assist those involved to understand what works well and what needs attention.

Data and methods

SAWS provides weather and climate information across a range of timescales, from nowcasting which is defined as 0–2 h, short-range forecasting (up to 3 days), medium range (up to 2 weeks) to seasonal and multi-decadal timescales. The focus of this study was on the nowcasting and short-range early warning systems.

Nowcasting systems

A weather radar (Radio Detection and Ranging) provides a three-dimensional view of the atmosphere in near real time.[22] SAWS operates 10 S-band Doppler radars, with one dual-polarised (placed in Bethlehem), and two mobile X-band radars. The radar information is received every 6 min and extrapolations of storm position up to 2 h ahead are made using the TITAN (Thunderstorm Identification, Tracking, Analysis and Nowcasting) software.[23] Quantitative precipitation estimates are also produced from the radar data for South Africa for stratiform and convective rainfall.[24]

SAWS receives satellite data from the Meteosat Second Generation (MSG) which provides 12 channels every 15 min across the whole of Africa. SAWS uses the nowcasting and very short-range forecasting Satellite Application Facility developed in Spain, Europe, which was first installed at SAWS in 2014 to produce the Rapidly Developing Thunderstorms (RDT) product. The RDT product uses data from the geostationary MSG satellite and Numerical Weather Prediction (NWP) models (discussed later) to provide information on significant convective systems. The RDT product distinguishes different phases of the thunderstorm, namely, growing, mature and decaying.[25] SAWS also uses satellite-based quantitative precipitation estimations namely the Hydroestimator and the Convective Rainfall Rate which is a product of the nowcasting Satellite Application Facility. These products are combined with the stratiform rainfall field from the NWP model running at SAWS to produce a comprehensive rainfall field.

Rainfall estimates from radar as well as from satellite are used as input towards the South African Flash Flood Guidance System (SAFFGS).[26] The satellite estimates are used when radar data are not available over South Africa. The Southern African Region Flash Flood Guidance System (SARFFGS) is available for SADC and uses only rainfall estimates from satellite data. For forecasting up to 6 h ahead, both the SAFFGS and SARFFGS employ NWP data described below.

Numerical weather prediction

To be able to issue forecasts or predictions beyond the nowcasting timescale, SAWS uses numerical models that are either run at SAWS or by other international meteorological organisations. The main operational NWP model that is utilised by SAWS is the Unified Model (UM) of the United Kingdom Met Office.[27,28] The model is run across most of SADC with a grid spacing of 4.4 km (UM_SA4), producing up to 72-h forecasts, and across the whole of South Africa with a grid spacing of 1.5 km (UM_SA1p5), up to 36 h ahead. The UM forecasts are updated four times a day at 00h00 UTC, 06h00 UTC, 12h00 UTC and 18h00 UTC. The simulations are made on the SAWS CRAY XC30 high-performance computing system. The UM_SA4 is used as input to the SARFFGS and SAFFGS for purposes of forecasting possible flash floods over river basins. Beyond 3 days up to 10 days, SAWS purchases data from the European Centre for Medium-Range Weather Forecasts (ECMWF)[29], where a deterministic forecast with a grid spacing of 16 km is provided for selected variables for South Africa.

Other observations

The description of the event is based on the ERA5 data set, which is the fifth major global reanalysis produced by ECMWF, and it covers the period 1950 to the present.[30] ERA5 was developed through the Copernicus Climate Change Service and has a grid spacing of approximately 31 km globally. The re-analyses were produced using 4D-Var data assimilation in the ECMWF's Integrated Forecast System. Atmospheric data are available on the surface or single level as well as on pressure, potential temperature and potential vorticity levels.

The rainfall simulations were also compared with the Integrated Multi-satellitE Retrievals for GPM (Global Precipitation Measurement) (IMERG) which is the unified US algorithm that provides the multi-satellite precipitation product for the US GPM team.[31] The precipitation estimates from the various precipitation-relevant satellite passive microwave sensors comprising the GPM constellation were computed using the 2017 version of the Goddard Profiling Algorithm, then gridded, intercalibrated to the GPM Combined Ku Radar-Radiometer Algorithm product, and merged into half-hourly 0.1°x0.1° (roughly 10x10 km) fields.

The rainfall simulations were also compared with the Tropical Applications of Meteorology using SATellite data and ground-based observations (TAMSAT). TAMSAT produces daily rainfall estimates for all of Africa at a 4-km grid length.[32] The TAMSAT data are available from 1983 to the present. Data from SAWS automatic weather stations and

automatic rainfall stations over KZN were also used to study the KZN floods.

Objective verification

An objective verification analysis was performed on the UM simulations for the KZN flood case based on three contingency table statistics. The UM simulations are coarse-grained from 4.4 km and 1.5 km to the GPM grid length of 0.1°, with GPM used as the observation. The three measures that were used are the False Alarm Ration (FAR)(Equation1), the Probability of Detection (POD) (Equation 2) and Critical Success Index (CSI) (Equation 3).[33] FAR gives the fraction of the predicted events that did not occur; a perfect score for FAR is 0, while a value of 1 shows no skill. POD is a measure of the fraction or percentage of the predicted events that occurred, while CSI provides the fraction of observed and forecasted events that were correctly predicted. A perfect score for POD and CSI is given by a value of 1, while 0 means there is no skill.

$$\text{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} \quad \text{(Equation 1)}$$

$$\text{POD} = \frac{\text{hits}}{\text{hits} + \text{misses}} \quad \text{(Equation 2)}$$

$$\text{CSI} = \frac{\text{hits}}{\text{hits} + \text{misses} + \text{false alarms}} \quad \text{(Equation 3)}$$

Forecasts and dissemination

Dissemination and effective communication of SAWS services is critical for the safety of lives and property. Weather forecasts that are communicated to the media are issued by the Disaster Risk Reduction Centre which was previously known as the National Forecasting Centre. The Disaster Risk Reduction Centre is based at the SAWS head office in Pretoria and is responsible for national guidance on potentially hazardous weather. In order to alert or warn the public, the DRR centre cooperates with the National Disaster Management Centre and the flood forecasting service of the South African Department of Water and Sanitation.[9] Regional forecasting offices at the major airports are responsible for the detailed forecasts and warnings within their regions and they also liaise with the Provincial and Municipal Disaster Management Centres. SAWS also issues warnings through Facebook (@WeatherServic) and Twitter (@SAWeatherServic) accounts, and a WeatherSMART mobile app. As an RSMC, SAWS makes its forecast products available through the password controlled RSMC Pretoria website (http://rsmc.weathersa.co.za)[34] for meteorological services within SADC.

Results

Idai tropical cyclone

Event description

Tropical cyclone Idai is reported to have impacted Mozambique, Madagascar, Malawi and Zimbabwe.[35] The cyclone started developing on 4 March 2019, initially as a tropical depression, which made landfall on the same day over the north of Quelimane as shown in Figure 1a. The system caused heavy rainfall over southern Malawi and some provinces of Mozambique for a period of 4 days (Figure 1b). The system then moved back over the Mozambique channel on 9 March (Figure 1c). Figure 1d shows a now intensified and mature tropical cyclone on 12 March over the Mozambique channel. Warm waters provide the needed energy to strengthen and sustain tropical cyclones because of increased latent heat fluxes.[36] The system moved towards the west and made landfall on 14 March in the northern vicinity of Beira as an intense tropical cyclone.

The ERA5 reanalysis shows that the minimum pressure over the area between 14° and 24° south and 32° to 44° east, was recorded on 14 March. The maximum wind speed and rainfall over the same area were also recorded on the same day. The highest hourly rainfall total (21 mm) occurred at 20h00 UTC according to ERA5 reanalysis. Figure 2 shows the observed/estimated rainfall from ERA5, GPM, Convective Rainfall Rate and TAMSAT on 15 March. Although the amount and rainfall pattern associated with each of the data sets are different, all the observations indicate that heavy rainfall was received over Mozambique and the eastern parts of Zimbabwe on 15 March. All the observations except TAMSAT show rainfall amounts exceeding 200 mm over at least one area in Mozambique. A study that compared TAMSAT with other rainfall estimations showed that TAMSAT generally underestimated rainfall over Southern Africa.[37] Convective Rainfall Rate, which is a combination of the estimated Convective Rainfall Rate from MSG and stratiform rainfall from the UM, shows rainfall peaks over the outskirts of the storm. The peaks are similar to those found in the UM simulations.

Idai moved inland and caused heavy rainfall over central Mozambique, and the eastern and central parts of Zimbabwe, while weakening at the same time. Tropical cyclones generally weaken and dissipate over land due to different surface properties which include topography, and the surface heat, moisture and momentum fluxes compared to those over the ocean.[38] The rainfall associated with Idai and its remnants continued over the next few days and its remnants finally dissipated on 21 March. The World Bank estimated the direct economic losses from Cyclone Idai amounted to about USD622 million, including damage to infrastructure, properties, crops and livestock.[39] The World Bank approved USD72 million to help with recovery needs of those affected in Zimbabwe; USD70 million in Malawi[40], and USD130 million for Mozambique for both Idai and Kenneth[41].

Numerical weather prediction – Idai

According to WMO, the reversing of the system back towards the Mozambican coastline was suggested by the European ensemble forecast from 6 March 2019.[35] The track forecast uncertainty is reported to have reduced by 12 March, indicating that Idai would hit Beira. The WMO further reports that forecasters in the Mozambique National Meteorology Institute (INAM) have access to the NWP products through global centres such as the ECMWF, the UK Met Office, and the RSMC La Reunion. INAM also runs an inhouse Weather Research and Forecasting model with a grid spacing of 14 km. As already discussed, Figure 2 shows 24-h rainfall totals on the 15 March from different observations, representing a full day of torrential rainfall following landfall the night before. There is general agreement that a large amount of rainfall was received between 32° and 35° east and 18° to 21° south, with some differences across the observations.

SAWS does not have a mandate to issue warnings directly to the public of neighbouring countries; however, as an RSMC, SAWS made NWP products from the UM produced locally available through the RSMC Pretoria website. Figure 3a–d shows UM simulations with different lead times ranging from 48 h to 12 h. A lead time represents the amount of time between the time at which the forecast is made to the forecast period. All the simulations agree on the area where Idai made landfall, and the position did not change as the lead time decreased. Figure 3e shows UM simulations of hourly area-averaged rainfall in the rectangle shown over Mozambique from 12h00 UTC on 14 March to 23h00 UTC on 15 March, with ERA5 (red) and GPM (black) data which represent observations. The rainfall curve from ERA5 follows a similar pattern to all the UM simulations with a rainfall maximum around 06h00 on 15 March. Generally, an improvement in skill is expected as the lead time reduces[42]; however, in this case, the UM simulations do not seem to change with the lead time. The consistency in the forecast with different lead times provides confidence to the forecasters and it also means the forecast statements did not need to be altered significantly when forecasts were updated. The first peak in GPM rainfall estimates occurs around 06h00, similar to the simulations; however, the rainfall is overestimated in the simulations, except in the 13 March 00Z which has the longest lead time of all the simulations. GPM shows a second peak at 11h00 that is not captured by the majority of the simulations except the 13 March 00Z forecast which has a later peak. The amount of rainfall as estimated by GPM increased significantly from around 17h00 on 15 March; however, this is not shown in the ERA5 reanalyses nor in any of the simulations.

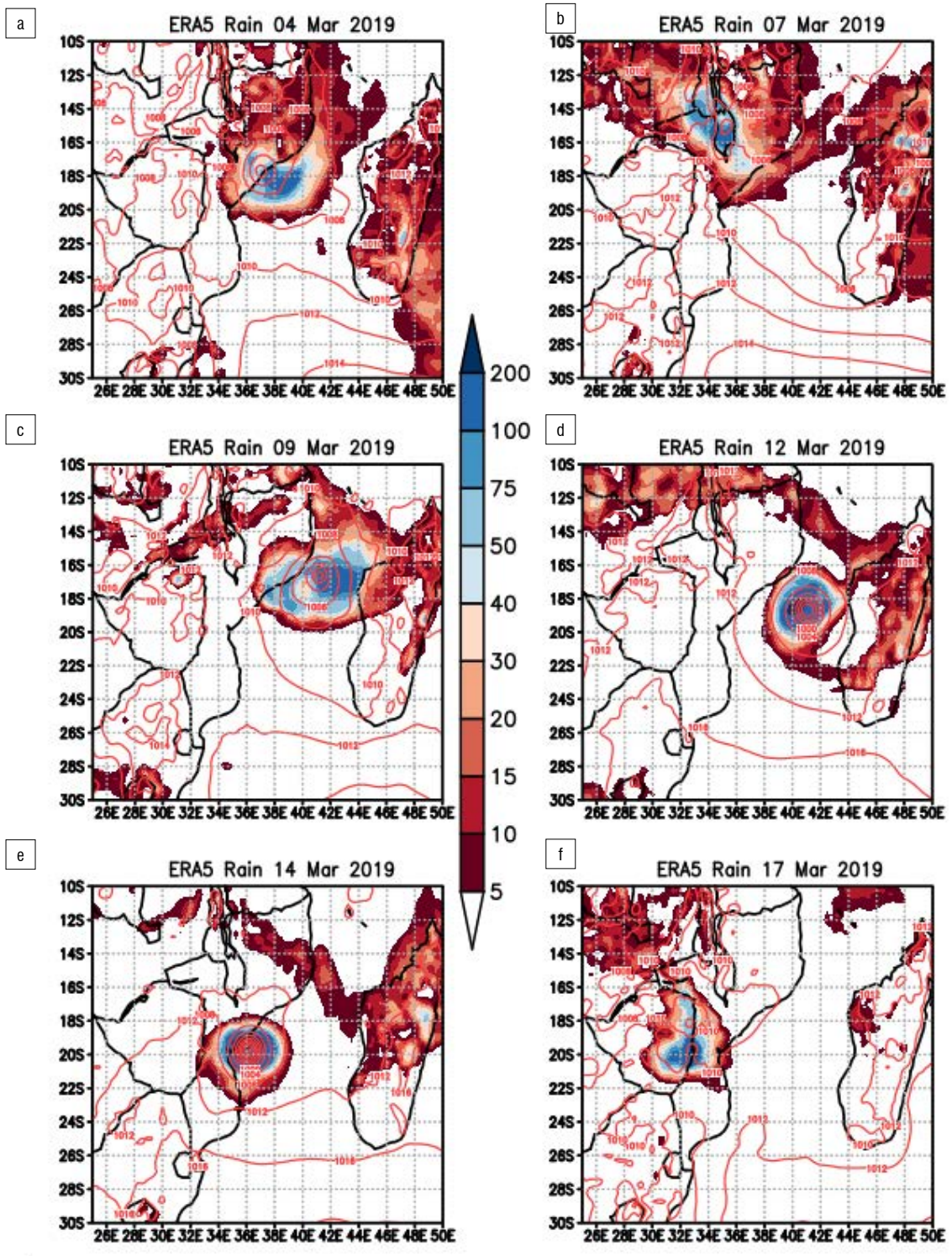


Figure 1: ERA5 reanalysis of 24-h rainfall (mm) overlaid with sea level pressure (hPa) on (a) 4 March 2019, (b) 07 March 2019, (c) 09 March 2019, (d) 12 March 2019, (e) 14 March 2019 and (f) 17 March 2019.

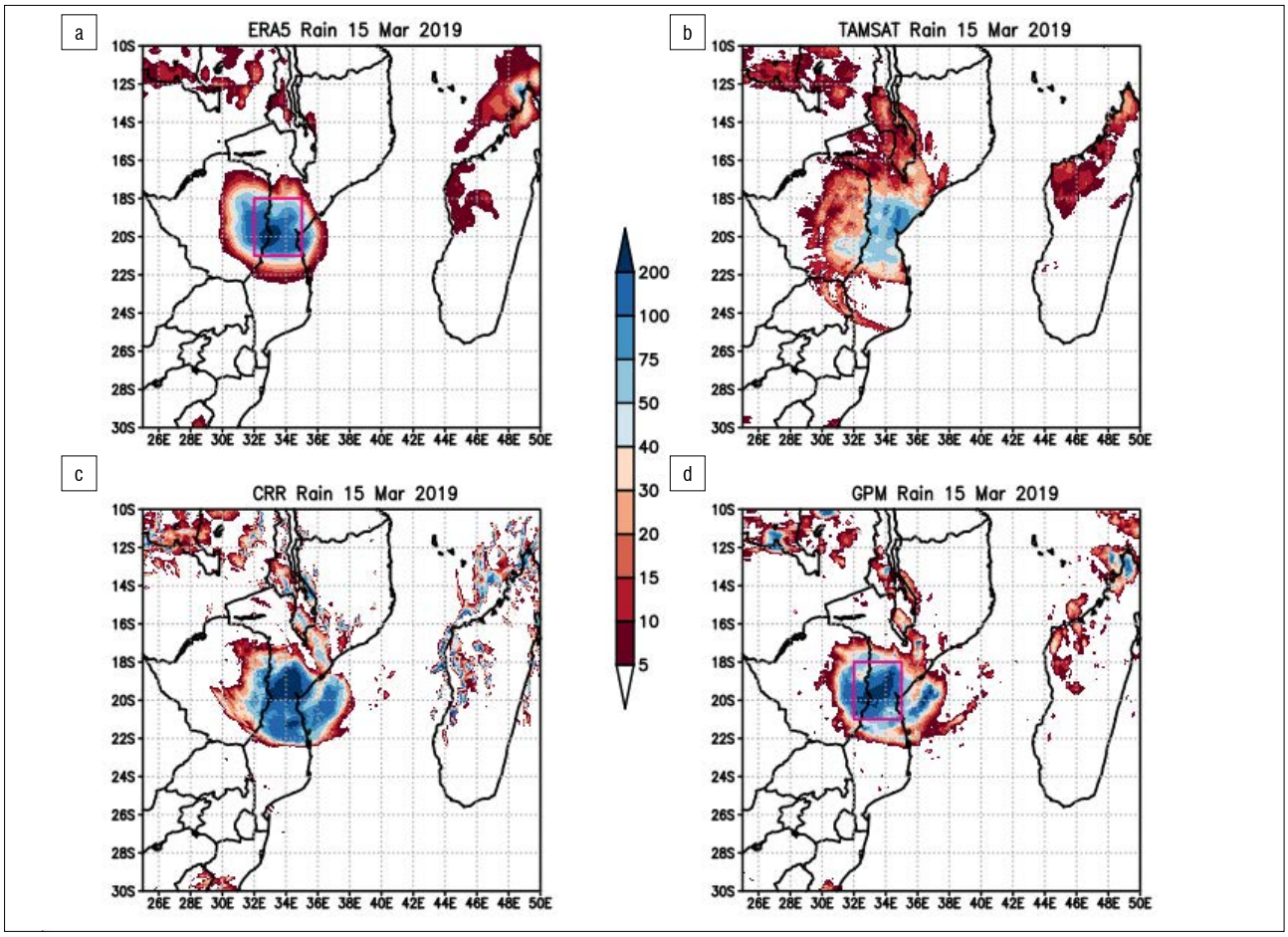


Figure 2: Total rainfall estimation (mm) from (a) ERA5 reanalysis, (b) TAMSAT, (c) Convective Rainfall Rate (CRR) combined with stratiform rainfall from UM simulations with a 4.4-km grid spacing, and (d) GPM on 15 March 2019.

Nowcasting to very short-range forecasting

Near real-time observations are essential for forecasters to be able to warn the public in real time as the event is occurring. Radar is considered the most important tool for nowcasting purposes[22]; however, there is currently no working radar in Mozambique. According to WMO, INAM has two radars which were installed in 2004 at Xai Xai and Beira[35]; however, these have not been operational for many years. Satellite data are also used for nowcasting purposes; INAM has access to satellite data from EUMETSAT and various data and products broadcast by EUMETSAT through EUMETCast.

INAM and other meteorological services within southern Africa have access to SARFFGS information, which relies on the Hydroestimator produced at SAWS (i.e. satellite data rainfall estimation). The SARFFGS uses the quantitative rainfall estimates from satellites to determine the amount of rain needed over a basin to cause flooding.[26] The SARFFGS uses the UM rainfall data to provide very short-range forecasts for the river basins. Figure 4 shows an example of output from SARFFGS produced on 15 March 06h00, for 1-h, 3-h and 6-h lead times for Mozambican basins. The availability of the figure shows that the system was available to forecasters in the SADC region to warn officials with specific information on basins at risk of flooding in a few hours. The SARFFGS shows that smaller amounts of rainfall were needed to flood Beira and sounding areas. The amount needed to flood most of Mozambique reduces by forecast hour 6. The information provided by the SARFFGS can help resource managers make decisions quicker, and prioritise basins with a high probability of flooding sooner. The SARFFGS was available during the Idai tropical cyclone, but according to the WMO report, it was not used.[35]

Dissemination

SAWS issued a media release on the tropical cyclone on 12 March 2019, where the location of the tropical cyclone within the Mozambique channel was described together with the expected impact. The media release indicated that the tropical cyclone would move in a southwesterly direction and make landfall over the Mozambican coast. The media release referred to the La Reunion tropical cyclone RSMC forecast, which indicated that the cyclone would make landfall around Beira in the following 48–60 h (around Thursday night – 14 March 2019). The information provided included the expected rainfall, flooding and damaging winds. The locations that were likely to be impacted were mentioned, namely the Sofala and Manica provinces in Mozambique, extreme eastern parts of Zimbabwe and the southern parts of Malawi. SAWS continued to keep the public updated on the event using its Facebook and Twitter accounts, as well as through interviews requested by the media.

SAWS provides information through the WMO Severe Weather Forecasting Demonstration Project, and SARFFGS to meteorological services in the region. Figure 5 shows information provided by RSMC Pretoria for 15 March with different lead times. The green line shows areas where heavy rainfall/ flooding was expected. The 4-day and 3-day lead times indicated that heavy rainfall exceeding 50 mm was expected over the areas where the tropical cyclone made landfall. The forecast issued on 14 March increased this amount from 50 mm to 100 mm. The area covered by the green line started out smaller with a 4-day lead time, and increased to also cover the eastern parts of Zimbabwe when the forecast was issued the following day. This indicates that there was enough information provided days in advance of Idai making landfall. This information was available to all meteorological services in the region.

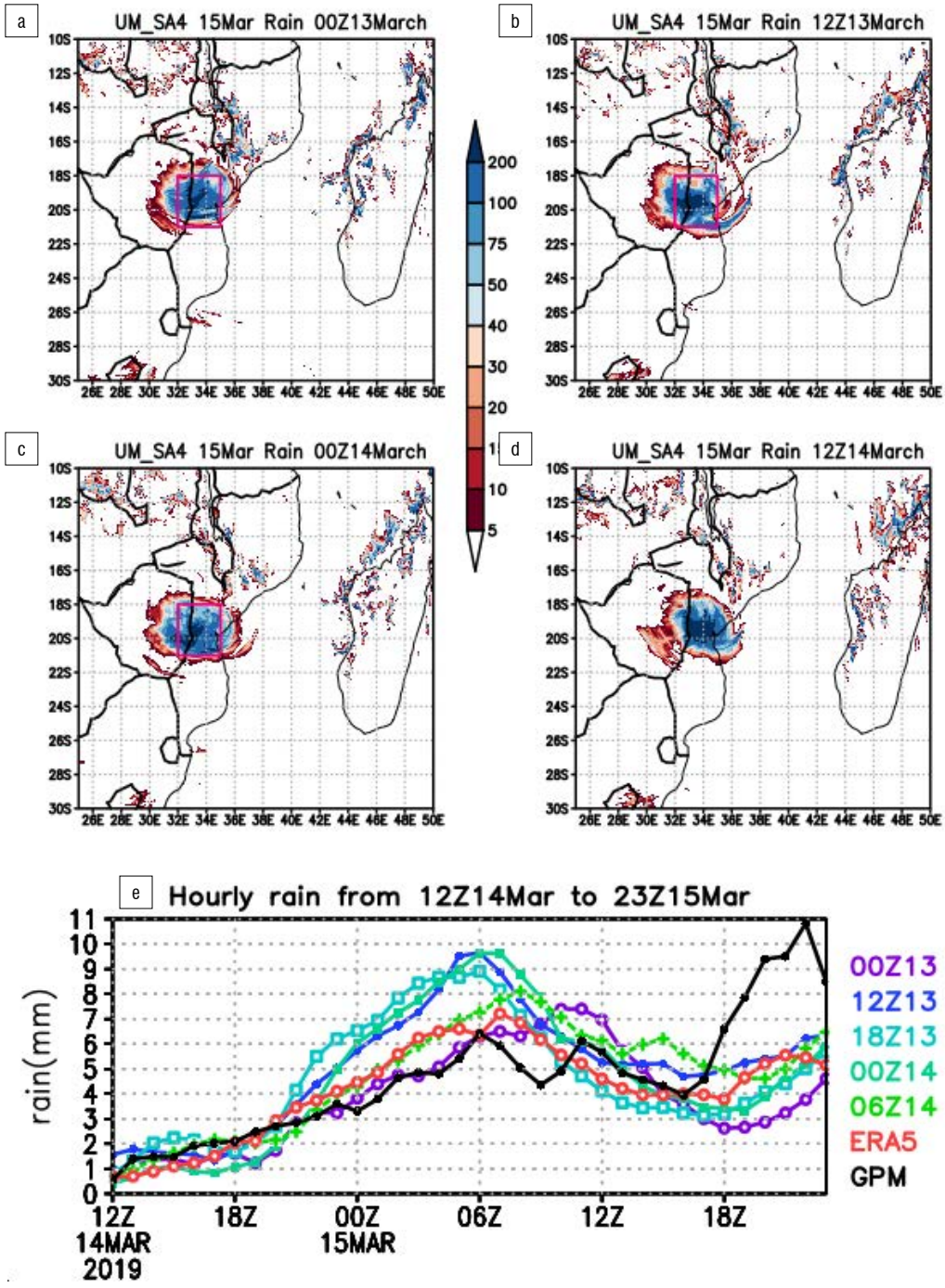


Figure 3: 24-h rainfall on 15 March as simulated by the UM_SA4 with (a) 48-h, (b) 36-h, (c) 24-h and (d) 12-h lead times as well as (e) hourly rainfall with 6-, 12-, 18-, 24- and 36-h lead times, and ERA5 reanalysis and GPM satellite rainfall estimation for 14–15 March 2019.

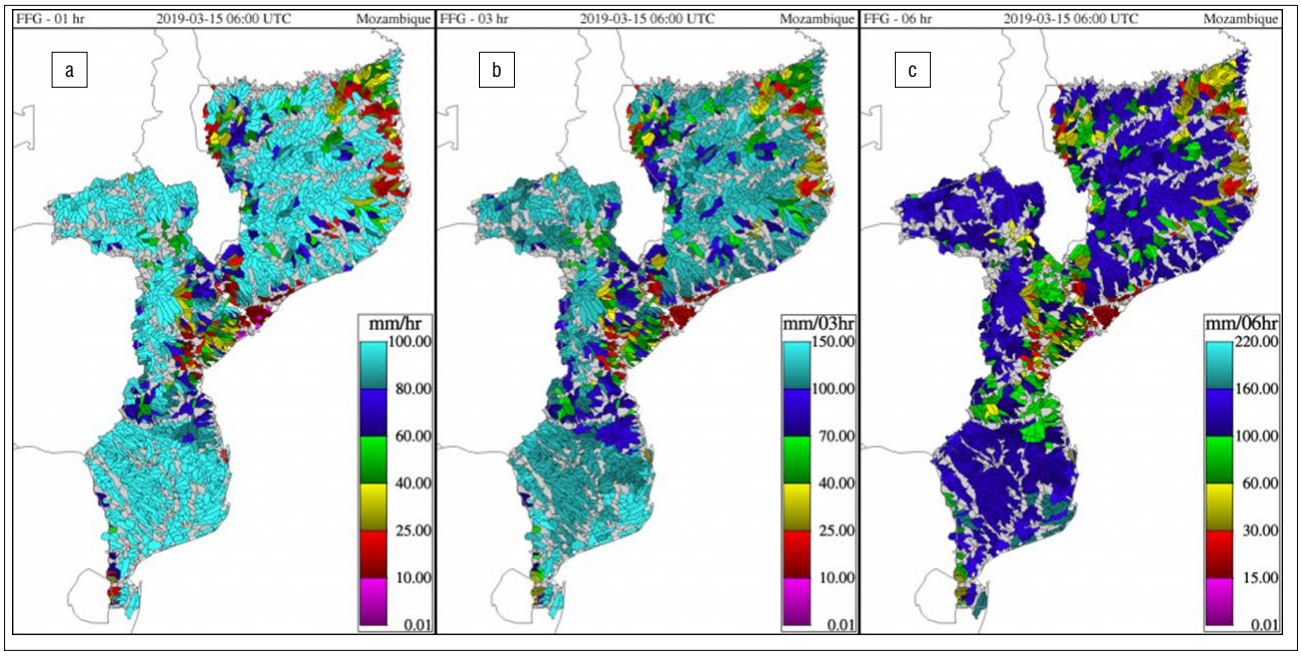


Figure 4: Southern Africa Regional Flash Flood Guidance issued at 06h00 UTC on 15 March 2019 for (a) 1-h, (b) 3-h and (c) 6-h lead times for river basins in Mozambique.

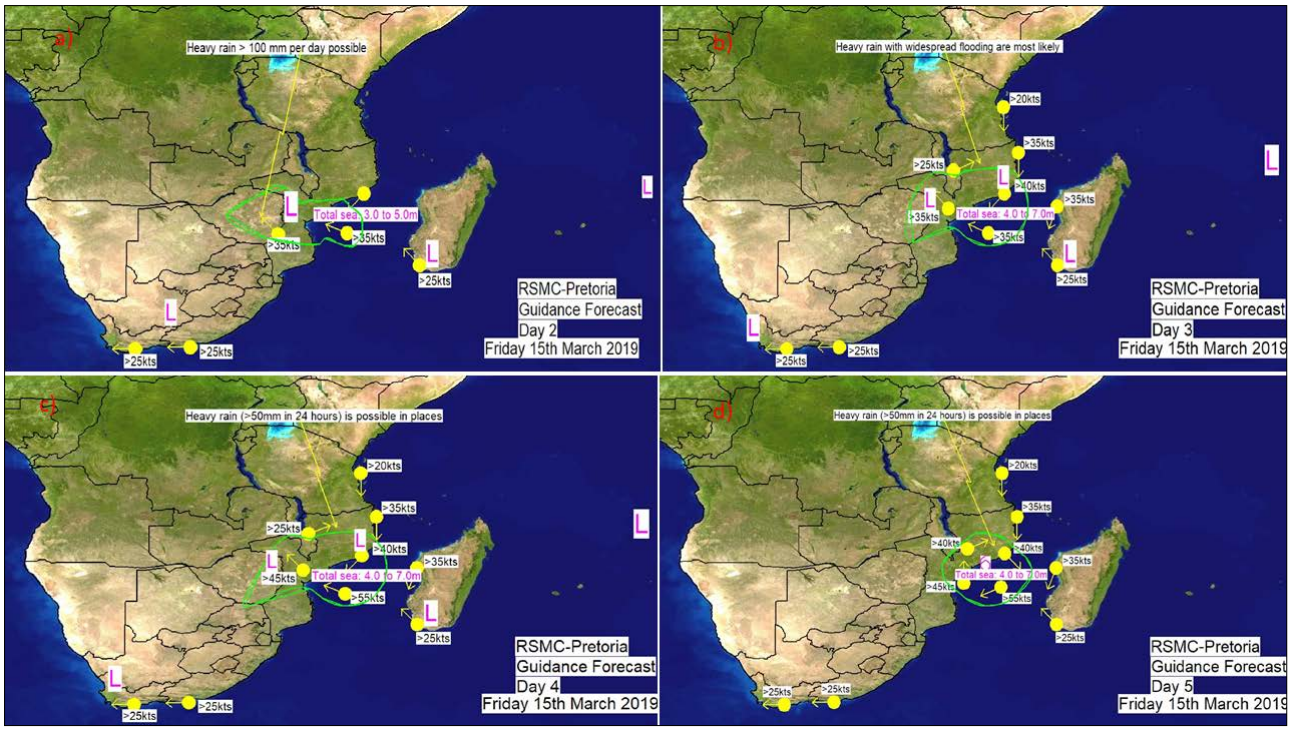


Figure 5: The Pretoria Regional Specialised Meteorological Centre Guidance Forecast issued on (a) 14 March 2019, (b) 13 March 2019, (c) 12 March 2019 and (d) 11 March 2019 for 15 March 2019.

KZN floods

Event description

Due to its location, South Africa usually experiences mid-latitude systems that move towards the east. The cut-off lows can result in heavy rainfall flooding over parts of South Africa depending on their location. One such event occurred in April 2019 and resulted in widespread rainfall over a large part of the country, with heavy rainfall and flooding over KZN. The floods were devastating to the extent that the President of South Africa, Mr Cyril Ramaphosa, visited the affected areas, and the floods were declared as a provincial disaster. Figure 6 shows the ERA5 reanalysis rainfall overlaid with the geopotential height at 500 hPa level. On 20 April, there was a perturbation south of the country, with some rainfall over parts of the Northern Cape, and Free State provinces (Figure 6). The perturbation extended further into the country, and resulted in widespread precipitation which affected seven of South Africa's nine provinces at different levels. By 22 April 2019, a closed low is visible and located over the Northern Cape with large amounts of rainfall east of the centre of the system. The SAWS synoptic chart of the day shows that there was an associated ridging high which transported moisture from the Indian ocean overland.[43] The moisture increase coupled with strong uplift due to the cut-off low resulted in large amounts of rainfall that caused flooding.

Large amounts of rainfall are shown over the southern parts of the KZN province and northern parts of the Eastern Cape province (Figure 6d, e). Fourteen weather stations in the southern parts of KZN reported over 100 mm of rainfall on 22 April (Figure 7a), with four of them continuing to report over 100 mm of rainfall the following day on 23 April. The Durban Kenneth Steinbank station reported 326 mm of rainfall, while Paddock reported 235 mm. The closed low persisted on 23 April, and caused more rainfall over the southern parts of KZN. By 24 April, the closed low was located over the eastern parts of the Western Cape and parts of the Eastern and Northern Cape. The largest amount of rainfall now fell over the ocean. The ERA5 reanalysis agrees with SAWS observations (Figure 7a), which can also be viewed on monthly or 10-day intervals on the SAWS website for the whole country.[44]

Numerical weather prediction

Figure 7 (b–i) shows the simulated precipitation from the UM with a grid spacing of 1.5 km and 4.4 km with different lead times, as well as from the ECMWF Integrated Forecast System model with a grid spacing of 16 km. The ECMWF provides the longest lead time and is plotted here up to a 4-day lead time. The ECWMF was able to capture the likelihood of the occurrence of the event well in advance, and this can be seen with the 4-day lead (Figure 7i); however, the amount of rainfall is underestimated.

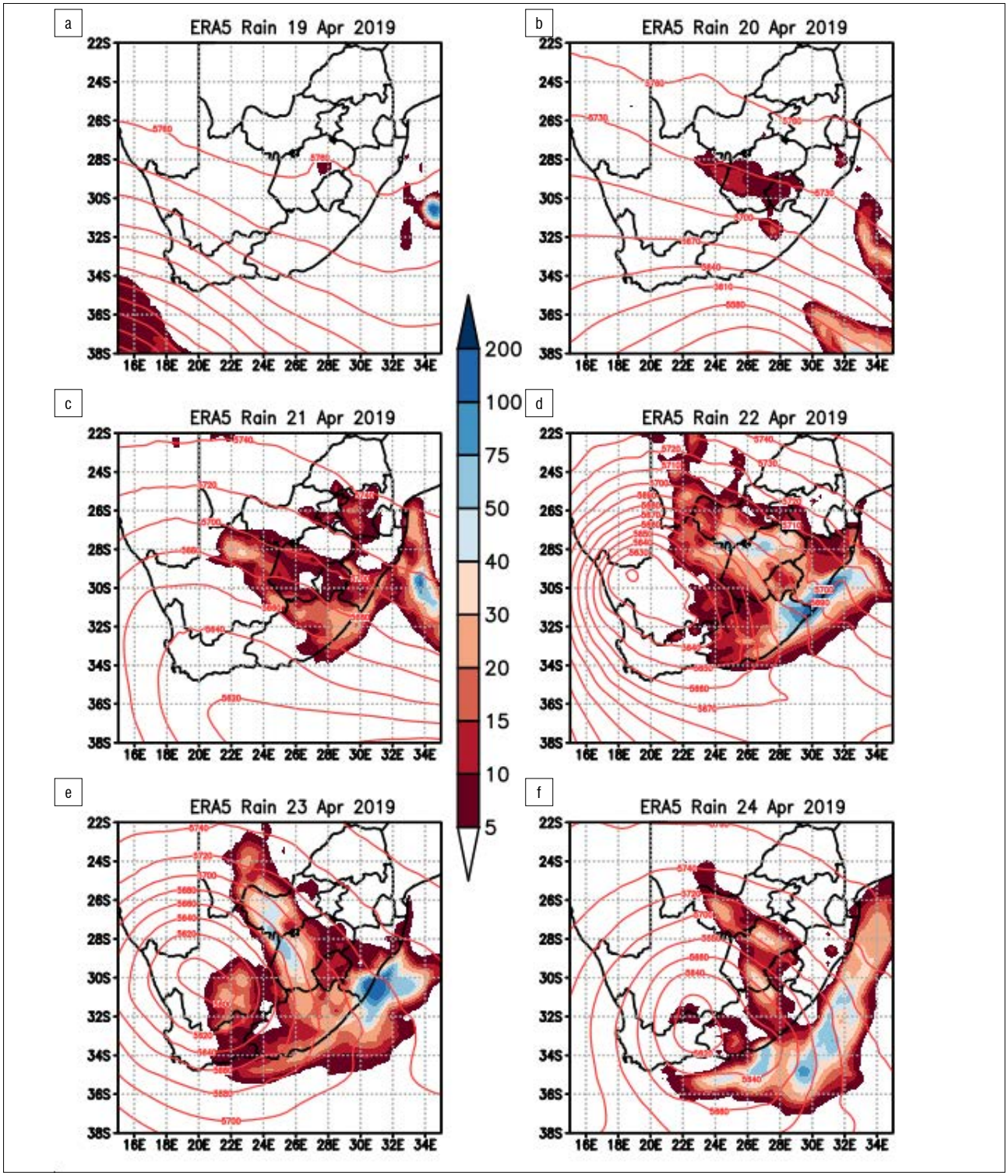


Figure 6: ERA5 reanalysis of 24-h rainfall (mm) overlaid with 500 hPa geopotential heights on (a) 19 April 2019, (b) 20 April 2019, (c) 21 April 2019, (d) 22 April 2019, (e) 23 April 2019 and (f) 24 April 2019.

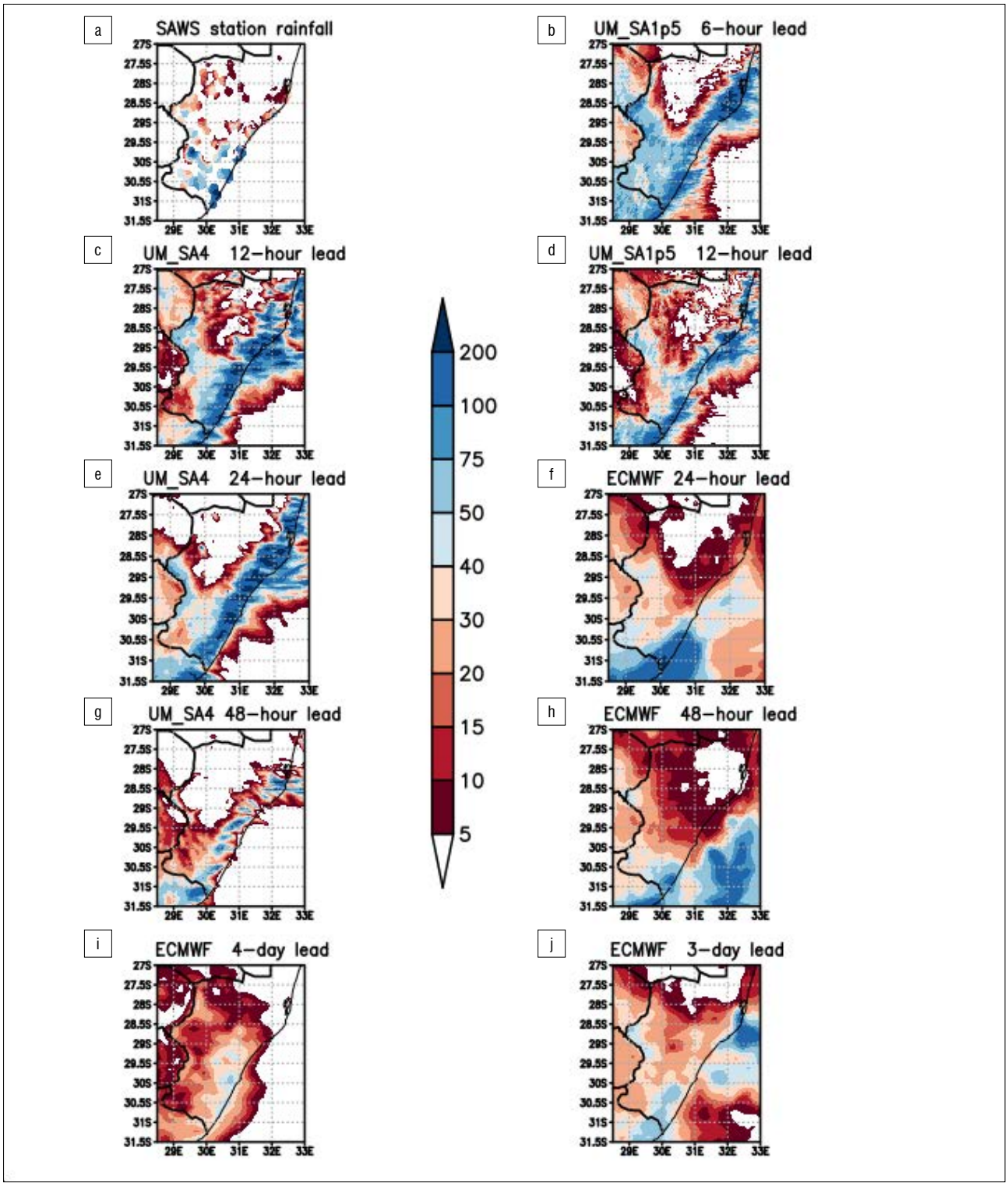


Figure 7: (a) SAWS station rainfall observations and rainfall simulations produced by the UM_SA4 with (c) 12-h, (e) 24-h and (g) 48-h lead times and UM_SA1p5 with (b) 6-h and (d) 12-h lead times and ECMWF with (f) 24-h, (h) 48-h, (j) 72-h and (i) 96-h lead times.

As the forecast time approaches, the amount of rainfall that is predicted increases. The longest lead time for a full 24-h total precipitation that we can obtain from the UM_SA4 running locally is 48 h. The UM was able to capture the likelihood of the occurrence of large amounts of rainfall at this 2-day lead time; however, the forecast of heavy rainfall extended further north along the coast of KZN. This behaviour persisted and can also be seen with a 6-h lead time. The longest lead time that can be provided by the UM_SA1p5 for a full-day total precipitation is 12 h. The performance of the UM_SA1p5 is similar to that of UM_SA4, with both extending the heavy rainfall along the whole KZN coast.

Table 1 shows three statistical measures, namely POD and CSI which range from 0 to 1, with 0 indicating no skill and 1 indicating a perfect score, as well as FAR which has the same range, with 0 indicating a perfect score and 1 no skill. The scores were calculated for the UM_SA4 and UM_SA1p5 for simulations with 12-h (12h00 initialisation on 21 April) and 6-h (18h00 initialisation on 21 April) lead times. The measures

are calculated against GPM satellite rainfall estimates, and the model data were coarse-grained to the satellite data resolution. The UM_SA4 scores suggest that the 12-h lead simulation performed slightly better than the 6-h lead simulation, across all three scores. The opposite is true for the UM_SA1p5 simulation, where the simulation with the shorter lead time of 6 h performed slightly better. When comparing different resolutions, the UM_SA4(UM_SA1p5) performed slightly better than the UM_SA1p5 (UM_SA4) with a 12-h lead (6-hour lead). These results confirm the results shown in Figure 7 that the performance across the two UM configurations with different lead times is similar. All the UM configurations extended the heavy rainfall amounts across the whole of the KZN coast. If forecasters only relied on the UM rainfall to issue warnings, they would have extended the heavy rainfall warning across the whole of the KZN coast, which would have been a false alarm in the northern parts. These results show that shortcomings remain in models, and as a result, models need to be improved further with data assimilation, improved physics, and dynamics.

Table 1: The Probability of Detection, Critical Success Index, and False Alarm Ratio for the UM_SA4 and UM_SA1p5 configurations with 6-h and 12-h lead times

UM Configuration	Probability of Detection	Critical Success Index	False Alarm Ratio
UM_SA4 12-h lead	0.90698	0.80161	0.12658
UM_SA1p5 12-h lead	0.90142	0.7928	0.13194
UM_SA4 18-h lead	0.88372	0.77689	0.13465
UM_SA1p5 18-h lead	0.91304	0.80267	0.1309

Nowcasting products

SAWS systems for nowcasting are based on radar, satellite and the lightning detection network. The RDT is based on satellite data and shows the location and phase of the storms, as well as the direction in which the storm is moving. Figure 8a shows an example of the RDT output which was made available to forecasters for 17h00 UTC on 22 April 2019. Figure 8a shows a large storm which is growing and expected to move southwards over KZN. These images are updated every 15 min, which is the baseline provided by MSG. It may be noted that countries that develop satellites receive updates more frequently than the African continent. For example, MSG provides updates every 5 min for Europe using rapid scanners.[45]

In South Africa, radar provides updates much quicker than satellite, every 6 min. Figure 8b shows output from a storm tracking system (TITAN), where the position of the storm and the direction the storm will take are indicated. The intense cells are captured much more easily by radar compared to the RDT satellite product. The red lines show the storm direction of these intense storms, and therefore provide forecasters with far more detailed information to use for issuing warnings in real time. This shows that it is important that radars remain operational and are maintained to provide the necessary information that can reduce the impact of severe weather events. Many African countries do not have radar systems and therefore rely on satellite data. The SAFFGS was also available for SAWS forecasters to use to provide relevant officials with information on river basins most at risk of flooding. For South Africa, the rainfall estimate is provided by radar when available.

Dissemination

The SAWS Disaster Risk Reduction Centre is responsible for issuing forecasts which include watches, alerts and warnings. Figure 9 shows a forecast that went out on 21 April 2019 in the afternoon (around 14h00 UTC) to the media for the following day. Figure 9 shows that SAWS issued a warning for parts of KZN and the Eastern Cape that experienced large amounts of rainfall. The warning was not extended across the whole KZN coast as was predicted by the UM simulations, but only showed for areas indicated by the ECMWF model. Those who watch the weather forecast on television stations that broadcast forecasts/warnings from SAWS would have seen this warning the day before the event occurred. As the event was taking place, SAWS continued to communicate with the public through its Twitter and Facebook accounts.

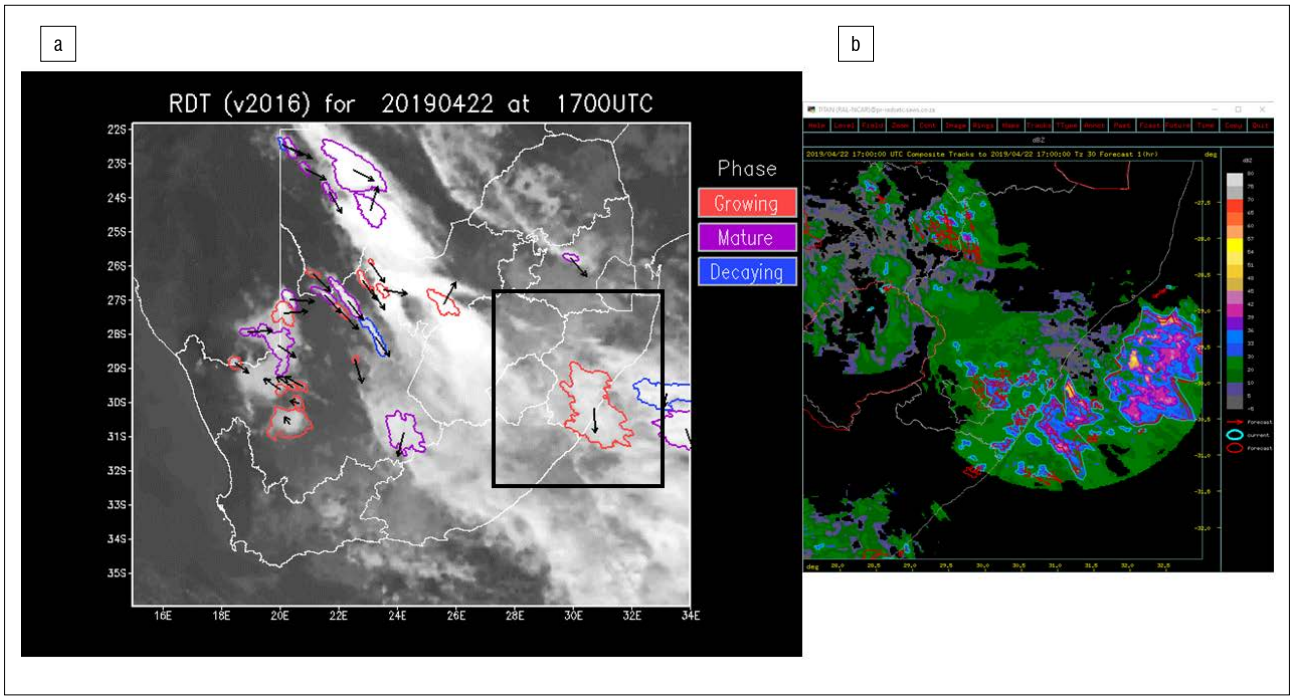


Figure 8: (a) Rapidly Developing Thunderstorm with a box showing the area shown on radar and (b) radar reflectivity of storms over KwaZulu-Natal and storm nowcasting on 22 April 2019 at 17h00 UTC.

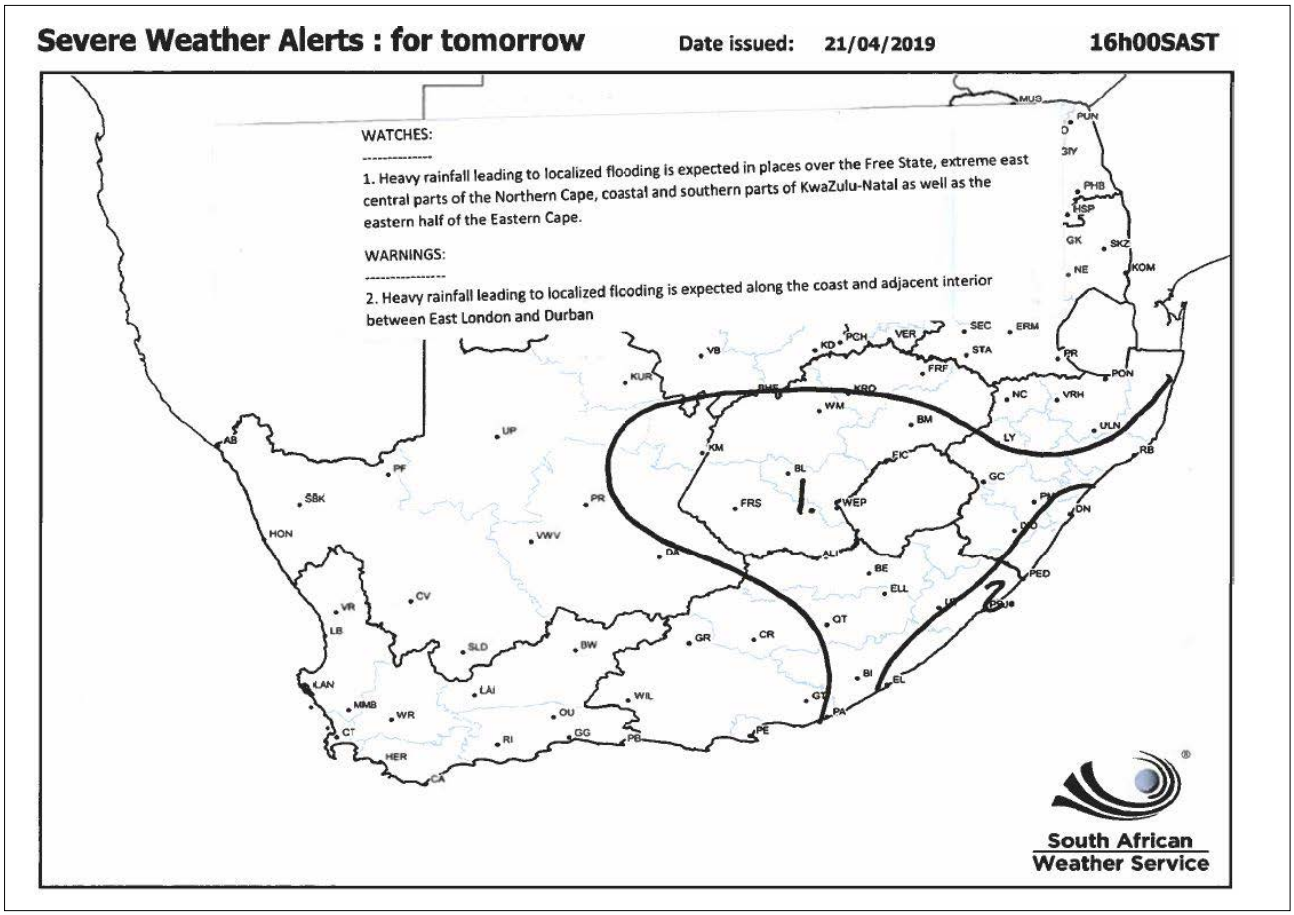


Figure 9: Severe weather alerts issued by the Disaster Risk Reduction division of SAWS to the media on 21 April 2019 at 14h00 UTC.

As with this event, SAWS has also noticed that despite warnings going out in advance, impacts of adverse weather events continue to ravage the country. SAWS is working actively to increase its reach in South Africa by using different platforms. One of these platforms is a WeatherSmart app, which is a free mobile app that shows the location of the storm and direction in which the storm is moving. Information on storm movements that is provided by the app is based on radar information. Further, SAWS is also embarking on a project to reach a larger listenership across South Africa. The number of community radio stations receiving information from SAWS increased from 30 in the 2016/2017 financial year to 86 during 2017/2018. This number further increased to 140 community radio stations during the last quarter of the 2018/2019 financial year. This indicates an increase in listenership from 984 000 in 2017/2018 to 4 326 000 in the first quarter of 2019.

Summary and conclusions

In this study, we evaluated the availability and performance of SAWS weather information during the Idai tropical cyclone which resulted in over 1000 deaths, as well as the floods that took place in KZN province in April 2019, resulting in over 70 fatalities. We have looked at output from NWP models as well as nowcasting systems. With the Idai tropical cyclone, we show that there was confidence in the forecasts 2 days before Idai made landfall over Mozambique to the extent that a media release was issued by SAWS. The UM, which is the main NWP model of SAWS, was consistent with its forecast, and the forecast did not change much as the lead time changed. The RSMC Pretoria issued warnings which should have been available for all forecasters in the SADC region to see to issue warnings within their countries. We also showed output from the SARFFGS which provides flash flood guidance for river basins which was also available to be used by forecasters in countries affected by Idai.

We also discuss the Durban floods that were caused by a cut-off low which encouraged uplift, accompanied by a ridging high-pressure system which transported moisture from the Indian Ocean into the country. We show that warnings were issued and communicated to the local media at least a day in advance. The UM and the ECMWF were both able to capture the event; however, the ECMWF captured the location of the heavy rainfall better. Neither model was able to capture the actual rainfall amounts associated with the event, and this shows that models need further improvement. In terms of dissemination of information, we discuss the radar and satellite products that are used for the nowcasting timescale, and which forecasters use and communicate with the media as well as the disaster risk management centres at different levels (i.e. national, provincial and local). We also discuss the WeatherSmart app which uses radar information as input and which the community can use to check storm occurrence. To reach more South Africans, SAWS is also increasing the number of community radio stations that receive weather information.

The study shows that more needs to be done to improve models further, to increase community weather awareness and to reduce community vulnerability to extreme weather events. Suarez[46] pointed to the underutilisation of the available science by humanitarian organisations and people at risk, with extreme events being allowed to become deadly events due to a failure to act. This he attributed to humanitarian organisations not being able to build institutional and stakeholder capacity to use the available tools, and he recommended a transition from action-based to knowledge-based entities that can effectively use newly available tools. De Perez[47] developed a forecast-based finance system that can trigger action based on weather/climate forecasts or observations, while accounting for the possibility of false alarms. In their classification of African countries' Disaster Risk Management policies, Tall[48] classified both South Africa and Mozambique as disaster averters (best category of 3); however, the two events described in this paper

show that more needs to be done to improve disaster preparedness and reduce risk in both countries.

Acknowledgements

We acknowledge SAWS for supporting this study as well as the anonymous reviewers who helped improve the paper. This work was also partially supported through the Climate Research for Development (CR4D) Postdoctoral Fellowship CR4D-19-11 implemented by the African Academy of Sciences (AAS) in partnership with the United Kingdom's Department for International Development (DfID) Weather and Climate Information Services for Africa (WISER) programme and the African Climate Policy Center (ACPC) of the United Nations Economic Commission for Africa (UNECA).

Competing interests

We declare that there are no competing interests.

Authors' contributions

M.M.B.: Conceptualisation, data collection, data analysis, writing. E.S., B.M.: Conceptualisation, data collection. T.N.: Conceptualisation, design of the study, writing. M.N., E.P.: Data collection and analysis. M.G., S.L., G.R., L.v.H., M.M.: Data collection.

References

1. De Carli M, Olesen BW, Zarrella A, Zecchin R. People's clothing behaviour according to external weather and indoor environment. Build Environ. 2007;42(12):3965–3973. https://doi.org/10.1016/j.buildenv.2006.06.038
2. Akin D, Sisiopiku VP, Skabardonis A. Impacts of weather on traffic flow characteristics of urban freeways in Istanbul. Procedia-Soc Behav Sci. 2011;16:89–99. https://doi.org/10.1016/j.sbspro.2011.04.432
3. Staffella I, Pfenninger S. The increasing impact of weather on electricity supply and demand. Energy. 2018;145:65–78. https://doi.org/10.1016/j.energy.2017.12.051
4. Centre for Research on the Epidemiology of Disasters (CRED). The human cost of weather related disasters 1995–2015 [document on the Internet]. c2015 [cited 2020 Jun 09]. Available from: https://reliefweb.int/sites/reliefweb.int/files/resources/COP21_WeatherDisastersReport_2015_FINAL.pdf
5. Lubchenco J, Karl TR. Predicting and managing extreme weather events. Physics Today. 2012;65(3):31. https://doi.org/10.1063/PT.3.1475
6. Pielke R, Carbone RE. Weather impacts, forecasts, and policy. Bull Amer Meteor Soc. 2002;83:393–406. https://doi.org/10.1175/1520-0477(2002)083<0393:WIFAP>2.3.CO;2
7. Tyson PD, Preston-Whyte RA. The weather and climate of southern Africa. 2nd ed. Cape Town: Oxford University Press; 2000.
8. Lukamba MT. Natural disasters in African countries: What can we learn about them. J Transdiscip Res South Afr. 2010;6(2):478–495. https://doi.org/10.4102/td.v6i2.266
9. South African Department of Environmental Affairs (DEA). Climate information and early warning systems for supporting the disaster risk reduction and management sector in South Africa under future climates. In: Munzhedzi SM, Khavhagali VP, Midgley GM, de Abreu P, Scorgie S, Braun M, et al., editors. Long-Term Adaptation Scenarios Flagship Research Programme. Pretoria: DEA; 2016. p. 31–43. Available from: https://www.environment.gov.za/sites/default/files/reports/ltasbook2of7_climateinformationandearlywarningsystemsforsupportingtheDRR.pdf
10. Harrison MSJ. A generalized classification of South African summer rain-bearing synoptic systems. J Climatol. 1984;4:547–560. https://doi.org/10.1002/joc.3370040510
11. Hart NCG, Reason CJC, Fauchereau N. Cloud bands over southern Africa: Seasonality, contribution to rainfall variability and modulation by the MJO. Clim Dyn. 2013;41:1199–1212. https://doi.org/10.1007/s00382-012-1589-4
12. Favre A, Hewitson B, Tadross M, Lennard C, Cerezo-Mota R. Relationships between cut-off lows and the semiannual and southern oscillations. Clim Dyn. 2011;38:1473–1487. https://doi.org/10.1007/s00382-011-1030-4
13. Malherbe J, Engelbrecht FA, Landman WA, Engelbrecht CJ. Tropical systems from the southwest Indian Ocean making landfall over the Limpopo River Basin, southern Africa: A historical perspective. Int J Climatol. 2012;32:1018–1032. https://doi.org/10.1002/joc.2320
14. Changnon SA, Pielke RA, Changnon D, Sylves RT, Pulwarty R. Human factors explain the increased losses from weather and climate extremes. Bull Amer Meteor Soc. 2000;81:437–442. https://doi.org/10.1175/1520-0477(2000)081<0437:HFETIL>2.3.CO;2
15. Shapiro A, Jonathan L. He thought his city was prepared for big storms. Then Cyclone Idai hit. NPR. 2019 April 10. Available from: https://www.npr.org/sections/goatsandsoda/2019/04/10/711576607/he-thought-his-city-was-prepared-for-big-storms-then-cyclone-Idai-hit
16. Al Jazeera and New Agencies. Death toll from devastating Cyclone Idai rises above 1000. Aljazeera Media Network. 2019 April 10. https://www.aljazeera.com/news/2019/04/death-toll-devastating-cyclone-Idai-rises-1000-190410155136406.html
17. BBC News. South Africa floods: Death toll after Durban rains rises to 60. BBC. 2019 April 24. https://www.bbc.com/news/world-africa-48036252
18. Singh O. Killer KZN floods declared a provincial disaster. TimesLive. 2019 May 02. https://www.timeslive.co.za/news/south-africa/2019-05-02-killer-kzn-floods-are-declared-a-provincial-disaster/
19. Sai F, Cumiskey L, Weerts A, Bhattacharya B, Khan R. Towards impact-based flood forecasting and warning in Bangladesh: A case study at the local level in Sirajganj district. Nat Hazard Earth Sys. 2018:1–20. https://doi.org/10.5194/nhess-2018-26
20. Weyrich P, Scolobig A, Bresch DN, Patt A. Effects of impact-based warnings and behavioral recommendations for extreme weather events. Weather Clim Soc. 2018;10:781–796. https://doi.org/10.1175/WCAS-D-18-0038.1
21. Silvestro F, Rossi L, Campo L, Parodi A, Fiori E, Rudari R, et al. Impact-based flash-flood forecasting system: Sensitivity to high resolution numerical weather prediction systems and soil moisture. J Hydrol. 2019;572:388–402. https://doi.org/10.1016/j.jhydrol.2019.02.055
22. Browning KA. The role of radar in weather forecasting. Phys Technol. 1983;14(3):140. https://doi.org/10.1088/0305-4624/14/3/I03
23. Dixon M, Wiener G. TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting – A radar-based methodology. J Atmos Oceanic Technol. 1993;10:785–797. https://doi.org/10.1175/1520-0426(1993)010<0785:TTITAA>2.0.CO;2
24. Becker EH. Application of a quantitative precipitation estimation algorithm for the S-Band radar at Irene, South Africa [MSc dissertation]. Durban: University of KwaZulu-Natal; 2014. http://hdl.handle.net/10413/12795
25. De Coning E, Gijben M, Maseko B, Van Hemert L. Using satellite data to identify and track intense thunderstorms in South and southern Africa. S Afr J Sci. 2015;111(7/8), Art. #2014-0402, 9 pages. http://dx.doi.org/10.17159/SAJS.2015/20140402
26. De Coning E, Poolman E. South African Weather Service operational satellite based precipitation estimation technique: Applications and improvement. Hydrol Earth Syst Sci. 2011;15:1131–1145. https://doi.org/10.5194/hess-15-1131-2011
27. Davies T, Cullen MJP, Malcolm AJ, Mawson MH, Staniforth A, White AA, et al. A new dynamical core for the MetOffice's global and regional modelling of the atmosphere. Quart J R Meteorol Soc. 2005;131:1759–1782. https://doi.org/10.1256/qj.04.101
28. Landman S, Engelbrecht FA, Engelbrecht CJ, Dyson LL, Landman WA. A short-range weather prediction system for South Africa based on a multi-model approach. Water SA. 2012;38(5):765–774. http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S1816-79502012000500016
29. Haiden T, Janousek M, Bidlot J-R, Buizza R, Ferranti L, Prates F, et al. Evaluation of ECMWF forecasts, including the 2018 upgrade. ECMWF Technical Memoranda. 2018. https://doi.org/10.21957/ldw15ckqi
30. Hersbach, H, Dee D. ERA5 reanalysis is in production. ECMWF Newsletter 147. Reading: ECMWF; 2016. Available from: https://www.ecmwf.int/en/newsletter/147/news/era5-reanalysis-production

31. Huffman GJ, Stocker EF, Bolvin DT, Nelkin EJ, Jackson T. GPM IMERG final precipitation L3 half hourly 0.1 degree x 0.1 degree V06 [data set]. Greenbelt, MD: Goddard Earth Sciences Data and Information Services Center (GES DISC); 2019 [cited 2020 Jun 09]. https://doi.org/10.5067/GPM/IMERG/3B-HH/06

32. Maidment R, Grimes D, Black E, Tarnavsky E, Young M, Greatrex H, et al. A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa. Sci Data. 2017;4, Art. #170063. https://doi.org/10.1038/sdata.2017.63

33. Jolliffe IT, Stephenson DB. Forecast verification. A practitioner's guide in atmospheric science. Hoboken, NJ: John Wiley & Sons Ltd.; 2003.

34. South African Weather Service (SAWS). Regional Specialized Meteorological Centre (RSMC) [webpage on the Internet]. No date [cited 2020 Jun 09]. Available from: http://rsmc.weathersa.co.za

35. World Meteorological Organization. Reducing vulnerability to extreme hydro-meteorological hazards in Mozambique after Cyclone IDAI. WMO mission report following tropical Cyclone IDAI (29 April – 7 May 2019) [webpage on the Internet]. c2019 [cited 2020 Jun 09]. Available from: https://public.wmo.int/en/resources/library/reducing-vulnerability-extreme-hydro-meteorological-hazards-mozambique-after

36. Gao S, Jia S, Wan Y, Li T, Zhai S, Shen X. The role of latent heat flux in tropical cyclogenesis over the western North Pacific: Comparison of developing versus non-developing disturbances. J Mar Sci Eng. 2019;7:28. https://doi.org/10.3390/jmse7020028

37. Seyama S, Masocha M, Dube T. Evaluation of TAMSAT satellite rainfall estimates for southern Africa: An inter-product comparison study. Phys Chem Earth A/B/C. 2019;112:141–153. https://doi.org/10.1016/j.pce.2019.02.008

38. Wong ML, Chan JC. Tropical cyclone motion in response to land surface friction. J Atmos Sci. 2006;63:1324–1337. https://doi.org/10.1175/JAS3683.1

39. Verner D. Restoring Zimbabwe's livelihoods, infrastructure after Cyclone Idai [webpage on the Internet]. c2019 [cited 2020 Jun 09]. Available from: https://www.worldbank.org/en/news/feature/2019/09/16/restoring-zimbabwes-livelihoods-infrastructure-after-cyclone-idai

40. Nabeta L. World Bank helps Malawi recover from Cyclone Idai [webpage on the Internet]. c2019 [cited 2020 Jun 09]. Available from: https://www.worldbank.org/en/news/press-release/2019/06/06/world-bank-helps-malawi-recover-from-cyclone-idai

41. Saute R. World Bank injects $130 million in support of recovery efforts in cyclones affected communities [webpage on the Internet]. c2019 [cited 2020 Jun 09]. Available from: https://reliefweb.int/report/mozambique/world-bank-injects-130-million-support-recovery-efforts-cyclones-affected

42. Stern H, Davidson NE. Trends in the skill of weather prediction at lead times of 1–14 days. Q J R Meteorol Soc. 2015;141:2726–2736. https://doi.org/10.1002/qj.2559

43. South African Weather Service. Synoptic weather map [document on the Internet]. c2019 [cited 2020 Jun 09]. Available from: http://www.weathersa.co.za/Documents/Publications/20190422.pdf

44. South African Weather Service. Historical rain [webpage on the Internet]. c2020 [cited 2020 Jun 09]. Available from: http://www.weathersa.co.za/home/historicalrain

45. Gallucci D; Romano F, Cimini D, Di Paola F, Gentile S, Larosa S, et al. Improvement of hourly surface solar irradiance estimation using MSG rapid scanning service. Remote Sens. 2019;11:66. https://doi.org/10.3390/rs11010066

46. Suarez P. Linking climate knowledge and decisions: Humanitarian challenges. Boston, MA: Boston University Frederick S. Pardee Center for the Study of the Longer-Range Future; 2009. Available from: http://www.bu.edu/pardee/publications/pardee-paper-007-climate

47. De Perez EC, Van den Hurk B, Van Aalst MK, Jongman B, Klose T, Suarez P. Forecast-based financing: An approach for catalyzing humanitarian action based on extreme weather and climate forecasts. Nat Hazard Earth Sys. 2015;15:895–904. https://doi.org/10.5194/nhess–15-895-2015

48. Tall A, Patt A, Fritz S. Reducing vulnerability to hydro-meteorological extremes in Africa. A qualitative assessment of national climate disaster management policies: Accounting for heterogeneity. Weather Clim Extremes. 2013;1:4–16. https://doi.org/10.1016/j.wace.2013.07.007

AUTHORS:
Annah V. Bengesai[1]
Jonathan Pocock[2]

AFFILIATIONS:
[1]Teaching and Learning Unit, University of KwaZulu-Natal, Durban, South Africa
[2]School of Engineering, University of KwaZulu-Natal, Durban, South Africa

CORRESPONDENCE TO:
Annah Bengesai

EMAIL:
bengesai@ukzn.ac.za

DATES:
Received: 10 Dec. 2019
Revised: 07 July 2020
Accepted: 28 July 2020
Published: 29 Mar. 2021

HOW TO CITE:
Bengesai AV, Pocock J. Patterns of persistence among engineering students at a South African university: A decision tree analysis. S Afr J Sci. 2021;117(3/4), Art. #7712. https://doi.org/10.17159/sajs.2021/7712

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Jennifer Case

KEYWORDS:
graduation, engineering education, major switching, data mining, classification algorithms

FUNDING:
None

Patterns of persistence among engineering students at a South African university: A decision tree analysis

Globally, there is growing concern about student progression in most higher education institutions. In this study, we examined patterns of persistence among students who began their engineering degree at the University of KwaZulu-Natal (UKZN) in 2012 and 2013. The sample was restricted to 1370 incoming students who were tracked to 2019, allowing for a 7-year graduation period for the initial cohort. The data were analysed using descriptive statistics as well as the decision tree approach – a highly visual data-mining technique which helps identify subgroups and relationships that are often difficult to detect through traditional statistical methods. The results from these analyses indicate that up to 50% of students enrolled in the School of Engineering had chosen engineering as their first choice. Approximately 40% had persisted in engineering, 50% had withdrawn by the time of this survey, while the remaining 10% were still registered in the engineering programme. Departure from engineering occurs most in the first year, while graduation most likely occurs after 5 years of registration. Student persistence in engineering can also be classified based on first-year accumulated credits, admission point scores, race, and financial aid, of which first-year accumulated credits is the most critical factor. Overall, our study suggests that understanding failure in the first year might be the missing link in our understanding of student persistence in engineering.

Significance:

- The study makes several contributions to the field. First, the issue of student persistence is of concern to higher education institutions, not only in South Africa but globally. For institutions to improve student outcomes, they need to understand the reasons behind attrition. Second, in our analysis, we separate students who withdraw from the university and those who switch to other programmes within the same university. Most studies on student progression treat withdrawing students as a single population, which might lead to inaccurate prediction of student outcomes. Third, we draw on data-mining methods and present an interesting way of classifying students using both enrolment data as well as the rules derived from each node of the classification tree. The classification tree analysis method is highly visual and helps identify subgroups and relationships, which might be difficult to detect through traditional statistical methods. The information derived from the classification trees can be used to identify students who might be at risk of failing timeously and come up with interventions that will support them.

Introduction

Science, technology, engineering, and mathematics (STEM) graduates are critical for economic growth and sustainable development, especially in a developing country like South Africa.[1-3] South African higher education institutions are aware of this, and great strides have been made to increase the number of enrolments across all disciplines, specifically in STEM majors.[4] The shift in focus from elitism to the 'massification' of higher education has also resulted in a diverse student body with a wide range of social, cultural and educational backgrounds[2] for which institutions are not prepared.[5] Consequently, the widening of access has proved to be a double-edged sword characterised by high dropout rates and increasing calls to improve the quality of education.[5,6] The student protests of 2015/2016, which called for, among other things, the decolonisation of the curriculum are also reminders that higher education institutions must continue to invest in both access and student success.[7]

The concern with student retention and non-completion is not unique to South Africa. It has been the subject of research for many years, with studies consistently showing that it is a significant problem in most institutions globally.[8-10] Most of this research, which is dominated by the USA, has investigated the trends as well as the various factors that influence student retention, attrition and persistence.[11,12] For instance, in the USA, it is estimated that the 8-year graduation rate at public universities is 60%, although this varies by institution and degree programme.[9] Similar trends have also been observed in the Organisation for Economic Co-operation and Development (OECD) countries, where female students and students of colour were also found to have lower graduation rates than their male and/or white counterparts.[10,13] The Council of Higher Education reported that approximately 50% of all first-time students entering South African higher education institutions are likely to leave without obtaining a degree.[4] In a review of the progression and non-completion rates in engineering at a South African university, Pocock[14] found that completion rates in the region of 50–72% of entering students were not uncommon, while attrition in the first year of study was approximately 15–20%. The Engineering Council of South Africa's throughput study also found that between 10% and about 45% of students completed their engineering degrees in minimum time at different institutions, while total completion rates ranged between 35% and 60%.[15]

Clearly, these earlier studies have provided numerous insights into the nature of student persistence both globally and in South Africa. However, and especially in South Africa, most of the earlier studies on student persistence have used indicators which are often overstated, without accounting for the complexity in student progression.

For instance, it is possible that a student might enrol for a particular major at an institution, persist for a few years before switching to another major. Yet, most studies on student retention and dropout in South Africa do not pay attention to student enrolment changes. Second, of the quantitative studies that have sought to understand factors associated with persistence, most have relied on descriptive work with fewer or no controls[6,14,15], as well as traditional regression analyses[16-18]. While regression methods have many benefits in predictive analysis, one of their weakness is the inability to capture non-monotonic relationships as well as unspecified patterns across factors[19], which can be resolved through classification methods.

The main aim of this study, therefore, was to add to the scholarship by providing a longitudinal examination of the patterns of selecting, persisting in, switching and leaving majors among students enrolled in the School of Engineering at the UKZN, South Africa. A student is considered to have persisted in an engineering major if they continue their studies from one year to the next and ultimately graduate in their initially registered specialisation. We use the term 'major' to refer to a specific engineering field or specialisation. Those who did not re-register within 2 years from the time of their initial departure to the time of the survey are considered to have withdrawn from the engineering major.

We acknowledge that student progression in higher education is rarely straightforward and can take many pathways beyond what we have included in our analysis. Students who leave a particular qualification might transfer to another institution. It is also not uncommon for students to return to their qualification after a couple of years – a phenomenon referred to as stop-out in the literature.[12,14] However, data on institutional transfer are rare, and tracking stop-outs might require interviewing the students who withdrew to understand their reasons for leaving, which was beyond the scope of this study.

Our study also seeks to make a specific contribution to the scholarship by utilising a classification (decision) tree approach to data analysis. Decision tree analysis is a data-mining approach which has the potential to identify hidden as well as simplify complex patterns and relationships found in a data set.[20,21] Unlike other discriminatory models such as regression analysis, a decision tree is a highly visual and transparent model which makes explicit all the possible outcomes and patterns in a data set.[20] While decision tree algorithms are increasingly being used in many fields such as medicine and public health[22], they remain less common in higher education research, although an emerging body of literature[23-25] is showing their potential in identifying hidden patterns in institutional data.

Factors influencing student persistence

Past studies have identified several factors that influence student persistence at university.[11,12] The conclusion from these studies is that the determinants of student persistence are numerous and complex and perhaps even context dependent. Factors that have been identified in South Africa include school-level factors[26]; career counselling[27]; high-school grades, especially for Mathematics and Science[28,29]; high-school rank; as well as institutional factors such as financial aid and academic support[17,30]. While some of these factors are generic to all academic disciplines, some are more specific to STEM degree majors such as engineering.

School-level factors steering students in or out of university in South Africa are well documented.[26,31,32] The consensus is that the schooling system does not adequately prepare students and has become a 'dominant learning-related cause of the poor performance patterns in higher education'[32]. Related to this is the school ranking system which divides schools into five quintiles based on socio-economic status. Quintile 1 is the most disadvantaged and found in rural communities, while Quintile 5 is the most advantaged.[33] This dualistic nature of the education landscape is also apparent in the large performance gap between students from Quintile 5 schools and those from the rest of the education system.[34,35] Some scholars have further identified the poor quality of science and maths teaching, language underpreparedness and lack of resources, especially in rural schools, as other factors impacting on student outcomes.[29,32] Maree[27] also criticised the lack of career counselling at high schools (particularly in lower quintile schools), where it has been placed in the Life Orientation curriculum at the National Senior Certificate level. Both the extent of career counselling (20% of the curriculum) and the lack of specialist training for teachers in the subject were highlighted as areas that needed improvement. Maree[27] further argues that the traditional approach to career counselling at the university level fails many first-generation students as it leads them into inappropriate field choices, with the consequence of higher dropout rates.

Several studies have also investigated the role of some of these factors in determining student success in engineering programmes in South Africa.[14,17,18,26,30,36] Pocock[14] interviewed students at a South African university who had not re-registered to complete their engineering degrees 18 months post-leaving. He found that, in combination with other factors, financial support was one of the primary reasons for leaving for 48% of the students. Zewotir et al.[18] benchmarked success rates across all faculties of the same university against a variety of factors and separated dropout statistics into students failing or choosing to leave. For engineering students, the significant findings were wide-ranging, with students with a lower Mathematics score at high school found to be more likely to leave, and Indian South Africans more likely to leave than white South African students. In terms of the economics of continuation, the probability of persisting was higher for students with financial aid than those without, although different patterns of persistence were found among students with different forms of financial aid. Additional to this, those without a place in university residence were more likely to fail than those in residence. Bengesai and Paideya[17] found that students who pass 75% of their first-year course credits were up to eight times more likely to persist in engineering and graduate in regulation time or regulation time plus one year. Other factors peculiar to engineering – but not well researched in South Africa – include stereotypes about gender, lack of role models, course workloads as well as attitudinal characteristics that are harder to determine but might lead to weeding out.[37]

Choosing a major at UKZN

Applications to study engineering programmes at UKZN (along with all programmes of study) are handled centrally through a Central Admissions Office. On application, prospective undergraduate students specify their preferred major as well as other majors that they would like to be considered for admission as first to sixth choices on the application. Pre-selection of students for admission is carried out at the middle to end of the year prior to their potential admission. Prospective students are offered places conditional on their high-school (National Senior Certificate) grades meeting specified levels upon their matriculation from high school. Once final school results are available, firm offers are made to those scholars who have met the conditions, including students who have met the requirements but were not necessarily given the first-choice conditional offer. Once first choices are exhausted, should there be places still available, these would then be offered to students who meet the minimum requirements but might have specified the major as a second or third choice etc. While this does lead to a dynamic system of offers being made, taken or declined, in most of the engineering majors, the scholars selected for admission have chosen a specified engineering field as their major (e.g. Chemical Engineering, Mechanical Engineering) prior to admission. In two of the engineering disciplines (Agricultural Engineering and Computer Engineering), there are occasions where students have chosen these as a second or third choice after other engineering majors. In a few cases, scholars have chosen Medical School (which has the most competitive entry requirements) as a first choice, with an engineering major as a second choice.

Table 1 shows the degree choice patterns among students who were enrolled in the School of Engineering at UKZN in 2012 and 2013. As can be seen from Table 1, the majority (58%, $n=789$) of the students chose engineering as a major, including the specialisation they eventually registered for, as their first choice. In comparison, 17% ($n=234$), 7% ($n=92$) and 4% ($n=55$) were enrolled in their second, third and fourth choices, respectively. Approximately 15% ($n=200$) of the enrolled

students chose engineering as a fifth or sixth choice at the application stage. The average admission point (AP) score for students who chose an engineering major, including the specialisation, as a first choice was 39, while the average AP score for all students in the sample was 37.

Table 1: Degree major choice patterns (n=1370)

Qualification choice at the application stage	*N* (%)	Mean AP score
First	789 (57.6	39
Second	234 (17.0)	38
Third	92 (7.0)	37
Fourth+	55 (4.0)	35
Other	200 (14.6)	34
Total	1370 (100)	37

Materials and methods

The data used in this analysis were obtained from the cohort data archived in the Department of Institutional Intelligence at UKZN. This data set captures students' biographical information (race, gender), AP scores, academic performance in the first year, as well as information about financial aid and whether the student resides on the university campus. The sample used in the descriptive analysis consisted of 1370 first-time entry students who began their academic year in 2012 or 2013, thus allowing for a 6-year graduation rate for the latter cohort. For the decision tree method, we excluded 116 students who were still enrolled in their initially declared major. The following research questions guided this study:

1. What are the patterns of selecting, persisting and switching qualifications among a cohort of engineering students?
2. Do gender, race, AP scores, school quintile, first-year accumulated credit load, financial aid and campus residence influence persistence in an engineering major?
3. Which of these factors (in Question 2) can efficiently classify students' likelihood of persisting in an engineering major?

To understand students' persistence patterns, we considered the students' initial registered engineering specialisation, and tracked them through graduation, withdrawal or any subsequent majors in which they enrolled. The conceptual model is shown in Figure 1.

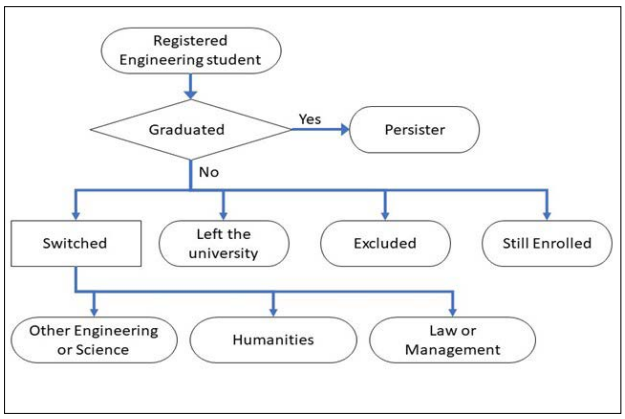


Figure 1: Conceptual model for understanding students' persistence patterns.

The starting point of this model is students registered in an engineering major. Students who graduated from the initially registered engineering specialisation were defined as 'persisters'. The next step was to look at those students who had not graduated and to track their pathways of leaving the initially registered qualification (that is, withdrawal from the engineering specialisation, exclusion or still enrolled). In our analysis, we first present the analysis of students who switched to other programmes and those who left the university as one category, 'withdrawn' (Table 2 and Figure 2), and then as the specific subcategories (Figure 3 and Table 4). Finally, we built the decision tree model in SPSS v. 26 to classify student persistence according to the given set of variables specified below.

Outcome variables

Our main outcome variable of interest was persistence to graduation (1=yes; 0=other). This was a measure of a student's persistence in the engineering specialisation in which they initially registered at the beginning of their studies, and was used in both the descriptive and classification tree analyses.

Explanatory variables

Gender was categorised as male (=0) or female (=1) and race as black South African (=0) or 'Other' (=1), with 'Other' being a combination of coloured, Indian and white students. We collapsed race into two categories due to the small sample sizes of two of the race categories (white, 85 and coloured, 18). We acknowledge that this categorisation has the potential to mask significant heterogeneity as it does not differentiate between all racial groups. However, given the data limitations, we worked on the plausible assumption that the academic performance of black South African students has been shown to fall behind that of students from other racial groupings.[4,6,15] We also rationalised that race is only a factor in South Africa because people have been historically disadvantaged based on race, with black South Africans suffering the extreme consequences.[38] Moreover, black South African students constitute more than 70% of the student population.[39] Therefore, comparing their persistence patterns relative to a composite race category seemed reasonable as it would at least give us information about this particular risk group.

Financial aid is increasingly important in South Africa given the growing enrolments of students from disadvantaged backgrounds.[40] In fact, financial aid is so important that the 2015/2016 #FeesMustFall student protests that took the nation by storm highlighted its influence on student persistence.[41] Thus, we also considered whether a student had financial aid (0=no; 1=yes) as a proxy for socioeconomic status.

The students' AP scores were coded as a four-category variable (0=alternative; 1=30–35; 2=36–39; 3=40 and above). We derived this classification from the patterns detected in the enrolment data for the various engineering specialisations. For instance, specialisations such as Agricultural and Civil Engineering tend to have a majority of students with AP scores between 36 and 39, while those with AP scores above 40 tend to enrol in Chemical Engineering programmes. Students who achieve a grade of between 90% and 100% in a National Senior Certificate subject are awarded 8 points; hence, the maximum possible AP score for six National Senior Certificate subjects is 48. We categorised students who wrote a different school leaving examination from the National Senior Certificate written in government-funded schools as 'alternative'. Nine of the students in our sample had inaccurate AP scores (below 20). To avoid bias in our analysis, we decided to treat this group as a missing or floating category[21] that would intuitively merge with its most similar AP scores category.

At UKZN, students across all engineering specialisations follow a generic first year consisting of 13 courses in mathematics, chemistry, physics and engineering drawing, with one exception for the Chemical Engineering students who take a different chemistry course. This structure of introductory courses is common in most engineering programmes globally and is designed to provide students with an overview of and the basic skills required in subsequent engineering courses. However, there is a common perception in the literature that these courses often act as gatekeepers, weeding out students who cannot perform at the expectations of faculty.[37] For this reason, we also included first-year

accumulated credits (FYAC) as an indicator of the student's academic performance in the first year.

To progress in a major at UKZN, a student must achieve a minimum number of credits per semester of study from a fixed curriculum. Should a student fail to achieve 75% of their maximum expected credit load at each level of study, they are coded as being at risk.[42] This assignation is based on the assumption that students who pass at least 75% of their accumulated credit load are on track for degree completion in regulation time or regulation time plus one year. Thus, the classification lays the benchmark for acceptable performance levels as well as early identification of underperforming students.[42] Following the same logic predetermined in the data set, we accordingly coded FYAC as: $<75=1$; $\geq 75=2$. We had 62 students whose FYAC was difficult to determine in the data set due to a break taken in the middle of their first year of study. As with AP scores, we opted to treat this group as a missing or floating category[21] that would merge with its most comparable FYAC category. Other covariates included in our model were university residence (0=no; 1=yes) and school quintile (1–5 for Quintiles 1–5).

Decision tree model building

Although there are several statistical algorithms that can be used to build classification trees, in this study, we opted for the decision tree algorithm due to its ability to map hierarchical decisions in an easily interpretable flowchart diagram.[20] Given our focus was more on description, classification and interpretation, we opted to use the chi-square automatic interaction detector (CHAID), which uses the chi-square test to stop tree growth.[21,43]

The decision tree approach aims to create a model that can be used to classify predictor variables using simple decision rules inferred from existing data (also called training data).[21,22] Building a decision tree model starts with a single node or root node whose outcome has a known class, for instance, persistence (yes/no). This node has the highest discriminative power and represents a choice that will result in the partitioning of all cases in the data into two or more mutually exclusive subsets.[19,20] The decision tree algorithm also uses a recursive technique to further partition cases from the root node into branches or subtrees until they cannot be split further, resulting in a leaf or terminal node. A unique feature of the CHAID decision tree approach is that it uses *p*-values with a Bonferroni correction as a splitting criterion.[19] Thus the stopping rules applied in this method automatically account for statistical significance.[19]

Once the tree has been constructed, the performance of the model can be evaluated.[20] SPSS allows for two validation approaches: k-fold cross-validation and the split-sample validation.[43] In this study, we opted for the k-fold cross-validation technique, which divides a data set into several 'n' subsets of equal size; in our case, we chose 10 folds. The decision tree algorithm iteratively classifies each of these subsets using 90% of the data as training data, while 10% is treated as the test sample. In other words, each of the 10% subsets is used once as the test data set, and nine times as part of the training sample, while the rules derived from the 90% training sample are applied to the 10% test sample. A misclassification error is estimated for the training data, while the overall cross-validation risk is calculated as an average of the risk across the 10 test subsamples.[40,41]

Results

Table 2 presents the descriptive statistics of the distribution of outcome variables. The results show that 40% ($n=238$) and 35% ($n=273$) of the students from the 2012 and 2013 cohorts, respectively, graduated. Between 40% ($n=310$; 2013) and 47% ($n=282$; 2012) withdrew from an engineering major, while 8% ($n=50$) and 13% ($n=101$) were excluded due to underperformance. Approximately 5% ($n=27$; 2012) and 12% ($n=89$; 2013) of the students were still registered in the engineering programme in which they initially registered.

Table 2: Descriptive statistics in counts and percentages for persistence patterns in an engineering major ($n=1370$)

Cohort	*n*	Graduated	Withdrawn	Excluded	Currently registered
2012	597	238 (39.9%)	282 (47.2%)	50 (8.3%)	27 (4.5%)
2013	773	273 (35.3%)	310 (40.1%)	101 (13.1%)	89 (11.5%)
Total	1370	511	592	151	116

Note: The withdrawn category is made up of students who either switched or left the university.

Research has shown that the two groups of students with low throughput rates in STEM degrees are likely to be female and black South African students.[17,18] Figure 2 presents two-way histograms for graduation (our persistence variable), withdrawal and exclusion according to these demographic factors. From Figure 2, we can see that for both gender and race, graduation is most likely to occur in year 5, followed by year 4 and year 6 in that order. This shows that the majority of students who graduate from the engineering major do so in regulation time plus one year. However, while the histograms for both male and female students, as well as black South Africans and the 'Other' race category, follow a similar pattern, it is also clear that female and black South African students are less likely to graduate in minimum time than are male students or their peers from the 'Other' race category.

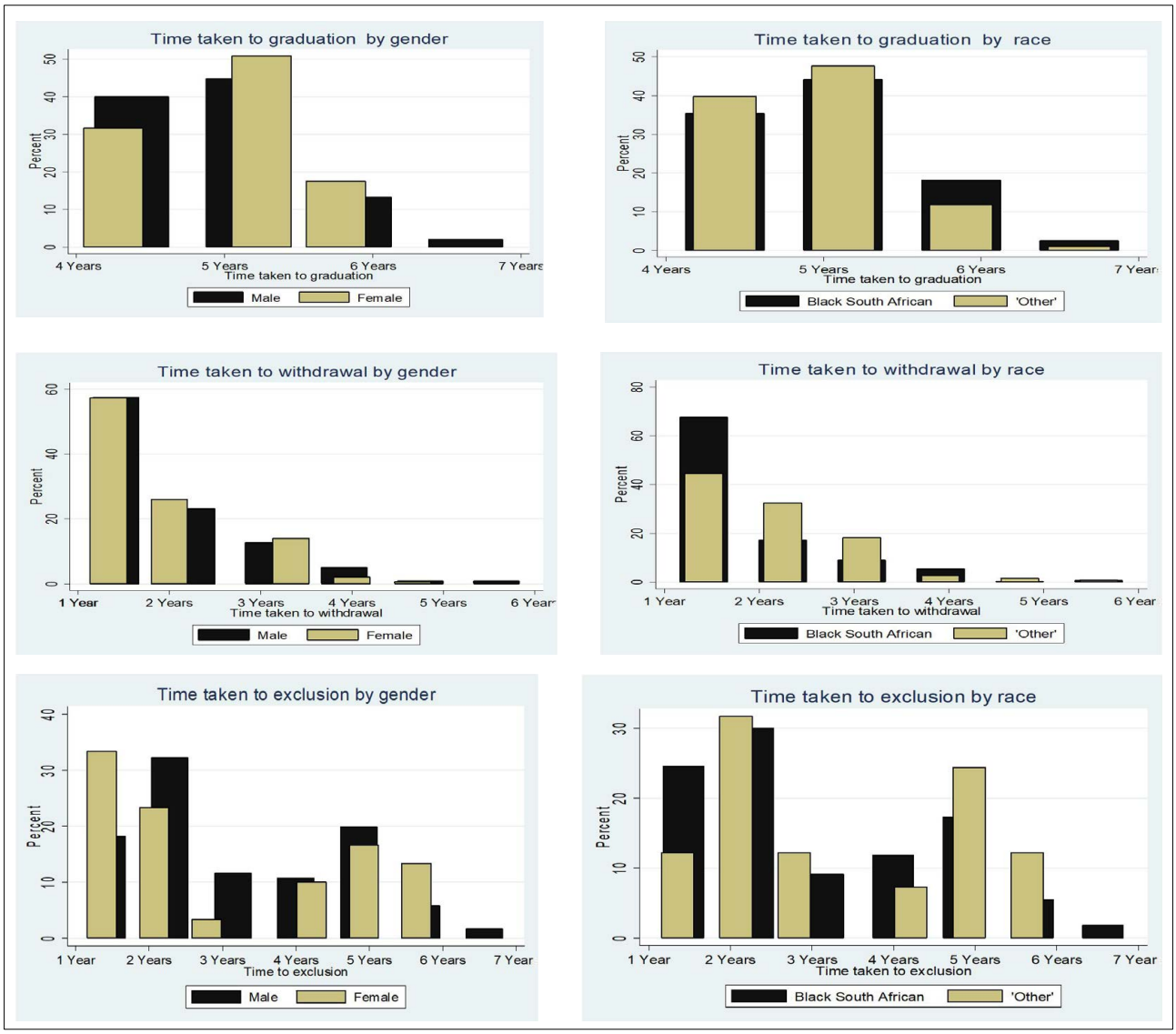
Regarding departure, Figure 2 indicates that most students who leave an engineering major do so at the end of the first year. There are no significant differences between male and female students in the first year. At the same time, black South African students have a higher propensity for dropping out at this level than students from the 'Other' race category. In terms of exclusion, female and Black South African students are more likely to be excluded at the end of the first year than their male and 'Other' peers. This pattern changes in the second year, with more male than female students being excluded, while for race, there is a small difference in the rate of exclusion.

Having ascertained the proportion of students who graduate and drop out, as well as the timing of these events, the next step in our analysis was to explore the trajectories of those who dropped out from the engineering major.

Figure 3 reveals that 69% (407) of the students did not register for any subsequent qualifications. In the absence of additional information on stop-out, these students are taken as dropouts from UKZN. Therefore only 31% (185) of the students registered for alternative qualifications. This translates to 13.5% of the whole sample. We refer to the students who registered for alternative qualifications at the same university as 'switchers'. Tables 3 and 4 present descriptive statistics for (1) withdrawal from an engineering major and (2) switching by gender and race.

Table 3 shows that there was no difference in the withdrawing pattern between male and female students. However, regarding race, 47% ($n=330$) of the black South African students withdrew from the engineering major relative to 39% ($n=260$) of their peers from the 'Other' race groups.

From Table 4, we notice that 53% ($n=71$) of the male students and 56% ($n=28$) of the female students switched to science-related fields such as health sciences, agriculture and chemistry. Likewise, 49% ($n=48$) of black South African students and 59% ($n=52$) of the 'Other' students also moved to science-related fields. A greater percentage of black South African students switched between engineering majors (40%, $n=39$) compared with 'Other' students (6%, $n=5$). Similarly, male students were more likely to switch between engineering majors (26%, $n=35$) than their female counterparts (18%, $n=9$).



Note: Category 'withdrawn' includes students who either switched to other qualifications or left the university.

Figure 2: Two-way histograms for graduation (persistence), withdrawal and exclusion from the engineering major.

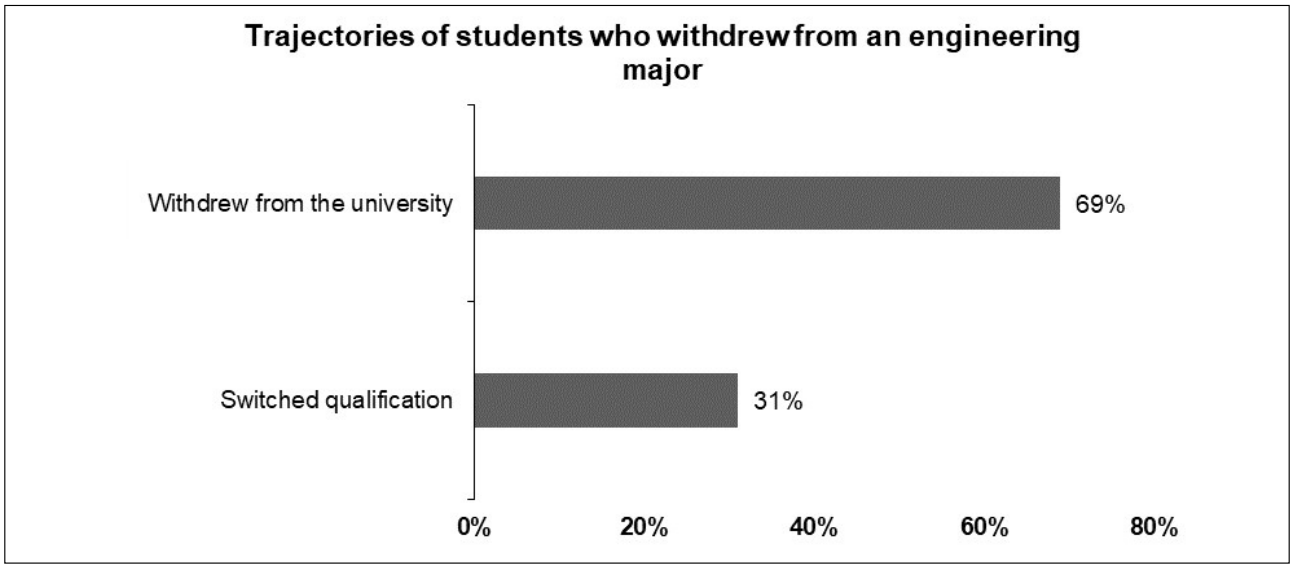


Figure 3: Trajectories of students who withdrew from the engineering major, $n=592$.

Table 3: Proportion of students who withdrew from engineering, by race and gender ($n=1370$)

	Number of students	Number of students who withdrew (%)
Gender		
Male	1031	443 (43%)
Female	339	149 (44%)
Race		
Black South African	703	330 (47%)
Other	667	260 (39%)

Table 4: Destinations of the students who switched from the engineering major ($n=185$)

Variable	*n*	Engineering	Other science	Humanities	Law or Management Studies	Excluded
Gender						
Male	135	35 (26%)	71 (53%)	7 (5%)	22 (16%)	34 (25%)
Female	50	9 (18%)	28 (56%)	3 (6%)	10 (20%)	3 (6%)
Race						
Black South African	97	39 (40%)	48 (49%)	2 (2%)	8 (8%)	33 (34%)
Other	88	5 (6%)	52 (59%)	8 (9%)	23 (26%)	4 (5%)

Approximately 27% ($n=24$) of students in the 'Other' race category switched to a Law or Management Studies major relative to 8% ($n=8$) of black South African students, while 16% ($n=22$) and 20% ($n=10$) of the male and female students, respectively, also switched to major in Law or Management Studies. Only a few of the switching students subsequently enrolled in a Humanities major (less than 10% for both gender and race categories). Of the 185 switchers we had in our sample, 20% ($n=37$) were excluded from the subsequent qualifications in which they registered. The last column in Table 4 shows the disaggregation of the students excluded from the subsequent qualification by race and gender.

Decision tree analysis

Figure 4 presents the results from the decision tree analysis using the CHAID growing method. Figure 4 shows that all the cases were partitioned into nine subgroups from the root node to terminal nodes. The findings can be summarised as follows:

- FYAC is the most critical indicator of persistence amongst the engineering cohort (Nodes 1 and 2). Students who took a break in the middle of their first year of study were also intuitively merged with the category (FYAC=>75=2), while those with inaccurate AP scores were combined with the category 36–39 AP scores.
- Among students who passed 75% of their FYAC, 63.8% persisted, relative to only 11.1% of those who failed to acquire 75% of the FYAC or were in the floating category.
- Among students who passed at least 75% of their FYAC load and had AP scores in the range 40 and above, 79.9% persisted (Node 5).
- Among students who passed at least 75% of their FYAC load and had AP scores in the range 36–39 or missing, 61.7% persisted in their first declared engineering major (Node 4).
- Among students with 75% of FYAC and alternative AP scores or AP scores in the range 30–35, only 40.3% persisted (Node 3).
- For those students who did not achieve 75% of their FYAC load or had taken a break in the middle of their first year of study, and were coded as 'Other', 14.8% persisted (Node 7).
- Among students who did not achieve 75% of their FYAC load or had taken a break in the middle of their first year of study, and were black South Africans, only 8.2% persisted (Node 6).
- Among students coded as 'Other', who failed to achieve at least 75% of the FYAC load or had taken a break in the middle of their first year of study, and had financial aid, 22.7% persisted (Node 9) relative to only 11.3% (Node 8) with the same characteristics but who did not have financial aid.

Table 5 presents the risk error estimates of the CHAID model. The training or re-substitution error, which is a measure of how well the classifier adapts to the training data, was 0.225 (22.5%). This error is generally considered to be optimistically biased and likely to underestimate the misclassification error because it uses only the training data to evaluate the model.[44] Hence, it should be interpreted with care. On the contrary, the cross-validation approach, which uses a resampling approach, estimates the true prediction error in assigning group membership in the model.[44] As shown in Table 5, the cross-validation risk error for our model was 0.234 (23.4%), suggesting that the average risk of misclassifying students (based on all 10 sub-samples using 10-fold cross-validation) was 23%. In other words, our model performed reasonably well in classifying students' persistence patterns.

Table 5: Risk estimate of the classifier

	Estimate	Standard error
Re-substitution	0.225	0.012
Cross-validation	0.234	0.12

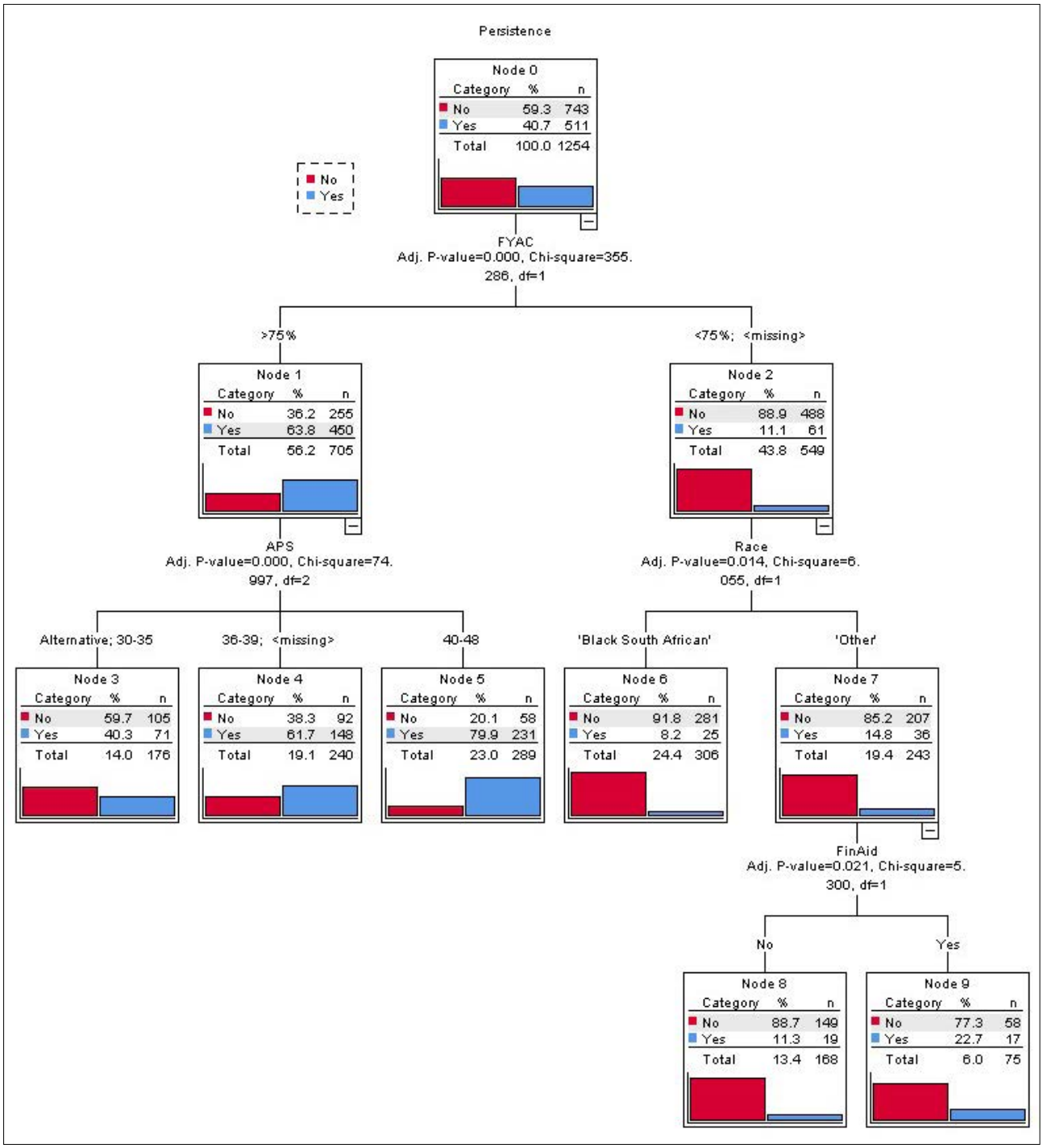


Figure 4: Classification tree model of persistence in engineering, $n=1254$.

Discussion

We tracked a cohort of students beginning in 2012 or 2013 over 7 or 6 years, respectively, to examine their persistence patterns. Our results show that approximately up to 40% of the students persist in the engineering specialisation in which they initially registered. Of those who persisted, the majority graduated after 5 years, which is regulation time plus one year, while withdrawal from an engineering major was most likely to occur in the first year. Thus, our results from the descriptive analysis confirm those of earlier studies which have shown similar persistence trends in South Africa.[4,14-17] We also found that most of the students who left the engineering programme did not register for any subsequent degrees at the same institution, at least within the 6–7-year period under investigation. While there is a possibility that these students might have transferred to other institutions, in the worst-case scenario, they might have been lost to higher education altogether. Of those that switched majors, the majority took up studies in other science-related fields while switching into the humanities was a relatively infrequent occurrence.

Research has shown that switching courses is a natural process for most undergraduate students who are often at the stage of finding the right fit for their academic studies[45] and is more pronounced in STEM disciplines such as engineering[44,45]. For instance, analyses from the USA suggest that 30% of students do not persist in their first declared major[9] and that switching is highest amongst female and African-American students[37,45-47]. However, our findings suggest that, while a significant

proportion of students does not persist in engineering, switching is less common, at least within the same institution, as only 13.5% of our sample switched to other majors. There is a need for further research investigating what happens to the majority of students who depart from an engineering major.

Our results from the CHAID analysis suggest that FYAC and AP scores, race and financial aid status were the most important variables associated with persistence. Gender, school quintile and campus residence were not used in building the tree, suggesting that these factors had the least discriminative power in classifying students at risk of not persisting. Of all our input variables, FYAC load was the most critical factor in determining persistence, reinforcing prior studies which have shown that the first year is critical for on-track graduation.[17,26] For engineering students, the first year is especially important and serves as an initial roadblock to persistence for several reasons. First, the engineering degree is more sequential and structured, such that failure in the first year can derail a student completely, threatening their chances of completing the degree on time.[17,37] Second, first-year introductory science and mathematics courses that provide the background needed for an engineering major hold the distinction of gatekeepers.[37,46] King comments that this weed-out culture 'suggests that grades are used to send a message to low achieving students that they do not belong in the major'[46]. Hence, students respond by either leaving or switching to other programmes.

The role of AP scores in determining success in university has been a subject of much debate in South Africa[26,28,29] with some studies suggesting that they are an unreliable indicator of future success, and others showing the opposite. However, our findings suggest that students who pass 75% of their FYAC load and have high AP scores (above 40) have the most favourable outcomes. Put differently, our findings suggest that the higher the AP score, the better the chance of persisting. Given that the majority of the students enrolled in the engineering programmes at UKZN have AP scores below 40, underpreparedness might be a norm rather than an exception.

Although prior studies have documented an achievement gap in STEM subjects between black South African students and their peers from other racial groups[4,6,15], the decision tree method in our analysis suggests that this effect is more important for those students who fail to acquire at least 75% of their FYAC load. Similarly, we also found that financial aid was an important factor for non-black South African students who did not achieve 75% of their FYAC loads. Thus, our findings suggest that the relationship between race, financial aid and persistence is not as straightforward as often reported in previous studies[16,17]; instead, these factors interact with first-year performance in a critical way. Put differently, focusing on race or access to financial aid without taking into account the whole student experience, can disguise potential disparities that might distinguish students, especially in the first year of study.

The finding that financial aid (Nodes 8 and 9) explained persistence more for students in our composite race category than for black South Africans was rather unexpected. This is because financial aid is one of the strategies that has been put in place to redress past inequalities which have disproportionately affected the black South African students, both in terms of access and success.[40] Hence, we would have expected that financial aid would be an important factor for these students. However, from a methodological perspective, and considering that black South African students make up the majority of financial aid recipients[47], it is highly possible that the algorithm might have been sensitive to the data distribution. Thus, it would not have been a good discriminator of their performance. Again, there is also the possibility of noise in the terminal nodes (8 and 9) as they contain only a few of the remaining cases.[48] In other words, financial aid is the least informative factor in our model, and hence, caution should be taken when interpreting this finding.

Limitations

Although this study advances the way we should think about student persistence, there are some limitations which might affect how the results should be interpreted. The first limitation relates to our choice of covariates. Our analysis was limited to only the covariates that were found in the data we used. Other factors that might influence student persistence, such as family background or language, were not explored. Second, we did not consider a departure from engineering due to transfer to other institutions. It is possible that some of the students who did not re-register in other programmes at UKZN might have enrolled in other institutions. While tracking transfer-outs is currently a challenge, the growing use of institutional data and the recent growth in institutional research and data analytics suggest that this will soon be possible. Potential improvements in the data analysis might also include more predictors to lower the risk of misclassifying students. Third, our study presents a case study of one school in a single institution. Hence, while the results can be informative for other schools within and outside UKZN, we do not claim generalisability beyond the School of Engineering at UKZN. Therefore, context should be taken into consideration when interpreting these results.

Despite these limitations, our findings reinforce the argument that not all students who depart from an academic programme are lost to higher education. Some do switch to other programmes. This study has also shown that performance in the first year can be extremely helpful in identifying students at risk of not persisting to graduation. Thus, academic support should be strongest in the first year of registration and universities should not wait until the end of the first year to assist these students.

Overall, while our study confirms the findings of earlier studies, our main contribution is in showing the viability of decision tree analysis as an alternative way to understand patterns of persistence amongst a cohort of engineering students. Specifically, our analysis revealed that by using machine-learning algorithms such as decision trees, researchers can detect the specific combinations of factors that influence student persistence. Such a model can be used to classify students using both enrolment data as well as the rules derived from each node of the classification tree. For instance, for Node 5, the rules can be written as follows:

If FYAC='>75%' and AP score='40+' THEN outcome='persist'. This information can then be used to timeously identify students who might be at risk of failing and come up with interventions that will support them.

Competing interests

We declare that there are no competing interests.

Authors' contributions

A.V.B. undertook the initial conceptualisation. J.P. provided critical feedback on the conceptualisation, which led to a revised focus. A.V.B. devised the methodology and performed the analysis. Both authors discussed the results and contributed to the final manuscript. A.V.B. took the lead in the write-up of the manuscript, with support from J.P. Both authors provided critical feedback on the literature review, analysis and the discussion.

References

1. Clark JV. Introduction. In: Clark JV, editor. Closing the achievement gap from an international perspective: Transforming STEM for effective education. London: Springer; 2014. p.3–6.
2. Grayson D, Collier-Reed B, Pearce H, Shay S. A curriculum framework for flexible engineering degrees in South Africa. Paper presented at: Enhancement and Innovation in Higher Education Conference; 2013 June 11–13; Glasgow, UK. p.343-352.
3. Fisher G, Scott I. The role of higher education in closing the skills gap in South Africa. Background paper 3 for 'Closing the skills and technology gap in South Africa'. Washington DC: The World Bank; 2011.
4. Council on Higher Education (CHE). A proposal for undergraduate curriculum reform in South Africa: The case for a flexible curriculum structure. Pretoria: CHE; 2013.
5. Shay S. Educational investment towards the ideal future: South Africa's strategic choices. S Afr J Sci. 2017;113(1/2), Art. #2016-0227. http://dx.doi.org/10.17159/sajs.2017/20160227

6. Scott I, Yeld N, Hendry J. A case for improving teaching and learning in South African higher education. Higher Education Monitor no. 6. Pretoria: Council on Higher Education; 2007. Available from: http://www.che.ac.za/documents/d000155/index.php

7. Costandius E, Nell I, Alexander N, McKay M, Blackie RM, Setati E. #Feesmustfall and decolonising the curriculum: Stellenbosch university students' and lecturers' reactions. S Afr J High Educ. 2018;32(2):65–85. https://doi.org/10.20853/32-2-2435

8. Ost B. The role of peers and grades in determining major persistence in sciences. Econ Edu Rev. 2010;29(6):923–934. https://doi.org/10.1016/j.econedurev.2010.06.011

9. McFarland J, Hussar B, Zhang J, Wang X, Wang K, Hein S, et al. The condition of education 2019 (NCES 2019-144). Washington DC: National Center for Education Statistics, US Department of Education; 2019. Available from: https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2019144

10. Weedon E, Riddell S Higher education in Europe: Widening participation. In: Shah M, Bennett A, Southgate E, editors. Widening higher education participation: A global perspective. Kidlington: Chandos Publishing; 2016. p. 49–62.

11. Pascarella ET, Terenzini PT. How college affects students: Findings and insights from twenty years of research. San Francisco, CA: Jossey-Bass; 1991.

12. Tinto V. Dropout from higher education: A theoretical synthesis of recent research. Rev Educ Res. 1975;45:89–125. http://dx.doi.org/10.3102/00346543045001089

13. Organisation for Economic Co-operation and Development (OECD). Education at a glance: OECD indicators. Paris: OECD Publishing; 2017. https://doi.org/10.1787/eag-2017-en

14. Pocock J. Leaving rates and reasons for leaving in an engineering faculty in South Africa: A case study. S Afr J Sci. 2012;108(3–4), Art. #634. http://dx.doi.org/10.4102/sajs.v108i3/4.634

15. Fisher G. Improving throughput in the engineering bachelor's degree: Report to the Engineering Council of South Africa [document on the Internet]. c2011 [cited 2020 Jul 07]. Available from: https://www.ecsa.co.za/about/pdfs/091211_ECSA_Throughput_Report.pdf

16. Murray M. Factors affecting graduation and student dropout rates at the University of KwaZulu–Natal. S Afr J Sci. 2014;110(11/12), Art. #2014-0008. https://doi.org/10.1590/sajs.2014/20140008

17. Bengesai AV, Paideya V. An analysis of academic and institutional factors affecting graduation among engineering students at a South African university. Afr J Res Math Sci Technol Educ. 2018;22(2):137–148. https://doi.org/10.1080/18117295.2018.1456770

18. Zewotir T, North D, Murray M. Student success in entry-level modules at the University of KwaZulu-Natal. S Afr J High Educ. 2011;25(6):1233–1244.

19. Antipov E, Pokryshevskaya E. Applying CHAID for logistic regression diagnostics and classification accuracy improvement. J Target Meas Anal Mark. 2010;18:109–117. https://doi.org/10.1057/jt.2010.3

20. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Belmont, CA: Wadsworth Inc.; 1984.

21. Ritschard G. CHAID and earlier supervised tree methods. In: McArdle JJ, Ritschard G, editors. Contemporary issues in exploratory data mining in behavioral sciences. New York: Routledge; 2013. p.48–74.

22. Uddin S, Khan A, Hossain M, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inform Decis Mak. 2019;19:281. https://doi.org/10.1186/s12911-019-1004-8

23. Kirby NF, Dempster ER. Using decision tree analysis to understand foundation science student performance. insight gained at one South African university. Int J Sci Educ. 2014;36(17):2825–2847. https://doi.org/10.1080/09500693.2014.936921

24. Nudelman Z, Moodley D, Berman S. Using Bayesian networks and machine learning to predict computer science success. In: Kabanda S, Suleman H, Gruner S, editors. ICT education. SACLA 2018. Communications in computer and information science. Cham: Springer; 2019. p. 207–222. https://doi.org/10.1007/978-3-030-05813-5_14

25. Delen D. Predicting student attrition with data mining methods. J Coll Stud Ret. 2011;13(1):17–35. https://doi.org/10.2190/CS.13.1.b

26. Jawitz J. Performance in first- and second-year engineering at the University of Cape Town. S Afr J High Educ. 1995;9:101–108.

27. Maree JG. Career counselling in the 21st century: South African institutions of higher education at the crossroads. S Afr J High Educ. 2009;23(3):436–458.

28. Naidoo P, Motala N, Joubert R. Matriculation scores as an indicator of academic success in an occupational therapy education programme. S Afr J Occup Ther. 2013;43(1):21–25.

29. Mthimunye KDT, Daniels FM. Performance in grade 12 mathematics and science predicts student nurses' performance in first-year science modules at a university in the Western Cape. Curationis. 2017;40(1), Art. #1776. https://doi.org/10.4102/curationis.v40i1.1776

30. Bengesai A. Engineering students' experiences of supplemental instruction: A case study. Alternation. 2011;18(2):59–77.

31. Case J, Marshall D, Grayson D. Mind the gap: Science and engineering education at the secondary–tertiary interface. S Afr J Sci. 2013;109(7/8), Art. #2012-0113. http://dx.doi.org/10.1590/sajs.2013/20120113

32. Monnapula-Mapesela M. Students' perception of own preparedness for higher education. Int J Educ Sci. 2015;9(2):255–264. https://doi.org/10.1080/09751122.2015.11890315

33. South Africa. Amended national norms and standards for school funding. Government Gazette 31496(1087). Pretoria: Government Printers; 2008.

34. Spaull N. Schooling in South Africa: How low-quality education becomes a poverty trap. In: De Lannoy A, Swartz S, Lake L, Smith C, editors. South African child gauge. Cape Town: Children's Institute, University of Cape Town; 2015. p. 34–40.

35. Van der Berg S. How effective are poor schools? Poverty and educational outcomes in South Africa. Stud Educ Eval. 2008;34:145–154. https://doi.org/10.1016/j.stueduc.2008.07.005

36. Pocock J, Bengesai AV, Moodley M. Barriers to progression: A longitudinal study of an access programme for engineering. Proceedings of the SASEE Conference. Stellenbosch: SASEE; 2011. p. 164–173.

37. Sanabria T, Penner A. Weeded out? Gendered responses to failing calculus. Soc Sci (Basel). 2017;6(2):47. https://doi.org/10.3390/socsci6020047

38. Keswell M. Education and racial inequality in post-apartheid South Africa. In: Attewell P, Newman KS, editors. Growing gaps: Educational inequality around the world. Oxford: Oxford University Press; 2010. p. 82–104.

39. University of KwaZulu-Natal. UKZN at a glance [document on the Internet]. c2017 [cited 2020 Jul 07]. Available from: https://www.ukzn.ac.za/wp-content/uploads/2017/10/At-a-Glance-2017.pdf

40. Sehoole C, Adeyemo KS. Access to, and success in, higher education in post-apartheid South Africa: Social justice analysis. J High Educ Afr. 2016;14:1–18.

41. Naidoo A, McKay TJM. Student funding and student success: A case study of a South African university. S Afr J High Educ. 2018;32:158–172. http://dx.doi.org/10.20853/32-5-2565

42. University of KwaZulu-Natal. University of KwaZulu-Natal: Academic monitoring and exclusions policy and procedures [webpage on the Internet]. c2009 [cited 2020 Jul 07]. Available from: http://registrar.ukzn.ac.za/Libraries/policies/Academic_Monitoring_Exclusion_Policy_Procedure_-_CO06041209.sflb.ashx

43. IBM Corporation. IBM SPSS Decision Trees 25. New York: IBM Corporation; 2015.

44. Berrar D. Cross-validation. Encyclopedia of bioinformatics and computational biology. 2018;1:542–545. https://doi.org/10.1016/B978-0-12-809633-8.20349-X

45. Astoni-Figari C, Speer JD. Drop out, switch majors, or persist? The contrasting gender gaps. Econ Lett. 2018;164:82–85. https://doi.org/10.1016/j.econlet.2018.01.010

46. King B. Changing college majors: Does it happen more in STEM and do grades matter? J Coll Sci Teach. 2015;44(3):44–51.

47. Geisinger BN, Raman D. Why they leave: Understanding student attrition from engineering majors. Int J Eng Educ. 2013;29:914–925.

48. Perner P. How to interpret decision trees? In: Perner P, editor. Advances in data mining: Applications and theoretical aspects. ICDM 2011. Lecture Notes in Computer Science. Berlin/Heidelberg: Springer; 2011. p. 40–55.

AUTHORS:
Shariefa Hendricks[1]
Duncan J. Cartwright[2]
Richard G. Cowden[3]

AFFILIATIONS:
[1]College of Law and Management Studies, University of KwaZulu-Natal, Durban, South Africa
[2]School of Applied Human Sciences, University of KwaZulu-Natal, Durban, South Africa
[3]Department of Psychology, University of the Free State, Bloemfontein, South Africa

CORRESPONDENCE TO:
Shariefa Hendricks

EMAIL:
Hendrs@unisa.ac.za

DATES:
Received: 25 Sep. 2019
Revised: 26 May 2020
Accepted: 20 Sep. 2020
Published: 29 Mar. 2021

HOW TO CITE:
Hendricks S, Cartwright DJ, Cowden RG. Clinical supervision in South Africa: Perceptions of supervision training, practices, and professional competencies. S Afr J Sci. 2021;117(3/4), Art. #7428. https://doi.org/10.17159/sajs.2021/7428

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:
competency, intern, professional development, counselling supervision, accredited training

FUNDING:
UKZN University Teaching and Learning Office (TLCRG15)

Clinical supervision in South Africa: Perceptions of supervision training, practices, and professional competencies

We investigated South African clinical and counselling psychology supervisors' ($n=44$) perceptions of supervision training, their supervision experiences, and their perceived competence, confidence and effectiveness in providing supervision. Results indicated that many supervisors prematurely engage in supervision responsibilities and initiate supervision prior to receiving formal training in supervision. With limited regulatory guidelines available on supervision training and practices in South Africa, the findings indicate a need for the South African psychology profession to establish a formal regulatory framework on supervision training and practices. This includes identifying supervision training needs, developing training programmes, and instituting formal training requirements for practitioners who participate in clinical supervision.

Significance:

- The sample of South African clinical and counselling psychologists involved in the supervision of trainee psychologists tended to engage in clinical supervision in advance of obtaining three years of independent clinical practice and prior to receiving appropriate training in providing supervision.
- There is a need for the Professional Board for Psychology of the Health Professions Council of South Africa (PBP-HPCSA) to appropriately monitor and enforce ethical obligations of psychologists who engage in supervision of trainee psychologists.
- Psychologists who provide clinical supervision to trainee psychologists ought to take personal responsibility for ensuring that they are appropriately trained and have acquired the necessary competencies to provide supervision before deciding to engage in supervision activities.
- Formal guidelines and policies regulating clinical supervision are necessary for ensuring psychology supervisors obtain appropriate training in supervision and fulfil mandatory HPCSA-accredited supervision training requirements.

Introduction

In recent years, considerable effort has been dedicated to frameworks, guidelines and regulations on competency-related requirements (e.g. training) of psychologists who engage in clinical supervision.[1] In comparison to more developed countries (e.g. the UK, USA), South Africa has lingered behind global initiatives to improve standards of supervision practices.[2] In South Africa, the directed master's degree in clinical, counselling, education or industrial psychology comprises two years. All trainee psychologists must complete a one-year coursework master's degree at an accredited university. In the second year of their studies, as psychology *interns*, trainee psychologists are required to undertake a 12-month supervised internship at a Health Professions Council of South Africa (HPCSA)-accredited institution.[3]

At present, the single competency prerequisite for psychologists prescribed by the regulatory body of health professions in South Africa (i.e. the Professional Board for Psychology of the Health Professions Council of South Africa (PBP-HPCSA)) is completion of at least three years of independent practice prior to supervising interns. The PBP-HPCSA[4] considers a psychology *supervisor* or *supervising psychologist* as 'a senior psychologist who has been registered as a psychologist with the Board for more than three years'[4(p.1)]. As such, supervision training represents a desirable qualification rather than an essential one. Presently, there appears to be an implicit expectation that supervisors draw on past supervision experiences and their skills as clinicians. This raises ethical concerns about competency and commencing supervisory responsibilities without relevant training.[5]

As per the PBP-HPCSA's clinical guidelines for internship, supervisors are required to work with psychology interns as they commence their clinical work during their 12-month internship by providing structured clinical supervision for at least two hours per week. The PBP-HPCSA offers clear ethical guidelines on the role of supervisors in (1) the professional development of the trainee and (2) ensuring the welfare of their trainees' clients. Psychologists undertaking supervisory responsibilities are obligated to provide supervision within the boundaries of their competence, training, education and supervised experience.[3,6] Supervisors also serve as gatekeepers to the profession by ensuring that trainee psychologists are deemed competent to practice.

Supervision is a critical teaching method and a core component of training mental health professionals.[7] Clinical supervision has emerged as a distinct field comprising unique theoretical perspectives, processes and skills, out of which recognised international standards of practice have developed.[1,8,9] However, very little research on clinical supervision has been conducted globally, and much of the existing body of work has focused on developed nations. Since the publication of the *Standards for Counselling Supervisors*[10], the transformation and evolution from traditional supervision practice to competency-based supervision has spotlighted both the recognition and importance of supervision training[2,11]. The heightened interest in competency-based clinical supervision has

also been stimulated by the formal development and implementation of regulatory guidelines and best practices in a number of developed nations, including the British Psychological Society, New Zealand Psychologists Board, Psychology Board of Australia, and the American Psychological Association.[1,11]

Salience of supervision training

The importance of training and competence in clinical supervision has gained considerable interest[12], with specific attention devoted to the effects of training on supervisor and supervisee development[13]. Supervisor development is regarded as a developmental process that emphasises the acquisition of skills with structured training.[14] Despite the importance of supervision in the professional training and development of psychology interns, formal training in supervision has generally been neglected as a core competency area in psychology.[12] Several concerns involving competency to supervise have been raised, particularly supervisors' clinical competence and clinical supervision practices.[15,16] In fact, a growing body of research has highlighted the negative implications of ineffective supervision experiences and the challenges faced by supervisors in the professional development of trainees[17,18], which underscores the need for training in supervision. A number of studies have reported an increase in the incidence of inadequate, ineffective, and harmful supervision experienced during psychology training placements.[19] In one recent study involving a sample of South African trainee psychologists, Hendricks and Cartwright[20] found evidence of inadequate and potentially harmful supervision among psychology interns in South Africa. Research on supervision training continues to grow, yet research output on clinical supervision has been disproportionately lower in less developed nations, including South Africa.[21] While numerous studies have focused on challenges facing clinical psychology training and mental health service provision in the country[22,23], insufficient effort has been directed to understanding the extent and quality of training for supervisors[24].

Some scholars have suggested that while training in supervision is far more readily and widely recognised, it remains an anomaly.[9] In South Africa, supervision training is not considered a mandatory requirement for South African psychologists who supervise psychology trainees, and explicit regulations on supervision practices have not been promulgated by the PBP-HPCSA. This is surprising, given the compelling evidence in support of the potential efficacy of supervision training. First, it enhances supervisors' theoretical and conceptual knowledge, psychological readiness (e.g. self-confidence, self-awareness), and development of complex supervisory skills and techniques.[25] This is purported to improve supervisors' understanding of their own supervision experiences and the supervision services offered to future psychologists.[2] Second, there has been some support for the notion that supervision training increases motivation to supervise.[26] Third, supervision training increases supervisory self-efficacy and lowers anxiety about the supervisory role.[27-30] Conversely, a lack of education and training in supervision theory and practice may lead to insufficient supervisory preparation, which can compromise the supervisory relationship and inhibit trainee growth and clinical competence.[31] Without training, novice supervisors may be disadvantaged by their limited experience, may be more inclined to perpetuate mistakes made by their own supervisors[32], may not be aware of the need to develop proper supervision contracts, or may fail to participate in the requisite monitoring and evaluation of trainees.[33]

Developments in supervision training

Training and competence in supervision augment the supervisory process by ensuring supervisors are equipped with the necessary theoretical, organisational and management skills.[34] However, the importance of supervision training has historically been underemphasised.[2,35] Recently, Counselling and Psychology Boards (e.g. American Psychological Association's Division of Psychotherapy) have advanced the notion of *competent* practice[1,25] and begun to mandate training and acquisition of skills germane to providing supervision. For example, the Council for Accreditation of Counseling and Related Educational Programs in the USA has made supervision training a mandatory requirement for counsellors at the doctoral level. Some international psychology boards (e.g. Psychology Board of Australia) also require psychologists to complete an approved training programme in the provision of supervision before being permitted to obtain board-level approval to offer supervision. Some scholars have stipulated mandatory peer supervision, professional accreditation, and licensure as *prerequisites* to supervise.[12,36]

More recent emphasis on competency-based practices of supervision represent a 'culture change' that as developed from reflections on the present status of professional psychology education and training.[37] The competencies framework proposed by Falender et al.[38] offers an intentional, systematic approach to supervision practice. The competency-based model is strengths-based and adopts a reflective, mindful approach to supervision. The model specifies a number of elements that should be addressed to ensure adequate training and development of the trainee and the role of the supervisor in that process.[1] Competence also includes preventing and managing supervisee vicarious traumatisation, engaging in self-care, providing corrective and positive feedback, and managing and evaluating impaired or incompetent supervisees.[1] Of the six core competencies outlined by Falender et al.[38] – *knowledge, skills, values, social context, overarching issues,* and *assessment of supervision competencies* – five assert that providing competent supervision is an ongoing, complex process. Although this framework represents an important point of reference, more discourse is needed to determine the factors that contribute to competent clinical supervision.[39] Training in supervision is markedly different from training in the provision of psychotherapy; it requires a substantive shift in vision and perspective.[9] However, much like other areas of clinical training, training in supervision also requires designated curricula that guide supervision training and practices. In South Africa, training in the provision of supervision has not been prioritised. Presumably, work demands, teaching and lecturing commitments, high caseloads, limited staff capacity, resources, and budgetary or time constraints in public institutions limit opportunities for and availability of supervision training. Offering evidence in support of this, a qualitative study involving eight South African psychology internship supervisors found that none of the participants had received any formal training in supervision.[40] These findings warrant further examination, as perceptions of psychologists' ability to provide supervision in South Africa have historically relied on the assumption that skills and knowledge gained from the clinical domain will automatically transfer to the supervisory domain.[41] Although emergent research has speculated that inadequate training in supervision is tied to reduced confidence and self-perceived competence in providing supervision[40], a comprehensive examination of South African internship supervisors' confidence, competence and effectiveness in the practice of supervision is necessary.

The present study

There have been several recommendations to methodologically examine and apply evidence-based practices to psychotherapy supervisor training.[17,42] To date, there have been few attempts to examine the state of supervision training and practices in South Africa. Research suggests that the importance of training in supervision is gaining recognition, with many international credentialing and licensing boards monitoring the practice of supervision.[1,43] Given that formal training in supervision is not currently implemented in South Africa and there are no formal requirements for training in supervision for clinicians mandated by the PBP-HPCSA in order to supervise psychology trainees, an improved understanding of the current state of supervision training and practices in South Africa is needed. Thus, the purpose of this study was to examine South African internship supervisors' perceptions of (1) supervision training and practices in South Africa and (2) their own supervision training, supervision experiences, and perceived abilities (e.g. competence, confidence, and effectiveness) in providing supervision.

Method

Participants

The sample consisted of 44 clinical and counselling psychology supervisors who were licensed to practise in South Africa. Participants were stratified into the following age groups: 25 to 34 (15.91%), 35 to 45 (45.45%), and 46 to 65 (38.64%) years of age. The majority of the sample consisted of female participants (68.18%). Participants self-identified as

black (11.36%), coloured (15.91%), Indian (25.00%), white (40.91%), or 'other' (2.27%, unspecified race = 4.55%). Supervisors were employed at local universities (54.55%) or hospitals (36.36%), or were in private practice (9.09%) and had been practising for between 1 and 36 years (mean=14.28 years, s.d.=7.97). Participants were required to have engaged in intern supervision for a minimum of a year (mean=10.33 years, s.d.=7.23) in order to participate in the study.

Materials

The instrument used in this study contained a range of items that formed part of a broader project on intern training experiences and supervisors' experiences of supervision in South Africa. The items included in the current study were developed to measure different aspects of supervision training based on content areas (detailed below) explored in prior research.[44] Participants were asked to respond to the items with reference to their supervision experience as psychologists and their experience as internship supervisors. After developing the initial set of items, a pilot study was performed to assess for content validity. The pilot sample (n=5, female=60%) consisted of private (40%) and public sector clinical (60%) and counselling psychologists ($mean_{age}$=40.40, $s.d._{age}$=3.56) with a mean of 10.60 years (s.d.=5.94) of intern supervision experience. They provided qualitative feedback to several open-ended questions that was used to make appropriate revisions to the survey items.

The final set of items included in this study assessed a range of domains, including supervision *training and experience* (14 items; e.g. 'Had you had any formal training in supervision prior to assuming supervisory responsibilities?'), perceptions of the *supervisory relationship and intern training* (3 items; e.g. 'Which of the following form part of your training techniques in supervision?'),_*supervisor self-ratings on various supervision skills* (4 items; e.g. 'How would you rate the priority of supervision in comparison to your other professional tasks?'), Supervisor *perceived suitability of interns* (1 item; e.g. 'How confident are you in supervising interns perceived as incompetent?'), and intern evaluation and feedback (2 items; e.g., 'Rate your competence in evaluating interns'). Participants rated the items using the rating scales that were specific to each item (see Tables 1 to 4 for details on response options for each item).

Procedure

Ethical clearance to conduct the study was obtained from the University of KwaZulu-Natal's Human Social Sciences Research Ethics Committee (ethical clearance number HSS/1350/013D). We contacted and obtained permission from professional bodies and organisations, including the South African Association for Counselling and Development in Higher Education, HPCSA-accredited higher education institutions, public hospitals, and the Psychological Society of South Africa, to recruit participants from their respective databases of psychologists who met criteria for inclusion in this study. To be eligible, participants needed to (1) have professional registration within the clinical or counselling psychology scope of practise, (2) have at least one year's experience as a clinical supervisor, (3) have actively engaged in the supervision of psychology interns, and (4) be employed in either the private or public sector. A combination of purposive and snowball sampling techniques was used to recruit participants. Using this approach, 152 eligible supervisors working in the public and private sector were initially contacted to participate in this study. Although 75 supervisors agreed to participate, a total of 44 complete responses were received. Based on those who participated, the final response rate was 29%. Prospective participants were provided with a weblink that directed them to a secure data collection website, which ensured that participation was entirely anonymous. After providing online consent to participate in the study, participants completed the survey items. The items included in this study were a subset of items that formed part of a larger project. Participation was voluntary and participants were informed that they could withdraw from the study at any time without any penalties. All data were stored in password-protected electronic format.

Data analyses

Missing data diagnostics revealed 3.66% of missing values across 29.55% of cases. Analyses were computed using a pairwise deletion approach. Descriptive statistics were produced for all study variables. Bivariate associations were estimated using Spearman correlations, and group differences were detected using Wilcoxon rank-sum tests. Estimates of effect size are reported alongside bivariate inferential analyses. Values of 0.20, 0.50 and 0.80 for r_s, and values of 0.04, 0.25 and 0.64 for epsilon squared (ε^2), represent small, medium and large effects, respectively.[45] All statistical analyses were performed in R[46], with alpha set to 0.05.

Results

Univariate analyses

Supervision training

The results on supervision training are reported in Table 1. Only a small portion of participants (16.28%) indicated that they had received formal training in supervision prior to assuming supervision responsibilities, whereas a larger proportion (46.51%) reported receiving some form of training in supervision after assuming supervisory responsibilities. Among those who had received training in supervision, the most frequently cited mode of training was via a workshop (90.48%). The highest proportion (38.10%) reported being *mostly satisfied* with the training in supervision they had received. As an indicator of self-initiated continued professional development in supervision, a large proportion of the participants (44.19%) reported reading one scholarly article or book on supervision each month. However, 34.88% noted they did not engage in any scholarly reading on supervision. The highest proportion of participants (41.86%) also reported being *somewhat prepared* for supervision based on their master's-level training. Supervisors tended to rate the importance of receiving training in supervision as *extremely important* (53.85%), and the majority (88.64%) indicated that training in supervision should be mandatory.

To further examine areas of supervision training considered important to participants, those who reported that supervision training should be mandatory completed additional training component items they considered important to include in a supervision training programme (see Table 2). The training components that received the highest ratings of importance included ethics in supervision, managing supervisee resistance, conflict and power issues in supervision, dealing with boundary violations, and assessing and evaluating competencies. Comparably, components of supervisor training that were given lower ratings of importance were models of supervision, supervisory styles, and contracting in supervision.

Supervision experience

The average number of years of independent practice experience reported by the participants prior to engaging in supervisory responsibilities was 2.14 years (s.d.=0.94). Most participants (79.07%) had less than the HPCSA's designated three-year minimum independent practice experience requirement before initiating supervision responsibilities. On average, participants reported supervising more than 30 (s.d.=34.87) trainee psychologists during their careers and currently supervised 2 trainee psychologists (s.d.=1.50) for an average of 2.35 (s.d.=1.24) hours a week.

Supervision abilities and practices

The majority of supervisors (65.12%) indicated that their own positive experiences as supervisees *often* influence the manner in which they supervise, and the highest proportion of participants (38.64%) also reported negative experiences *often* influencing their supervision practices (see Table 3). All participants felt they had become more competent (100%) and confident (95.35%) in supervision over time, and the majority of supervisors (55.81%) perceived themselves to be *somewhat effective* in providing supervision. Most participants (63.41%) indicated that they did not make use of a formal model of supervision, although most (81.82%) noted that they sought supervision relating to their own supervisory performance.

Table 1: Frequency statistics for supervision training items

Item	*n* (%)
Training in supervision (before assuming supervisory responsibilities)	
No	36 (83.72%)
Yes	7 (16.28%)
Training in supervision (after assuming supervisory responsibilities)	
No	23 (53.49%)
Yes	20 (46.51%)
Type of training received in supervisionab	
Degree or diploma	1 (4.76%)
Certificate course	1 (4.76%)
Coursework module	0 (0.00%)
Workshop	19 (90.48%)
Lecture series	3 (14.29%)
Online training	1 (4.76%)
Satisfaction with received training in supervisiona	
Extremely satisfied	0 (0.00%)
Very satisfied	6 (28.57%)
Mostly satisfied	8 (38.10%)
Mildly satisfied	7 (33.33%)
Quite dissatisfied	0 (0.00%)
Number of articles read a month	
0	15 (34.88%)
1	19 (44.19%)
2	2 (4.65%)
3	7 (16.28%)
Preparation of master's degree for supervision	
Extremely well prepared	9 (20.93%)
Very well prepared	15 (34.88%)
Somewhat prepared	18 (41.86%)
A little prepared	1 (2.33%)
Not at all prepared	0 (0.00%)
Importance of training in supervision	
Extremely important	21 (53.85%)
Very important	13 (33.33%)
Moderately important	3 (7.69%)
Slightly important	2 (5.13%)
Not at all important	0 (0.00%)
Supervision training mandatory	
Yes	39 (88.64%)
No	2 (4.55%)
Don't know	3 (6.82%)

*Note: [a]Includes only those participants who reported receiving training in supervision either before or after assuming supervisory responsibilities (*n*=21); [b]Sum of subcategory percentages may exceed 100%.*

Table 2: Frequency statistics (*n* (%)) for perceptions of the importance of supervisor training components

Item	Item responses			
	Extremely important	Very important	Somewhat important	Not at all important
Theory and practice of supervision	25 (62.50%)	10 (25.00%)	5 (12.50%)	0 (0.00%)
Models of supervision	15 (37.50%)	17 (42.50%)	7 (17.50%)	1 (2.50%)
Supervisory styles	19 (47.50%)	16 (40.00%)	5 (12.50%)	0 (0.00%)
Ethics in supervision	28 (70.00%)	12 (30.00%)	0 (0.00%)	0 (0.00%)
Managing supervisee resistance, conflict and power issues in supervision	28 (70.00%)	10 (25.00%)	2 (5.00%)	0 (0.00%)
Managing transference and countertransference	24 (60.00%)	12 (30.00%)	4 (10.00%)	0 (0.00%)
Dealing with boundary violations	27 (67.50%)	10 (25.00%)	2 (5.00%)	1 (2.50%)
Assessing and evaluating competencies	27 (67.50%)	12 (30.00%)	1 (2.50%)	0 (0.00%)
Contracting in supervision	19 (47.50%)	15 (37.50%)	6 (15.00%)	0 (0.00%)
Report writing	20 (50.00%)	16 (40.00%)	4 (10.00%)	0 (0.00%)
Diversity/multicultural aspects of supervision	23 (57.50%)	12 (30.00%)	5 (12.50%)	0 (0.00%)

The highest proportion of supervisors (48.84%) indicated that supervision was a *high priority* responsibility relative to other professional tasks. Almost half of the participants (46.51%) considered supervision to be *moderately challenging*.

With regard to the types of formal training techniques that participants used while providing supervision (see Table 4), the most commonly utilised supervision training technique was in-session observation (72.73%). The majority of supervisors also used audio taping (63.64%), role playing (56.82%), and supervision contracts (56.82%) as part of their training protocols. A minority of supervisors reported using video recordings (45.45%) and a two-way mirror (40.91%) to facilitate training of supervisees.

Bivariate analyses

There were no significant differences in the quantity of scholarly articles and books on supervision read among participants who had received formal training in supervision, as compared to those who had not (see Table 5). There were no differences in perceived competence, confidence, or effectiveness in supervision between participants who had received formal training in supervision and those who had not. Supervisors who reported becoming more competent (r_s=0.52, 95% CI [0.26, 0.71], $p<0.05$) and confident (r_s=0.35, 95% CI [0.05, 0.59], $p<0.05$) in providing supervision over time tended to rate themselves as more effective in providing supervision. Participants who reported higher levels of prioritisation of supervision reported higher levels of self-perceived effectiveness in providing supervision (r_s=0.40, 95% CI [0.11, 0.63], $p<0.05$). There was no evidence to suggest that confidence in supervising interns perceived as incompetent (r_s = 0.08, 95% CI [-0.22, 0.37], $p>0.05$), or self-perceived competence in evaluating interns (r_s = 0.02, 95% CI [-0.28, 0.31], $p>0.05$), are associated with years of supervisory experience. The extent to which participants perceived supervision as more challenging was unrelated to perceived effectiveness in providing supervision (r_s = -0.16, 95% CI [-0.44, 0.15], $p>0.05$).

Discussion

In this study, we investigated supervisors' perceptions of supervision training, their supervision experiences, and their perceived abilities in providing supervision to obtain an understanding of the current state of supervision training and practice in South Africa. Overall, findings suggest that South African clinical supervisors perceive themselves to be at least somewhat effective in their supervisory engagements and appear to have developed confidence and competence over time. While internship supervisors considered supervision important and a high-priority professional responsibility, a negligible number of participants undertook formal training in supervision prior to assuming their supervisory responsibilities.

Formal training in supervision

A major finding of this study is the relative lack of supervision training reported by internship supervisors. Falender et al.[12] consider training in supervision beneficial to perceived competence and effectiveness in providing supervision. Evidence suggests that supervisors are prone to perceiving themselves as underprepared for the supervisory role.[47] The findings of this study are concerning, given research suggesting that practitioners with more training in supervision have a stronger understanding of the supervision process and the supervisory relationship, are more psychologically prepared (e.g. confident and motivated) and effective in providing supervision, and engage supervisees with less criticism and dogmatism.[48,49] Therefore, training or preparation for supervision may enhance trainees' clinical training experiences, thereby increasing the likelihood that emerging practitioners acquire necessary professional skills and competencies via supervisors who have relevant knowledge and skills in the provision of supervision.[20]

Supervision training opportunities

Although the majority of supervisors included in this study rated training in supervision as highly important and indicated that supervision training should be mandated by the HPCSA, many had not completed formal training prior to assuming supervision responsibilities. This finding is consistent with supervision training patterns evidenced in prior studies involving non-South African samples[48,50], but likely stems from the paucity of formal training opportunities available to supervisors in South Africa. Of those supervisors who had received some form of supervision training prior to initiating supervision practices, many reported attending professional workshops.

Table 3: Frequency statistics for perceptions of supervision experience, training, and abilities items

Item	*n* (%)
Influence of positive experiences as supervisee	
Not at all	1 (2.32%)
Rarely	1 (2.32%)
Occasionally	4 (9.30%)
Sometimes	9 (20.93%)
Often	28 (65.12%)
Influence of negative experiences as supervisee	
Not at all	3 (6.82%)
Rarely	6 (13.64%)
Occasionally	9 (20.45%)
Sometimes	9 (20.45%)
Often	17 (38.64%)
Competence in supervision over time	
Yes, definitely	24 (54.55%)
Yes, I think so	20 (45.45%)
No, I don't think so	0 (0.00%)
No, definitely not	0 (0.00%)
Confidence in supervision over time	
Yes, definitely	20 (46.51%)
Yes, I think so	21 (48.84%)
No, I don't think so	1 (2.33%)
No, definitely not	1 (2.33%)
Effectiveness in providing supervision	
Very effective	11 (25.58%)
Somewhat effective	24 (55.81%)
Somewhat ineffective	8 (18.60%)
Very ineffective	0 (0.00%)
Formal model of supervision	
No	26 (63.41%)
Yes	15 (36.59%)
Sought supervision	
No	8 (18.18%)
Yes	36 (81.82%)
Prioritisation of supervision	
Very high priority	16 (37.21%)
High priority	21 (48.84%)
Moderate priority	4 (9.30%)
Low priority	0 (0.00%)
Very low priority	2 (4.65%)
Role as a supervisor challenging	
Extremely challenging	4 (9.30%)
Very challenging	13 (30.23%)
Moderately challenging	20 (46.51%)
Slightly challenging	6 (13.95%)
Not at all challenging	0 (0.00%)
Confidence in supervising interns perceived as incompetent	
Extremely confident	2 (4.55%)
Very confident	12 (27.27%)
Moderately confident	16 (36.36%)
Somewhat confident	9 (20.45%)
Not at all confident	5 (11.36%)
Competence in evaluating interns	
Extremely competent	5 (11.63%)
Very competent	20 (46.51%)
Moderately competent	16 (37.21%)
Somewhat competent	2 (4.65%)
Not at all competent	0 (0.00%)

Table 4: Frequency statistics (*n* (%)) for supervision training technique items

Item	Item responses	
	No	Yes
Audio taping	16 (36.36%)	28 (63.64%)
Video recording	24 (54.55%)	20 (45.45%)
Role playing	19 (43.18%)	25 (56.82%)
In-session observation†	12 (27.27%)	32 (72.73%)
Two-way mirror	26 (59.09%)	18 (40.91%)
Supervision contract	19 (43.18%)	25 (56.82%)

†Supervisor sits in during supervisee session with client

Table 5: Summary statistics for Wilcoxon rank-sum tests on training in supervision

Item	Training in supervision (before assuming supervisory responsibilities)								Training in supervision (after assuming supervisory responsibilities)							
	No			Yes			Wilcoxon *W*	ε^2	No			Yes			Wilcoxon *W*	ε^2
	n	*Mdn*	*IQR*	*n*	*Mdn*	*IQR*			*n*	*Mdn*	*IQR*	*n*	*Mdn*	*IQR*		
Number of articles read a month	35	1	1	7	1	1.5	180.50	0.10	23	1	1.50	19	1	0.50	224	0.00
Competence in supervision over time	36	4	1	7	4	0.50	149.50	0.02	23	4	1	20	4	1	248	0.01
Confidence in supervision over time	35	3	1	7	3	1	119.50	0.00	22	3	1	20	3.5	1	229.50	0.00
Effectiveness in providing supervision	35	3	0	7	3	1	163	0.05	23	3	.50	19	3	1	261	0.03
Confidence in supervising interns perceived as incompetent	36	3	2	7	3	1	129	0.00	23	3	2	20	3	2	247	0.00
Competence in evaluating trainees	35	4	1	7	4	0.50	141	0.01	23	4	1	19	4	1	262	0.03

Even though supervision training choices may be directed by financial and time constraints, brief trainings (e.g. workshops) are often held intermittently and have restricted periods of knowledge transfer and acquisition[34], ultimately affecting the quality of learning outcomes. Selected evidence supports the efficacy of short-term training initiatives (e.g. workshops) in promoting supervisor competence[2,51], but there have been few studies examining the effectiveness of such in the South African context. The findings of this study highlight a need for the PBP-HPCSA to formulate formal supervision training guidelines and engage higher education institutions to develop training curricula for practitioners to complete prior to initiating supervision responsibilities.

Competence and confidence in supervision

The majority of supervisors included in this study felt confident and competent in their ability to supervise, which associated positively with perceived effectiveness in supervision. However, formal training in supervision was unrelated to perceptions of competence, confidence and effectiveness in supervising trainees. As many of the supervisors who completed training in supervision had attended short-term training programmes (e.g. workshops), it is plausible that South African supervisors develop their supervision skills and competencies through informal experiences, non-training avenues and clinical practice experience. However, clinical experience is not a sufficient prerequisite for ensuring competence in supervision[49,52], and evidence suggests that formal didactic training in supervision is essential for appropriate supervisor development[53,54]. More importantly, self-perceived competency does not mean that a clinician has demonstrated competency as a supervisor.

Compliance with ethical guidelines

The finding that the majority of internship supervisors in this study prematurely commenced supervision of psychology interns in advance of obtaining three years of independent practice experience raises ethical concerns about the current practice of supervision in accredited training hospitals and higher education institutions in South Africa. It is speculated that this may sometimes arise out of limited supervision training opportunities, employment expectations imposed by employers or limited staff capacity and the absence of mandated supervision training by regulatory bodies (e.g. HPCSA). For example, Singh-Pillay found that South African supervisors employed in the public sector perceived supervision as an 'imposition and an unavoidable obligation'[40(p.117)]. Further, supervisors often felt pressured into providing supervision without being adequately prepared – a process that necessitates an *automatic transition* from being a novice therapist to supervising trainees.[40] Hence, there appears to be a false assumption that clinicians who supervise are competent to do so.[55] This practice

also inadvertently supports the assumption that clinical skills are a sufficient prerequisite for providing competent supervision[36,56] and fails to recognise supervision as a distinct professional competence[57].

Thus, South African internship supervisors may be conflicted between adhering to necessary employment requirements in the public sector and the ethical guidelines that govern the profession. For this reason, some internship supervisors may be unable to fully comply with the HPCSA's regulatory and professional practice guidelines and may compromise unwittingly to accommodate a system that has yet to completely appreciate and value the importance of training as a prerequisite to effective supervision practice.

Implications for supervision training and practice

Drawing on the findings of this study, there are a number of relevant practical implications for the training and practice of intern supervision in South Africa. First, the PBP-HPCSA may be providing inadequate monitoring and oversight to internship supervision practices in the country. Currently, there are no policies regulating psychologists to (1) obtain supervision training and accreditation and (2) fulfil mandatory, HPCSA-approved supervision training. This is inconsistent with international trends and regulations enforced by regulatory bodies in other countries (e.g. Psychology Board of Australia). Because the HPCSA has a critical role in ensuring mental health professionals provide services in accordance with ethical regulations, and that supervisors are charged with gatekeeping the profession and protecting those who seek mental health support, it is imperative that the HPCSA is at the forefront of developing, monitoring and enforcing supervision training and supervision practice guidelines.

Practitioners are also obligated to provide professional services in an ethically responsible manner. Engaging in supervision practices without requisite competencies represents a violation of ethical obligations and professional code of conduct. Acquiring supervision training should be a shared responsibility and supervisors should take personal responsibility for ensuring they are appropriately trained to provide supervision before deciding to engage in supervision activities. Considering supervisors' training may affect clinical and non-clinical (e.g. conflict resolution) interactions with trainees, the growth and professional development of interns may be compromised by supervisors' lack of training and inadequate competence in providing supervision. As such, there is an urgent need for the South African psychology profession to develop and implement a competency-based framework of supervisory training.[57]

Limitations and future research directions

The findings of this study should be considered alongside selected limitations. First, the low sample size may have contributed to truncated statistical power, lowering the likelihood of detecting significant effects in inferential analyses. Second, a combination of non-probability sampling techniques was used to recruit participants, which affects the representativeness of the sample and generalisability of the results to the larger professional population of practitioners engaged in supervision within South Africa. Further research is needed to determine how the pattern of South African internship supervisors' training and practices compare across (1) the different categories of psychology professions governed by the HPCSA in South Africa (e.g. clinical versus educational) and (2) other countries in which psychological services are offered and governed by professional bodies. Third, the cross-sectional design precludes determinations of causality. Research is needed to examine changes in supervisors' perceptions of competence and confidence over time, particularly as a function of access to and participation in training activities. Along similar lines, researchers are encouraged to examine the impact of supervision training on supervisors, their trainees' experiences in supervision, their professional development and clinical training outcomes.

Conclusion

In this study, we explored South African internship supervisors' perspectives on supervision training, their supervision experiences, and their perceived abilities in providing supervision. Overall, findings indicate that a large proportion of intern supervisors prematurely initiate supervision responsibilities and engage in supervision without receiving formal training. However, supervisors tended to become more confident and competent in providing supervision over time, with many perceiving themselves as effective in their supervisory responsibilities. Considering regulatory guidelines on supervision training requirements are not currently provided, coupled with the dearth of available supervision training opportunities in the country, there is a need for further study on intern supervision training and practices in South Africa, including access to training opportunities, evaluating supervision training needs, designing a training framework and curriculum, and instituting formal training requirements for practitioners who participate in clinical supervision.

Acknowledgements

This research was supported by the UKZN University Teaching and Learning Office (TLCRG15).

Competing interests

We declare that there are no competing interests.

Authors' contributions

The data reported here come from a PhD project that was conceptualised and designed by S.H. S.H. developed the online survey and was responsible for managing data collection. Input, critical review and supervision was provided by D.J.C. S.H. and R.G.C. contributed to the data analysis and interpretation of results. S.H. wrote the initial draft of the article. S.H. and R.G.C. provided critical commentary and revisions in preparing the article for publication. Funding acquisition for the study was the responsibility of S.H.

References

1. Falender CA. Clinical supervision in a competency-based era. S Afr J Psychol. 2014;44:6–17. https://doi.org/10.1177/0081246313516260
2. McMahon M, Simons R. Supervision training for professional counselors: An exploratory study. Couns Educ Superv. 2004;43:301–309. https://doi.org/10.1002/j.1556-6978.2004.tb01854.x
3. Health Professions Council of South Africa. Health Professions Act, 1974 (Act no. 56 of 1974). Ethical rules of conduct for practitioners registered under the Health Professions Act, 1974; 2006.
4. Health Professions Council of South Africa (HPCSA). The professional board for psychology policy regarding intern psychologists: Guidelines for universities, internship training institutions and intern psychologists. Pretoria: HPCSA; 2014.
5. Russell RK, Petrie TA. Issues in training effective supervisors. Appl Prev Psychol. 1994;3:27–42. https://doi.org/10.1016/s0962-1849(05)80106-4
6. American Psychological Association (APA). Ethical principles of psychologists and code of conduct [Revised]. Section 2.01. Washington DC: APA; 2017.
7. Barnett J, Molzon C. Clinical supervision of psychotherapy: Essential ethics issues for supervisors and supervisees. J Clin Psychol. 2014;70:1051–1061. https://doi.org/10.1002/jclp.22126
8. Ellis MV. Critical incidents in clinical supervision and in supervisor supervision: Assessing supervisory issues. Train Educ Prof Psyc. 2006;S(2):122–132. https://doi.org/10.1037/1931-3918.s.2.122
9. Watkins CE. Being and becoming a psychotherapy supervisor: The crucial triad of learning difficulties. Am J Psychother. 2013;67:135–151. https://doi.org/10.1176/appi.psychotherapy.2013.67.2.135
10. Association for Counseling and Education. Standards for counseling supervisors. J Couns Dev. 1990;69:30–32. http://dx.doi.org/10.1002/j.1556-6676.1990.tb01450.x
11. Borders LD, Glosoff HL, Welfare LE, Hays DG, DeKruyf L, Fernando DM, et al. Best practices in clinical supervision: Evolution of a counseling specialty. Clin Superv. 2014;33:26–44. https://doi.org/10.1080/07325223.2014.905225
12. Falender CA, Cornish JAE, Goodyear R, Hatcher R, Kaslow NJ, Leventhal G, et al. Defining competencies in psychology supervision: A consensus statement. J Clin Psychol. 2004;60:771–785. https://doi.org/10.1002/jclp.20013

13. Ybrandt H, Armelius K. Changes in self-image in a psychotherapy supervisor training program. Clin Superv. 2009;28:113–123. https://doi.org/10.1080/07325220903343819

14. Borders LD, Fong ML. Cognitions of supervisors-in-training: An exploratory study. Couns Educ Superv. 1994;33:280–293. https://doi.org/10.1002/j.1556-6978.1994.tb00294.x

15. Barnett JE. Ethical issues in clinical supervision. Clin Psych. 2011;64:14–20. https://doi.org/10.1037/e582572011-006

16. Barnett J, Johnson W. Ethics desk reference for psychologists. Washington DC: American Psychological Association; 2008. https://doi.org/10.5860/choice.46-1823

17. Watkins CE. Handbook of psychotherapy supervision. New York: Wiley; 1997.

18. Magnuson S, Wilcoxon SA, Norem K. A profile of lousy supervision: Experienced counselors' perspectives. Couns Educ Superv. 2000;39:189–202. https://doi.org/10.1002/j.1556-6978.2000.tb01231.x

19. Ellis MV, Taylor EJ, Corp DA, Hutman H, Kangos KA. Narratives of harmful clinical supervision. Introduction to the special issue. Clin Superv. 2017;36:4–19. https://doi.org/10.1080/07325223.2017.1297753

20. Hendricks S, Cartwright DJ. A cross-sectional survey of South African psychology interns' perceptions of negative supervision events. S Afr J Psychol. 2018;48:86–98. https://doi.org/10.1177/0081246317698858

21. Watkins CE Jr. Psychoanalytic developmental psychology and the supervision of psychotherapy supervisor trainees. Psychodyn Prac. 2010;16:393–407. https://doi.org/10.1080/14753634.2010.510345

22. Pillay AL, Ahmed R, Bawa U. Clinical psychology training in South Africa: A call to action. S Afr J Psychol. 2013;43:46–58. https://doi.org/10.1177/0081246312474411

23. Pillay AL, Kramers-Olen AL. The changing face of clinical psychology intern training: A 30-year analysis of a programme in KwaZulu-Natal, South Africa. S Afr J Psychol. 2014;44:364–374. https://doi.org/10.1177/0081246314535683

24. Milne DL, James IA. The observed impact of training on competence in clinical supervision. Br J Clin Psychol. 2002;41:55–72. https://doi.org/10.1348/014466502163796

25. Falender CA, Shafranske EP. Clinical supervision: The state of the art. J Clin Psychol. 2014;70:1030–1041. https://doi.org/10.1002/jclp.22124

26. Merlin-Knoblich C, Harris PN, Chung SY, Gareis CR. Reported experiences of school counseling site supervisors in a supervision training program. J Sch Couns. 2018;16:1–33.

27. Nelson KW, Oliver M, Capps F. Becoming a supervisor: Doctoral student perceptions of the training experience. Couns Educ Superv. 2006;46:17–31. https://doi.org/10.1002/j.1556-6978.2006.tb00009.x

28. Lorenz D. Counseling self-efficacy in practicum students: Contributions of supervision [doctoral dissertation]. Pennsylvania: Pennsylvania State University; 2009.

29. Motley V, Reese MK, Campos P. Evaluating corrective feedback self-efficacy changes among counselor educators and site supervisors. Couns Educ Superv. 2014;53:34–46. https://doi.org/10.1002/j.1556-6978.2014.00047.x

30. Wheeler S, Richards K. The impact of clinical supervision on counsellors and therapists, their practice and their clients. A systematic review of the literature. Couns Psychother Res. 2007;7:54–65. https://doi.org/10.1080/14733140601185274

31. Magnuson S, Wilcoxon SA, Norem K. A profile of lousy supervision: Experienced counselors' perspectives. Couns Educ Superv. 2000;39:189. https://doi.org/10.1002/j.1556-6978.2000.tb01231.x

32. Worthington EL. Changes in supervision as counselors and supervisors gain experience: A review. Prof Psychol. 1987;18:189–208. https://doi.org/10.1037/1931-3918.s.2.133

33. Nielsen JA, Jacobsen CH, Mathiesen BB. Novice supervisors' tasks and training – A descriptive study. Nord Psychol. 2012;64:182–191. https://doi.org/10.1080/19012276.2012.731312

34. Bernard JM, Goodyear RK. Fundamentals of clinical supervision. 5th ed. Boston, MA: Allyn & Bacon; 2014.

35. Milne DL, James IA. The observed impact of training on competence in clinical supervision. Br J Clin Psychol. 2002;41:55. https://doi.org/10.1348/014466502163796

36. Milne DL, A.I. S, Pattison S, Wilkinson A. Evidence-based training for clinical supervisors: A systematic review of 11 controlled studies. Clin Superv. 2011;30:53. https://doi.org/10.1080/07325223.2011.564955

37. Kaslow NJ, Falender CA. Valuing and practicing competency-based supervision: A transformational leadership perspective. Train Educ Prof Psyc. 2012;6:47–54. https://doi.org/10.1037/a0026704

38. Falender CA, Shafranske EP, Ofek A. Competent clinical supervision: Emerging effective practices. Couns Psychol Q. 2014;27:393–408. http://dx.doi.org/10.1080/09515070.2014.934785

39. Rings JA, Genuchi MC, Hall MD, Angelo M-A, Cornish M, A EJ. Is there consensus among predoctoral internship training directors regarding clinical supervision competencies? A descriptive analysis. Train Educ Prof Psyc. 2009;3:140–147. https://doi.org/10.1037/a0015054

40. Singh-Pillay ND. The unsaid: An interpretative phenomenological analysis to understanding non-disclosure in clinical supervision from trainee and supervisor perspectives [dissertation]. Durban: University of KwaZulu-Natal; 2016.

41. Hess AK. Psychotherapy supervision: Stages, Buber, and a theory of relationship. Prof Psychol-Res Pr. 1987;18:251–259. https://doi.org/10.1037//0735-7028.18.3.251

42. Ellis MV, Ladany N. Inferences concerning supervisees and clients in clinical supervision: An integrative review. In Watkins CE Jr, editor. Handbook of psychotherapy supervision,. Hoboken, NJ: John Wiley & Sons Inc.; 1997.

43. Desmond KJ, Rapisarda CA, Nelson JR. A qualitative study of doctoral student supervisory development. J Int Couns Educ. 2011;3:39–54.

44. Robiner WN, Saltzman SR, Hoberman HM, Schirvar JA. Psychology supervisors' training, experiences, supervisory evaluation and self–rated competence. Clin Superv. 1997;16:117–144. https://doi.org/10.1300/j001v16n01_07

45. Ferguson CJ. An effect size primer: A guide for clinicians and researchers. Prof Psychol-Res Pr. 2009;40:532–538. https://doi.org/10.1037/14805-020

46. R CoreTeam. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2019.

47. Uellendahl GE, Tenenbaum MN. Supervision training, practices, and interests of California site supervisors. Couns Educ Superv. 2015;54:274–287. https://doi.org/10.1002/ceas.12026

48. DeKruyf L, Pehrsson DE. School counseling site supervisor training: An exploratory study. Couns Educ Superv. 2011;50:314–327. https://doi.org/10.1002/j.1556-6978.2011.tb01918.x

49. Ladany N, Bradley LJ. Counselor supervision: Principles, process and practice. 4th ed. New York: Routledge; 2011.

50. Studer JR, Oberman A. The use of the ASCA National Model in supervision. Prof Sch Couns. 2006;10:82–87. https://doi.org/10.5330/prsc.10.1.f82t11475451422m

51. Culloty T, Milne D, Sheikh AI. Evaluating the training of clinical supervisors: A pilot study using the fidelity framework. Cog Beh Ther. 2010;3:132–144. https://doi.org/10.1017/s1754470x10000139

52. Lyon RC, Heppler A, Leavitt L, Fisher L. Supervisory training experiences and overall supervisory development in predoctoral interns. Clin Superv. 2008;27:268–284. https://doi.org/10.1080/07325220802490877

53. McMahon M, Errity D. From new vistas to life lines: Psychologists' satisfaction with supervision and confidence in supervising. Clin Psychol Psychot. 2014;21:264–275. https://doi.org/10.1002/cpp.1835

54. Stevens D, Goodyear R, Robertson P. Supervisor development. Clin Superv. 1998;16:73–88. https://doi.org/10.1300/J001v16n02_05

55. Falender CA. Clinical supervision—the missing ingredient. Am Psychol. 2018;73:1240–1250. http://dx.doi.org/10.1037/amp0000385

56. Gonsalvez CJ, Milne DL. Clinical supervisor training in Australia: A review of current problems and possible solutions. Aust Psychol. 2010;45:233–242. https://doi.org/10.1080/00050067.2010.512612

57. Falender CA, Shafranske E. Clinical supervision: A competency-based approach. Washington DC: American Psychological Association; 2004. https://doi.org/10.1037/10806-000

AUTHORS:
Andrew Cooke[1]
Michael N. Bruton[2]
Minosoa Ravolololoharinjara[3]

AFFILIATIONS:
[1]Resolve sarl, Ivandry Business Center, Antananarivo, Madagascar
[2]Honorary Research Associate, South African Institute for Aquatic Biodiversity, Makhanda, South Africa
[3]Resolve sarl, Ivandry Business Center, Antananarivo, Madagascar

CORRESPONDENCE TO:
Andrew Cooke

EMAIL:
andrew@resolve.mg

DATES:
Received: 23 June 2020
Revised: 02 Oct. 2020
Accepted: 20 Oct. 2020
Published: 29 Mar. 2021

HOW TO CITE:
Cooke A, Bruton MN, Ravolololoharinjara M. Coelacanth discoveries in Madagascar, with recommendations on research and conservation. S Afr J Sci. 2021;117(3/4), Art. #8541. https://doi.org/10.17159/sajs.2021/8541

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:
coelacanth distribution, Western Indian Ocean, Onilahy marine canyon, coelacanth capture, gillnet fishing, *jarifa*, *Latimeria chalumnae*

FUNDING:
Resolve sarl

Coelacanth discoveries in Madagascar, with recommendations on research and conservation

The presence of populations of the Western Indian Ocean coelacanth (*Latimeria chalumnae*) in Madagascar is not surprising considering the vast range of habitats which the ancient island offers. The discovery of a substantial population of coelacanths through handline fishing on the steep volcanic slopes of Comoros archipelago initially provided an important source of museum specimens and was the main focus of coelacanth research for almost 40 years. The advent of deep-set gillnets, or *jarifa*, for catching sharks, driven by the demand for shark fins and oil from China in the mid- to late 1980s, resulted in an explosion of coelacanth captures in Madagascar and other countries in the Western Indian Ocean. We review coelacanth catches in Madagascar and present evidence for the existence of one or more populations of *L. chalumnae* distributed along about 1000 km of the southern and western coasts of the island. We also hypothesise that coelacanths are likely to occur around the whole continental margin of Madagascar, making it the epicentre of coelacanth distribution in the Western Indian Ocean and the likely progenitor of the younger Comoros coelacanth population. Finally, we discuss the importance and vulnerability of the population of coelacanths inhabiting the submarine slopes of the Onilahy canyon in southwest Madagascar and make recommendations for further research and conservation.

Significance:

- The paper contributes significantly to knowledge of the distribution and ecology of the Indian Ocean coelacanth, *Latimeria chalumnae*.
- The paper provides the first comprehensive account of Madagascar coelacanths and demonstrates the existence of a regionally important population and extensive suitable habitat, correcting an earlier hypothesis that coelacanths in southwest Madagascar were strays from the Comoros.
- The results have application in the study of the evolution, biology, ecology and conservation of the species.
- The significant threat posed by gillnet fishing to coelacanths and other species is highlighted as are the negative effects of the shark-fin trade.
- The paper emphasises the importance of the Onilahy marine canyon in southwest Madagascar as an especially important habitat and provides the basis for the development of a national programme of research and conservation.

Introduction

When a living coelacanth was caught off the coast of South Africa in December 1938 it caused an international sensation. J.L.B. Smith named the new species *Latimeria chalumnae*[1] and predicted[2] that it was a stray from warmer rocky reefs in the tropical Western Indian Ocean. Over the next 14 years, Smith and his wife Margaret scoured the coasts of Mozambique, Tanzania and Kenya looking for coelacanths but also collecting other fishes.[2,3] Their searches were confined to the mainland coast and to islands near the coast as they did not have the resources to explore the Comoros or Madagascar. Eventually a coelacanth caught by a traditional fisher off Anjouan (Nzwani) island in the Comoros in December 1952 was brought to Smith's attention[4] and, in one of the most remarkable episodes in the history of ichthyology, he rushed to fetch 'his' fish from a foreign country in a South African military aeroplane[2,5].

The French government, which held sovereignty over the Comoros and Madagascar at the time, was piqued at Smith's 'fishjacking' and banned research on the coelacanth (and other fishes) by foreign scientists in the Comoros; this ban lasted until the Comoros (except Mayotte) declared independence from France in 1975. A third coelacanth was caught off Anjouan Island in the Comoros in 1953[6] and a further six specimens off Grande Comore or Anjouan in 1954[7,8]. All these specimens, except one which was lost, as well as the next 15 specimens, all caught in the Comoros, were acquired by French scientists and lodged in the Museum National d'Histoire Naturelle in Paris and in other French museums. Thereafter, coelacanth specimens were sent to other museums, although by far the largest number of holdings is in museums in France (45 specimens by 2011[8]).

Early coelacanth research in Madagascar

In 1947, Jacques Millot, a French scientist based in Madagascar, was appointed as director of the Institut Scientifique de Madagascar. In 1948, he became president of the *Académie Malgache* and, in 1952, was placed in charge of French research on the coelacanth. The third coelacanth caught off Anjouan in 1953 was transported by air to the Tsimbazaza Museum in Antananarivo where it was examined and described in detail by Millot[6] in *Naturaliste Malgache*. This specimen is currently on display in the Department of Animal Biology at the University of Antananarivo. Millot[9] was also the first scientist to examine a live coelacanth when he briefly observed a dying female which had been caught in November 1954 and placed in a sunken boat in Mutsamudu, where it survived for over 19 hours. Madagascar therefore played an important role in early coelacanth research.

Fossil coelacanths

Fossils of extinct coelacanths have been known from Madagascar for over 100 years. *Coelacanthus madagascariense* was described by Woodward in 1910, *Whiteia woodwardi* and *W. tuberculatae* by Moy-Thomas in 1935 and *Piveteauia madagascariensis* by Lehman in 1952, all from Lower Triassic deposits.[10] The African mainland has also yielded an abundance of fossil coelacanths from both coastal and inland localities as many extinct coelacanths lived in fresh waters. Fossil coelacanth discoveries have been made in the Congo, Egypt, Morocco, Niger, South Africa, Zaire (now the Democratic Republic of the Congo) and Zimbabwe.[10]

Distribution of living coelacanths

The distribution of *L. chalumnae* includes South Africa (first recorded in 1938)[1], Comoros (1952)[4], Madagascar (1987)[11,12], Mozambique (1991)[13], Kenya (2001)[14] and Tanzania (2003)[15]. The recent sighting of a live *L. chalumnae* off the south coast of KwaZulu-Natal, 325 km south of the iSimangaliso Wetland Park where the main South African population is located, suggests that *L. chalumnae* is more widespread along the South African coast than previously thought and that the first specimen caught off East London may not have been a stray.[16] Another species of living coelacanth, *L. menadoensis*, has been found off Indonesia.[8,17,18]

Although the terrestrial fauna of Madagascar is characterised by high levels of endemism, with 84% of land vertebrates being endemic[19] due to its long separation from the African continent since the splitting of the supercontinent Gondwana 88 million years ago, its marine fauna is essentially continuous with the marine life of the east coast of Africa and other Western Indian Ocean islands and shows much lower levels of endemism than its terrestrial biota.

Coelacanth inventory

Since 1972, an inventory of all *Latimeria* specimens known to science has been compiled and maintained in an internationally collaborative effort.[7,8] Through the Coelacanth Conservation Council/Conseil pour la Conservation du Coelacanthe (CCC), established in 1987[5], CCC numbers have been allocated to all *Latimeria* specimens. To date (May 2020), 334 coelacanth captures have been documented, making it one of the most comprehensive inventories of all the known specimens of a species ever compiled.

In addition to the inventory of dead coelacanth specimens, H. Fricke and his team have compiled an inventory of 68 individual live *L. chalumnae* which they have observed over a 21-year period off Grande Comore Island in the Comoros.[20] Fricke and the African Coelacanth Ecosystem Programme team have also compiled an inventory of 32 individuals in South Africa in the iSimangaliso Wetland Park in KwaZulu-Natal.[21] As all living coelacanths have unique patterns of white spots on their bodies, which are effectively 'individual fingerprints', individuals can be distinguished visually from one another by divers.

Live coelacanths caught off Madagascar

In his book 'Old Fourlegs'[2], J.L.B. Smith predicted that coelacanths live off Madagascar. 'There must be stretches of coast there that no enlightened scientific eye has ever seen', and even suggested that local people 'feasting unsuspected on succulent coelacanth steaks on a Madagascan shore did not seem too fantastic'.

In 1969, 28-year-old Hans Fricke, who would later become the most distinguished researcher on the living coelacanth, visited Madagascar to study the fish in its natural environment. During his initial scuba dives to a depth of 85 m off Nosy Iranja, a small island in northwest Madagascar, he did not find any coelacanths (Nosy Iranja and nearby Ankazoberavina Islands are now part of the Ankivonjo Marine Protected Area and the nearby drop-offs are potential coelacanth habitats). When Fricke returned to the Western Indian Ocean in 1987 with a research submersible, he chose to dive off Grande Comore (Ngazidja) in the Comoros where he and his team carried out a detailed study of the living coelacanth in its natural habitat and proclaimed the Comoros to be the home of the coelacanth.[20]

Although coelacanths may have been caught previously by fishers off Madagascar, at least since the arrival of gillnets in the 1980s, the first specimen caught off the island which was known to Western scientists until recently (CCC 173) was landed off St Augustin, south of Toliara, on 5 August 1995.[11] This specimen was bought by Dominque Coutin from fishers for USD6.00 and taken by boat to Toliara. However, an earlier specimen (CCC 300) has recently come to our attention which was caught in 1987 off Anakao in southwest Madagascar (Table 1, Figure 1). This specimen is on display in the Museo Civico di Storia Naturale in Comis in Italy, and is described by Insacco et al.[12] Since then, at least 32 additional specimens known to science have been landed off Madagascar, although others have been caught but lost. Sufficient coelacanths have been caught in Madagascar for it to have a common name. In the Toliara area it is known as *fiandolo* ('ghost fish'). It is called *gombessa* ('taboo') in the Comoros and the Indonesian species is known as *raja laut* ('king of the sea').

Figure 1: Coelacanth CCC 300, the first coelacanth from Madagascar known to science, caught in 1987 in St Augustin Bay and now on display in the Museo Civico di Storia Naturale in Coviso, Italy (photo: Gianni Insacco).

In their 1996 paper, Heemstra et al.[11] surmise that the 1995 specimen caught in Madagascar (CCC 173) was most likely a stray from the Comoros population, based on fishers' lack of local knowledge of the coelacanth and the genetic similarity of the pups with the Comoros population. We argue that the coelacanth populations in Madagascar are ancestral to those in the Comoros and that Comorian coelacanths are the descendants of those in Madagascar. We go beyond Hans Fricke and reinforce Green et al.[22] in predicting that coelacanths are likely to be distributed around the entire coast of Madagascar and that, with its >5000-km coastline, Madagascar is likely to harbour the largest populations of coelacanths in the Western Indian Ocean.

Inventory of coelacanths caught off Madagascar

The inventory of coelacanth specimens caught off Madagascar has been updated using data from the official CCC Coelacanth Inventory[8], and subsequent updates have been made from supplementary information collected on coelacanth specimens during a survey in Madagascar in November 2019 by one of the authors (M.R.) under the auspices of the consultancy company Resolve sarl in collaboration with Clemence Ravelo from the Institut Halieutique et des Sciences Marines (IHSM) at the University of Toliara[23], and from other publications[24-27]. During the survey, meetings were held with the director of the IHSM, Dr Jamal Mahafina, representatives of the fishing companies Copefrito and Murex, staff of the Regional Fisheries Directorate in Toliara, staff of the Jardin de la Mer in Toliara, and with Mr Tinard, an experienced fisherman. Frozen, formalin-preserved and dried coelacanths in various locations were also photographed during this survey. Specimens CCC 251 (Figure 2) and 317 (Figure 3) were photographed and credited to D. Stanwell-Smith and T. Cordenos, respectively.

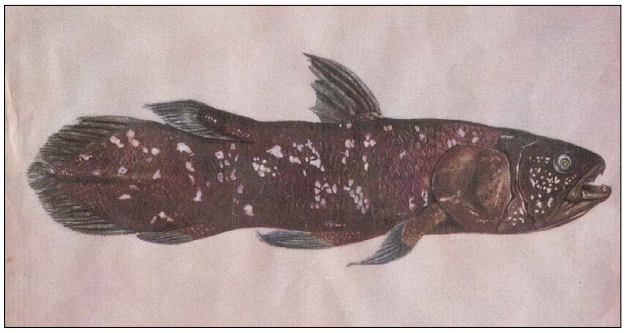


Figure 2: Coelacanth CCC 251 caught in March 2001 at Fiherenemasay and now in the collection of the IHSM Museum, Toliara (photo provided by D. Stanwell-Smith).

Figure 3: Coelacanth CCC 317 caught on 13 May 2010 at Nosy Ve, identified by IHSM Museum staff, and sold in the market at Toliara (photo: T. Cordenos).

The 34 specimens known to have been caught in Madagascar are listed in Table 1; further details on these specimens are available online at https://www.resolve.mg/download/MadagascarCoelacanthInventory_23Sept2020.pdf.[28] The surviving specimens (or parts thereof) are currently located in a variety of institutions and locations including museums, universities, commercial fishing companies (Copefrito and Murex, both in Toliara), the Résidence Eden Ecolodge, the Jardin de la Mer (an exhibition centre on Malagasy plants and animals in Toliara), NGOs and the regional fisheries directorate of Toliara. One specimen (CCC 176) is on display in the Tolagnaro (Fort Dauphin) Town Hall.

Location of captures

Capture locations where known are shown on bathymetric maps of western and southern Madagascar (Figure 4) and off the Onilahy River mouth, Toliara, in southwest Madagascar (Figure 5). Capture points are approximate and represent the best estimate of capture location based on available information on depth, distance and direction from any reference point and the location of other captures by fishers from the same village. The captures were made over a wide geographical range extending from 80 km to the southwest of Cap Ste Marie, the southernmost point of Madagascar, as far north as three sites near Maintirano in northwest Madagascar (590 km to the north of Toliara) – a range of almost 1000 km. Of the 34 specimens, 21 form a cluster in southwestern Madagascar in the vicinity of the Onilahy canyon (Anakao, Lovokampy, Soalara, Nosy Ve, St Augustin, Andanora and Sarodrano; Figure 5). Nonetheless, the capture of coelacanths at Fiherenamasay and Tsiandamba, respectively 40 km and 85 km north of Toliara (Figure 4) where the shelf is narrow and no canyons are present, suggests that their association with undersea canyons is not exclusive; depth and slope may be the primary determinants for the occurrence of coelacanths, as also suggested by Green et al.[22]

Despite the widespread practice of shark fishing using gillnets set at depths of 100 m or more throughout the island[29-31], no coelacanths have as yet been reported from the northwest around Nosy Be or the northernmost point of Madagascar near Antsiranana, which is just 652 km from the Comoros. There are also no coelacanth records from the east coast of Madagascar, despite the presence there of a steeply shelving continental slope and at least one undersea canyon (at Maningory, south of Ile Ste Marie), although this may be an artefact of the much lower levels of fishing effort there and the absence of sailing pirogues capable of reaching the continental slope. Further research is necessary to ascertain whether coelacanths do live in these unexplored areas along the east coast.

The distance from shore at which coelacanths were estimated to have been caught off Madagascar, all along the west coast, ranged from 0.8 km to 80 km (average 9 km), which is further offshore than in Grande Comore (where 85% of coelacanths are caught less than 1.5 km from shore[32,33]) and Tanzania (average 6.9 km, range 0.5–8 km)[15]. The wide range of distance from shore at which coelacanths have been caught in Madagascar corresponds with the widely varying width of the continental shelf, which is as narrow as 1 km at St Augustin in the south, where most captures have been recorded, extending to over 100 km off Cape St André in the west. The estimated depth of capture in Madagascar ranged from 60 m to 500 m (average 191 m), shallower than in the Comoros[8,32,33] but deeper than in Tanzania (average 141 m, range 50–250 m[15]).

Madagascar originally occupied a landlocked position at the centre of Gondwana until the supercontinent began to break up about 160 million years ago. About 88 million years ago, India split off from Madagascar, moving northwards to join Asia. Madagascar then shifted more slowly northwards to its current isolated position in the Western Indian Ocean about 40 million years ago, since when it has experienced relatively stable climatic and oceanographic conditions.[34] This can be compared with the young Comoro islands whose ages range from 15–10 million years (Mayotte) to just 130 000 years (Grande Comore).[35] It is likely that either Madagascar or the African mainland represent the more ancestral habitat of coelacanths before they colonised the Comoros in relatively recent geological time, but this proposal needs to be tested using genetic evidence. DNA studies on coelacanths that have previously been caught in Madagascar would have to be carried out on the frozen specimens as all the dried and formalin-preserved specimens have been exposed to formalin.

Madagascar's ancient continental margin is cut at several locations by deep canyons which were created during previous ice ages including the Pleistocene (which started about 2.6 million years BP), when sea levels dropped by 100 m or more, and when Madagascar's major rivers would have cascaded off cliffs and down steep slopes into the sea.[36] The closest of these canyons to the existing shoreline is the Onilahy canyon, site of most coelacanth captures to date, but other deep canyons also exist on the east coast.[22,37] Furthermore, Madagascar's southwestern and eastern continental margins are steeply shelving, potentially providing 2000–3000 km of suitable habitat for coelacanths, as suggested by Green et al.[22]

Table 1: Information on the coelacanths known to science caught in Madagascar between 1987 and 2019. Additional information on these specimens is given in the text and full details are given in Cooke et al.[28]

No.	CCC number	Date of capture	Site of capture	Distance from shore (km)	Depth of capture (m)	Weight (kg)	Length (cm)	Sex	No. of eggs	No. of pups	Fishing gear	Current holding
1	300	1987	St Augustin	Nd	nd	30-35	121	M			nd	MCSN
2	173	03.08.1995	Onilahy	4–9	<190	34.98	134	M			Jarifa	IHSM
3	176	06.12.1997	Onilahy	2–3	60	90	190	F	13		Jarifa	Tolagnaro Town Hall
4	177	03.03.2001	Fiherenamasay	3–4	100	75	160	F	9		nd	IHSM
5	179	21.07.2001	Tsiandamba	5–6	>100	73	160	F	4	2	Jarifa	IHSM
6	205	18.05.2006	Nosy Lava	Nd	140	Nd	171	Nd			Nd	IHSM
7	231	18.02.2009	Fiherenamasay	Nd	200	nd	nd	nd			Jarifa	IHSM
8	232	July 2002	Toliara	nd	nd	35	150	nd			Jarifa	Not kept
9	244	20.09.2008	Cap Ste Marie	80	nd	40	150	nd			Jarifa	Copefrito
10	245	April 2008	Maintirano	nd	nd	nd	nd	nd			nd	Not kept
11	251	22-29.03.2001	Fiherenamasay	nd	120	80	180	F	2		Jarifa	IHSM
12	252	12.07.2005	Fiherenamasay	nd	nd	60.3	155	F	2		nd	Not kept
13	284	22-23.09.2010	W of Nosy Ve	>2	>150	85	187	F			Jarifa	IHSM
14	285	27.11.2010	W of Nosy Ve	1–2	250	41	134	M			nd	IHSM
15	310	April 2011	Onilahy Canyon	nd	>300	nd	nd	nd			nd	Not kept
16	288	5 May 2010	W of Nosy Ve	>1	150	nd	nd	F			Jarifa	Not kept
17	289	10.02.2011	W of Nosy Ve	20	200	80	175	F			Jarifa	IHSM
18	290	11.02.2011	W of Nosy Ve	2	200–300	60	149	F			Jarifa	IHSM
19	291	13.02.2011	NW Sarodrano	7	200–300	75	170	F			Jarifa	Copefrito
20	292	12.03.2011	NW Sarodrano	3	200–300	75.2	182	F			Jarifa	Copefrito
21	293	21.05.2011	W of Nosy Ve	2	150–200	29.45	130	M			Jarifa	Copefrito
22	294	02.07.2011	Andanora	7	150–200	84.64	170	F			Jarifa	Copefrito
23	295	03.08.2011	W of Nosy Ve	10	150–200	32	140	F			Jarifa	Murex
24	296	25.08.2011	W of Nosy Ve	2	150–200	62	170	F			Jarifa	Not kept
25	297	Jan 2012	St Augustin	nd	nd	nd	nd	nd			Jarifa	MHNH
26	311	31.05.2012	Fiherenamasay	nd	100–200	36	Nf	F			Jarifa	Not kept
27	298	July 2012	Toliara	nd	nd	nd	'Very large'	F			Jarifa	Not kept
28	312	2012	Maintirano	nd	100–200	36	nd	nd			nd	IHSM
29	301	Feb 2013	Ambanilia	6	500	35	130	M			Jarifa	Eden Ecolodge
30	313	Feb 2013	St Augustin	nd	nd	35–50	150	nd			nd	Not kept
31	314	2015	Anakao	nd	nd	nd	nd	nd			nd	Jardin de la Mer
32	315	11.06.2011	Toliara	nd	nd	nd	nd	nd			nd	DRAEP
33	316	23.03.2019	Barnhill Point	0.8–1	200	79	150	F			Jarifa	DRAEP
34	317	13.05.2010	W of Nosy Ve	nd	nd	nd	nd	nd			Jarifa	Not kept

Sources: Bruton and Coutouvidis[7], Nulens et al.[8], Heemstra et al.[11], Insacco et al.[12], Ravolololoharinjara[23], Vicente[24], Niaina[25], Anon[26], Houssen[27] and personal communications as indicated in the text.

Holdings: Copefrito, Compagnie de Pêche Frigorifique de Toliara fishing company, Toliara; DRAEP, Direction Régionale de l'Agriculture, de l'Elevage et de la Pêche, Toliara; Eden Ecolodge, Eden Ecolodge, Sarodrano; IHSM, Institut Halieutique et des Sciences Marines, University of Toliara; MCSN, Museo Civico di Storia Naturale, Comiso, Italy; MNHN, Muséum National d'Histoire Naturelle, Paris, France; Murex, Murex International fishing company, Toliara

Jarifa, large-mesh gillnet laid for sharks

nd, no data

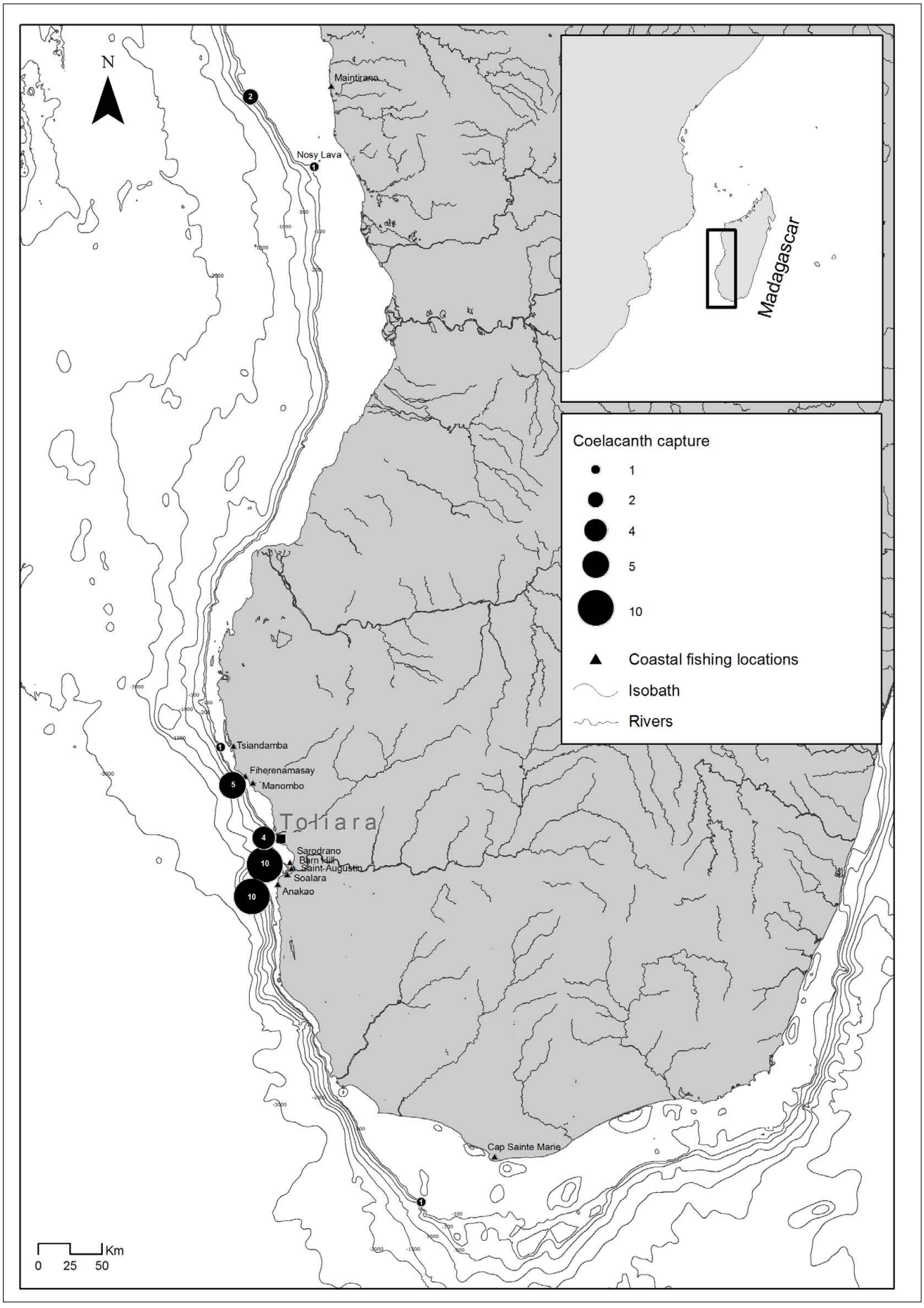


Figure 4: Map showing the location of captures from 1987 to 2019 of all coelacanths in Madagascar whose capture location is known.

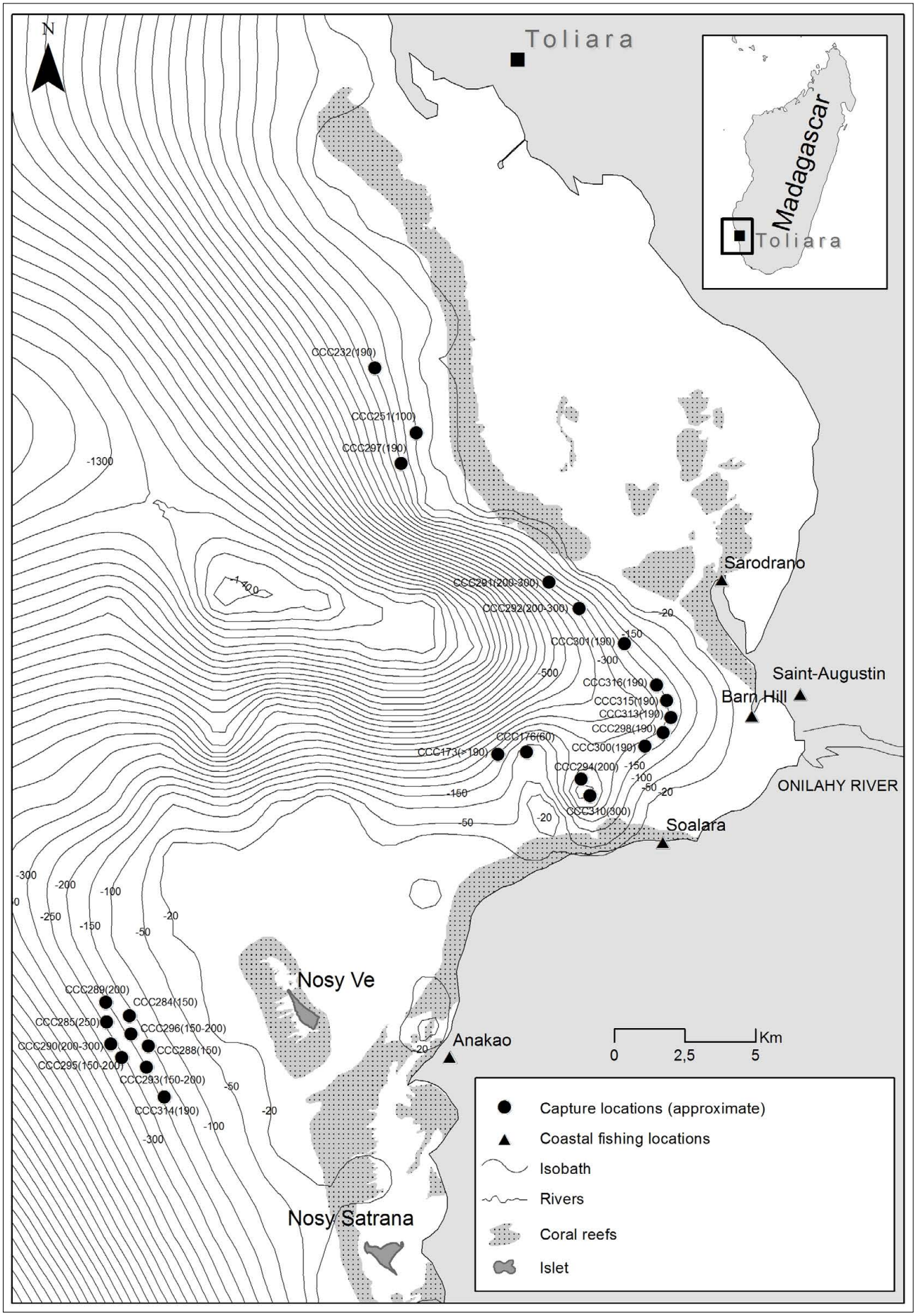


Figure 5: Map of St Augustin Bay, southwestern Madagascar showing the locations of the cluster of coelacanths caught from 1987 to 2019 near the Onilahy canyon.

Along the west coast, near the Onilahy canyon, the continental shelf is very narrow.[38] Northwards from the Manombo River, the shelf broadens and suitable habitats for coelacanths at depths of 150 m or more are typically found far offshore, such as off the Barren Islands, to the west of Maintirano, where migrant *vezo* fishers regularly establish seasonal camps about 40 km offshore of the main island.[39] In contrast, on Grande Comore, a young island with an active volcano (Karthala), the shores are steep sloping and deep water can be reached within a few hundred metres from shore using wooden dugouts (*galawas*). Handline fisheries, to which coelacanths are vulnerable, are therefore better developed in the Comoros than in Madagascar.[32,33,40]

Capture frequency, seasonality and demography

Many coelacanths, other than those recorded in Table 1, are likely to have been caught by artisanal fishers in Madagascar. In December 1998, a commercial fishing crew was reported to have caught a female coelacanth (180 cm, 85 kg) off Toliara which contained seven embryos, but the specimen was not kept nor authenticated.[23] An experienced fisherman, Mr Tinard, from Lovokampy and his team were reported to have caught 'dozens of coelacanths in a single week, but they were all consumed' in 2010[23] and none was documented. In 2016, Tinard caught another undocumented coelacanth and, in May 2016, he landed a small female and found 'a few small coelacanths in the belly'[23]. He also reported to Ravololoharinjara[23] that his family did not use *jarifa* gillnets in 2018 and caught no coelacanths in that year. Another fisherman, Tine Hoe Julien, is reported to have caught seven coelacanths off Nosy Ve, Sarodrano and Andanora between May 2010 and July 2011.[23]

Coelacanth catches were made in every month of the year in Madagascar except October, with most catches in February (6), May and July (5) and March (4), although the sample size (31) for which the month of capture is known is too small to show real trends. In the Comoros, coelacanth catches were also made throughout the year with a peak from November to March[8,33], whereas catches peaked in September (21%) and August (17%) in Tanzania[15]. All coelacanth captures in Madagascar for which the capture method is known (23) were made using deep-set *jarifa* gillnets targeting sharks[11,23] (Table 1). In the Comoros, all coelacanth catches were made using handlines until the arrival of gillnets in the 1990s.[33] In Tanzania, 35 (87.5%) of the 40 coelacanths for which the capture method is known between 2003 and 2015 were caught using 15-cm *jarifa* gillnets except for two caught on handlines, two moribund specimens found floating on the water surface and one caught in a ring net.[8,15]

The coelacanths caught in Madagascar which were measured ranged in weight from 29.45 kg to 90 kg (average 57.2 kg) and length from 121 cm to 190 cm (average 156.9 cm; Table 1), equivalent to the parameters for female coelacanth catches in the Comoros[8,32,33], which reflects the higher proportion of females (which grow larger than males[8,40]) caught off Madagascar. The size of the 40 coelacanths caught off Tanzania (including Zanzibar) between 2003 and 2015 averaged 41.3 kg (range 5.8–105 kg) and 133.7 cm (range 70–184 cm), substantially smaller than the coelacanths caught off Madagascar, although the weight range is wider.[8,15] Although aspects of the demography of coelacanths in Madagascar, the Comoros and the African mainland differ, these differences are more likely to be a function of the different fishing methods that are used, and the different locations and depths that are fished, rather than representing discrete demographic groups of coelacanths, but this supposition needs to be tested using genetic evidence. Such evidence could also be used to determine whether gene flows exist between the different coelacanth populations and to support or reject the hypothesis that Madagascar or the African mainland is the ancestral home of coelacanths on the basis that older (and larger) populations would be expected to have higher genetic diversity. Studies carried out so far suggest that coelacanths in the Western Indian Ocean share very similar genetic material.[11,41]

Threat of *jarifa* gillnets

Sharks have been targeted by artisanal fishers for shark fin and oil in the Western Indian Ocean for more than a century. Petit[42] found that shark fishing and shark fin exports from Madagascar were already established by 1900 and Schaeffer[43] reports that shark fin exports started as early as 1919 and that 6.6 tons of shark fins had been exported from Zanzibar by 1923. Shark fishing intensified significantly with the rapid growth of the Chinese economy in the 1980s and the resulting demand for shark fin and continues today.[30,44,45]

Shark meat is widely eaten by fishing communities in Zanzibar[43,46] and Madagascar[45], although it is not as prized as the flesh of bony fish. A benefit of shark meat to rural people is that it preserves well unrefrigerated when it has been dried and salted but shark meat consumption and export are sometimes discouraged due to the risk of poisoning caused by toxicity originating from dinoflagellates.[29]

The *jarifa* gillnets used to catch sharks are a relatively new and more deadly innovation as they are large and can be set in deep water. There are two kinds of *jarifa* nets: those with large meshes (15 cm or 24 cm stretched mesh) which are often baited with small fish, and those with smaller meshes (10 cm; called 'ZZ nets') which are not baited (Nulens R, personal communication, 9 March 2020). Large-mesh *jarifa* gillnets are used in Madagascar and Tanzania, with fishers from the former country using 24-cm stretched-mesh nets[8] and, from the latter, 15-cm nets[15].

The introduction of market forces from abroad has often resulted in much greater pressure being placed on a natural resource that was once exploited sustainably for local use[43], and this appears to be the case in Madagascar. There is little doubt that large mesh *jarifa* gillnets are now the biggest threat to the survival of coelacanths in Madagascar. The nets are set in deep water, generally between 100 m and 300 m, within the preferred habitat range of coelacanths, and, unlike trawl nets, can be deployed in the rugged, rocky environments which coelacanths prefer. They would be difficult to detect by the fishes as they are static and do not produce a pressure wave like active gear, such as a trawl net. Furthermore, coelacanths hunt at night and have poor eyesight and their main sense organ, electro-reception[47], may not be triggered by the thin strands of a gillnet. In fact, coelacanths may be attracted to the nets as they are typically baited with small fish in Madagascar. A significant number of coelacanths has also been caught in *jarifa* gillnets off Tanga in Tanzania where 19 were caught in 6 months in 2004/2005, including 6 in one night.[8,15,48]

The incidental capture of coelacanths in *jarifa* gillnets off Madagascar is not a disincentive for shark fishers because of the high scientific interest in the fish which inevitably commands a price, even in the absence of a true market. The presence in Toliara of a marine research institute (IHSM) has increased fishers' awareness of the coelacanth's significance and value. Baker-Médard and Faber[45] report payments of 150 000–400 000 ariary (USD40–110) for coelacanth specimens caught in the Toliara region.

The illegal trafficking of coelacanths may be taking place in Madagascar. When the Centre de Surveillance des Pêches checked the cargo of the factory ship, *El Amine*, on 20 September 2008, following a tip-off from the Coast Guard, they found an undeclared coelacanth (CCC 244) hidden on board, and several other infringements of the fishing laws were revealed. The *El Amine* was escorted to Toliara where the coelacanth was handed over to the authorities and stored in the cold room of Copefrito, where it is still housed. On 21 October 2008, the newspaper *Les Nouvelles* headlined an article on the incident which reported that over 300 kg of coelacanths had been captured by *El Amine's jarifa* nets and stating that they suspected that coelacanth trafficking had been taking place. After paying a fine, the *El Amine* was allowed to leave Toliara on 30 October 2008.

Furthermore, a fish biologist, Dr Faratiana Ratsifandrihamanana, reported to one of us (M.R.) that she had seen a cartful of dead coelacanths in the yard of the IHSM in April 2012. They comprised 4 adults (about 1.5 m in length) and 5–7 juveniles (60–70 cm). These specimens had been caught by fishers from St Augustin who told her that they had been deliberately targeted as they could sell them to *vazaha* (foreigners) at 100 000 ariary per fish.[23] This is clear evidence that fishers can target coelacanths and that there is an informal market for them. The IHSM refused to buy the fish and they were taken away.

The coelacanth bycatch fishery is significant as their populations are unlikely to be able to survive high exploitation rates as they have all the attributes of species that are vulnerable to extinction, including rarity, large size, high trophic level in the food pyramid, low dispersal rates, few offspring, high longevity and high levels of specialisation.[5,32] In addition, coelacanth populations may be small. The best studied population is that off Grande Comore where Fricke and his team estimated that the population size in 1990 was no more than 300 adults.[20,47] Coelacanths may also be susceptible to capture in the snagging meshes of a gillnet as they have large mouths with sharp teeth, large opercula, eight spines on the first dorsal fin and paired lobed fins.

Gillnets are deadly for another reason – if they are lost or abandoned at sea they continue to catch fishes and shellfishes for months, or even years, as the synthetic fibres from which they are made do not rot quickly. This 'ghost fishing' can be very harmful to fish and shellfish stocks.[48] *Jarifa* gillnets (in use or lost) are also known to catch dugongs and turtles in Madagascar.[49,50]

Outside Madagascar, the biggest threats to coelacanths, other than *jarifa* gillnets, are the use of explosives by fishers, recorded in Tanzania and the Comoros, and insecticide residues and plastic litter in the oceans.[8,15,51,52-55] The use of explosives has not been reported in Madagascar, whilst the presence of insecticide residues and plastic litter has not been assessed.

Coelacanths as food

The capture of coelacanths as a source of food is hard to justify as its flesh is rancid and contains large amounts of urea, which coelacanths store in their tissues like elasmobranchs, as well as oils, wax esters and other compounds that are difficult to digest. Madagascar is one of the few places where coelacanth flesh is regularly eaten. Of the 34 specimens listed in Table 1, 10 were sold by fishers at a market or eaten (or used as bait) after they had been documented. Ravololoharinjara[23] was told that a 32-kg male fish (CCC 295) which had been housed in the cold room of the fishing company Murex in Toliara, had been 'shared with company personnel during the passage of Cyclone Haruna in 2013'. Coelacanth flesh is occasionally eaten in Tanzania[15] and in Anjouan in the Comoros (Fricke H, personal communication, June 2020).

Future coelacanth research

Although most coelacanth specimens known to have been caught in Madagascar have resulted from chance catches by artisanal fishers, who are mainly targeting sharks, rather than from a structured scientific research programme, the available evidence suggests that Madagascar does have a permanent and widespread population of breeding coelacanths. As the coelacanth is such an important species from ecological, conservation and historical perspectives (see below) it makes sense to take advantage of this opportunity to mount a structured international research programme on the species, based not only on chance catches but also on live observations of the fishes in their natural habitat, as has been done in the Comoros, Tanzania and South Africa.

The Madagascar coelacanth programme would build on the African Coelacanth Ecosystem Programme and the former collaboration between this Programme and the Agulhas & Somali Currents Large Marine Ecosystem Programme, linking coelacanth and ocean ecosystem research, of which Madagascar is a participating country. African Coelacanth Ecosystem Programme (and formerly the Agulhas & Somali Currents Large Marine Ecosystem Programme) is based out of the South African Institute for Aquatic Biodiversity.

The most practical scenario for live observations of coelacanths in Madagascar would be to use a remotely operated submersible such as the *Sea-Eye* owned by the South African Institute for Aquatic Biodiversity which has already been used with great success to document the distribution, abundance and behaviour of coelacanths in the iSimangaliso Wetland Park in KwaZulu-Natal and elsewhere.[21] This research could initially focus on determining the distribution, abundance, habitat preferences, depth range and diel activity patterns of coelacanths – information which is needed for their management.

The study of dead coelacanths derived from the artisanal fishery can also continue to yield useful information if the collection of data and the preservation of the specimens are carried out professionally. A standardised questionnaire is required which captures as much information as possible (as per the categories in the CCC Coelacanth Inventory) on each caught specimen. This information should then be included in the official inventory and made available to the international community via publications. An awareness campaign among artisanal fishers also needs to be launched to encourage them to share information on their coelacanth catches with the authorities.

Whenever practical, caught coelacanths that are in good condition should be deep frozen rather than preserved in formalin so that tissue samples can be taken for further analyses. The only Madagascan specimen that has so far been subject to detailed tissue analysis is CCC 177 caught in March 2001 which was taken by P.C. Heemstra to the J.L.B. Smith Institute of Ichthyology (now South African Institute for Aquatic Biodiversity) in South Africa.[11] Samples of scales and of muscle and dorsal fin tissue were used for stable isotope analyses.

It is important that this study is pursued further using genetic methodologies with tissues taken from frozen specimens. At present, this research would have to be performed outside Madagascar as the only DNA analysis machine available (in the capital Antananarivo) can only extract and conduct DNA hybridisation assays but cannot sequence the genome. mtDNA and full genome tests would be useful to assess the extent of divergence of coelacanth populations, and the genetic diversity among regional populations, perhaps using the methods adopted in the EDGE programme. Interestingly, genetic research on dugongs has revealed that the Madagascan and Comorian populations of dugongs are genetically distinct from those of the East African coast, which suggests that the Mozambique Channel can be a barrier to the movement of primarily coastal shelf species.[55]

Coelacanth conservation

Both the Western Indian Ocean coelacanth (*L. chalumnae*) and the Indonesian coelacanth (*L. menadoensis)* are listed on Appendix I of CITES (may not be traded for commercial gain). *L. chalumnae* is rated as 'Critically endangered' by the IUCN (very highly vulnerable to extinction) and *L. menadoensis* as 'Vulnerable'. The two living coelacanths may also be considered 'EDGE species' (Evolutionarily Distinct and Globally Endangered) which have a high global conservation priority due to the significant level of unique evolutionary history that they embody. Considering the international significance of *L. chalumnae*, and the fact that Madagascar is one of only four countries known to host breeding populations (with the Comoros, South Africa and Tanzania), although single specimens have so far been caught off Mozambique[13] and Kenya[14], it is very important for Madagascar to contribute to the conservation of the coelacanth.

The exact conservation status of coelacanths in Madagascar cannot be determined until we have better information on their distribution and population densities around the entire island. If coelacanths occur around the whole coast, as we predict, then the total population could be regarded as stable as the catches made on the southwest and west coasts, even if several times greater than the documented rate (about one per year for 33 years) would probably be trivial in relation to the size of the population. However, we have anecdotal evidence that 'dozens more' coelacanths have been caught off southwest Madagascar in recent years compared to the number that has been officially recorded[8,23], so the true catch rate may be substantial. If coelacanths are only found at or near the currently known sites, or at only a small number of other sites, then there would be reason for concern.

The results of demographic studies on coelacanth populations off Grande Comores and Anjouan islands in the Comoros demonstrate that the known catch rates of 3.5 fish per year in the 1960s, 1970s and 1980s were insignificant compared to natural mortality rates that were calculated to be between 137 and 174 individuals per annum. The main source of natural mortality was considered to be predation by sharks.[40]

A worrying trend in coelacanth catches in Madagascar is the relatively high proportion of pregnant females which is landed. The breeding mode of coelacanths – a very long gestation period (36 months) with the live bearing of a few, very large young (33 cm, 500 grams[5,56]) rather than the production of a large number of small eggs as in most teleost fishes – means that they invest a large amount of energy in each of a few young. Killing of pregnant females carrying eggs or unborn pups is therefore a major setback for the population. There is evidence from Madagascar and other localities that pregnant coelacanths may be relatively vulnerable to gillnet and trawl net catches.

Of the 22 coelacanths caught off Madagascar whose sex is known, only 5 were male and 17 (77%) female; 5 of the 15 female individuals (33%) carried eggs and/or unborn pups. Of the 26 coelacanths caught off Tanzania between 2003 and 2015 for which the sex is known, 10 were male and 16 (61.5%) were female; half of the 16 female individuals caught were carrying eggs or unborn pups.[8,15] The only coelacanth caught so far off Mozambique (CCC 162) was a 98-kg, 179-cm female fish carrying 26 late-term pups which was landed using a trawl net.[13] The only coelacanth known from Kenya (CCC 178) was a 77-kg, 170-cm female individual carrying 17 eggs caught in a trawl net in April 2001.[14] An 86.5-kg, 176-cm female coelacanth caught in a net off Unguja Island, Zanzibar, Tanzania, in July 2009 (CCC 253) was carrying 23 fully developed juveniles.[8,15] It is important to note that over 90% of all coelacanths larger than 50 kg are female[8,40] and that these larger female fish may be more susceptible than the smaller male fish to capture by large-mesh gillnets set for sharks. The continued capture of pregnant female coelacanths in Madagascar and elsewhere is a serious concern as Fricke et al.[52,54] have estimated that they produce only 140 young during their entire life cycle.

Although it is tempting, from the perspective of the conservation of the whole marine megafauna, especially sawfish, sharks, coelacanths, turtles, dugongs and dolphins, to call for a complete ban on the importation, transport, manufacture, sale and/or use of *jarifa* gillnets in Madagascar, such a ban would have wide socio-economic implications for the many people who rely on marine resources for their livelihood. It is therefore necessary to include the human dimension into conservation recommendations, otherwise these recommendations would be ignored and/or the fishing activities would be carried out illicitly. Instead, the use of *jarifa* gillnets in fisheries management areas and marine protected areas should be strictly controlled and their use should be restricted to areas where they do not pose a significant threat to threatened species.

Coelacanth conservation measures which should be introduced in Madagascar include:

- Passing legislation adding *L. chalumnae* to the list of integrally protected species under Madagascar's wildlife laws, which forbid the capture, holding, transport or sale of such species.
- Establishing a strictly protected coelacanth sanctuary in the Onilahy canyon near Anakao where the highest concentration of coelacanths in Madagascar is known to occur. (In Tanzania the Tanga Coelacanth Marine Park has been established along 100 km of coastline from the Pangani River estuary to Mafuriko village north of Tanga City but *jarifa* gillnets continue to be used in this marine reserve, which results in mortalities to coelacanths and other marine life[5,15,48]).
- Extending or reinforcing marine protected areas, or areas under regional fisheries management plans, where coelacanth populations occur. In these areas, bottom fishing with demersal *jarifa* gillnets or longlines should be banned, although controlled pelagic fishing could continue, as in the iSimangaliso Wetland Park in KwaZulu-Natal.
- Continuing to enforce a strict ban on the export of coelacanth specimens or body parts in accordance with CITES regulations.
- Implementing an awareness raising campaign targeting fishing communities in areas where coelacanths may occur to discourage their capture.
- Providing incentives for fishers to release caught coelacanths which are still alive, as in the Comoros. The option of tagging and photographing caught coelacanths, in collaboration with registered fishers, in return for an incentive payment, should also be considered.
- Examining, more broadly, the pros and cons of allowing the continuation of the shark fishery using *jarifa* gillnets by assessing its impact on other artisanal fisheries and on threatened marine megafauna such as sawfish, coelacanths, turtles and dugongs. This assessment should take account of the ecological roles of these species.
- Continuing to mount a nationwide public awareness campaign, including displays, media releases, TV and radio interviews, public talks, talks at schools, and popular publications, on the importance of conserving the coelacanth to build its value as an iconic species.
- Encouraging traditional leaders to support coelacanth conservation, taking inspiration from the venerated status which the species already enjoys in the migrant v*ezo* fishing culture.
- Encouraging museums, zoological gardens, research institutions and tourism facilities to mount new and improved displays on the coelacanth in ecological dioramas, using accurate fibreglass replicas rather than real specimens, which are scientifically valuable and deteriorate under display conditions.
- Further developing the genetics laboratory at the University of Antananarivo so that the genomes of coelacanths and other endangered Madagascan species can be sequenced.
- Encouraging the Madagascar government to develop a National Strategy for the Conservation of the Coelacanth in consultation with scientists and natural resource managers and to implement the recommended conservation actions in terms of this National Strategy.
- Recommending that the Madagascar government should use the coelacanth as a flagship for marine conservation.

Conclusions

L. chalumnae is highly significant from several points of view. It belongs to an ancient group of fishes whose origins can be traced back 420 million years and which was close to the important evolutionary transition from water onto land about 320 million years ago.[5] The survival of living coelacanths has therefore provided a unique window into the past. Their enormous longevity; ability to survive four major extinction events over hundreds of millions of years; early adoption of advanced life-history traits such as live bearing; an extraordinarily long gestation period; unusual swimming, feeding, hunting and social behaviour; and their unique combination of physiological and anatomical characters, some of which they share with bony and cartilaginous fishes and others with tetrapods, set them apart from all other animals.[5,52,57,58] They are among the most valuable animals on the planet due to the unique messages about the past which they carry in their DNA.

Coelacanths have also played a key role in promoting public understanding of the theory of evolution and have become important flagship species for science. Coelacanths also have a rich symbolic history, probably more than any other fish. Their iconic image has been adopted by institutions, artists and craftspeople and has appeared on money and postage stamps. They are the emblem of the Department of Animal Biology, University of Antananarivo, and the mascot of the Comorian national football team. Their phoenix-like 'resurrection' from the past has inspired poetry, prose, songs, films, figures of speech and political metaphors.[5,52,57]

Madagascar may have the largest population of *L. chalumnae* in the world, much of it still to be discovered. Madagascar also has a research infrastructure comparable to other Western Indian Ocean countries which harbour coelacanth populations and an historical connection to coelacanth research. It is therefore appropriate that Madagascar should

play a key role in marine research and conservation, not only for the coelacanth, but also for the entire ecosystem which they share with other marine organisms.

Acknowledgements

We are grateful to Dr Jamal Mahafina for his permission to work on the IHSM's coelacanth collection and to Clemence Ravelo, coordinator of the IHSM Museum, for her assistance with the survey. Funding for the field survey was provided by Resolve sarl. We thank Gianni Insacco of the Natural History Museum in Comiso, Italy, for information on, and an image of, the first coelacanth specimen known to science caught in Madagascar. We thank Damon Stanwell-Smith and Thierry Cordenos for their photographs and the staff of the fisheries directorate of Toliara, fishing companies and hotels who provided access to coelacanths in their possession. We thank Tantely Tianarisoa for the preparation of the maps and Rik Nulens and Hans Fricke for useful comments on the manuscript. We also thank the anonymous reviewers for useful comments on the manuscript.

Competing interests

We declare that there are no competing interests.

Authors' contributions

A.C. was responsible for project leadership and management and contributed to conceptualisation, methodology, sample analysis, data analysis, validation and writing the initial draft and subsequent revisions. M.N.B. took the lead with the drafting of the initial manuscript and contributed to the conceptualisation, methodology, data collection, sample and data analysis, validation, data curation and the writing of revisions. M.R. conducted the field survey and contributed to data collection and sample analysis.

References

1. Smith JLB. A living fish of Mesozoic type. Nature. 1939;143(3620):455–456. https://doi.org/10.1038/143455a0
2. Smith JLB. Old Fourlegs. The story of the coelacanth. London: Longmans, Green & Co.; 1956.
3. Bruton MN. The fishy Smiths. A biography of JLB and Margaret Smith. Cape Town: Penguin; 2017.
4. Smith JLB. The second coelacanth. Nature. 1953;171:99–101. https://doi.org/10.1038/171099a0
5. Bruton MN. The annotated Old Fourlegs. The updated story of the coelacanth. Cape Town: Struik; 2017.
6. Millot J. Le troisième coelacanthe [The third coelacanth]. Le Naturaliste Malgache, Premier Supplément 1954:1–26. French.
7. Bruton MN, Coutouvidis SE. An inventory of all known specimens of the coelacanth, *Latimeria chalumnae*, with comments on trends in the catches. Env Biol Fish. 1991;32:371–390. https://doi.org/10.1007/978-94-011-3194-0_25
8. Nulens R, Scott L, Herbin M. An updated inventory of all known specimens of the coelacanth, *Latimeria* spp. Smithiana Special Publication. 2011;3:1–52.
9. Millot J. First observations on a living coelacanth. Nature. 1955;175(4452):362–363. https://doi.org/10.1038/175362a0
10. Cloutier R, Forey PL. Diversity of extinct and living actinistian fishes (Sarcopterygii). Env Biol Fish. 1991;32:59–74. https://doi.org/10.1007/978-94-011-3194-0_4
11. Heemstra PH, Freeman ALJ, Yan Wong H, Hensley DA, Rabesandratana HD. First authentic capture of a coelacanth, *Latimeria chalumnae* (Pisces: Latimeriidae), off Madagascar. S Afr J Sci. 1996;92:150–151.
12. Insacco G, Nulens R, Zava B. The coelacanth, *Latimeria chalumnae* Smith, 1939 at the Natural History Museum of Comiso, taxidermic preservation and notes on the other world specimens. Natura Rerum. 2016;4(1):25–38.
13. Bruton MN, Cabral AJP, Fricke H. First capture of a coelacanth, *Latimeria chalumnae* (Pisces, Latimeriidae) off Mozambique. S Afr J Sci. 1992;88:225–227.
14. De Vos L, Oyugi D. First capture of a coelacanth, *Latimeria chalumnae* Smith, 1939 (Pisces: Latimeriidae) off Kenya. S Afr J Sci. 2002;98(7/8):345–347.
15. Benno B, Verheij E, Stapley J, Rumisha C, Ngatunga B, Abdallah A, et al. Coelacanth (*Latimeria chalumnae* Smith 1939) discoveries and conservation in Tanzania. S Afr J Sci. 2006;102(9–10):486–490.
16. Fraser MD, Henderson BAS, Carstens PB, Fraser AD, Henderson BG, Dukes MD, et al. Live coelacanth discovered of the KwaZulu-Natal south coast, South Africa. S Afr J Sci. 2020;116(3/4), Art. #7806. https://doi.org/10.17159/sajs.2020/7806
17. Pouyaud L, Wirjoatmodjo S, Rachmatika I, Tjakrawidjaja A, Hadiaty RK, Hadie W. Une nouvelle espèce de coelacanthe. Preuves génétiques et morphologiques. [A new species of coelacanth. Genetic and morphological evidence.]. Comptes Rendus des Sciences Naturelles. 1999;322(3):261–267. https://doi.org/10.1016/S0764-4469(99)80061-4 French.
18. Erdmann MV. An account of the first living coelacanth known to scientists from Indonesian waters. Env Biol Fish. 1999;54:439–443. https://doi.org/10.1023/A:1007584227315
19. Goodman SM, Benstead JP. Updated estimates of biotic diversity and endemism in Madagascar. Oryx. 2005;39(1):73–77. https://doi.org/10.1017/S0030605305000128
20. Fricke H, Hissmann K, Schauer J, Reinicke O, Kasang L, Plante R. Habitat and population size of the coelacanth *Latimeria chalumnae* at Grande Comore. Env Biol Fish. 1991;32:287–300. https://doi.org/10.1007/978-94-011-3194-0_20
21. Hissmann K, Fricke H, Schauer H, Ribbink AJ, Roberts MJ, Sink K, et al. The South African coelacanths – an account of what is known after three submersible expeditions. S Afr J Sci. 2006;102:491–500.
22. Green A, Uken R, Ramsay P, Leuci R, Perritt S. Potential sites for suitable coelacanth habitat using bathymetric data from the western Indian Ocean. S Afr J Sci. 2009;105(3/4):151–154. https://doi.org/10.4102/sajs.v105i3/4.68
23. Ravololoharinjara M. Rapport de Mission à Toliara pour l'inventaire des captures et spécimens de coelacanthes [Mission report to Toliara for the inventory of catches and specimens of coelacanths]. Resolve sarl internal report, November 2019. French.
24. Vicente N. Un coelacanthe à Madagascar [A coelacanth in Madagascar]. Océanorama. 1997;27:11–15. French.
25. Niaina N. A peine capturé un coelacanthe rejoint la marmite – capture d'un spécimen de coelacanthe au large de Maintirano [As soon as a coelacanth has been captured, it joins the cooking pot – capture of a specimen of coelacanth off Maintirano]. Les Nouvelles. 2006 July 01. French.
26. Anon. Le coelacanthe, icône de la Journée mondiale du Tourisme [The coelacanth, icon of World Tourism Day]. Les Nouvelles. 2005 October 05. French.
27. Houssen S. Le coelacanthe [The coelacanth]. Flaque et cours d'eau: Le journal du Collège Etienne de Flacourt, Tuléar, Madagascar. 2012;11 January 10. Available from: http://www.collegetulear.fr/fichiers_utiles/college_francais__journal_N11.pdf French.
28. Cooke AJ, Bruton MN, Ravololoharinjara M. Detailed table of coelacanths known to have been caught in Madagascar: 1987 to 2019 [document on the Internet]. [updated 2020 Sep 23; cited 2020 Oct 02]. Available from: https://www.resolve.mg/download/MadagascarCoelacanthInventory_23Sept2020.pdf
29. Cooke AJ. Survey of elasmobranch fisheries and trade in Madagascar. In: Marshall NT, Barnett R, editors. The trade in sharks and shark products in the Western Indian and Southern Indian and South East Atlantic Oceans. Nairobi: TRAFFIC East/Southern Africa; 1997. p. 101–130.
30. Cripps G, Harris A, Humber F, Harding S, Thomas T. A preliminary value chain analysis of shark fisheries in Madagascar. Unpublished report to the Indian Ocean Commission SF/2015/34; 2015.
31. Du Feu TA. Fisheries statistics for the large meshed gill net fishery, north west Madagascar. Internal report, Promotion de la Pêche Maritime Traditionnelle et Artisanale; September 1998.
32. Bruton MN, Stobbs RE. The ecology and conservation of the coelacanth *Latimeria chalumnae*. Env Biol Fish. 1991;32:313–339. https://doi.org/10.1007/BF00007464

33. Stobbs RE, Bruton MN. The fishery of the Comoros, with comments on its possible impact on coelacanth survival. Env Biol Fish. 1991;32:341–359. https://doi.org/10.1007/BF00007465

34. University of Berkeley. Where did all of Madagascar's species come from? Understanding evolution. Evo in the news. 2009 October. Available from: http://evolution.berkeley.edu/evolibrary/news/091001_madagascar

35. Ali JR, Aitchison J. Gondwana to Asia: Plate tectonics, paleogeography and the biological connectivity of the Indian sub-continent from the Middle Jurassic through latest Eocene (166–35 Ma). *Earth Sci Rev.* 2008;88:145–166. https://doi.org/10.1016/j.earscirev.2008.01.007

36. Ludt WB, Rocha LA. Shifting seas: The impacts of Pleistocene sea-level fluctuations on the evolution of tropical marine taxa. J Biogeogr. 2015;42:25–38. https://doi.org/10.1111/jbi.12416

37. Cooke AJ, Brand J. Madagascar – a guide to marine biodiversity. New York: Wildlife Conservation Society; 2012.

38. Battistini R, Jouannic C, Mauget LA, Castellato G, Vernier RE. Morphologie et sédimentologie du canyon sous-marin de l'Onilahy (S-W de Madagascar) [Morphology and sedimentology of the Onilahy submarine canyon (S-W of Madagascar)]. Cahiers de l'ORSTOM Série Géol II. 1975;2:95–110. French.

39. Cripps G, Gardner CJ. Human migration and marine protected areas: Insights from Vezo fishers in Madagascar. Geoforum. 2016;74:49–62. https://doi.org/10.1016/j.geoforum.2016.05.010

40. Bruton MN, Armstrong MJ. The demography of the coelacanth *Latimeria chalumnae*. Env Biol Fish. 1991;32:301–311. https://doi.org/10.1007/BF00007463

41. Schartl M, Hornung U, Hissmann K, Schauer J, Fricke H. Relatedness among East African coelacanths. Nature. 2005; 435:901. https://doi.org/10.1038/435901a

42. Petit C. L'industrie des pêches à Madagascar [The fishing industry in Madagascar]. Paris: Société d›éditions géographiques, maritimes et colonials, Bibliothèque de la faune des Colonies Françaises; 1930. French.

43. Schaeffer D. Assessment of the artisanal shark fishery and local shark fin trade on Unguja Island, Zanzibar. Independent Study Project (ISP) Collection. 2004;536:1–29.

44. Dockerty T. International trade in shark fins. Cambridge, UK: Wildlife Trade Monitoring Unit, World Conservation Monitoring Centre; 1992.

45. Baker-Médard M, Faber J. Fins and (mis)fortunes: Managing shark populations for sustainability and food sovereignty. Mar Policy. 2020(113);103–805. https://doi.org/10.1016/j.marpol.2019.103805

46. Barnett R. The shark trade in mainland Tanzania and Zanzibar. In: Marshall NT, Barnett R, editors. Trade review: The trade in sharks and shark products in the Western Indian and Southeast Atlantic Oceans. Nairobi: Traffic-East/Southern Africa; 1997.

47. Fricke H, Hissmann K. Feeding ecology and evolutionary survival of the living coelacanth *Latimeria chalumnae*. Mar Biol. 2000;136:379–386. https://doi.org/10.1007/s002270050697

48. Bruton MN. Traditional fishing methods of Africa. Cape Town: Cambridge University Press; 2017.

49. Kiszka JJ, Muir C, Poonian C, Cox T, Amir OA, Bourjea J, et al. Marine mammal bycatch in the Southwest Indian Ocean: Review and need for a comprehensive status assessment. West Indian Ocean J Mar Sci.2008;7(2):119–136.

50. Rakotonirina BP, Cooke A. Sea turtles of Madagascar – their status, exploitation and conservation. Oryx. 1994;28:51–61.

51. Bruton MN. When I was a fish. Tales of an ichthyologist. Cape Town: Jacana Media; 2015.

52. Fricke H. Die jagt nach dem Quastenflosser, der Fisch, der aus der Urzeit kam [The hunt for the coelacanth, the fish that came from prehistoric times]. Munich: Verlag CH Beck; 2007. German.

53. Hale RC, Greaves J, Gundersen JL, Mothershead II RF. Occurrence of organochlorine contaminants in tissues of the coelacanth *Latimeria chalumnae*. Env Biol Fish. 1991;32:361–367. https://doi.org/10.1007/BF00007466

54. Fricke H, Hissmann K, Schauer J, Plante R. Yet more danger for coelacanths. Nature. 1995;374:314. https://doi.org/10.1038/374314a0

55. Plon S, Thakur V, Parr L, Lavery SD. Phylogeography of the dugong (*Dugong dugon*) based on historical samples identifies vulnerable Indian Ocean populations. PLoS ONE. 2019;14(9), e0219350. https://doi.org/10.1371/journal.pone.0219350

56. Froese R, Palomares MLD. Growth, natural mortality, length-weight relationship, maximum length and length-at-first-maturity of the coelacanth, *Latimeria chalumnae*. Env Biol Fish. 2000; 58:45–52. https://doi.org/10.1023/A:1007602613607

57. Fricke H. Living coelacanths: Values, eco-ethics and human responsivity. Mar Ecol Prog Ser. 1997;161:1–15. https://doi.org/10.3354/meps161001

58. Cloutier R. Patterns, trends, and rates of evolution within the Actinistia. Env Biol Fish. 1991;32:23–58. https://doi.org/10.1007/BF00007444

AUTHORS:
Roger L. Gibson[1]
Timothy Marais[1,2]
Lewis D. Ashwal[1]
Fenitra Andriampenomanana[1]
Andriamiranto Raveloson[1]
Andry H. Ramanatsoa[3]
Matthias Laubenstein[4]
Alexander Ziegler[5]
Warrick C. Fuchsloch[1,2]

AFFILIATIONS:
[1]School of Geosciences, University of the Witwatersrand, Johannesburg, South Africa
[2]AfriTin Mining, Johannesburg, South Africa
[3]Seismology and Infrasound Laboratory, Antananarivo Institute and Geophysical Observatory, University of Antananarivo, Antananarivo, Madagascar
[4]Gran Sasso National Laboratory, National Institute of Nuclear Physics, L'Aquila, Italy
[5]Microscopy and Microanalysis Unit, University of the Witwatersrand, Johannesburg, South Africa

CORRESPONDENCE TO:
Roger Gibson

EMAIL:
roger.gibson@wits.ac.za

DATES:
Received: 21 Apr. 2020
Revised: 16 July 2020
Accepted: 21 Aug. 2020
Published: 29 Mar. 2021

HOW TO CITE:
Gibson RL, Marais T, Ashwal LD, Andriampenomanana F, Raveloson A, Ramanatsoa AH, et al. An investigation of the 27 July 2018 bolide and meteorite fall over Benenitra, southwestern Madagascar. S Afr J Sci. 2021;117(3/4), Art. #8200. https://doi.org/10.17159/sajs.2021/8200

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:
meteor fireball, chondrite, infrasound, cosmogenic radionuclide analysis, seismic

FUNDING:
None

An investigation of the 27 July 2018 bolide and meteorite fall over Benenitra, southwestern Madagascar

Several dozen stones of an ordinary chondrite meteorite fell in and around the town of Benenitra in southwestern Madagascar during the early evening of 27 July 2018, minutes after a widely observed meteor fireball (bolide) transit and detonation. The event was confirmed by low-frequency infrasound recordings received at ~17h15 UTC (Coordinated Universal Time; 19h15 local time) at the Comprehensive Nuclear Test Ban Treaty Organization (CTBTO) infrasound station I33MG near Antananarivo, 542 km north-northeast of Benenitra. An energy release equivalent to 2.038 kt of TNT was calculated from the infrasound signals. Seismograph readings at the SKRH station 77 km north-northwest of Benenitra recorded a two-stage signal consistent with the arrivals of an initial air-coupled ground wave at 16h48:08 UTC and a stronger pulse at 16h49:22 UTC linked directly to the atmospheric pressure wave. The infrasound and seismic signal arrival times suggest that the bolide entry and detonation occurred at approximately 18h46 local time (16h46 UTC), entry was from the northwest, and the detonation hypocentre was located within ~20 km of Benenitra. Despite meteorite debris being found among buildings within Benenitra, there was no damage to structures or injuries reported. Eyewitness accounts and photographic records indicate that approximately 75 mostly intact stones were collected; however, the remoteness of the area, the rugged nature of the terrain and sales of fragments to meteorite collectors have limited scientific analysis of the fall and the extent of the strewn field. The total mass of recovered stones is estimated at between 20 kg and 30 kg, with one fragment of 11.2 kg and several of ~1 kg. Petrographic and mineral chemical analyses indicate that the stones belong to the L6 class of ordinary chondrites. Cosmogenic radionuclide analysis confirms that the fall is linked to the bolide event. The name *Benenitra* has been officially accepted by the Meteoritical Bulletin Database.

Significance:

- Eyewitness reports, CTBTO infrasound records, seismograph records and cosmogenic radionuclide analysis confirm a meteorite fall over Benenitra on 27 July 2018.
- Petrographic and geochemical analyses confirm that the meteorite is an L6 ordinary chondrite.
- Recovery of meteorite falls is rare; this is Madagascar's second known meteorite fall and the first that can be linked to a bolide.
- Regional and global science monitoring networks can be interrogated to improve the understanding of bolide events.
- Interaction of scientists with local communities is important to dispel misunderstandings around scientific phenomena, and can improve collection of data.

Introduction

Although the widespread availability of both static and mobile digital camera equipment for surveillance and social media purposes in the past two decades has substantially increased the documentation of meteor fireball events and the consequent chances of retrieval of debris (stones), meteorite falls still constitute fewer than 5% of recovered meteorites.[1,2] In addition, the chances of recovering stones from a fall in remote areas away from surveillance technology remain exceedingly low. In this article, we document a meteorite fall that fortuitously deposited stones with characteristic fused crusts in and around Benenitra in the Atsimo-Andrefana District of the Toliara Province of southwestern Madagascar (Figure 1). The fall can be linked via cosmogenic radionuclides, infrasound and seismograph recordings and eyewitness accounts to a bright meteor fireball (bolide) event that occurred at approximately 18h46 (local time) on 27 July 2018. This is Madagascar's second known meteorite fall, and the first that can be linked to a bolide.

Background of the area

The small town of Benenitra lies in the east of the Toliara Province, in one of the most sparsely populated parts of Madagascar (fewer than five people per square kilometre). The nearest significant town (commune population >20 000) is Bezaha, approximately 110 km by gravel road to the west (Figure 1). Topographically, Benenitra is located at 220 masl (metres above sea level) on the west bank of a large meander of the Onilahy River. Rounded hills dissected by perennial streams rise to 250–270 masl in the vicinity of the Onilahy River, but a deeply dissected plateau escarpment rises to 700–750 masl north of the town. The area is underlain by predominantly horizontal to shallowly west-dipping siliciclastic Palaeozoic sediments of Karoo Supergroup equivalence, which unconformably overlie older deformed and more steeply dipping strata that crop out to the east. The town lies on the southern edge of a broader area of artisanal alluvial sapphire mining that exploits the local rivers and ancient river palaeochannels.

The area receives an annual average rainfall of 730 mm, mostly in the summer months. Vegetation comprises mostly semi-arid grassland with isolated trees; however, dense forest is found along watercourses and the escarpment slopes. Most homestead settlements are located close to the river and its tributaries where crop fields (cassava, maize) and rice paddies occur; subsistence farming also involves cattle and goats.

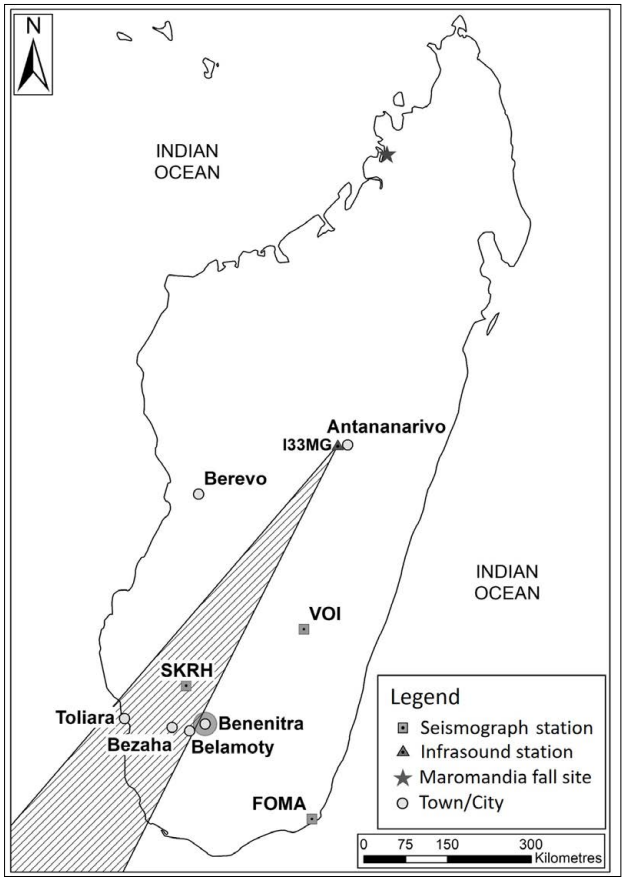


Figure 1: Map of Madagascar showing the locations of Benenitra, the Comprehensive Nuclear Test Ban Treaty Organization station I33MG and the SKRH, VOI and FOMA seismograph stations. The star indicates the location of the Maromandia fall of 5 July 2002. The hatched region indicates the vector range of the infrasound sources detected at I33MG (Figure 3b) and the dark grey circle marks the 20-km radius from Benenitra as used in the model in Figure 5.

History of the Benenitra Event and its investigation

During the early evening of 27 July 2018, residents in various towns and farming settlements in southwestern Madagascar observed a rapidly moving bright light in the sky for a few seconds before it exploded into multiple smaller glowing fragments with smoke trails that descended more steeply before extinguishing. In the vicinity of Benenitra, a few residents standing or sitting outside saw a rapidly approaching fireball, too bright to look at directly, that exploded in a bright flash. Other residents who were indoors reported a brief bright flash of light from outside. Many residents reported hearing a loud noise like rolling thunder a few minutes later, while a few mentioned feeling a ground tremor and/or slight wind accompanying the noise. A small group reported seeing rocks falling between the buildings in town and bouncing across the ground until they came to rest.

Some Benenitra residents who witnessed both phenomena were concerned that these were somehow linked to the imminent total lunar eclipse, which was due to commence at 20h15 local time. Consequently, they immediately contacted the local authorities. Furthermore, 3 days later, they also described the events to a geophysical team from the Antananarivo Institute and Geophysical Observatory (IOGA) of the University of Antananarivo, who were shown several of the collected fragments. A short article with an accompanying photograph of an ~1 kg stone with a smooth black surface was published in the *Triatra Gazette* in the provincial capital Toliara on 4 August 2018. The article reported that stones had been collected from locations up to 3 km apart in and east of the town. It quoted an unnamed IOGA geophysicist as saying that 'stones falling from the sky', whilst potentially dangerous for anyone unfortunate enough to be struck, were a well-known, albeit relatively rare, natural phenomenon. The article ended by speculating on whether the falling rocks were indeed parts of a meteorite or were related to a secret military exercise or failed rocket test, or some sort of prank; however, it made no mention of the fireball phenomenon.

Although it is likely that some of the Benenitra stones would have eventually found their way to meteorite dealers and private collectors, the reporting of the events of 27 July would probably have ended with the single *Triatra Gazette* article (no further follow-up was conducted) were it not for the fortuitous timing of a visit to the area by one of our number (Marais). Marais was overnighting with his driver-interpreter, Rene Robinson, in Benenitra on 29 July on his way to a geological project in the area east of the town and heard rumours of stones having fallen from the sky. He was shown a fragment with a black crust enclosing a greyish-white interior that was exposed on fractured surfaces (Figure 2). Although clearly not an iron meteorite, he noted millimetre-sized metallic grains that proved to be magnetic. Prior to departure the following day, Marais and Robinson were able to speak to several eyewitnesses who showed them photographs of multiple fragments collected after the event as well as locations in the town of where two of the larger stones, reputed to each weigh >0.5 kg, landed. The coordinates of these two fall sites were recorded, one of which had narrowly missed the town's electricity power station. He also noted that there were rumours of a larger stone having fallen 'east of the river'. At least three main fall sites in the town were reported by locals: the aforementioned site next to the power station (at 23°26.786'S, 45°4.692'E), another ~250 m west of the power station alongside the District Office, and a third ~400 m south of the power station next to the Catholic Church. At least one stone generated a 40-cm wide crater in the soil (Figure 2a), which was filled in after the stone was removed.

Whilst purchasing supplies in the town market in preparation for departure, Marais noticed several broken fragments lying on a chair behind the store counter and managed to successfully negotiate their purchase. These were delivered to the School of Geosciences at the University of the Witwatersrand for analysis upon his return to Johannesburg on 3 August, where hand specimen and petrographic analysis confirmed a chondrite meteorite type (see next section).

Prior to leaving Madagascar, Marais had already contacted the International Meteor Organisation (IMO) to enquire if any recent meteor events had been recorded over southern Africa, but was informed that no reports had been received in the previous month. In view of this, a three-pronged approach to collect more information was immediately adopted: first, in the absence of any audiovisual records, eyewitness accounts needed to be more thoroughly documented via a standardised questionnaire; second, any hard scientific data relating to the timing and nature of the bolide needed to be found; and third, scientific confirmation was needed to tie the meteorite unequivocally to the bolide.

The primary objective of the questionnaire was to obtain unbiased data to help constrain the bolide trajectory and the extent and size of the fall, such as was possible after the 2002 Thuate fall in Lesotho.[3,4] Without the option of scientists on the ground, this task was assumed by R. Robinson, who used an adapted version of a questionnaire developed for the 6 June 2018 bolide in Botswana[5] to conduct face-to-face interviews with witnesses not only from Benenitra but also from the nearby towns of Belamoty, Bezaha and Berevo (Figure 1). These provided first- and second-hand accounts of the event, as well as a range of perspectives of the sequence of events. One of the interviewees reported that a '15 or 20 kg' stone had landed 3 km east of the town, but that this had disappeared, possibly to a dealer; residents also reported that a similar-sized stone was collected 8 km west of the town and was apparently etched with the number '144' that could not be erased, despite repeated rubbing (Robinson R 2018, personal communication).

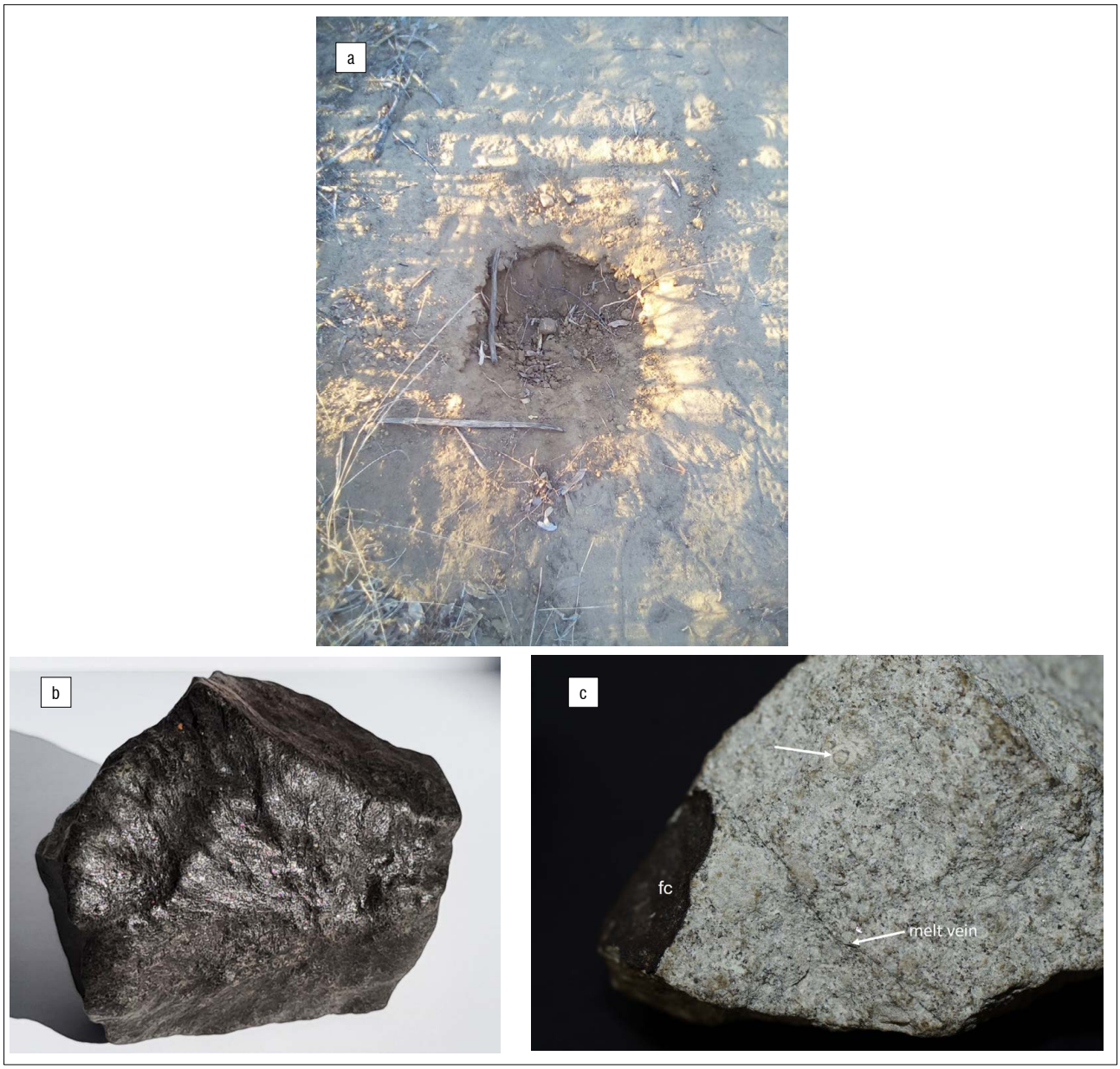


Figure 2: (a) Photograph of plunge pit in sandy soil reportedly caused by an ~1 kg stone at a site on the outskirts of Benenitra town, provided to R. Robinson by an anonymous Benenitra resident. The pit is ~40 cm in diameter and appears slightly elongated in a southeast to south-southeast direction (top of image). (b) Typical Benenitra meteorite stone showing the well-preserved black fusion crust with flow striae and relatively angular shape with sharp edges. Base is 43 mm. Photograph: T. Marais. (c) Close-up of stone 2018Ben-1 showing thin remnant fusion crust (fc), and interior comprising millimetre-sized rounded chondrules (arrow), metal-sulfide aggregates (dark or reflective spots), and thin melt veinlet. Sample was likely broken open by the person who found it. Base of photo is 35 mm.

In themselves, eyewitness accounts of bolides can be both contradictory and relatively unreliable as the unexpected and transient nature of the events challenges witness perspectives of direction, proximity, scale and time. Fortunately, the extremely energetic nature of bolide events makes their detection by a variety of ground- and satellite-based sensors possible. One such sensor array is the global network of 60 infrasound stations maintained by the Comprehensive Nuclear Test Ban Treaty Organization (CTBTO) through its International Monitoring System. Whilst the International Monitoring System network is primarily directed at recording and locating atmospheric or ground-based nuclear detonations through detection of low-energy sound waves that are able to travel considerable distances through the atmosphere, it has also proved invaluable in recording and investigating large fireballs and bolides entering Earth's atmosphere.[6] The specific design of infrasound stations allows both the direction and size of the energy source to be measured. The Seismology and Infrasound Laboratory (LSI) at the IOGA monitors the I33MG station located west-southwest of Antananarivo (Figure 1). Following an enquiry, staff at the LSI were able to confirm that a significant energy release event had been detected by the infrasound sensors at ~17h15 UTC (19h15 local time) from a source lying southwest of the station (Figures 1 and 3).

The infrasound data could not precisely constrain the energy release event to the Benenitra vicinity. Additional information was thus sought through the AfricaArray seismograph network. Records for 27 July from the three active Malagasy seismographs, SKRH, VOI and FOMA (Figure 1) were examined. Only SKRH, located 77 km north-northwest of Benenitra, recorded a distinct signal (Figure 4) that further analysis (see below) showed to be compatible with an atmospheric, rather than tectonic, source (e.g. Roelofse and Saunders[7]).

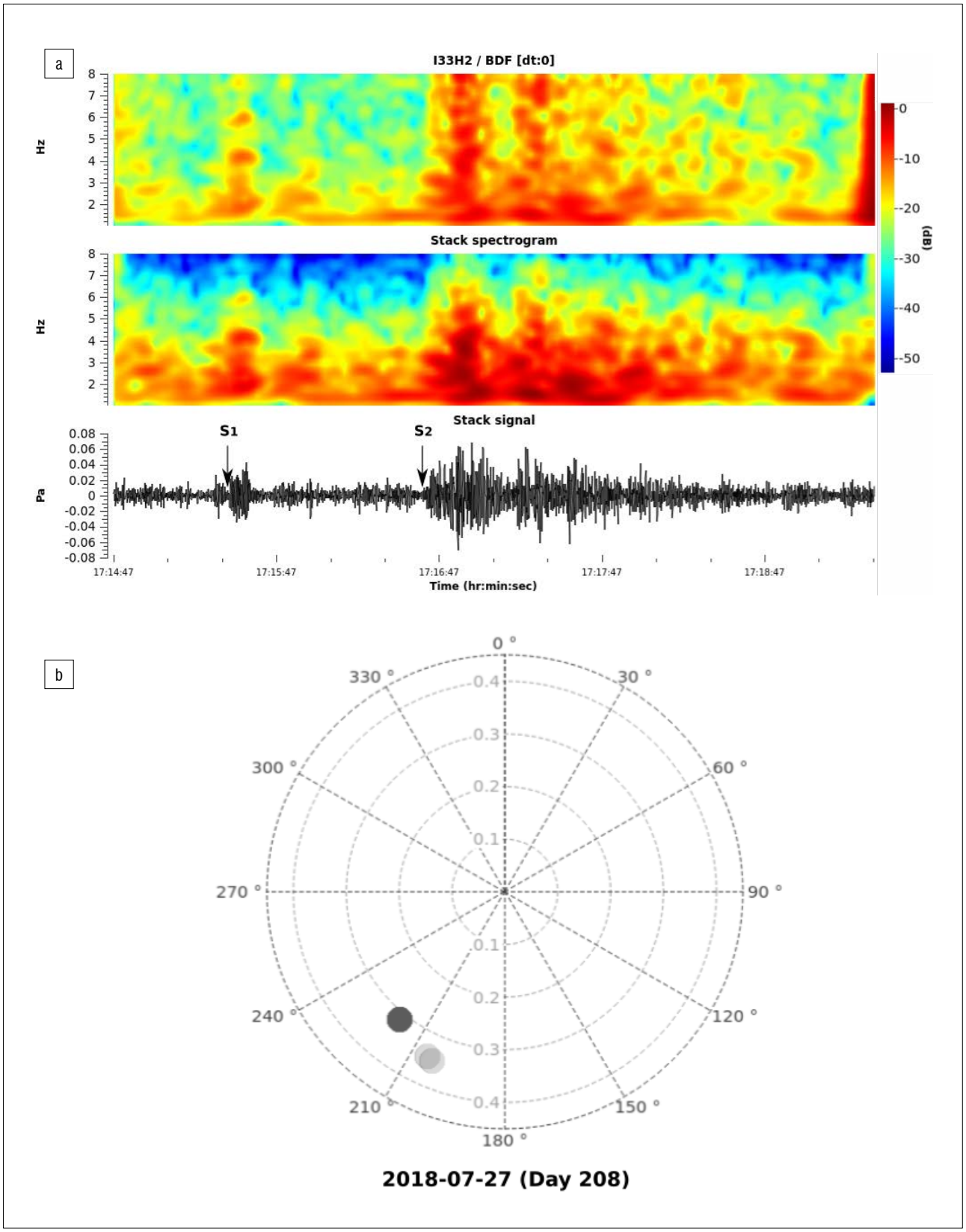


Figure 3: (a) Infrasound record of Station I33MG (19°00.65'S, 47°18.3'E; Figure 1) on 27 July 2018 showing a low-frequency (4 Hz) disturbance (S1) commencing at 17h15:27 UTC (19h15:27 local time) for 10 s and a second disturbance (S2) commencing at 17h16:24 UTC for 150 s. The upper spectrogram indicates the energy of the signal at one of the four elements (I33H2) at the station. The lower spectrogram is the signal stack of all four elements of the station. The low frequency and the prolonged nature of the signal are consistent with bolide-generated infrasound waves.[21,29] (b) Polar diagram showing the azimuth of detection of the first (grey dots; at 206°) and second (black dot; at 220°) waves. Based on the bolide timing and hypocentre location established from the seismograph data and fall parameters, the calculated celerities of the two waves (0.305 km/s and 0.301 km/s, respectively) are consistent with stratospherically ducted waves.

The presence of the signal at SKRH and not at the other two stations provided independent confirmation of the location of the infrasound source and its proximity to Benenitra. Given the limited distance between Benenitra and the station, the arrival time of the signal at ~16h48 UTC (18h48 local time) constrained the bolide to no more than a few minutes earlier.

The final step was to prove that the stones were linked to the bolide. Because Earth's atmosphere acts as a shield against high-energy cosmic radiation that produces a range of short-lived isotopes in asteroids, the timing of a fall can be determined by measuring the amounts and proportions of these cosmogenic radionuclides still present within a meteorite. Analysis of fragment 2018Ben-1 at the Gran Sasso National Laboratory (Italy) in September 2018 established that short-lived radionuclides such as ^{48}V (half-life = 16 days), and ^{51}Cr (half-life = 27.7 days) were present in sufficient quantities to confirm a very recent fall.

Description of the Benenitra meteorite

Two broken fragments with small areas of fusion crust weighing 82 g (Sample 2018Ben-1) and 14 g (2018Ben-2), respectively, were available for initial inspection and petrographic analysis. Subsequent to this, two other stones, weighing 99.32 g (2018Ben-3) and 26.6 g (2018Ben-4), each preserving an almost complete fusion crust, were also made available for study. In total, as of February 2020, photographic or physical evidence exists of at least 75 stones, most of which exceed 50 g in mass, and almost all of which show >95% fusion crust. In February 2020, an 11.241 kg stone with complete fusion crust was advertised for sale on eBay[8], consistent with the prior rumours of the large stone that landed in a cassava field outside the town. Preliminary estimates suggest a minimum total mass of ~22 kg of recovered material (T. Marais, unpublished data).

The stones are moderately to strongly magnetic. They are mostly angular to subangular, with somewhat flattened rectangular or asymmetric pyramidal to conical shapes (Figure 2b). Faces may be flat, slightly curved and locally scalloped, and edges relatively sharp or slightly rounded. Apart from rare broken fragments of larger stones (Figure 2c), stones are mostly completely covered by a <1 mm thick bluish-black to brownish-black fusion crust that may display flow features (Figure 2b).

Broken fragments and chipped edges of stones reveal a slightly friable, grey-white rock with a granular appearance in which disseminated metal and sulfide grains up to 1 mm in size, and a few white, 1–5 mm, spherical to elliptical chondrules occur (Figure 2c). A very thin (<0.5 mm), subplanar, blue-grey to black veinlet occurs in Sample 2018Ben-1 (Figure 2c). Most of the exposure of stone interiors is attributed to damage upon hard impact with terrestrial rock outcrops but several stones were reportedly broken open manually. Given their immediate collection within minutes to days after the fall, the stones display a weathering level of W0.[9]

Transmitted and reflected light petrographic analysis of five thin sections from stones 2018Ben-1 (two), 2018Ben-3 (one) and 2018Ben-4 (two) confirmed the presence of 1–4.5 mm (mean of 2 mm) olivine and/or pyroxene-bearing chondrules and, thus, that Benenitra belongs to the chondrite class of stony meteorites. Electron microprobe analysis has indicated that the samples are dominated by olivine and low-Ca pyroxene, with minor feldspar, high-Ca pyroxene, sulfide and Fe-Ni metal[10], confirming an ordinary chondrite[11,12]. Based on its intermediate metal+sulfide content and its olivine and pyroxene mineral chemistry[10] it can be classified more precisely as an L-chondrite[13,14]. The homogeneous mineral compositions, slightly elevated CaO content of low-Ca pyroxene and evidence for relatively coarse (10–50 μm but locally to 100 μm), well-equilibrated, metamorphic textures confirm metamorphic stage M6.[10,11,15-17] The predominance of fractures and patchy to undulose extinction in the main silicate minerals suggests an overall relatively low shock state (stage S3).[10,18] Combining all criteria, Benenitra is classified as an L6 ordinary chondrite showing shock stage S3 and no weathering (W0).

Constraining the Benenitra bolide

A bolide is a bright fireball caused by the penetration of Earth's upper atmosphere by a decimetre- to metre-sized meteoroid travelling at hypersonic speed (>11 km/s). A meteoroid large enough to survive and maintain its hypersonic momentum deep enough into the

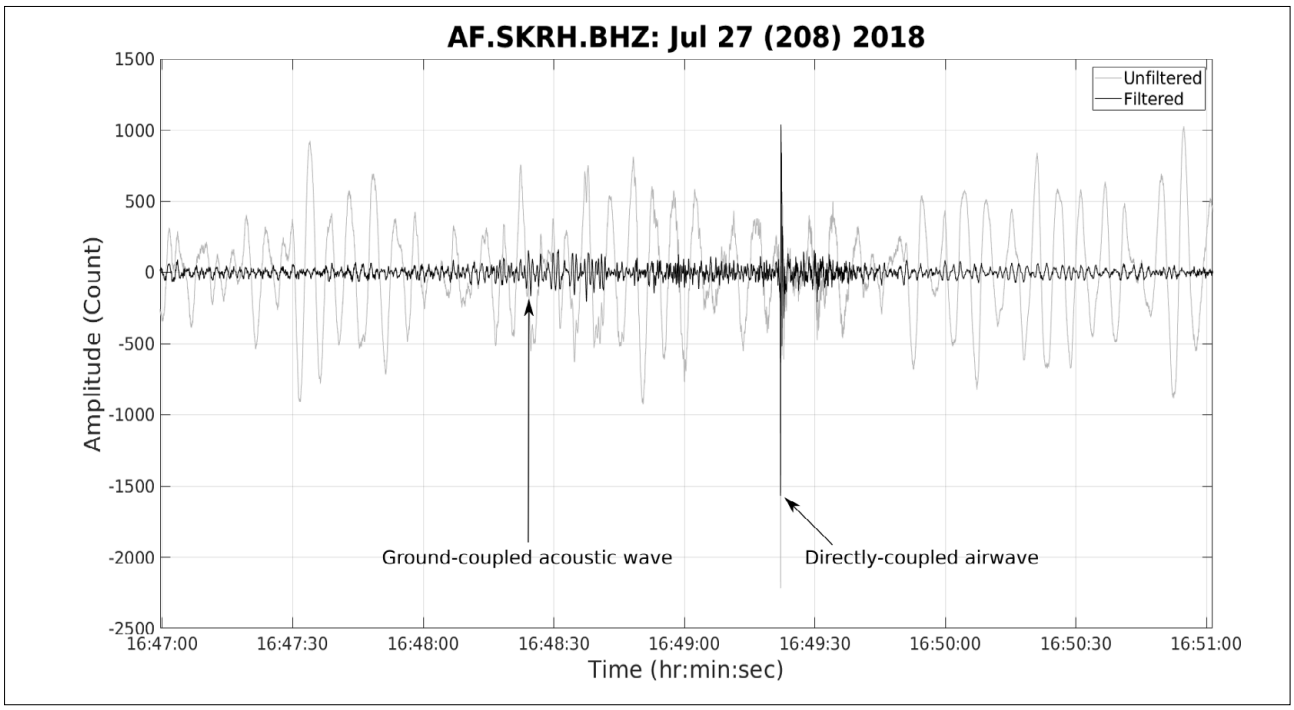


Figure 4: Seismic record from SKRH seismograph station, 77 km north-northwest of Benenitra (22.83°S, 44.75°E; Figure 1), on 27 July 2018, showing the unusual ground motion signature lasting ~100 s, consistent with a bolide source. Grey and black lines are the unfiltered and filtered (0.5–20 Hz) data, respectively. The air-coupled ground wave (see Figure 5) arrived at 16h48:08 UTC in the form of dispersed pulses within the frequency range of ~0.5–5 Hz, consistent with Rayleigh waves.[7,19] The sharp pulse at 16h49:22 UTC displays a W-shaped impulsive onset with downward first motion, consistent with overpressure, and has a frequency range of ~0.1–10 Hz, consistent with a directly coupled airwave.[30]

atmosphere is ultimately likely to explode spectacularly as stresses caused by increased atmospheric resistance and thermal effects force its catastrophic disintegration. If any meteoroid fragments survive this terminal burst, they will fall to Earth's surface at considerably slower terminal velocities (200–300 m/s). A bolide event (Figure 5) thus comprises: (1) an initial *luminous flight phase* typically lasting a few seconds as the meteoroid falls from ≥100 km altitude to below 50 km altitude at hypervelocity, becoming incandescent and ablating because of frictional heating, (2) the *terminal burst* at altitudes typically between 50 km and 30 km, and then (3) the *dark flight phase* that may last several minutes. Given that the hypersonic phase and airburst happen at extreme altitudes, the associated sonic booms, travelling at the speed of sound (0.343 km/s), would also only reach the ground up to several minutes after the fireball disappears.

Different observations and/or measurements relating to each of these three phases can provide information about key bolide characteristics, such as size, trajectory and terminal burst altitude. For instance, size may be estimated from fireball brightness (luminosity), total or terminal burst energy release, and radioisotope analysis of surviving fragments, whereas information about trajectory (azimuth and angle of inclination) may come from fixed camera footage and (reliable) eyewitness descriptions of the luminous flight phase as well as the distribution of stones within the strewn field[3,4], and the terminal burst altitude and hypocentre may be constrained from arrival times of different types of seismic signals[7,19]. As an example, with the Thuate fall of 21 July 2002, the delineation of the strewn field, and internal distribution of stones within it, were instrumental in helping to constrain the azimuth and angle of incidence of the bolide.[4] Conversely, astronomical, satellite, infrasound and fixed camera data were critical in first constraining the trajectory and airburst altitude of the Central Kalahari (Botswana) bolide of 2 June 2018; these, in turn, were used to predict the location of the strewn field that guided the meteorite ground search teams.[5]

In comparison to these two bolide-fall events, reliable visual data for the Benenitra event is limited, but the fall of stones within Benenitra places the town close to the hypocentre of the terminal burst (Figures 1 and 5). Furthermore, eyewitnesses noted that the fireball was visible for several seconds, suggesting a moderately steep trajectory, and that it approached Benenitra from a broadly northwesterly direction. The latter is consistent with the infrasound source azimuth of 206–220° registered at I33MG (Figures 1 and 3b). Based on the wave amplitude, the combined radiative energy release during phases 1 and 2 of the bolide was estimated at an equivalent of 2.038 kt of TNT by LSI staff. Studies of data from other bolide-fall events[6,19] suggest that the typical average entry velocity and angle of incidence of a bolide during phase 1 are likely to be ≤17 km/s and 45°, respectively. Thus, with the meteorite type (and therefore, its density) known, these parameters can be used to further constrain the Benenitra bolide via the Earth Impact Effects Program.[20] Assuming a spherical meteoroid, calculations suggest that an ~2 kt energy release would require an initial ~3 m diameter meteoroid and that the terminal burst would occur at an altitude of ~40 km. These figures must be regarded as a first-order approximation only. Accounts from the Benenitra witnesses (Robinson R 2018, written communication) that the fireball was 'bright/big like the sun' and 'bright like welding' and that the sound was 'as loud as a gunshot' might suggest a larger bolide or one that detonated at lower altitude either because of greater size or slower velocity. However, it is difficult to verify these observations without better-constrained audiovisual records.

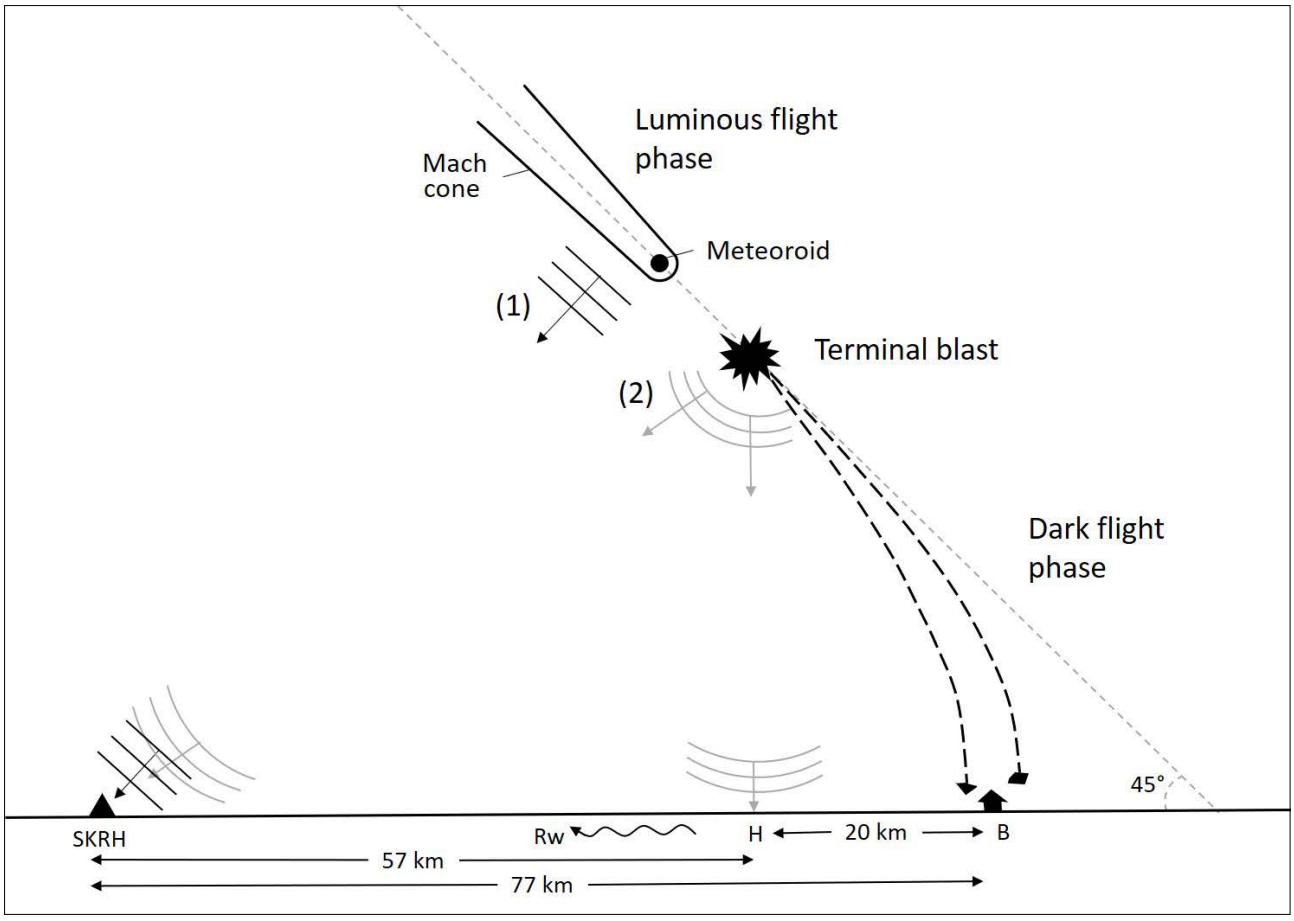


Figure 5: Model solution for the bolide-induced signal received at SKRH station (see Figure 4), involving a bolide travelling at a 45° angle and a terminal burst at 40 km altitude, 20 km north-northwest of Benenitra (B). The atmospheric pressure waves triggered by the bolide during hypervelocity flight (1) and/or the terminal burst (2) couple directly with the surface at the recording site (SKRH), but they arrive after the air-coupled Rayleigh waves (Rw). The latter are created when the acoustic waves of the terminal burst (2) couple with Earth's surface at the hypocentre (H) to form ground seismic waves which then propagate outwards. Because the velocity of the seismic waves in rocks is far higher (3–5 km/s) than in air (0.343 km/s), the air-coupled Rayleigh waves arrive before the direct airwave.

Combining the different arrival times of the ground-coupled seismic and acoustic waves at the SKRH station (Figure 4) with the results from the Earth Impact Effects Program can assist in constraining the bolide trajectory.[21] In the simplest scenario, the bolide is considered to have travelled in a south-southeast direction along a direct line between SKRH and Benenitra at a 45° angle before detonating at 40 km altitude and depositing stones in the town (Figure 5). Following the terminal burst, stones would have been scattered downrange of the hypocentre along parabolic trajectories whose shapes depend primarily on mass. As a rule of thumb, mid-sized stones in the 0.1–1-kg range could be expected to travel approximately halfway between the airburst hypocentre and the projection of the line of the luminous flight phase trajectory to ground level (Lyytenin E 2020, personal communication; Figure 5). In the scenario presented, a terminal burst located ~20 km north-northwest of Benenitra would still lead to a fall of mid-sized stones within the town. From this geometry, the acoustic wave from the terminal burst at 40 km altitude would take 117 s to reach the ground at the hypocentre and then a further 19 s (based on a surface Rayleigh wave velocity of ~3 km/s)[22] to reach SKRH. The acoustic wave from the terminal burst would travel hemispherically through the atmosphere at 0.343 km/s towards the SKRH station, a direct distance of 70 km (Figure 5), reaching the station after 204 s. The ballistic shock wave from the luminous flight phase, travelling essentially cylindrically perpendicular to the bolide path (Mach cone, Figure 5), would have to travel 69 km to reach the station and would take 200 s to arrive after being generated. Accepting that the sonic boom from the Mach cone actually originated no more than a few seconds before the terminal burst and, thus, that the difference in starting times of the various waves is negligible, the model predicts that the air-coupled ground wave should have arrived at SKRH 64–68 s before the direct airwaves. This timing is a surprisingly good fit to that observed in the SKRH record (74 s; Figure 4); the 6–10 s difference can be accounted for by a relatively small change in terminal burst altitude or shift in bolide trajectory. Although the model trajectory is unlikely to correspond exactly to the actual bolide trajectory, the good fit of the data suggests that it is a good first-order approximation of the bolide of 27 July 2018.

As Benenitra is a representative of the most common meteorite fall type and the velocity and incidence parameters are also the most commonly measured values for bolides, there is a high level of confidence in the proposed constraints on the Benenitra bolide's general trajectory and terminal burst altitude. A broadly northwest-to-southeast trajectory also places the bolide hypocentre over the largely unpopulated plateau north of Benenitra, explaining the lack of observational data of stronger bolide effects. This may also have restricted recovery of smaller stones, which would have fallen closer to the hypocentre.

Discussion

A total of 81 meteorites from southern Africa are listed in the Meteoritical Bulletin Database[23], with the overwhelming majority from South Africa (46), Namibia (19) and Botswana (12). Of these, falls constitute 20 of the South African meteorites, 2 each from Namibia, Zimbabwe and Madagascar, and 1 each from Botswana, Lesotho and Swaziland (Mozambique has no reported meteorites). In fact, the only meteorites recorded from Zimbabwe, Madagascar, Lesotho and Swaziland are falls. Owing to the limited nature of the reports, it is unclear whether these falls were directly linked to observed bolides. Furthermore, whilst bolides have been observed in recent years associated with falls in Lesotho (2002), Botswana (2018) and Madagascar (2018, this study), the last fall recovered in South Africa was in 1973. Disappointingly, southern Africa's largest recorded bolide, on 21 November 2009, which detonated in the vicinity of Alldays close to the border between South Africa, Botswana and Zimbabwe[7], failed to yield any stones despite extensive search efforts, particularly in Botswana (McKenzie R 2009, personal communication). In South Africa, bolide sightings are regularly reported in the social and news media, but with no associated discoveries of stones. Such statistics are not exceptional, as Graham et al.[2] estimated that only five or six falls per year were recovered globally, although in recent decades this has increased to an average of approximately a dozen per year. The increased success rate can be attributed to factors such as increased population density, dedicated scientific instrumentation and CCTV and other surveillance systems, as well as increased public awareness.[23] A total of 158 falls were recorded in Africa between 1801 and 2014[1], and a further 10 have since been added to the Meteoritical Bulletin Database[23]; however, it is likely that many more in the more remote parts of the continent have gone unreported. In this context, it is clear that the Benenitra bolide and fall is a rare and significant event.

According to the Meteoritical Bulletin Database[23], the proportion of falls among meteorite discoveries in southern Africa (35%) is significantly higher than the global average of ~5%. The global average has become progressively lowered by the large numbers of meteorite finds in recent decades in northwest Africa and Antarctica. In fact, the exceptional number of northwest Africa finds actually skews Africa's ratio to only 4% falls.[1]

Benenitra represents both Madagascar's second fall and second meteorite – the first being Maromandia on the northern coast on 5 July 2002 (Figure 1).[24] Whilst the single observer of the Maromandia fall reported seeing two stones with a combined mass of ~6 kg fall, no fireball or acoustic effects were reported. In contrast, the Central Kalahari bolide and fall of 2 June 2018 – less than 2 months before the Benenitra event – is one of the most comprehensively studied examples of these phenomena. Observational data related to the Central Kalahari event commenced with identification and tracking of its originating asteroid 2018 LA by the Catalina Sky Survey and ATLAS telescopes several hours before the fall. These accurately predicted that the asteroid was on a collision course with Earth, in addition to collecting important spectral and other data. Constraints on its actual atmospheric trajectory and terminal detonation altitude were obtained from CCTV and infrasound data that were then used to constrain its projected dispersal ellipse. Ultimately, several search expeditions were guided by this scientific trajectory analysis to locate and retrieve fragments of the meteorite.[5]

In the case of the Central Kalahari fall, it is indisputable that without the considerable data and resources to analyse the bolide trajectory, the meteorite itself would most likely never have been found. Benenitra and Central Kalahari thus represent two significantly different approaches in recovering meteorite falls. Between these extremes lies the Thuate event of 21 July 2002[3,4] where excellent local communication networks were rapidly able to ascertain that an extraordinarily loud, sustained, noise heard over a broadly circular area some 250 km in diameter in Lesotho and central South Africa was linked to a bolide and smoke trail that was seen by only a few observers (owing to there being 80% cloud cover at the time). News of the meteorite fall itself actually only surfaced 3 weeks later, once a regional public-service police newspaper published complaints by villagers of stones falling from the sky that damaged several structures.[3] The accessibility of the fall site to scientists from various universities and the Geological Survey of Lesotho then enabled extensive subsequent interaction between scientists and locals over the next 2 months, which raised public awareness and assisted in the collection of both observational data and more stones. This, in turn, allowed delineation of a 7.4 x 1.9 km, east-west-elongated, elliptical strewn field. Favourable field conditions (no rainfall, fallow winter fields and dry highveld grassland in a reasonably well-populated area) facilitated the collection over a 2-month period of at least 1029 stones, totalling 45.3 kg, one-third of which are smaller than 10 g.[4] The thoroughness of the collection of stones in this case allowed for the reconstruction of key elements of the bolide[4] that would otherwise not have been possible.

Deciphering the Benenitra mythology

Bolide and meteorite fall events are rare and highly unusual phenomena that disrupt, and can significantly challenge, the way in which witnesses view the natural world; simultaneously, as dramatic, widely and highly visible phenomena they present excellent opportunities for education. In the narrative of the field research conducted following the Thuate fall in Lesotho in 2002[3], it is clear that initial responses to the fall varied from anger and suspicion of criminal acts to fear of possible supernatural forces that led to reports to the local authorities. In the case of Benenitra, no stones appear to have actually hit any buildings and no

injuries were reported; however, interviews revealed several myths and misconceptions concerning the fall:

Harbinger of the Apocalypse: The timing of the bolide and fall relative to the well-publicised total lunar eclipse later that evening led to immediate alarm and speculation among witnesses that the bolide and stones may have been derived from the Moon whilst it was being 'devoured' by a dark shadow. This 'disintegration' of the Moon, in turn, was seen as a sign that Armageddon or the Apocalypse had arrived. Concerned townspeople thus contacted the local authorities and the visiting IOGA scientists, who tried to allay these fears. Whether their explanation was sufficient or not, the 'recovery' of the Moon following the eclipse may have ameliorated these fears. Certainly, by the time the matter was reported in the *Triatra Gazette* a week later, more prosaic explanations of a failed military experiment (perhaps aided by the '144' specimen that suggested some sort of human agency) or mischief were ventured and the matter was not speculated on further.

Stones too hot to touch: Eyewitness reports indicate that at least one of the large stones that was seen landing in the town was approached cautiously and not touched until 'at least two minutes' later. The eyewitness reports of the bolide confirmed that 'bright sparks like fireworks' that trailed smoke were seen descending from the final airburst. A direct connection was thus made that the objects that landed had 'been on fire' no more than a few minutes earlier, hence the caution. The first person to pick up the stone remarked that he was surprised at how 'icy cold' it felt, and that 'it was still cold more than an hour later'. This misconception about meteorites being hot is common. Despite the brief fireball in which the outer surface of the bolide reaches temperatures in excess of 1700 °C that cause it to melt[25], the hypersonic velocity ensures that the melt is very efficiently stripped (ablated) from the surface into the atmosphere, thus continuously exposing fresh material to the frictional heating and high pressures. Beyond limited heating of their surfaces by solar radiation, asteroids in space would have an internal temperature of approximately -100 °C.[26] Consequently, once the bolide fragments were decelerated to speeds too slow to generate frictional heating after the airburst, the core of the bolide remaining after ablation would have been capable of quenching and completely cooling the last vestiges of the surface melt within seconds, and still remain significantly colder than air temperature for hours after landing.

Precious gems in the stones: Photographic evidence reveals that fusion crust preservation on almost all the Benenitra stones is generally >95%. However, it appears that several of the stones, including 2018Ben-1 (Figure 2c), were broken open by locals. Similar behaviour was reported for some Thuate stones[3], with locals explaining that they believed the stones might have contained diamonds (much of Lesotho's income is derived from diamond mining within the country or from employment on mines in South Africa). Benenitra lies within an area of alluvial sapphire mining and locals are well accustomed to 'stones' as an income opportunity. The connection that the stones that fell to the ground were initially glowing brightly after the fireball disintegrated likely prompted some witnesses to investigate whether the source of the glow might have been a precious gem within its interior.

Contagion from the stones: In advertising a sizeable number of Benenitra stones for sale on its website, one commercial meteorite dealer noted claims that several locals reported experiencing mild irritation after being in close proximity to stones:

> *But to the dismay of all, it seems that the stones gave off a gas with multiple effects, in contact, according to the sensitivity of people. For some, the contact with the stone causes a small irritation to the eyes and makes tears fall, while for others, irritation is felt in the nostrils and causes a momentary runny nose.*[27]

This story was also relayed to Marais by locals a few days after the fall, but not thereafter (Robinson R 2018, personal communication). The claims are somewhat similar to respiratory, vomiting and diarrhoea ailments experienced by residents of Carancas (Peru) who rushed to examine a 14-m wide crater caused by a meteorite impact in clay soil near their village on 15 September 2007. These illnesses were linked to noxious gases emanating from the crater that had rapidly filled with water.[27] The initial claim that these were extraterrestrial toxins was later refuted by health experts who at first speculated that the impact may have liberated arsenic from the saturated soil. The most recent suggestion is that the large mass of the meteorite (estimated at 2 t) meant that it was still moving sufficiently fast when it impacted, that it had not lost all its frictional heat, and that sulfides in the hot surface layer then reacted with the water that flooded the crater, releasing toxic gases.[28]

Summary

The fall of stones of an L6 ordinary chondrite meteorite in and around the town of Benenitra in southwestern Madagascar is linked to a bolide witnessed across the region at approximately 18h46 local time (16h46 UTC). Despite multiple stones falling in the town, no damage or injuries were reported. This is Madagascar's second recorded fall, and second recovered meteorite. The bolide released an estimated 2.038 kt equivalent of energy on its transit through the upper atmosphere and final terminal burst, sufficient to be recorded by the I33MG infrasound station 542 km north-northeast of the town, and by the SKRH seismograph station 77 km to the north-northwest. Both signals show distinctive characteristics known to be associated with bolides. Data from both stations confirm eyewitness reports that the bolide was travelling in a broadly southeast to south-southeast direction. The reported recovered mass of stones of ~22 kg is likely to represent a small fraction of the total mass of the fall.

Acknowledgements

R. Robinson is thanked for his exceptional assistance in identifying and interviewing eyewitnesses of the bolide-fall event. C. Majola is thanked for preparing the thin sections. Helpful comments from E. Lyytenin improved the discussion of bolide trajectory constraints and R. Durrheim provided input on the seismic interpretation. J. Gibson is thanked for assistance with the initial translation of questionnaire responses.

Competing interests

We declare that there are no competing interests.

Authors' contributions

R.L.G.: Project leader; conceptualised and managed the project and recruited other members; described meteorite petrography and mineral chemistry for classification; developed bolide trajectory model; wrote initial manuscript draft. T.M.: Collected samples and liaised with W.C.F. regarding initial research; primary South African contact for eyewitness reports; co-wrote introductory section of manuscript and edited drafts; prepared interpretive diagrams. L.D.A.: Described meteorite petrography and mineral chemistry for classification; assisted in writing of initial draft and significant further edits. F.A.: Located and interpreted seismic data; developed bolide trajectory model; wrote section on seismic data and made editorial comments on manuscript. A.R.: Located and interpreted seismic data; translated eyewitness questionnaire and correspondence into/from Malagasy from/to English. A.H.R.: Located and interpreted infrasound data; wrote relevant section and made editorial comments on manuscript. M.L.: Analysed cosmogenic radionuclides, contributed section in manuscript and editorial comments. A.Z.: Generated mineral chemical data for meteorite classification. W.C.F.: Liaised with T.M. regarding initial sample-based research; prepared interpretive diagrams.

References

1. Khiri F, Ibhi A, Saint-Gerant T, Medjkane M, Ouknine L. Meteorite falls in Africa. J Afr Earth Sci. 2017;134:644–657. https://doi.org/10.1016/j.jafrearsci.2017.07.022
2. Graham AL, Bevan AWR, Hutchison R. Catalog of meteorites. 4th ed. Tucson, AZ: University of Arizona Press; 1985.
3. Ambrose DP, Reimold WU, Buchanan PC. The Thuathe meteorite of 21 July 2002, Lesotho: Mapping the strewn field and initial mineralogical classification. S Afr J Sci. 2003;99:153–159.

4. Reimold WU, Buchanan PC, Ambrose D, Koeberl C, Franchi I, Lalkhan C, et al. Thuate, a new H4.5 chondrite from Lesotho: History of the fall, petrography and geochemistry. Meteor Planet Sci. 2004;39(8):1321–1341. https://doi.org/10.1111/j.1945-5100.2004.tb00949.x

5. University of Helsinki. Fragment of impacting asteroid recovered in Botswana. ScienceDaily. 2018 July 06. https://www.sciencedaily.com/releases/2018/07/180706091720.htm

6. Edwards WN, Brown PG, Weryk RJ, ReVelle DO. Infrasonic observations of meteoroids: Preliminary results from a coordinated optical-radar-infrasound observing campaign. Earth Moon Plan. 2008;102:221–229. https://doi.org/10.1007/s11038-007-9154-6

7. Roelofse F, Saunders I. A first report on meteor-generated seismic signals as detected by the SANSN. S Afr J Sci. 2013;109(5/6), Art. #2012-0022. https://doi.org/10.1590/sajs.2013/20120022

8. Wesel R. Meteorite fall from Madagascar – 11,241 gram Main Mass – Benenitra Museum Piece. Ebay.com [website on the Internet]. c2020 [cited 2020 Feb 24]. Available from: https://www.ebay.com/itm/Meteorite-fall-from-Madagascar-11-241-gram-Main-Mass-Benenitra-Museum-Piece/392590554453

9. Wlotzka F. A weathering scale for the ordinary chondrites (abstract). Meteoritics. 1993;28:460.6. https://doi.org/10.1111/j.1945-5100.1993.tb00268.x

10. Benenitra. In: Meteoritical Bulletin Database [database on the Internet]. c2020 [cited 2020 Jul 16]. Available from: www.lpi.usra.edu/meteor/metbull.php?code=69345

11. Dodd RT, Grover JE, Brown GE. Pyroxenes in the Shaw (L-7) chondrite. Geochim Cosmochim Acta. 1975;39:1585–1586. https://doi.org/10.1016/0016-7037(75)90081-2

12. Wood JA. Chondrites: Their metallic minerals, thermal histories, and parent planets. Icarus. 1967;6:1–49. https://doi.org/10.1016/0019-1035(67)90002-4

13. Van Schmus WR, Wood JA. A chemical-petrologic classification for the chondritic meteorites. Geochim Cosmochim Acta; 1967;31:747–765. https://doi.org/10.1016/S0016-7037(67)80030-9

14. Brearley AJ, Jones RH. Chondritic meteorites. In: Papike JJ, editor. Planetary materials. Washington DC: Mineral Society of America. 1998. p. 313–398.

15. Huss GR, Rubin AE, Grossman JN. Thermal metamorphism in chondrites. In: Meteorites and the early solar system II. Tucson, AZ: University of Arizona Press; 2006. p. 567–586.

16. Mittlefehldt DW, Lindstrom, MM. Petrology and geochemistry of Patuxent Range 91501, a clast-poor impact-melt from the L chondrite parent body, and Lewis Cliff 88663, an L7 chondrite. Meteor Planet Sci. 2001;36:439–457. https://doi.org/10.1111/j.1945-5100.2001.tb01885.x

17. Weisberg MK, McCoy TJ, Krot AN. Systematics and evaluation of meteorite classification. In: Lauretta DS, McSween HR Jr, editors. Meteorites and the early solar system II. Tucson, AZ: University of Arizona Press; 2006. p. 19–52.

18. Stöffler D, Kei K, Scott ERD. Shock metamorphism of ordinary chondrites. Geochim Cosmochim Acta. 1991;55:3845–3867. https://doi.org/10.1016/0016-7037(91)90078-J

19. Gi N, Brown P. Refinement of bolide characteristics from infrasound measurements. Planet Space Sci. 2017;143:169–181. https://doi.org/10.1016/j.pss.2017.04.021

20. Marcus R, Melosh HJ, Collins G. Earth Impact Effects Program [webpage on the Internet]. c2010 [cited 2020 Feb 05]. Available from: https://impact.ese.ic.ac.uk/ImpactEarth/ImpactEffects/

21. Whitaker RW, Brown GP, ReVelle DO, Sandoval TD, Mutschlecner JP, Bueck NM. Bolides and other infrasound events. 23rd Seismic Research Review: Worldwide Monitoring of Nuclear Explosions. 2001:168–176.

22. Bormann P. New manual of seismological observatory practice (NMSOP-2). Potsdam: IASPEI, GFZ German Research Centre for Geosciences; 2012. https://doi.org/10.2312/GFZ.NMSOP-2

23. Maromandia. Meteoritical Bulletin Database [database on the Internet]. c2004 [cited 2020 Feb 24]. Available from: www.lpi.usra.edu/meteor/metbull.php?code=15430

24. Meteoritical Bulletin Database [database on the Internet]. c2020 [cited 2020 Feb 24]. Available from: www.lpi.usra.edu/meteor/metbull.php?

25. Genge M, Grady M. The fusion crusts of stony meteorites: Implications for the atmospheric reprocessing of extraterrestrial materials. Meteor Planet Sci. 1999;34:341–356. https://doi.org/10.1111/j.1945-5100.1999.tb01344.x

26. Low FJ, Beintema DA, Gautier TN, Gillett FC, Beichman CA, Neugebauer G, et al. Infrared cirrus – New components of the extended infrared emission. Astrophys J Lett. 1984;278:L19–L22. https://doi.org/10.1086/184213

27. Wesel R. History of a first fall of meteorites with eyewitnesses in Madagascar. Nakhla Dog Meteorites [website on the Internet]. c2020 [cited 2020 Feb 24]. Available from: http://www.nakhladogmeteorites.com/catalog/mada.htm

28. Brown P, ReVelle DO, Silber EA, Edwards WN, Arrowsmith S, Jackson LE Jr, et al. Analysis of a crater-forming meteorite impact in Peru. J Geophys Res. 2008;113, E09007. https://doi.org/10.1029/2008JE003105

29. Arrowsmith SJ, ReVelle D, Edwards W, Brown P. Global detection of infrasonic signals from three large bolides. Earth Moon Planet. 2008;102:357–363. https://doi.org/10.1007/s11038-007-9205-z

30. Edwards WN, Eaton DW, Brown PG. Seismic observations of meteors: Coupling theory and observations. Rev Geophys. 2008;46:RG4007. https://doi:10.1029/2007RG000253

AUTHORS:
Ian Towle[1]
Joel D. Irish[2,3]
Isabelle De Groote[4]
Christianne Fernée[5,6]
Carolina Loch[1]

AFFILIATIONS:
[1]Sir John Walsh Research Institute, Faculty of Dentistry, University of Otago, Dunedin, New Zealand
[2]Research Centre in Evolutionary Anthropology and Palaeoecology, Liverpool John Moores University, Liverpool, United Kingdom
[3]Evolutionary Studies Institute and Centre for Excellence in PaleoSciences, University of the Witwatersrand, Johannesburg, South Africa
[4]Department of Archaeology, Ghent University, Ghent, Belgium
[5]Department of Anthropology and Archaeology, University of Bristol, Bristol, United Kingdom
[6]Department of Archaeology, University of Southampton, Southampton, United Kingdom

CORRESPONDENCE TO:
Ian Towle

EMAIL:
ian.towle@otago.ac.nz

DATES:
Received: 28 July 2020
Revised: 01 Oct. 2020
Accepted: 08 Oct. 2020
Published: 29 Mar. 2021

HOW TO CITE:
Towle I, Irish JD, De Groote I, Fernée C, Loch C. Dental caries in South African fossil hominins. S Afr J Sci. 2021;117(3/4), Art. #8705. https://doi.org/10.17159/sajs.2021/8705

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Margaret Avery

KEYWORDS:
dental pathology, *Homo naledi*, *Paranthropus robustus*, diet, cariogenic bacteria

FUNDING:
Liverpool John Moores University, Sir Thomas Kay Sidey Postdoctoral Fellowship (University of Otago)

Dental caries in South African fossil hominins

Once considered rare in fossil hominins, caries has recently been reported in several hominin species, requiring a new assessment of this condition during human evolution. Caries prevalence and location on the teeth of South African fossil hominins were observed and compared with published data from other hominin samples. Teeth were viewed macroscopically, with lesion position and severity noted and described. For all South African fossil hominin specimens studied to date, a total of 10 carious teeth (14 lesions), including 4 described for the first time here, have been observed. These carious teeth were found in a minimum of seven individuals, including five *Paranthropus robustus*, one early *Homo*, and one *Homo naledi*. All 14 lesions affected posterior teeth. The results suggest cariogenic biofilms and foods may have been present in the oral environment of a wide variety of hominins. Caries prevalence in studied fossil hominins is similar to those in pre-agricultural human groups, in which 1–5% of teeth are typically affected.

Significance:

- This study adds to the growing evidence that dental caries was present throughout the course of human evolution. Caries prevalence in the fossil species studied is similar to those in non-agricultural human groups, with 1–5% of teeth displaying cavities.
- Differences in prevalence and position of dental caries can provide insight into dietary aspects of past populations. South African fossil hominins display lesions on their posterior teeth, suggesting sugary foods were consumed.

Introduction

A range of different intraoral bacteria may have been involved in caries formation in the past, including one of the most common forms today, *Streptococcus mutans*. *S. mutans* is generally associated with agriculture and population growth. Postindustrial populations, in particular, evidence a marked increase in cariogenic bacteria, as well as less diverse microbiotic ecosystems.[1,2] Recent studies suggest the entire microbiota has to be considered in caries research, with a variety of species acting collectively in creating carious lesions.[3] Types and proportions of bacteria today differ substantially between active lesions in different locations (e.g. white spot enamel caries vs. deep dentine lesions), as well as between individuals.[4] When acids from these bacterial consortiums cannot be neutralised by saliva, lesions form.[5] Some foods are especially cariogenic, such as those with high levels of refined carbohydrates and sugars[6-8], as well as fruits, honey, and some nuts and seeds.[9,10] Tough and fibrous foods are linked with low caries prevalence, as they create a more alkaline oral environment related to substantial saliva production.[7,8,11] Diets rich in meat and certain plants have also been associated with low caries prevalence.[10,11]

Environmental and genetic influences are also important to consider.[12,13] It is not yet clear how genetic differences between populations influence caries prevalence[13], but interspecific differences in dental morphology will affect lesion position and severity, as is evident when comparing distantly related groups[14]. Nevertheless, the tooth types, tissues, and tooth areas affected by caries can give further insight into diet and food processing behaviours of a population.[10,15,16] Caries lesions have been widely reported in archaeological and fossil samples, with prevalence varying substantially in agricultural groups.[10,17-19] Extant nonhuman primates are also reported to have caries, with captive individuals typically more affected than their wild counterparts.[20] It is often suggested that caries was scarce in pre-agricultural hominins.[21,22] Yet, evidence for such lesions in pre-agricultural groups is growing.[23-31]

In light of this evidence, we reanalysed South African fossil hominin material and *Homo naledi* recorded for the first time for the presence of caries. Comparisons of position and prevalence were then made with hominin samples from the literature. Ten carious lesions have already been recorded in South African hominins, including two in a mandible from early *Homo*, SK 15.[32] The others are attributed to *Paranthropus robustus*, with three lesions on SKX 5023[23], two on SK 55[33], two on SK 13/14[33], and one on DNH 40 from Drimolen[31]. The aim of this study was to compare the prevalence and characteristics of caries in South African fossil hominins. We hypothesised that there would be differences in caries prevalence and position in the dentition among hominid groups, given substantial dietary differences between great apes, including hominins.[34,35]

Materials and methods

The South African material that was analysed included specimens assigned to early *Homo*, *Homo naledi*, *Australopithecus sediba*, *Australopithecus africanus* and *Paranthropus robustus*, all of which are curated at the University of the Witwatersrand and Ditsong National Museum of Natural History. Only complete teeth were included in the analysis, with each tooth examined under a 10× hand lens for initial lesion classification. A carious lesion was recorded as present when clear enamel cavitation was evident; colour changes alone were not considered diagnostic. Severity was also recorded on a scale of 1 to 4, following Connell and Rauxloh[36]: (1) enamel destruction only; (2) dentine involvement but pulp chamber not exposed; (3) dentine destruction with pulp chamber exposed; and (4) gross destruction with the crown mostly affected. Lastly, location was scored as distal, buccal, occlusal, lingual, mesial, root, or a combination of these.

Tooth wear scores were recorded to provide some insight into the interaction with caries, as well as potential information on the role of individual age on lesion formation. Wear was recorded following Smith[37] for incisors, canines and premolars, and following Scott[38] for molars, as this method allows for more precise inferences on intracrown variability. Molars were split into four quadrants, with each given a wear score of 1 to 10; all other teeth were graded using a single value of 1 to 8. An average for the four molar quadrants was calculated. Like in Meinl et al.[16] and Larsen[39], no corrective methods were implemented and rates were reported as the per cent of carious teeth among the total number of teeth examined. Comparisons are made between teeth rather than individuals to allow insight into variation across the dentition, but also to maximise data, with fossil samples often represented by loose teeth only. This approach comes with a caveat, as an individual with one carious lesion is more likely to display multiple lesions due to a common oral environment. Thus, a well-preserved individual with multiple lesions would increase caries prevalence, although this potential limitation does not significantly affect the present results.

Micro-computed tomography (CT) scans of specific teeth were taken to clarify if a lesion had been carious. In a clinical setting, scans can differentiate between normal enamel and dentine; areas affected by caries display lower density, which is often identifiable.[40] Such techniques can aid in visualisation of a lesion's extent, even if the cavity on the surface is ambiguous.[41] Scans were provided by the Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology. They were obtained with a BIR ACTIS 225/300 (settings 100 kV, 100 mA, 0.25 brass filter) or a SkyScan 1172 (settings 100 kV, 94 mA, 2.0 mm aluminum and copper filter) micro-CT scanner. The isometric voxel sizes from these scans range between 15 μm and 50 μm (Skinner M 2018, personal communicaton).

Results

We added four new cases to the six carious teeth already described in the literature for South African hominins (Table 1). Of these four, two are from *P. robustus* and two are from one *H. naledi* individual. In total, a minimum of 14 lesions have been diagnosed in all South African material described (Table 1). These include five *P. robustus*, one *H. naledi*, and one early *Homo* individual. None was found in the deciduous teeth, or any tooth from *A. africanus* and *A. sediba*. The 14 lesions were observed on 10 teeth from seven individuals. Figures 1–3 highlight lesions previously reported, with more in-depth descriptions of the newly described cases below.

Caries on *H. naledi* specimen UW 101-001 appear the most severe of all yet described (Figure 4). They seem to have been active for an extended period given penetration deep into the dentine. No difference in crown wear in the antimeres suggests the lesions may not have affected normal mastication. Lesions formed interproximally between the right P4 and adjacent M1, so likely originated from the same microenvironment. Both likely started on the interproximal surface, and enlarged to affect the M1 occlusal surface, with both crown and root surfaces affected. Due to postmortem sediment present on the surface, it is unclear how deep into the dentine these lesions penetrated, or whether they reached the pulp chamber. A micro-CT scan of this specimen was not available at the time of study. Substantial wear and antemortem enamel fractures are also associated with the posterior teeth of this individual.

At least two teeth from *P. robustus* mandible SK 23 had occlusal caries. The left first molar and right second premolar exhibited larger and darker fissures on their occlusal surface than those present in their antimeres (Figure 5a). However, due to the presence of postmortem matrix in these depressions (Figure 5b), it was difficult to confirm the lesion extent through macroscopic observation. Micro-CT scans could not be used to assess caries presence, as taphonomic processes make the resolution of density differences in the enamel poor. However, enamel under the depression on the premolar occlusal surface appeared to have patches of less dense material, potentially supporting a diagnosis of occlusal caries (Figure 5d, e). There was also a low-density line extending down from the lesion, which was likely a postmortem artefact based on similar features elsewhere in the tooth. The difference in macroscopic appearance compared with its antimere (Figure 5a), large size of the depressions (Figure 5b), and potential demineralisation of enamel (Figure 5d,e), suggest the presence of a carious lesion on the occlusal fissures. These same features were associated with the lesion in the left M1 of SK23.

Caries prevalence for each sample is listed in Tables 2–3. *Homo naledi* and *P. robustus* have similar prevalence (1.36% and 1.75% of all permanent teeth analysed, respectively). Early *Homo* had the highest prevalence of all (4.55%). *Australopithecus sediba* and *A. africanus* had no detectable caries lesions, and the former was not included in the Tables due to small sample size. Four of seven carious teeth displayed more occlusal wear than the average for that tooth type (Table 4). Due to small sample size, comparisons for early *Homo* wear (SK 15) were not possible. However, most carious teeth were near the species wear mean scores (Table 4) and had much of the crown remaining.

Table 1: South African fossil hominin specimens with caries

Species	Specimen	Tooth	Position	Severity	Lesion #	Reference/source
Paranthropus robustus	SK 23	L M_1	Occlusal	1	1	This study
P. robustus	SK 23	R P_4	Occlusal	1	1	This study
P. robustus	SKX 5023	R M_1	Mesial/distal	2	3	Grine et al.[23]
P. robustus	SK 55	L M^1	Buccal	1	2	Robinson[33]
P. robustus	SK 13/14	L M^2	Occlusal	2	2	Robinson[33]
P. robustus	DNH 40	L M^3	Root/mesial	2	1	Towle et al.[31]
Early *Homo*	SK 15	L M_1	Mesial	2	1	Clement[32]
Early *Homo*	SK 15	R M_2	Mesial	2	1	Clement[32]
Homo naledi	UW 101-001	R P_4	Distal	2/3	1	This study
H. naledi	UW 101-001	R M_1	Mesial	2/3	1	This study

L, left; R, right

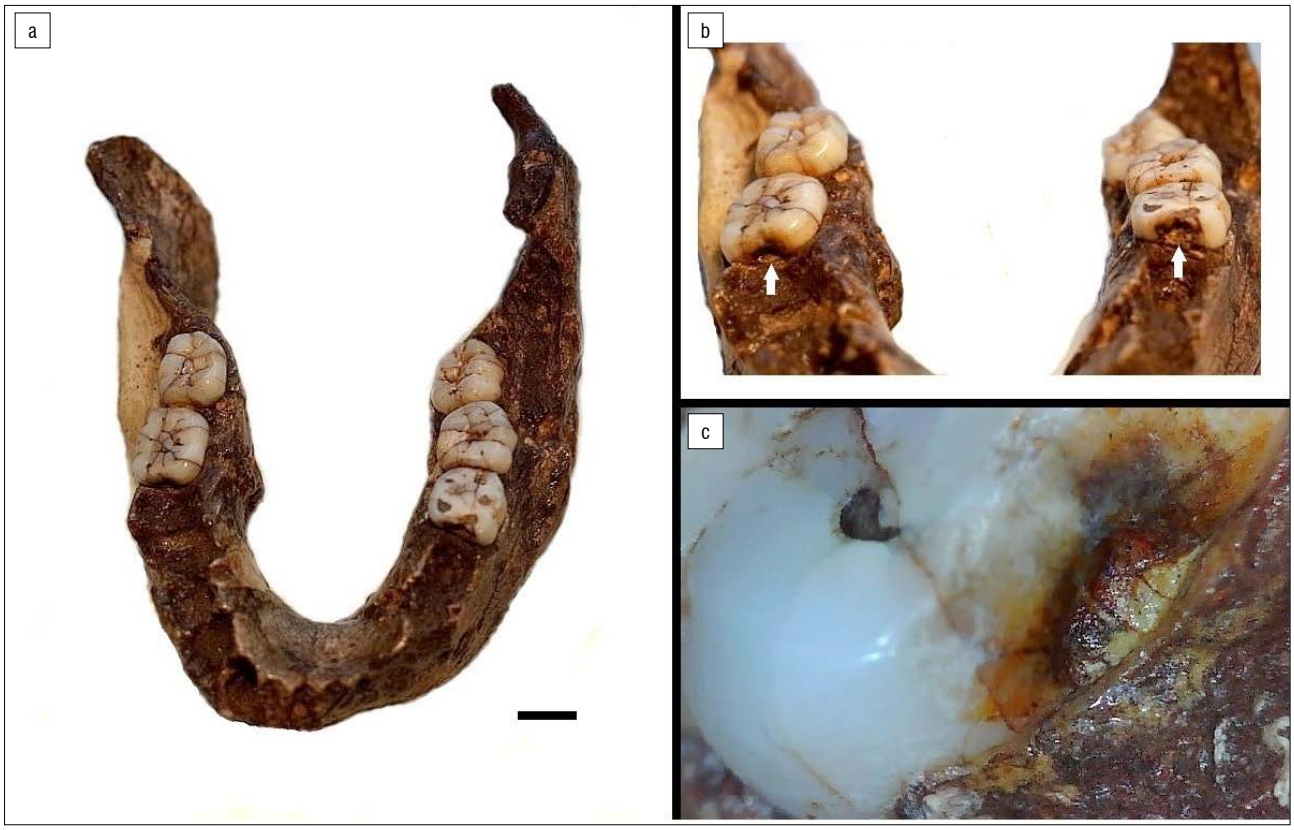


Figure 1: SK 15 (early *Homo*) interproximal caries on the lower right second molar and left first molar. (a) Overview of specimen; scale bar = 1 cm. (b) Mesial carious lesions (white arrows) and (c) close-up of the right second molar with carious lesion on the mesial surface.

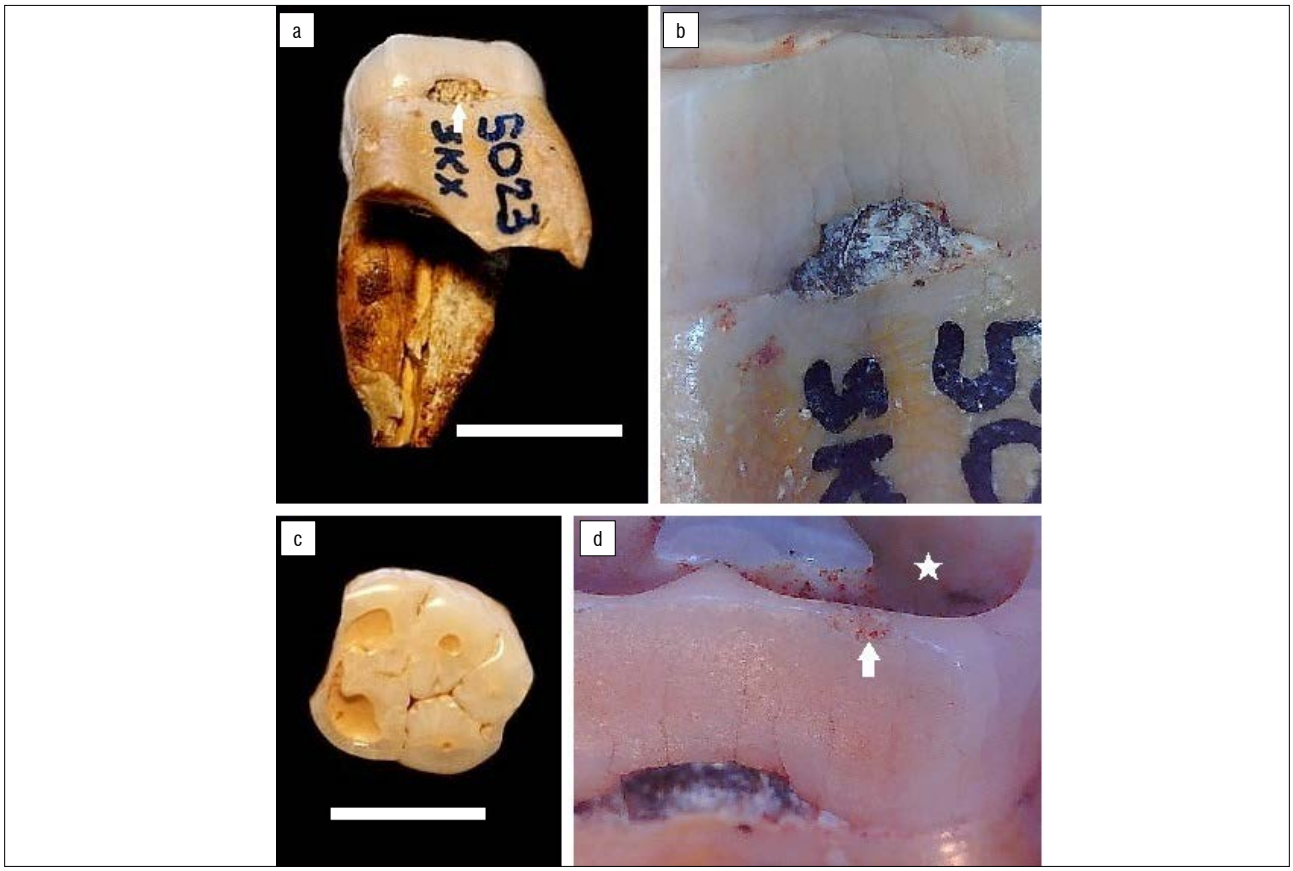


Figure 2: SKX 5023 (*Paranthropus robustus*) lower right first molar. (a) Overview of specimen with carious lesion on the mesial surface (white arrow). (b) Close-up of mesial lesion. (c) Occlusal view of the specimen. (d) Occlusal/mesial view of specimen showing the carious lesion, antemortem chip (white arrow) and cupping dentine wear (white star). Both scale bars = 1 cm.

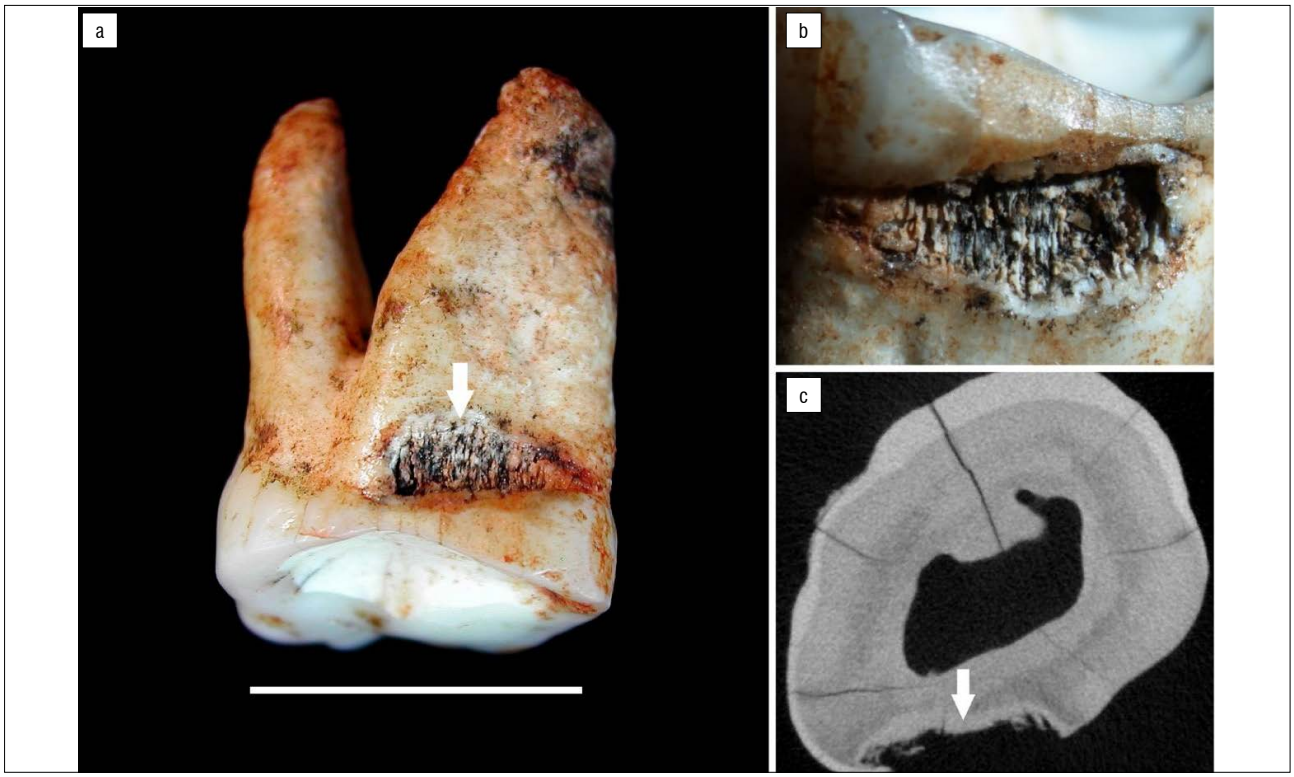


Figure 3: DNH 40 (*Paranthropus robustus*) upper left third molar. (a) Overview of the tooth, showing mesial and occlusal surfaces (carious lesion indicated by white arrow); scale bar = 1 cm. (b) Close-up of mesial lesion. (c) CT slice of the specimen; white arrow indicates the carious lesion.

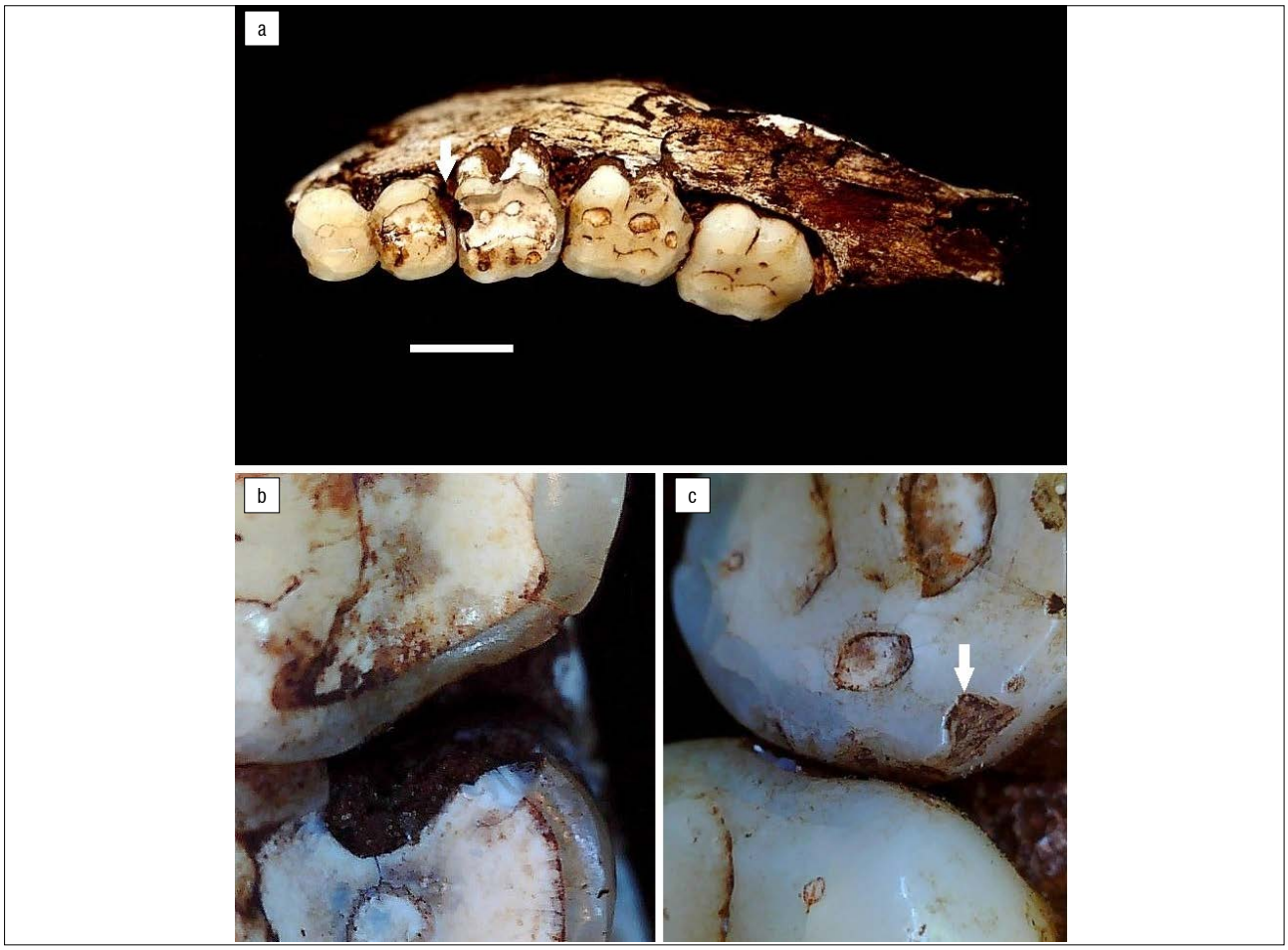


Figure 4: UW 101-001 (*Homo naledi*) carious lesions on the lower right second premolar (distal) and first molar (mesial). (a) Overview of specimen; white arrow shows location of the two interproximal carious lesions; scale bar = 1 cm; (b) Close-up of lesions. (c) Right second and third molars, with two antemortem chips on the mesial buccal corner (white arrow).

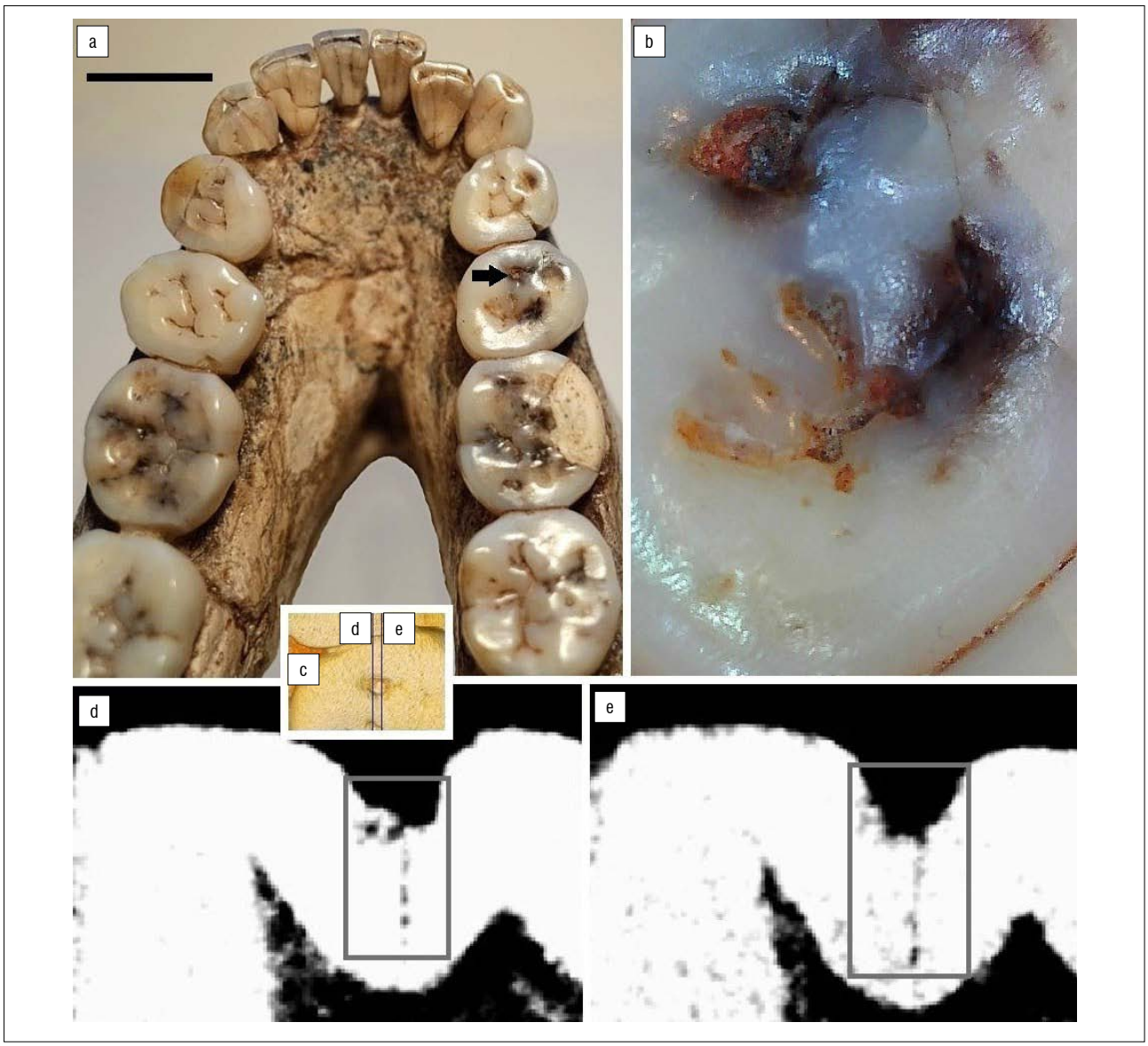


Figure 5: SK 23 (*Paranthropus robustus*). (a) Occlusal view of mandible, with the lesion on the right second premolar highlighted (black arrow). (b) Close-up of the occlusal surface of the right second premolar. (c) CT reconstruction with the position of the two slices highlighted. (d) CT slice toward the lingual part of the cavity. (e) CT slice toward the buccal portion of the cavity.

Table 2: Caries prevalence for permanent teeth for each species studied

Species	Anterior teeth			Posterior teeth			All teeth		
	# Teeth	Carious teeth	%	# Teeth	Carious teeth	%	# Teeth	Carious teeth	%
*Paranthropus robustus**	59	0	0	226	5	2.2	285	5	1.8
Australopithecus africanus	86	0	0	243	0	0	329	0	0
Homo naledi	50	0	0	97	2	2.1	147	2	1.4
Early *Homo*	15	0	0	29	2	6.9	44	2	4.6

**does not include specimens from Drimolen*

Table 3: Caries prevalence for deciduous teeth for each species studied

Species	# Teeth	Carious teeth	%
Paranthropus robustus	45	0	0
Early *Homo*	13	0	0
Australopithecus africanus	18	0	0
Homo naledi	20	0	0

Table 4: Wear stage for each carious tooth, based on Smith[37] for premolars, and Scott[38] for molars. Average wear is the average species wear for that tooth type.

Species	Specimen	Tooth	Wear stage	Average wear
Paranthropus robustus	SK 23	L M_1	4.3	4
P. robustus	SK 23	R P_4	3	3
P. robustus	SKX 5023	R M_1	5.3	4
P. robustus	SK 55	L M^1	3.3	4
P. robustus	SK 13/14	L M^2	3.5	3.6
Early *Homo*	SK 15	L M_1	5.8	–
Early *Homo*	SK 15	R M_2	4.3	–
Homo naledi	UW 101-001	R P_4	4	3.4
H. naledi	UW 101-001	R M_1	6.3	3.9

L, left; R, right

Discussion

The results suggest caries may have been more common in past populations than originally thought. In particular, caries is relatively common in several of the South African fossil hominin species, suggesting this pathology may be a common feature of hominins. In addition to macroscopic observations, X-ray imaging, exploration with dental probes and colouration changes are also commonly used to confirm caries lesions in the clinical setting. In archaeological and palaeoanthropological studies, these approaches are often difficult due to postmortem changes (e.g. tooth colouration or structure), or unavailability of diagnostic equipment. Additionally, lesions within interproximal areas may not be observable if teeth remain in situ within the mandible or maxilla. Only a small sample of teeth ($n=5$) were studied via micro-CT here. A comparative micro-CT study is needed to identify lesions with more accuracy, particularly to account for intersample differences in morphology and taphonomy. Preservation and sample bias are also important considerations, with diagnosis without clear large cavitation not possible in many fossil samples due to taphonomic changes to dental tissues. Therefore, caries prevalence in fossil hominins should be viewed with caution; nevertheless, presence and position can provide useful information concerning diet and behaviour in individuals and groups.

Given the results here and elsewhere[23-31], cariogenic bacteria seemed to be prevalent in a number of hominin species. This scenario is supported by research suggesting that a variety of bacteria contribute to a cariogenic microbiota; these same species are often part of the normal oral microbiome in non-pathological situations.[3,4] As noted by Sheiham and James[42], diet is the primary factor in caries formation and if enough cariogenic foods were consumed regularly by past groups, they would likely develop carious lesions. Differences in microbiota among hominin groups may have influenced caries prevalence, with the threshold required to create a lesion (e.g. regularity of sugar intake) likely varying depending on the prevailing oral microbiota.[22,43]

High prevalence of occlusal caries is associated with low attrition, while interproximal lesions are linked with high attrition.[44] However, interproximal caries may also be related to calculus, as a result of plaque deposition in these areas[45], although this link has not been extensively researched. There is no evidence for calculus-associated caries in South African hominins, although postmortem loss of calculus may have occurred. Enamel hypoplasia may also influence caries formation, acting as a site for lesion development because weakened enamel is more vulnerable to acid solubility.[8,46] Enamel hypoplasia likely influenced caries formation in *P. robustus* SK 55 and SK 13/14, with substantial hypoplastic pitting present; pitting is a common feature of this species.[47]

Other pathologies and wear can also create an environment in which caries is more likely to form, e.g. in response to unusual occlusion patterns.[14] It is noteworthy that several fossil teeth with interproximal caries also have substantial cupped occlusal wear and enamel chipping, either above the lesion or on interproximal surfaces of adjacent teeth (SKX 5023, DNH 40 and UW 101-001). Occlusal forces may therefore have been contributing factors, either directly through weakening dental tissues or indirectly through creating an interproximal gap where cariogenic bacteria and foods could accumulate.[48]

The mastication of hard items (e.g. grit) over time can lead to fatigue of the enamel[49,50]; as such, certain regions of the crown can become susceptible to pathology. In hominins, the interproximal regions of molars seem susceptible to different types of wear and pathologies, including chipping and non-carious cervical lesions (i.e. 'toothpick' or interproximal grooves). Therefore, higher interproximal stresses may lead to tissue fatigue, which could ultimately mean these regions are more prone to pathology and wear. In humans, caries tends to be more common in posterior teeth, with occlusal surfaces commonly affected.[10,17-19,51,52] This pattern is observed in the present hominin sample, with posterior teeth affected more commonly. However, in contrast to recent humans, most lesions were not on occlusal surfaces; this difference likely reflects differences in occlusal wear, in that surface fissures and pits were worn away before lesions could form.

Many human samples from the last 50 000 years show caries frequencies similar to those observed in this study.[15,27,39] With the development of agriculture, this frequency varied, and some populations had substantially greater caries rates.[10,17-19] It seems that a low to moderate occurrence of caries was common in these hominin populations, typically affecting 1–5% of teeth – as previously suggested for hunter-gatherer groups.[53] Lower or higher caries frequencies are usually associated with specific diets or behaviours. For example, high caries prevalence (i.e. >5% of teeth) is associated with agricultural diets[10,18,19,51,52], and hunter-gatherer groups reliant on cariogenic foods[9,54-56]. Low prevalence of caries (i.e. <1% of teeth) is often associated with a diet high in marine foods/terrestrial meat.[15,27,39]

It is unlikely that the oral microbiome composition was solely responsible for the lack of carious lesions in *A. africanus*. Instead, it may relate to dietary differences, or due to sampling bias. The large lesions on *P. robustus*, *H. naledi* and early *Homo* teeth suggest cariogenic foods were consumed. Most fossils assigned to the genus *Homo* also exhibit caries, including Neanderthals, early *Homo*, *H. erectus*, and Pleistocene age *H. sapiens*[25-30,57,58] Thus, caries was likely common throughout the evolutionary history of the genus, and may relate to the consumption of cariogenic foods such as tubers, nuts, plants or fruit.[9,54-56]

A recent study of caries in the African sub-continent through time found that hunter-gatherer groups had a caries prevalence similar to that of the South African fossil hominins analysed here.[59] This includes groups relevant to the present study because of geographic proximity, i.e. prehistoric and historic Khoesan peoples. The diet of these hunter-gatherers would have been diverse, but tubers rich in starch have been suggested as one possible contributor to caries occurrence.[44] Similarly, honey is consumed in many recent hunter-gatherer groups, thus it may have been present in the diet of early *Homo* and *Paranthropus*.[11] Cooking has also been proposed as a factor influencing caries prevalence[25], but there is currently insufficient evidence to conclude if the hominins in this study mastered the use of fire for cooking food. Further ecological/environmental comparisons may allow insight into the dietary items and food processing techniques facilitating the development of carious lesions. Caries has been present throughout human evolution, with fossil hominins from South Africa affected by this disease. Caries rates observed here are similar to those of pre-agricultural human groups, in which 1–5% of teeth are typically affected.

Acknowledgements

We thank L. Berger and B. Zipfel from the University of the Witwatersrand, and S. Potze from the Ditsong Museum of South Africa for access to their collections. CT scans were provided by the Ditsong Museum, and we thank M. Skinner for kindly providing technical assistance. Research was supported by a studentship to I.T. from Liverpool John Moores University, and a Sir Thomas Kay Sidey Postdoctoral fellowship from the University of Otago.

Competing interests

We declare that there are no competing interests.

Authors' contributions

I.T.: Formulation or evolution of overarching research goals and aims; application of statistical, mathematical, or computational techniques to analyse or synthesise study data. I.T., J.D.I.: Development or design of methodology. I.T., J.D.I., C.F., I.D.G., C.L.: Conducting a research and investigation process; preparation and creation of the published work, specifically writing the initial draft (including substantive translation). I.T., C.F.: Morphological or laboratory assessment of biological samples. J.D.I., I.D.G.: Supervision of student and oversight of study. J.D.I., C.L.: Management and coordination responsibility for the research activity planning and execution. C.F., C.L.: Development or design of methodology.

References

1. Adler CJ, Dobney K, Weyrich LS, Kaidonis J, Walker AW, Haak W, et al. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nat Genet. 2013;45(4):450–455. https://doi.org/10.1038/ng.2536
2. Cornejo OE, Lefébure T, Pavinski Bitar PD, Lang P, Richards VP, Eilertson K, et al. Evolutionary and population genomics of the cavity causing bacteria *Streptococcus mutans*. Mol Biol Evol. 2013;30(4):881–893. https://doi.org/10.1093/molbev/mss278
3. Simón-Soro A, Mira A. Solving the etiology of dental caries. Trends Microbiol. 2015;23(2):76–82. https://doi.org/10.1016/j.tim.2014.10.010
4. Aas JA, Griffen AL, Dardis SR, Lee AM, Olsen I, Dewhirst FE, et al. Bacteria of dental caries in primary and permanent teeth in children and young adults. J Clin Microbiol. 2008;46(4):1407–1417. https://doi.org/10.1128/JCM.01410-07
5. Gussy MG, Waters EG, Walsh O, Kilpatrick NM. Early childhood caries: Current evidence for aetiology and prevention. J Paediatr Child Health. 2006;42(1–2):37–43. https://doi.org/10.1111/j.1440-1754.2006.00777.x
6. Clarkson BH, Krell D, Wefel JS, Crall J, Feagin FF. In vitro caries-like lesion production by *Streptococcus mutans* and *Actinomyces viscosus* using sucrose and starch. J Dent Res. 1987;66(3):795–798. https://doi.org/10.1177/00220345870660031801
7. Prowse TL, Saunders SR, Schwarcz HP, Garnsey P, Macchiarelli R, Bondioli L. Isotopic and dental evidence for infant and young child feeding practices in an imperial Roman skeletal sample. Am J Phys Anthropol. 2008;137(3):294–308. https://doi.org/10.1002/ajpa.20870
8. Rohnbogner A, Lewis M. Dental caries as a measure of diet, health, and difference in non-adults from urban and rural Roman Britain. Dent Anthropol J. 2016;29(1):16–31. https://doi.org/10.26575/daj.v29i1.32
9. Humphrey LT, De Groote I, Morales J, Barton N, Collcutt S, Ramsey CB, et al. Earliest evidence for caries and exploitation of starchy plant foods in Pleistocene hunter-gatherers from Morocco. Proc Natl Acad Sci USA. 2014;111(3):954–959. https://doi.org/10.1073/pnas.1318176111
10. Novak M. Dental health and diet in early medieval Ireland. Arch Oral Biol. 2015;60(9):1299–1309. https://doi.org/10.1016/j.archoralbio.2015.06.004
11. Moynihan P. Foods and factors that protect against dental caries. Nutr Bull. 2000;25(4):281–286. https://doi.org/10.1046/j.1467-3010.2000.00033.x
12. Slade GD, Sanders AE, Do L, Roberts-Thomson K, Spencer AJ. Effects of fluoridated drinking water on dental caries in Australian adults. J Dent Res. 2013;92(4):376–382. https://doi.org/10.1177/0022034513481190
13. Haworth S, Shungin D, Van Der Tas JT, Vucic S, Medina-Gomez C, Yakimov V, et al. Consortium-based genome-wide meta-analysis for childhood dental caries traits. Hum Mol Genet. 2018;27(17):3113–3127. https://doi.org/10.1093/hmg/ddy237
14. Calcagno JM, Gibson KR. Selective compromise: Evolutionary trends and mechanisms in hominid tooth size. In: Kelley MA, Larsen CS, editors. Advances in dental anthropology. New York: Wiley-Liss; 1991. p. 59–76.
15. Kelley MA. Contrasting patterns of dental disease in five early northern Chilean groups. In: Kelley MA, Larsen CS, editors. Advances in dental anthropology. New York: Wiley-Liss; 1991. p. 59–76.
16. Meinl A, Rottensteiner GM, Huber CD, Tangl S, Watzak G, Watzek G. Caries frequency and distribution in an early medieval Avar population from Austria. Oral Dis. 2010;16(1):108–116. https://doi.org/10.1111/j.1601-0825.2009.01624.x
17. Watt ME, Lunt DA, Gilmour WH. Caries prevalence in the permanent dentition of a mediaeval population from the south-west of Scotland. Arch Oral Biol. 1997;42(9):601–620. https://doi.org/10.1016/s0003-9969(97)00061-7
18. Srejić MD. Dental paleopathology in a Serbian medieval population. Anthropol Anz. 2001;1:113–122. https://doi.org/10.1127/anthranz/59/2001/113
19. Šlaus M, Bedić Ž, Rajić Šikanjić P, Vodanović M, Domić Kunić A. Dental health at the transition from the Late Antique to the early Medieval period on Croatia's eastern Adriatic coast. Int J Osteoarchaeol. 2011;21(5):577–590. https://doi.org/10.1002/oa.1163
20. Crovella S, Ardito G. Frequencies of oral pathologies in a sample of 767 non-human primates. Primates. 1994;35(2):225–230. https://doi.org/10.1007/bf02382058
21. Tillier AM, Arensburg B, Rak Y, Vandermeersch B. Middle Palaeolithic dental caries: New evidence from Kebara (Mount Carmel, Israel). J Hum Evol. 1995;29(2):189–192. https://doi.org/10.1006/jhev.1995.1055
22. Adler CJ, Browne GV, Sukumar S, Hughes T. Evolution of the oral microbiome and dental caries. Curr Oral Health Rep. 2017;1;4(3):264–269. https://doi.org/10.1007/s40496-017-0151-1
23. Grine FE, Gwinnett AJ, Oaks JH. Early hominid dental pathology: Interproximal caries in 1.5 million-year-old *Paranthropus robustus* from Swartkrans. Arch Oral Biol. 1990;35(5):381–386. https://doi.org/10.1016/0003-9969(90)90185-d
24. Trinkaus E, Smith RJ, Lebel S. Dental caries in the Aubesier 5 Neandertal primary molar. J Archaeol Sci. 2000;27(11):1017–1021. https://doi.org/10.1006/jasc.1999.0512
25. Lanfranco LP, Eggers S. Caries through time: An anthropological overview. In: Li M-Y, editor. Contemporary approach to dental caries. Shangai: Intech Open; 2012. https://doi.org/10.5772/38059
26. Lacy SA, Wu XJ, Jin CZ, Qin DG, Cai YJ, Trinkaus E. Dentoalveolar paleopathology of the early modern humans from Zhirendong, South China. Int J Paleopathol. 2012;2(1):10–18. https://doi.org/10.1016/j.ijpp.2012.06.003
27. Lacy SA. Oral health and its implications in late Pleistocene Western Eurasian humans. St. Louis: Washington University; 2014.

28. Liu W, Martinón-Torres M, Cai YJ, Xing S, Tong HW, Pei SW, et al. The earliest unequivocally modern humans in southern China. Nature. 2015;526(7575):696–699. https://doi.org/10.1038/nature15696

29. Arnaud J, Benazzi S, Romandini M, Livraghi A, Panetta D, Salvadori PA, et al. Neanderthal deciduous human molar with incipient carious infection from the Middle Palaeolithic De Nadale cave, Italy. Am J Phys Anthropol. 2017;162(2):370–376. https://doi.org/10.1002/ajpa.23111

30. Margvelashvili A, Zollikofer CP, Lordkipanidze D, Tafforeau P, Ponce de León MS. Comparative analysis of dentognathic pathologies in the Dmanisi mandibles. Am J Phys Anthropol. 2016;160(2):229–253. https://doi.org/10.1002/ajpa.22966

31. Towle I, Riga A, Irish JD, Dori I, Menter C, Moggi-Cecchi J. Root caries on a *Paranthropus robustus* third molar from Drimolen. Am J Phys Anthropol. 2019;170(2):319–323. https://doi.org/10.1002/ajpa.23891

32. Clement AJ. Caries in the South African apeman: Some examples of undoubted pathological authenticity believed to be 800,000 years old. Brit Dent J. 1956;101:4–7.

33. Robinson JT. Some hominid features of the apeman dentition. J Dent Assoc Series Africa. 1952;7:102–107.

34. Kupczik K, Toro-Ibacache V, Macho GA. On the relationship between maxillary molar root shape and jaw kinematics in *Australopithecus africanus* and *Paranthropus robustus*. R Soc Open Sci. 2018;5(8):180825. https://doi.org/10.1098/rsos.180825

35. Towle I, Irish JD, De Groote I. Behavioral inferences from the high levels of dental chipping in *Homo naledi*. Am J Phys Anthropol. 2017;164(1):184–192. https://doi.org/10.1002/ajpa.23250

36. Connell B, Rauxloh P. A rapid method for recording human skeletal data. London: Museum of London; 2003.

37. Smith BH. Patterns of molar wear in hunter–gatherers and agriculturalists. Am J Phys Anthropol. 1984;63(1):39–56. https://doi.org/10.1002/ajpa.1330630107

38. Scott EC. Dental wear scoring technique. Am J Phys Anthropol. 1979;51(2):213–217. https://doi.org/10.1002/ajpa.1330510208

39. Larsen CS. Dental caries evidence for dietary change: An archaeological context. Adv Dent Anthropol. 1991:179–202.

40. Swain MV, Xue J. State of the art of micro-CT applications in dental research. Int J Oral Sci. 2009;1(4):177–188. https://doi.org/10.4248/IJOS09031

41. Rossi M, Casali F, Romani D, Bondioli L, Macchiarelli R, Rook L. MicroCT scan in paleobiology: Application to the study of dental tissues. Nucl Instrum Methods Phys Res B. 2004;213:747–750. https://doi.org/10.1016/S0168-583X(03)01697-5

42. Sheiham A, James WP. Diet and dental caries: The pivotal role of free sugars reemphasized. J Dent Res. 2015;94(10):1341–1347. https://doi.org/10.1177/0022034515590377

43. Simon-Soro A, Tomás I, Cabrera-Rubio R, Catalan MD, Nyvad B, Mira A. Microbial geography of the oral cavity. J Dent Res. 2013;92(7):616–621. https://doi.org/10.1177/0022034513488119

44. Hillson S. The current state of dental decay. In: Irish JD, Nelson GC, editors. Technique and application in dental anthropology. Cambridge: Cambridge University Press; 2008. p. 53–111.

45. Tomczyk J, Szostek K, Komarnitki I, Mańkowska-Pliszka H, Zalewska M. Dental caries and chemical analyses in reconstruction of diet, health and hygienic behaviour in the Middle Euphrates valley (Syria). Arch Oral Biol. 2013;58(6):740–751. https://doi.org/10.1016/j.archoralbio.2012.12.014

46. Hong L, Levy SM, Warren JJ, Broffitt B. Association between enamel hypoplasia and dental caries in primary second molars: A cohort study. Caries Res. 2009;43(5):345–353. https://doi.org/10.1159/000231571

47. Towle I, Irish JD. A probable genetic origin for pitting enamel hypoplasia on the molars of *Paranthropus robustus*. J Hum Evol. 2019;129:54–61. https://doi.org/10.1016/j.jhevol.2019.01.002

48. Hinton RJ. Differences in interproximal and occlusal tooth wear among prehistoric Tennessee Indians: Implications for masticatory function. Am J Phys Anthropol. 1982;57(1):103–115. https://doi.org/10.1002/ajpa.1330570111

49. Gao SS, An BB, Yahyazadehfar M, Zhang D, Arola DD. Contact fatigue of human enamel: Experiments, mechanisms and modeling. J Mech Behav Biomed Mater. 2016;60:438–450. https://doi.org/10.1016/j.jmbbm.2016.02.030

50. Sanchez-Gonzalez E, Pinilla-Cienfuegos E, Borrero-Lopez O, Rodríguez-Rojas F, Guiberteau F. Contact damage of human dental enamel under cyclic axial loading with abrasive particles. J Mech Behav Biomed Mater. 2020;102:103512. https://doi.org/10.1016/j.jmbbm.2019.103512

51. Varrela TM. Prevalence and distribution of dental caries in a late medieval population in Finland. Arch Oral Biol. 1991;36(8):553–559. https://doi.org/10.1016/0003-9969(91)90104-3

52. Vodanović M, Hrvoje B, Mario Š, Željko D. The frequency and distribution of caries in the mediaeval population of Bijelo Brdo in Croatia (10th–11th century). Arch Oral Biol. 2005;50(7):669–680. https://doi.org/10.1016/j.archoralbio.2004.11.014

53. Turner CG. Dental anthropological indications of agriculture among the Jomon people of central Japan. X. Peopling of the Pacific. Am J Phys Anthropol. 1979;51(4):619–635. https://doi.org/10.1002/ajpa.1330510413

54. Walker PL, Hewlett BS. Dental health diet and social status among Central African foragers and farmers. Am Anthropol. 1990;92(2):383–398. https://doi.org/10.1525/aa.1990.92.2.02a00080

55. Sealy JC, Patrick MK, Morris AG, Alder D. Diet and dental caries among later stone age inhabitants of the Cape Province, South Africa. Am J Phys Anthropol. 1992;88(2):123–134. https://doi.org/10.1002/ajpa.1330880202

56. Nelson GC, Lukacs JR, Yule P. Dates, caries, and early tooth loss during the Iron Age of Oman. Am J Phys Anthropol. 1999;108(3):333–343. https://doi.org/10.1002/(sici)1096-8644(199903)108:3<333::aid-ajpa8>3.0.co;2-#

57. Brothwell DR. The macroscopic dental pathology of some earlier human populations. In: Dental anthropology. Pergamon Press; 1963. p. 271–288.

58. Clement AJ. The antiquity of caries. Brit Dent J. 1958;104:115–123.

59. Carter F, Irish JD. A sub-continent of caries: Prevalence and severity in early Holocene through recent Africans. Dent Anthropol. 2019;32(2):22–29. https://doi.org/10.26575/daj.v32i2.285

AUTHORS:
Charles W. Helm[1]
Martin G. Lockley[2]

AFFILIATIONS:
[1]African Centre for Coastal Palaeoscience, Nelson Mandela University, Port Elizabeth, South Africa
[2]Dinosaur Trackers Research Group, University of Colorado Denver, Denver, Colorado, USA

CORRESPONDENCE TO:
Charles Helm

EMAIL:
helm.c.w@gmail.com

DATES:
Received: 25 Aug. 2020
Revised: 22 Oct. 2020
Accepted: 30 Oct. 2020
Published: 29 Mar. 2021

HOW TO CITE:
Helm CW, Lockley MG. Pleistocene reptile swim traces confirmed from South Africa's Cape south coast. S Afr J Sci. 2021;117(3/4), Art. #8830. https://doi.org/10.17159/sajs.2021/8830

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Margaret Avery

KEYWORDS:
Pleistocene, reptiles, crocodylian swim traces, cemented foreshore deposits, trackmaker length

FUNDING:
None

Pleistocene reptile swim traces confirmed from South Africa's Cape south coast

Large Pleistocene reptile tracks and traces were described from the Cape south coast of South Africa in 2020, including 'probable swim traces'. These trace fossils were found on loose slabs and blocks of the Klein Brak Formation. Subsequently, another surface has become exposed on this coastline, also on a loose slab. It exhibits more definitive evidence of swim traces in epirelief, probably made by the Nile crocodile (*Crocodylus niloticus*) or water monitor (*Varanus niloticus*), although a chelonian origin cannot be excluded. Length of a possible crocodylian trackmaker was estimated from measurements of interdigital distance in the swim traces. These provide a compelling example of reptile swim traces from Africa.

Significance:

- Pleistocene reptile swim traces have now been confidently confirmed from the Cape south coast of South Africa.
- The findings complement the suite of recently identified large reptile tracks.
- Trackmaker size can be estimated from the dimensions of reptile swim traces.
- These are the first compelling non-dinosaurian reptile swim traces to be described from Africa.

The coastal portion of the Garden Route National Park, on South Africa's Cape south coast, has recently yielded large reptile trace fossils.[1] While evidence of tracks of the Nile crocodile (*Crocodylus niloticus*) and the water monitor (*Varanus niloticus*) was regarded as conclusive, the evidence for reptile swim traces was more equivocal. As a result, they were labelled 'probable swim traces'.[1] The tracks and probable swim traces occurred on the surfaces of loose blocks and slabs which lay near the high-tide mark, having become dislodged and fallen from overlying cliffs.[1]

To the best of our knowledge, unequivocal reptile swim traces had not previously been reported from the continent of Africa. The only potential example was from Niger: Mudroch et al.[2] reported novel theropod dinosaur tracks from a Jurassic site, and refuted the possibility of swim traces. However, Milner and Lockley[3] interpreted these features as dinosaur swim traces.

The cemented rock layers containing the reptile tracks and traces on the Cape south coast occur in the Klein Brak Formation of the Bredasdorp Group.[4] The Klein Brak Formation contains sediments representing a succession of shallow marine, beach, and estuarine or lagoonal deposits.[4] The track-bearing surfaces were interpreted as representing a back-barrier lagoon or interdune lake.[1] While optically stimulated luminescence dates were awaited for the fossil-bearing sediments, stratigraphic correlation to nearby dated deposits suggested a probable age of Marine Isotope Stage (MIS) 5e.[1,5] Large reptiles do not currently inhabit the Cape south coast; it was inferred that the reptile tracks and traces had been registered during a warm interglacial period during or close to the time of a sea high-stand, which allowed large reptile species to occupy an extended range.[1]

Reptile swim traces may be preserved as parallel or sub-parallel grooves on epirelief surfaces, or as natural cast ridges on hyporelief surfaces.[6-8] The spectrum of reptile swim traces is related to various factors, including current patterns, water depth, and trackmaker behaviour and degree of buoyancy: in very shallow sub-aqueous environments they may be virtually indistinguishable from tracks and traces made on dryland substrates, because the trackmaker can walk through shallow water without being buoyed up. In contrast, as water depth increases, causing buoyancy, typical walking tracks can be expected to become less evident, and swim traces to become more predominant. At a certain depth, only the tip of the longest digit will touch and scrape the bottom. If the water is deeper still, no trace will be registered, unless the trackmaker engaged in 'bottom-walking'. It is generally accepted that such swim tracks are often irregular.[3] Such irregular contact with the substrates has been referred to as 'punting'.[9] In deeper sub-aqueous settings, tail drag impressions are less common.[10]

We note that there are no reports in the global palaeo-ichnology record of tracks or swim traces of monitor lizards (Order: Squamata, Sub-order: Sauria, Family: Varanidae). However, fossil chelonian (Order: Chelonii) swim traces have been reported from multiple sites.[6,11] Although the typical morphotypes of crocodylian and chelonian swim traces may readily be distinguished[6], such traces are often variable in appearance, and it is not always possible to reliably make this distinction. However, some chelonians more often employ 'bottom-walking', and tracks are therefore more commonly encountered in addition to swim traces.[6] Grigg and Kirshner[12] noted three types of swimming in crocodylians: paraxial swimming, axial swimming and hybrid swimming, as well as bottom-walking. Axial swimming is used for rapid propulsion, and involves extensive sinuous movement of the muscular tail, while the limbs are held by the sides and are only used for steering. Swim traces of the kind recorded here are unlikely or impossible to result from such motion. Theoretically, true bottom-walking should result in a track pattern very similar to that of a trackway on dry land, and should not result in the kind of raking traces recorded here. (Proven examples of trackways registered while bottom-walking have not been documented, despite videos that show

this mode of locomotion.) Paraxial swimming and hybrid swimming are therefore the types of motion most likely to result in raking swim traces.[12]

Crocodylians are a major group of archosaurs with a fossil record that extends back to the Triassic.[7] Lockley et al.[7] proclaimed: 'Inasmuch as dinosaurs were archosaurs of the land, and pterosaurs archosaurs of the air, extant crocodylians are archosaurs of the water'. These authors also pointed out that crocodylian swim tracks were far more common than walking traces in the fossil record. Lockley[13] summarised the notion of crocodylian swim traces: 'They were inhabitants of countless prehistoric rivers, where their remains are common. Their tracks, however, are quite rare. Obviously, when in the water they did not produce many footprints. Although they would touch the bottom of river channels and ponds from time to time, tracks made when swimming are quite different from those made when walking, and often consist only of incomplete footprints and scrape marks, distributed with little sense of pattern or order. As a result, the interpretation of fossil crocodile tracks is controversial.' This conclusion has to some extent been superseded by multiple reports of unambiguous (non-controversial) crocodilian swim tracks, particularly in the Cretaceous of North America.[6] In contrast, with respect to the interpretation of monitor lizard swim traces, the palaeo-ichnological record is silent.

Subsequent to the description of these 'probable swim traces'[1], a further site has become exposed in the same area of the Garden Route National Park, also on a loose slab which had fallen down from the cliffs above and had come to rest near the high tide mark (Figure 1). It, too, contains tracks and traces, but in this case the evidence for swim traces is much less equivocal, and is consistent with descriptions of swim traces in the global ichnological record.[6] Measurements of the newly identified traces included interdigital distances, length, width, and depth. Results were recorded in centimetres. Tracings were made on clear acetate film. Photographs were taken, including photogrammetric analysis[14]: three-dimensional models were generated with Agisoft MetaShape Professional (v. 1.0.4) using an Olympus TG-5 camera (focal length 4.5 mm; resolution 4000 x 3000; pixel size 1.56 x 1.56 um). The final images were rendered using CloudCompare (v.2.10-beta).

The dimensions of the trace fossil surface are 180 cm x 70 cm. The slab is orientated in a shore-parallel, west-east direction, just above the high tide mark, with the upper surface facing gently northwards. The lithology appears identical to that of the previously reported track-bearing surfaces[1]: the surface stratum, containing the traces, consists of a thin (less than 2 cm) veneer of silt and fine sand, above and in contact with a heavily bioturbated layer. The swim traces, all of which are preserved in epirelief, are shown in Figures 2 and 3.

Recorded measurements are presented in Table 1. At the eastern end of the surface (to the left in Figures 1, 2 and 3), a probable manus-pes set of tracks is present in epirelief, with claw impressions. While these tracks are poorly preserved, as they penetrate into the underlying bioturbated layer, they bear a resemblance to the large reptile tracks described by Helm et al.[1] At the western end a raised white feature, 15 cm x 5 cm in size, is interpreted as a possible coprolite. Between these features a variety of swim traces is apparent. Depth varies from a maximum of 5 cm for the deepest track (labelled A1 in Table 1) to less than 1 cm for the smaller swim traces. The sets of swim traces, totaling at least eight (and labelled from B to I in Table 1), have an approximately unimodal orientation. They exhibit a range of digit patterns from monodactyl to pentadactyl. In some cases, they are contiguous with probable partial tracks, whereas in most cases they are consistent with 'raking traces' as described by Romano and Whyte[8], without any evidence of tracks.

Figure 1: The loose slab containing a newly identified trace fossil-bearing surface lies near the high tide mark; scale bar = 10 cm.

Figure 2: The trace fossil surface, facing south, with labels showing tracks, swim traces and possible coprolite; scale bar = 10 cm. Measurements of labelled features are presented in Table 1.

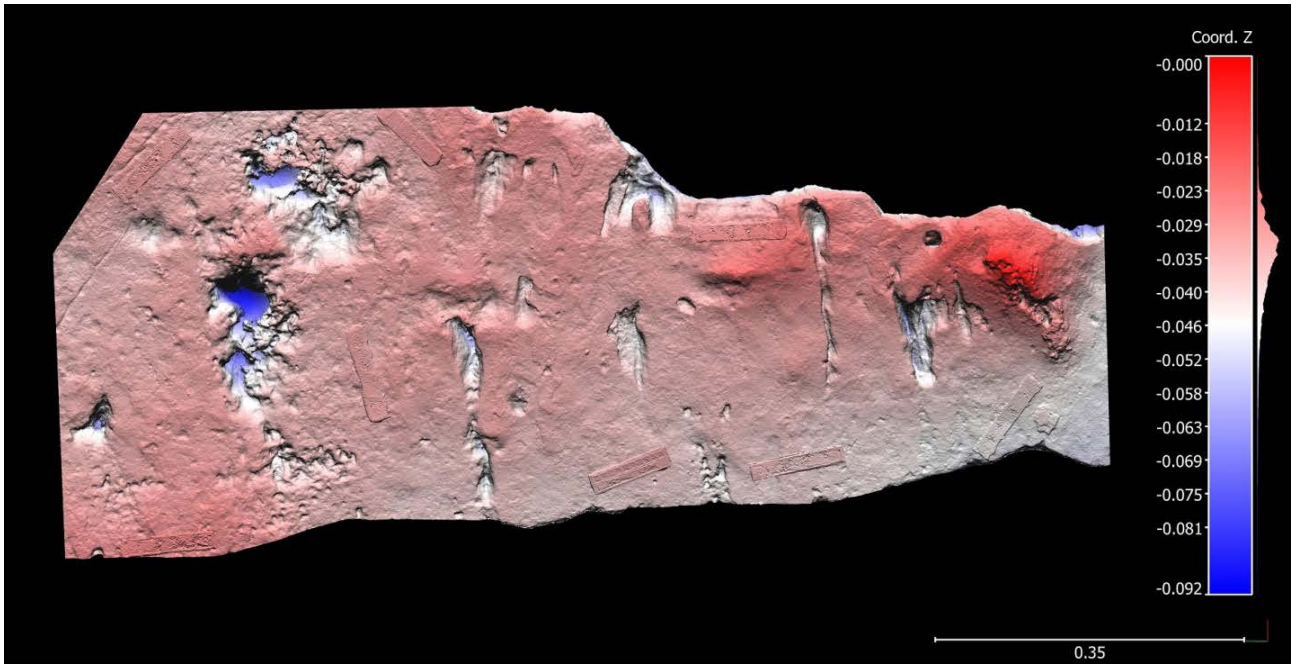


Figure 3: Photogrammetry colour mesh of tracks and swim traces, facing south. The three-dimensional model was generated with Agisoft MetaShape Professional (v. 1.0.4) using 108 images. Photos were taken an average 28.6 cm from the surface. The reprojection error is 0.381 pix. Vertical and horizontal scales are in metres.

Our interpretation is that the observed features may represent a deepening aqueous environment from east to west (left to right in Figures 1, 2 and 3), e.g. that the manus-pes set of tracks at the east end was thus made in shallower water, and that the traces further to the west were thus made in deeper water and were therefore registered as swim traces.

Such a deepening environment from east to west would help explain the long monodactyl trace further to the west (labelled 'H' in Figure 2), which is 22 cm long, with a slight convexity to the west. Mean interdigital distance, where measurable (no measurement was possible for the monodactyl trace), ranges from 1.5 cm for the smaller sets of traces to 2.2 cm for the larger sets of traces.

Table 1: Recorded measurements, in centimetres, on the surface containing tracks and traces

Label	Length	Width	Number of traces	Interdigital distance	Convexity	Comments
A	A1 = 16? A2 = 7.5?	A1 =10? A2 = 7.5?	n/a	n/a	n/a	pes and manus tracks
B	B1 = 8 B2 = 9	B1 = 1.5 B2 = 2	?1	n/a	to west	two traces in series
C	C1 = 1 C2 = 4.5 C3 = 1.5 C4 = 1.5 C5 = 1.5	C1 = 1 C2 =2 C3 =1 C4 = 1 C5 = 0.5	5	8.8/4 = 2.2	to west?	probably pentadactyl, with possible partial track
D	D1 = 2 D2 = 5.5	D1 = 1.5 D2 = 1.5	2	n/a	uncertain	two traces in series
E	E1 = 6 E2 = 7 E3 = 9	E1 = ? E2 = ? E3 =1?	3	3/2 = 1.5	to east	tridactyl, parallel, narrowly spaced traces
F	F1 = 7.5 F2 = 7.5 F3 = 6	F1 = 1 F2 = 1.5 F3 = 2	3	6.6/3 = 2.2	to east?	truncated at south edge, probably tridactyl, with possible track
G	G1 = 11 G2 = 11	G1 = <1 G2 = <1	2	variable	to east	non-parallel, intermittently registered traces, possibly truncated at north edge
H	22	0.5–2	1	n/a	to west	monodactyl
I	I1 = 14 I2 = 4.5 I3 = 4.5 I4 = 4.5 I5 = 6.5	I1 = 1.5 I2 = ? I3 = ? I4 = 1 I5 = 1	5	8/4 = 2	uncertain	probably pentadactyl
J	15	5	n/a	n/a	n/a	raised feature

In order to estimate the size of the reptiles that registered these traces, based on measured interdigital distance, data from a variety of sources was used. Thomson[15] investigated the ratio of pes (hindfoot) length to total body length for *C. niloticus*. A mean ratio of 1:12 was obtained.[15] Likewise, Milàn and Hedegaard[10] concluded from studies of extant crocodylians that pes length could be used to provide a useful approximation of total length. We are not aware of similar data that can be used to estimate length of monitor lizards (varanids) or chelonians. Illustrations of *C. niloticus* tracks in three reference works[16-18] on southern African tracks were compared , with attention to the ratio between pes length and interdigital distance. A mean interdigital distance to pes length ratio of 1:6.4 was obtained, and a significant difference was not apparent between interdigital distance of manus tracks and pes tracks.

In order to calculate total length from interdigital distance (if the traces were made by crocodylians), first the mean interdigital distance for a given set of swim traces was calculated, e.g. for a tetradactyl trace this involved calculating the mean of three measurements. This number was then multiplied by 77 (12 x 6.4 = ~77). It is acknowledged that the resulting estimate of total length is imprecise, as it does not take into account possible splaying or adduction of digits when the swim traces were registered, as described by Sadlok and Pawełczyk[19]. Nonetheless, we contend that it can be used to provide an approximate indication of size. Based on the above-mentioned methods, estimated total length (if the trackmakers were crocodylians) varied from 115 cm for the smallest swim traces to 170 cm for the largest swim traces.

In conclusion, while the interpretation of tetrapod swim traces remains difficult, the newly exposed surface indicates the presence of multiple compelling examples of large reptile swim traces in Africa, thereby supplementing the evidence that led to the conclusion of 'probable swim traces' reported by Helm et al.[1] Although interdigital distance recorded from swim traces may not be constant for an individual trackmaker, substantial differences in this measurement suggest that reptiles of varying size may have made the swim traces. While juvenile *C. niloticus* or juvenile *V. niloticus* may be the most likely trackmakers, a swimming chelonian trackmaker, such as a large terrapin, cannot be fully excluded on morphological grounds. Although crocodylian and chelonian swim traces are not uncommon in the global trace fossil record, swim traces of monitor lizards have not previously been reported. These appear to

be the first compelling examples of non-dinosaurian reptile swim traces from the African continent.

Acknowledgements

We appreciate the support of Jonathan Britton, Hayley Cawthra, Xander Combrink, Jan De Vynck, Carina Helm, Renee Rust, SANParks staff, Willo Stear, Chris Stuart, Mathilde Stuart, David Thesen, Guy Thesen, and Alex van den Heever.

Competing interests

We have no competing interests to declare.

Authors' contributions

C.W.H.: Lead author, corresponding author, conceptualisation, data collection, photogrammetry, data analysis, writing, project leadership, trace fossil analysis. M.G.L.: Conceptualisation, site visit, data analysis, trace fossil analysis and contribution on ichnology, review of drafts and revisions

References

1. Helm CW, Cawthra HC, Combrink X, Helm CJZ, Rust R, Stear W, Van Den Heever A. Pleistocene large reptile tracks and probable swim traces on South Africa's Cape south coast. S Afr J Sci. 2020;116(3/4), Art. #6542. https://doi.org/10.17159/sajs.2020/6542
2. Mudroch A, Richter U, Joger U, Kosma R, Idé O, Maga A. Didactyl tracks of paravian theropods (Maniraptora) from the Middle Jurassic of Africa. PLoS ONE. 2011;6(2), e14642. https://doi.org/10.1371/journal.pone.001464
3. Milner ARC, Lockley MG. Dinosaur swim track assemblages: Characteristics, contexts and ichnofacies implications. In: Falkingham PL, Marty D, Richter A, editors. Dinosaur tracks: The next steps. Bloomington, IN: Indiana University Press; 2016. p. 152–180.
4. Malan JA. Lithostratigraphy of the Klein Brak Formation (Bredasdorp Group). South African Committee for Stratigraphy Lithostratigraphic Series 13. Pretoria: Department of Mineral and Energy Affairs; 1991.
5. Bateman MD, Carr AS, Dunajko AC, Holmes PJ, Roberts DL, Mclaren SJ, et al. The evolution of coastal barrier systems: A case study of the Middle-Late Pleistocene Wilderness barriers, South Africa. Quat Sci Rev. 2011;30:63–81. https://doi.org/10.1016/j.quascirev.2010.10.003
6. Lockley MG, Cart K, Martin J, Prunty R, Houck K, Hups K, et al. A bonanza of new tetrapod tracksites from the Cretaceous Dakota Group, western Colorado: Implications for paleoecology. Bull N M Mus Nat Hist Sci. 2014;62:393–409.
7. Lockley MG, Lucas SG, Milàn J, Harris JD, Avanzini M, Foster JR, Spielmann JA. The fossil record of crocodylian tracks and traces: an overview. In: Milàn J, Lucas SG, Lockley MG, Spielmann JA, editors. Crocodyle tracks and traces. Bull N M Mus Nat Hist Sci. 2010;51:1–15.
8. Romano M, Whyte MA. Crocodilian and other non-dinosaurian tracks and trackways from the Ravenscar Group (Middle Jurassic) of the Cleveland Basin, Yorkshire, UK. In: Milàn J, Lucas SG, Lockley MG, Spielmann JA, editors. Crocodyle tracks and traces. Bull N M Mus Nat Hist Sci. 2010;51:69–81.
9. Mustoe GE. Lower Eocene footprints from northwest Washington, USA. Part 1: Reptile tracks. Geosciences. 2019;9(321):1–19. https://doi.org/10.3390/geosciences9070321
10. Milàn J, Hedegaard R. Interspecific variation in tracks and trackways from extant crocodylians. In: Milàn J, Lucas SG, Lockley MG, Spielmann JA, editors. Crocodyle tracks and traces. Bull N M Mus Nat Hist Sci. 2010;51:15–29.
11. McCrea RT, Buckley LG, Currie PJ, Plint AG, Helm CW, Haggart JW. A review of vertebrate track-bearing formations from the Mesozoic and earliest Cenozoic of western Canada with a description of a new theropod ichnospecies and reassignment of an avian ichnogenus. In: Lockley MG, Lucas S, editors. Tracking dinosaurs and other tetrapods in western North America. Bull N M Mus Nat Hist Sci. 2013;62:5–94.
12. Grigg G, Kirshner D. Biology and evolution of crocodylians. Ithaca, NY: Cornell University Press / CSIRO Australia; 2015.
13. Lockley M. The eternal trail. New York: Perseus Publishing; 1999.
14. Matthews NA, Noble TA, Breithaupt BH. Close-range photogrammetry for 3-D ichnology: The basics of photogrammetric ichnology. In: Falkingham PL, Marty D, Richter A, editors. Dinosaur tracks: The next steps. Bloomington, IN: Indiana University Press; 2016. p. 28–55.
15. Thomson WR. A field guide for estimating the overall length of the crocodile. S Afr J Wildl Res. 1972;2(1):27.
16. Liebenberg L. A photographic guide to tracks and tracking in Southern Africa. Cape Town; Struik Publishers; 2000.
17. Stuart C, Stuart T. A field guide to the tracks and signs of southern and East African wildlife. Cape Town: Struik Nature; 2000.
18. Van den Heever A, Mhlongo R, Benadie K. Tracker manual – a practical guide to animal tracking in southern Africa. Cape Town: Struik Nature; 2017.
19. Sadlok G, Pawełczyk K. Tetrapod swim techniques interpreted from swim trace fossils from the Lower Triassic Baranów Formation, Holy Cross Mountains, central Poland. PalZ. 2020. https://doi.org/10.1007/s12542-019-00510-w

AUTHORS:
Christopher Harris[1,2]
Robert W. Gess[3]
Cameron R. Penn-Clarke[2,4]
Bruce S. Rubidge[2]

AFFILIATIONS:
[1]School of Geosciences, University of the Witwatersrand, Johannesburg, South Africa
[2]Evolutionary Studies Institute, School of Geosciences, University of the Witwatersrand, Johannesburg, South Africa
[3]Devonian Lab, Albany Museum, Rhodes University, Makhanda, South Africa
[4]Council for Geoscience, Pretoria, South Africa

CORRESPONDENCE TO:
Christopher Harris

EMAIL:
ch.kriekkrok@gmail.com

DATES:
Received: 06 Nov. 2020
Revised: 09 Dec. 2020
Accepted: 13 Jan. 2021
Published: 29 Mar. 2021

HOW TO CITE:
Harris C, Gess RW, Penn-Clarke CR, Rubidge BS. Coombs Hill: A Late Devonian fossil locality in the Witpoort Formation (Witteberg Group, South Africa). S Afr J Sci. 2021;117(3/4), Art. #9190. https://doi.org/10.17159/sajs.2021/9190

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Margaret Avery

KEYWORDS:
Famennian, high-latitude, fossils, plants, shelly invertebrates

FUNDING:
DSI/NRF Centre of Excellence in Palaeosciences, DSI/NRF African Origins Platform, National Research Foundation (South Africa), the Palaeontological Scientific Trust (PAST), the Millennium Trust

Coombs Hill: A Late Devonian fossil locality in the Witpoort Formation (Witteberg Group, South Africa)

Coombs Hill, a new fossil locality in the Witpoort Formation (Witteberg Group) of South Africa, preserves a record of Famennian (Late Devonian) life in Gondwana. Fossil plants collected at Coombs Hill are preliminarily assigned to several classes. Shelly invertebrates include a variety of bivalve mollusc forms, some of which appear to be preserved in life position. Biodiversity at Coombs Hill is comparable to that of the well-known Waterloo Farm *lagerstätte* in ordinal diversity, but exhibits differences in species composition. Ongoing taxonomic analysis will provide a rare window into the ecology of high-latitude environments during this pivotal stage of Earth history, which immediately preceded the end-Devonian extinction. Sandstone dominated sedimentary facies at Coombs Hill suggest a high-energy coastal marine setting, with brackish back-barrier estuarine/lagoonally derived fossiliferous mudstones. Exact stratigraphic placement within the Witpoort Formation is hampered by structural deformation, and precise age comparisons with Waterloo Farm are currently tenuous.

Significance:

- A new fossil locality at Coombs Hill comprises the second known site with a suite of well-preserved continental and marginal marine fossils from the Witpoort Formation, providing an exceptionally rare example of high-latitude life during the critical latest Devonian Famennian age.
- Several new plant taxa will be diagnosed from this locality, which also gives important insights into the morphology of *Archaeopteris notosaria*, South Africa's earliest known tree.
- Discovery of a second palaeontologically significant site in the Witpoort Formation provides impetus for further structural and sedimentary facies analyses to align the unit with datable global eustatic events, and to clarify its internal chronology.

Introduction

Coombs Hill (33°17'51.86"S 26°45'28.70"E) is situated within the predominantly quartzitic Witpoort Formation (mid Witteberg Group, Cape Supergroup) east of Makhanda (Figure 1). The Witteberg Group accumulated in a shallow marine setting along the southern coastline of Gondwana, between the Givetian (Middle Devonian) and Visean (Early Carboniferous), with the Witpoort Formation being of Famennian age[1,2] – a time in which the region was within the Antarctic Circle[3] (Figure 1).

Strata of the Witpoort Formation, along with the rest of the Cape Supergroup, crop out for more than 1000 km along South Africa's southern and western coastlines, having been tectonically deformed during the formation of the late Palaeozoic Cape Fold Belt.[4]

The Witpoort Formation has, in recent decades, provided evidence for high-latitude latest Devonian life, although exclusively from a single *lagerstätte* at Waterloo Farm (Figure 1), which comprises the only known high-latitude marginal marine ecosystem from the Famennian. The 23 plant and animal taxa as yet diagnosed include the only known high-latitude Devonian tetrapods, the earliest known lampreys and Gondwana's only Devonian terrestrial invertebrates.[3,5,6] Here, we report the discovery of a new locality at Coombs Hill, 21 km to the east of Waterloo Farm.

The Late Devonian records the end of the Siluro-Devonian hothouse climate, with the appearance of continental glaciers and increasingly dramatic sea-level (eustatic) fluctuations[7] associated with a series of major biotic crises[8]. Cooling is considered by many authors to have resulted from drawdown of CO_2 associated with widespread vegetation of land.[9,10] Abundant vegetation at Coombs Hill and Waterloo Farm provides the only evidence for high-latitude forestation during the Late Devonian biocrisis.

Structural geology

The precise stratigraphic position of the Coombs Hill site in the Witpoort Formation is uncertain, as there is no continuous lithological section to formational contacts, and strata exhibit duplications of stratigraphic sequences at outcrop. This results from structural deformation characteristic of the eastern Cape Fold Belt[11] and is evident in a 350-m-long road cutting (Figure 2a), wherein indicators of younging direction reveal that most of the strata are overturned about a synclinal fold (Figure 2b). Strata are continuous around the fold hinge as is demonstrated by two distinct mudrock horizons, resulting in stratigraphic duplication of much of the overturned succession. Some minor faulting and brecciation occurs within the section, associated with kinking along the overturned limb, but it is insufficient to displace marker horizons.

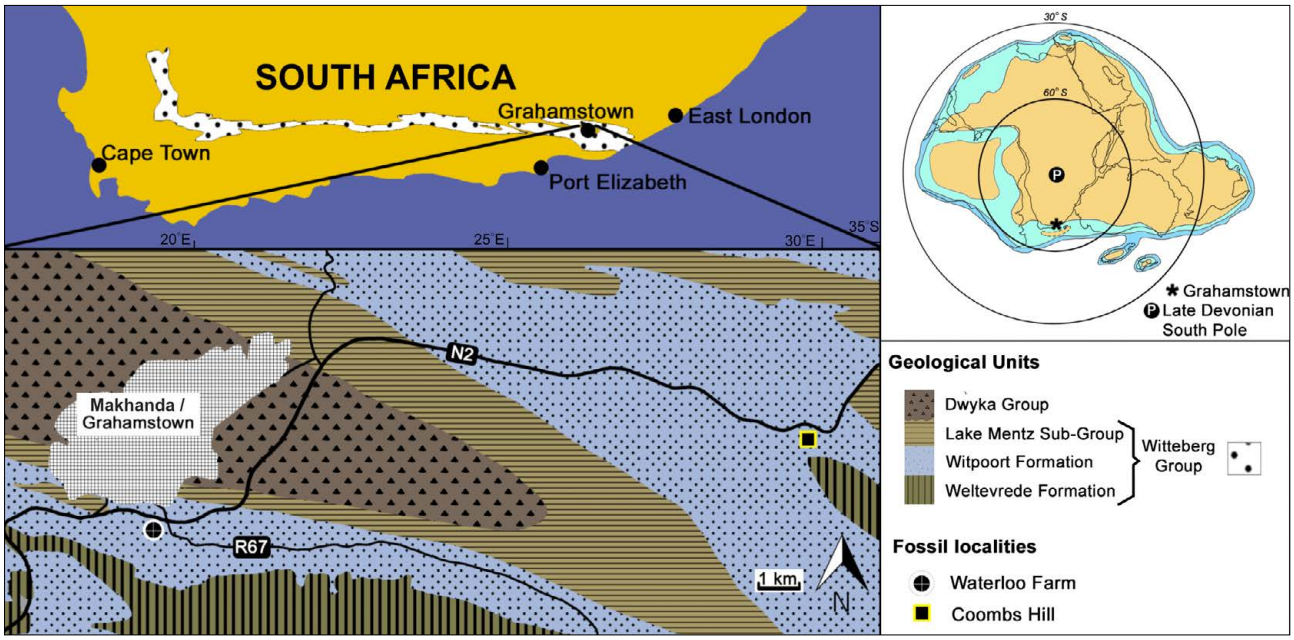


Figure 1: Geological map of the Makhanda (Grahamstown) region (based on Council for Geoscience, Grahamstown sheet 3326 (1995)), showing the locations of fossil localities at Waterloo Farm and Coombs Hill. A key to stratigraphic units is provided at the bottom right. Palaeogeographic reconstruction of Late Devonian Gondwana (top right) modified from Gess and Ahlberg[3].

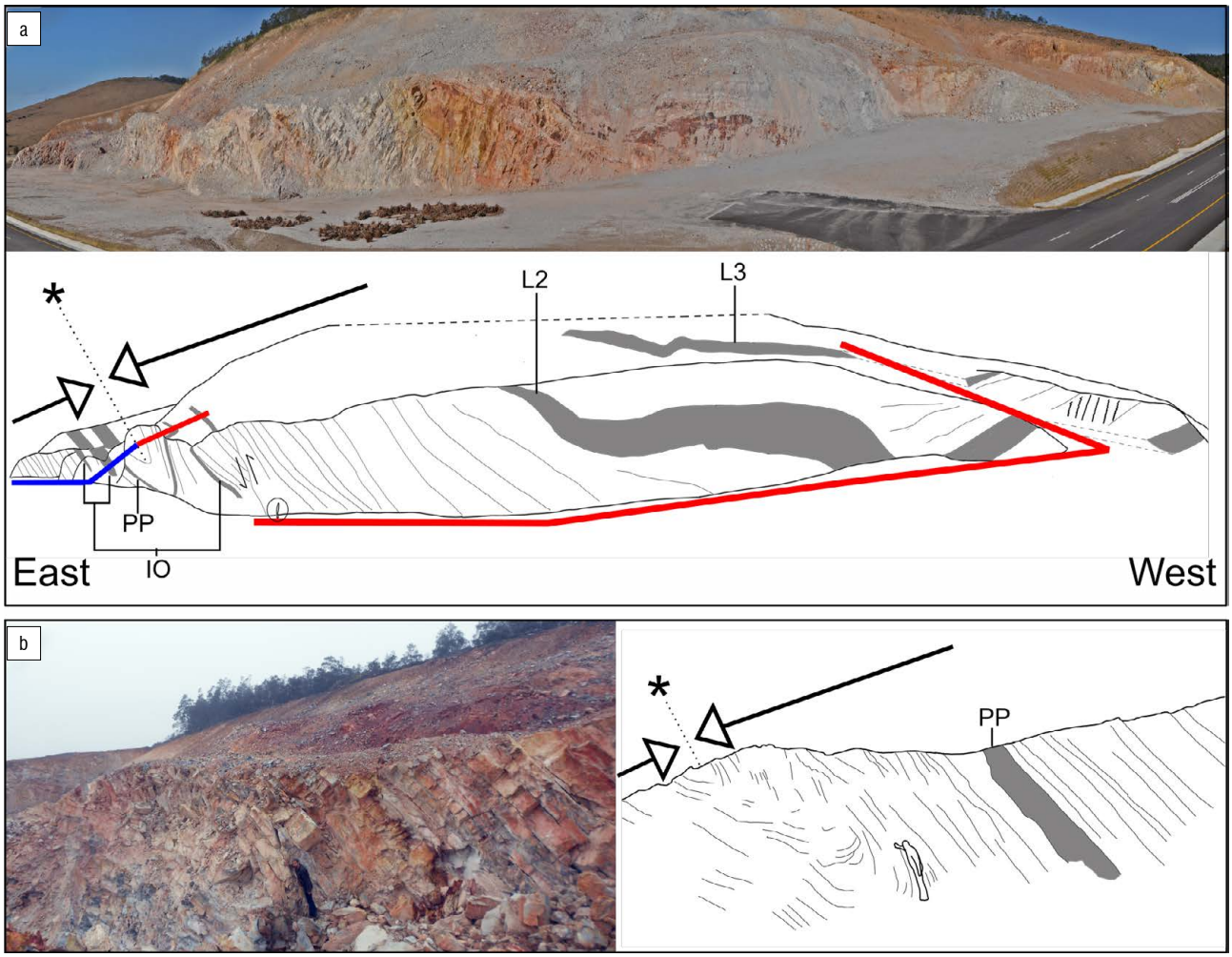


Figure 2: Geological structure of Coombs Hill. (a) Panoramic photograph of the road cutting with interpretive line drawing below. Arrows indicate younging direction of strata, which converge towards a fold axis (*); fossiliferous mudstone horizons are labelled (see Figure 3); the transect followed during systematic logging is indicated by a red line for the overturned strata and a blue line for the right-way-up strata; person for scale is circled. (b) Synclinal fold axis with interpretive line drawing to right, with arrows as in (a).

Stratigraphy and sedimentology

A stratigraphic log of exposures along the recent road cutting at Coombs Hill records roughly 80 m of stratigraphic sequence (Figure 3). The majority of the strata comprise clean blue-grey (when fresh; 5B 6/2–5B 7/1 (Munsell colour system)), medium-grained quartz arenites, which are stacked in a succession displaying only subtle lithological changes, and punctuated by discrete mudstone horizons, comprising ~10% of the succession. Many of the quartzitic strata superficially appear red to purple due to precipitation of minerals along joints. Weathered quartzites usually alter to light brown (5YR 5/6) or light grey (N8).

Quartzitic sedimentary facies include (1) planar cross-stratification (Figure 4a), (2) low-angle planar cross-stratification (Figure 4b), (3) trough cross-stratification (Figure 4d), (4) parabolic cross-stratification (Figure 4f), (5) low angle undulatory (c.f. swaley) cross-stratification (Figure 4c), (6) horizontal lamination, (7) ripple lamination and (8) internally structureless beds. Muddy facies include (9) horizontally laminated to structureless mudstone and (10) horizontally laminated heterolithic greywacke. Mudstone horizons exhibit lateral variations in thickness and texture, and are commonly eroded at the contact with overlying arenaceous strata (Figure 4e).

Palaeocurrent trends derived from cross-strata (n=47) are polymodal, with modes towards the east, south-southwest, and a minor component to the northwest (Figure 3). These trends, relative to a roughly east–west trending shoreline[12,13], indicate both offshore and longshore sediment transport with a minor landward component, consistent with the fluctuations of current, wave and storm energy in shallow marine rather than fluvial settings[14,15].

Many of the facies associations formerly described in the Witpoort Formation around Makhanda[13] are recognised at Coombs Hill, and three broad depositional environments are provisionally identified: (1) shoreface, (2) foreshore and (3) backshore (as illustrated in Figure 3). The new locality is similarly interpreted as representing a shallow to marginal-marine succession characterised by high wave energy and continual reworking of sediments.

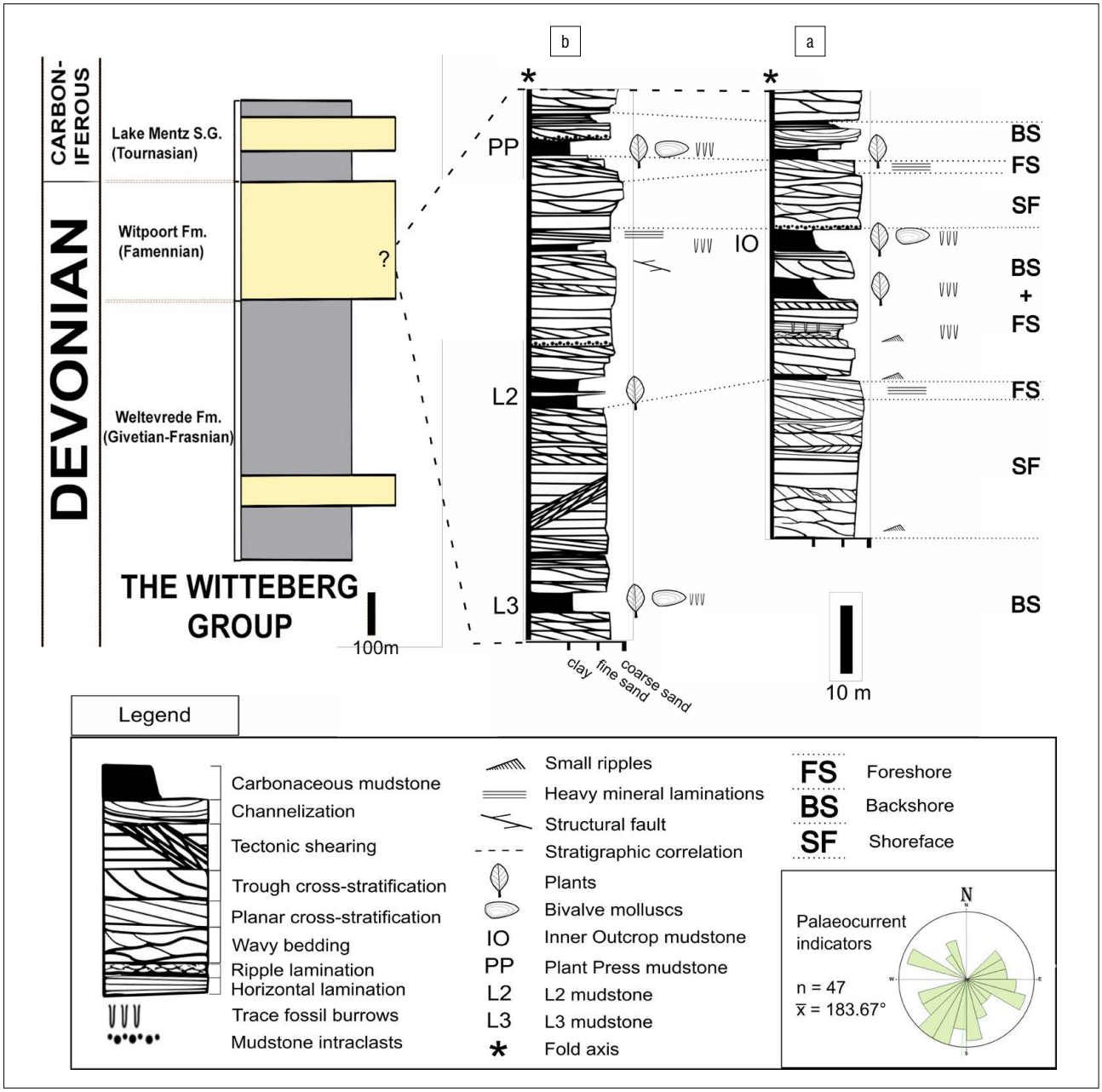


Figure 3: Stratigraphic log of the Coombs Hill road cutting showing (a) the right-way-up portion and (b) the overturned portion of the succession (refer to Figure 2). Cumulative palaeocurrent indicators for the whole succession are inset at the bottom right.

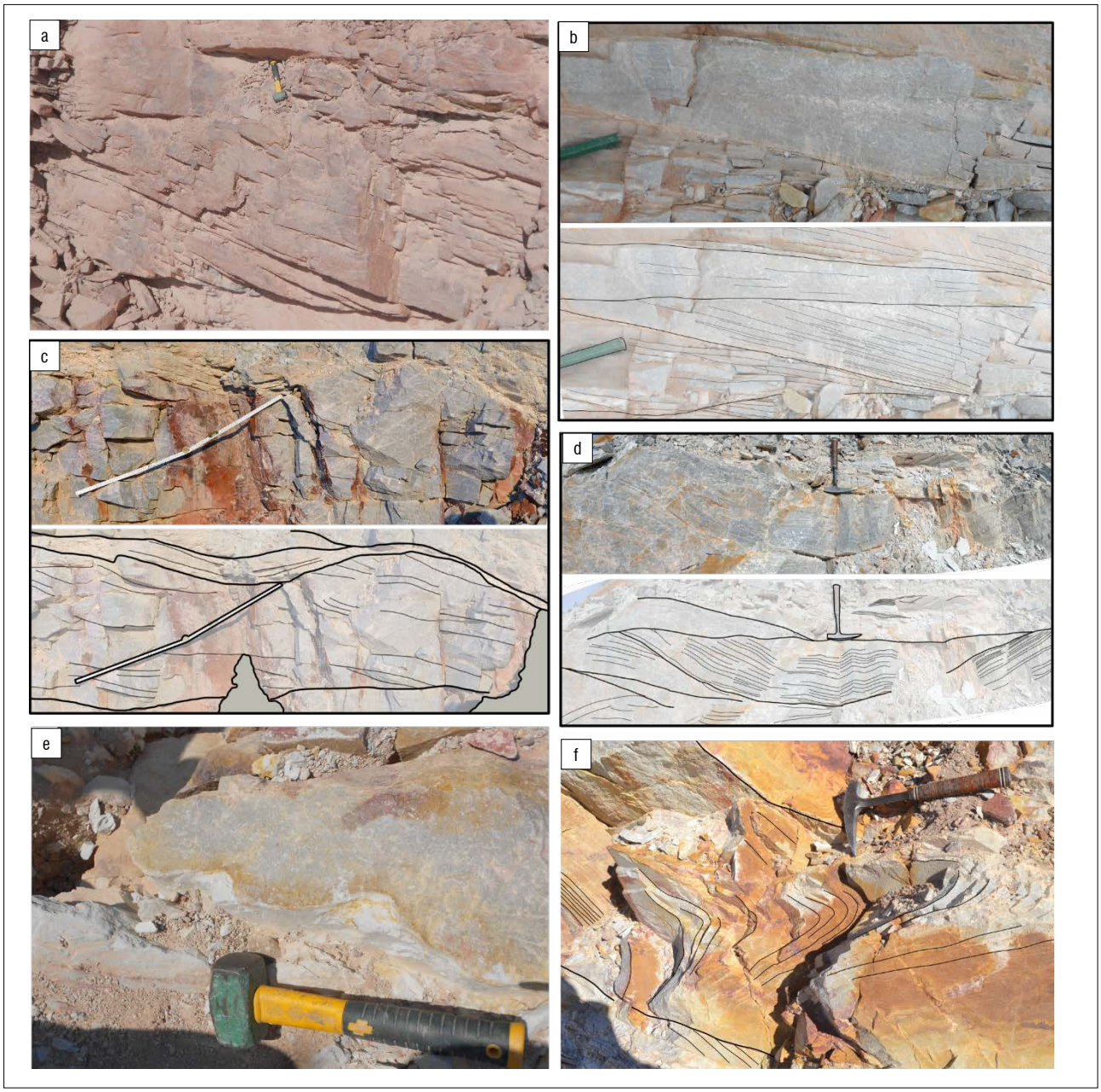


Figure 4: Sedimentary facies at Coombs Hill. (a) Thickly bedded planar cross-stratification (4-lb hammer for scale). (b) Low-angle planar cross-stratification (handle of chisel is approximately 2 cm wide). (c) Swaley cross-stratification (scale is 1 m). (d) Channelised trough cross-stratification (geological pick for scale). (e) Stepped channel bank eroded into mudstone (4-lb hammer for scale). (f) Parabolic cross-stratification (younging direction towards the bottom of the photograph, geological pick for scale). (b,c,d,f,) Interpretive line drawings showing erosive set bounding surfaces (thicker lines) and bedding planes (thinner lines).

The alternation between shoreface and backshore is the result of transgressive-regressive cycles, the chronology and driving mechanisms of which are not clearly understood. Further detailed facies studies are required to test the postulated depositional environments.

Fossil biota

Excavations by R.W.G. and C.H. in 2015 and 2016 produced a representative sample of the biotic remains preserved in the mudstones, including more than 200 specimens of relatively well-preserved plants and shelly invertebrates which are curated at the Albany Museum.

Mudstones preserve an association of putative marine algae, abundant terrestrial plant matter and bivalve remains. This evidence is in line with existing hypotheses which propose stagnant brackish-water lagoonal[13] or estuarine[16] depositional environments for the genesis of fossiliferous mudstone in the Witpoort Formation around Makhanda.

Fossil plants include remains provisionally ascribed to Cladoxylopsida, Progymnospermopsida (*Archaeopteris notosaria* Anderson et al. 1995), Lycopsida, Zosterophyllopsida, Gymnospermopsida and form taxa *Palaeostigma* Kräusel and Dolianti 1957, '*Dutoitia' alfreda* Plumstead 1967 and '*Dutoitia' maraisia* Plumstead 1967.[17]

Archaeopteris notosaria, hitherto only known from Waterloo Farm, provides the earliest evidence for trees at high palaeolatitudes.[18] Exquisite vegetative and fertile remains from Coombs Hill contribute significantly to taxonomic understanding of this species (Figure 5a, b). Lycopod remains comprise at least three new taxa, none of which is known to occur at Waterloo Farm.[17]

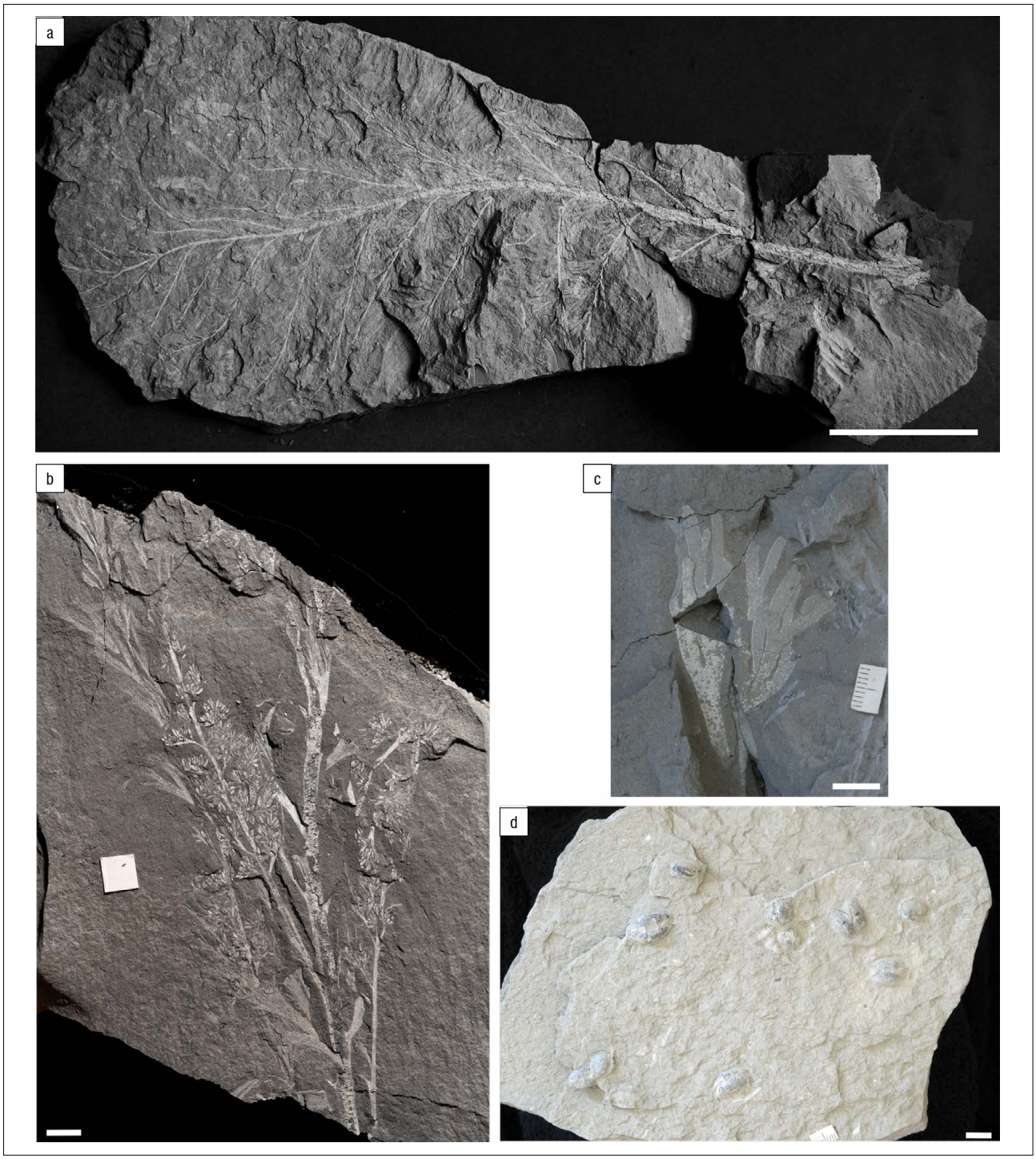


Figure 5: Fossil remains from Coombs Hill. (a) A near-complete lateral branch system of *Archaeopteris notosaria,* AM 7710a (scale bar = 10 cm). (b) Fertile branch of *A. notosaria,* AM 7714a (scale bar = 1 cm). (c) Putative marine alga, AM 7717 (cf. *Hungerfordia fionae*) (scale bar = 1 cm). (d) Articulated bivalve shells on a bedding plane, AM 7718a (scale bar = 1 cm).

Algal thalli, similar to *Hungerfordia fionae* Hiller and Gess 1995, are relatively common at Coombs Hill (Figure 5c), although they are more elongate and less bilaterally symmetrical, probably representing a new species.

Invertebrates are solely represented by abundant articulated bivalves consistently preserved with the umbo upwards (in apparent life position), usually in mudstones but sometimes in underlying sandstones (Figure 5d). These likely comprise several euryhaline genera, and their dominance of the muddy substrate suggests a brackish setting.[19]

Future research

The palaeontology of Coombs Hill is the subject of a series of taxonomic papers allowing palaeoecological comparisons with Waterloo Farm. Structural analysis will facilitate finer stratigraphic resolution of these localities and improved understanding of their relative ages. Detailed reconstructions of the depositional environments form part of a broader study on the Witpoort Formation and are important in understanding the relative influence of autocylic (i.e. subsidence and sediment supply) and allocyclic (eustatic) processes on environmental change that could elucidate major bioevents in the Late Devonian.

Acknowledgements

C.H. would like to thank Rhodes University's Geology Department for the use of a geological compass and Lechelle Goslin for assistance with the unfolding of strata using a stereonet. SANRAL are thanked for supporting the initial phases of fieldwork at Coombs Hill and for preserving this important section for posterity. This study has benefitted from financial assistance by the DSI/NRF Centre of Excellence in Palaeosciences, the DSI/NRF African Origins Platform, the National Research Foundation, the Palaeontological Scientific Trust (PAST) and the Millennium Trust. The opinions and conclusions expressed in this study are not attributed to the funding bodies mentioned above.

Competing interests

We declare that there are no competing interests.

Authors' contributions

C.H.: Conceptualisation, data collection, data analysis, writing – the initial draft, writing – revisions. R.W.G.: Conceptualisation, methodology, data collection, writing – revisions, student supervision. C.R.P.C.: Methodology, data collection, writing – revisions, student supervision. B.S.R.: Methodology, writing – revisions, project management, funding acquisition, student supervision.

References

1. Cooper MR. Facies shifts, sea-level changes and event stratigraphy in the Devonian of South Africa. S Afr J Sci. 1986;82(5):255.
2. Gess RW. Vertebrate biostratigraphy of the Witteberg Group and the Devonian-Carboniferous Boundary in South Africa. In: Linol B, De Wit MJ, editors. Origin and evolution of the Cape Mountains and Karoo Basin. Cham: Springer; 2016. p. 131–140. https://doi.org/10.1007/978-3-319-40859-0_13
3. Gess R, Ahlberg PE. A tetrapod fauna from within the Devonian Antarctic Circle. Science. 2018;360(6393):1120–1124. https://doi.org/10.1126/science.aaq1645
4. Shone RW, Booth PWK. The Cape Basin, South Africa: A review. J Afr Earth Sci. 2005;43(1–3):196–210. https://doi.org/10.1016/j.jafrearsci.2005.07.013
5. Gess RW, Coates MI, Rubidge BS. A lamprey from the Devonian period of South Africa. Nature. 2006;443(7114):981–984. https://doi.org/10.1038/nature05150
6. Gess RW. The earliest record of terrestrial animals in Gondwana: A scorpion from the Famennian (Late Devonian) Witpoort Formation of South Africa. Afr Invertebr. 2013;54(2):373–379. https://doi.org/10.5733/afin.054.0206
7. Eriksson KA, McClung WS, Simpson EL. Sequence stratigraphic expression of greenhouse, transitional and icehouse conditions in siliciclastic successions: Paleozoic examples from the central appalachian basin, USA. Earth-Sci Rev. 2019;188:176–189. https://doi.org/10.1016/j.earscirev.2018.11.010
8. Bond DP, Wignall PB. The role of sea-level change and marine anoxia in the Frasnian–Famennian (Late Devonian) mass extinction. Palaeogeogr Palaeoclimatol Palaeoecol. 2008;263(3–4):107–118. https://doi.org/10.1016/j.palaeo.2008.02.015
9. Algeo TJ, Scheckler SE, Maynard JB. Effects of the Middle to Late Devonian spread of vascular land plants on weathering regimes, marine biotas, and global climate. In: Gensel P, Edwards D, editors. Plants invade the land: Evolutionary and environmental perspectives. New York: Columbia University Press; 2001. p. 213–236. https://doi.org/10.7312/gens11160-013
10. Kenrick P, Wellman CH, Schneider H, Edgecombe GD. A timeline for terrestrialization: Consequences for the carbon cycle in the Palaeozoic. Philos Trans R Soc Lond *B* Biol Sci. 2012;367(1588):519–536. https://doi.org/10.1098/rstb.2011.0271
11. Booth PWK, Munro AJ, Shone RW. Lithological and structural characteristics of Cape Supergroup rocks at Port Alfred, Eastern Cape, South Africa. S Afr J Geol. 1999;102(4):391–404.
12. Theron JN, Loock JC. Devonian deltas of the Cape Supergroup, South Africa. In: Proceedings of the Second International Symposium on the Devonian System; Calgary, Canada. CSPG Special Publications; 1988.
13. Hiller N, Taylor FF. Late Devonian shoreline changes: An analysis of Witteberg Group stratigraphy in the Grahamstown area. S Afr J Geol. 1992;95(5–6):203–214.
14. Clifton HE. A reexamination of facies models for clastic shorelines. In: Posamentier HW, Walker RG, editors. Facies models revisited. Special Publication no. 84. Tulsa, OK: Society for Sedimentary Geology; 2006. p. 293–337.
15. Plint AG. Wave- and storm-dominated shoreline and shallow-marine systems. Facies Models. 2010;4:167–200.
16. Gess RW, Whitfield AK. Estuarine fish and tetrapod evolution: Insights from a Late Devonian (Famennian) Gondwanan estuarine lake and a southern African Holocene equivalent. Biol Rev. 2020;95(4):865–888. https://doi.org/10.1111/brv.12590
17. Harris C. Exploring biodiversity patterns among southern Gondwana's primeval forests [MSc dissertation]. Johannesburg: University of the Witwatersrand; 2019.
18. Anderson HM, Hiller N, Gess RW. *Archaeopteris* (Progymnospermopsida) from the Devonian of southern Africa. Bot J Linn Soc. 1995;117(4):305–320. https://doi.org/10.1111/j.1095-8339.1995.tb02593.x
19. Dalrymple RW, Choi K. Morphologic and facies trends through the fluvial–marine transition in tide-dominated depositional systems: A schematic framework for environmental and sequence-stratigraphic interpretation. Earth-Sci Rev. 2007;81(3–4):135–174. https://doi.org/10.1016/j.earscirev.2006.10.002

AUTHOR:
Romala Govender[1,2]

AFFILIATIONS:
[1]Research and Exhibitions, Iziko Museums of South Africa, Cape Town, South Africa
[2]Department of Biological Sciences, University of Cape Town, Cape Town, South Africa

CORRESPONDENCE TO:
Romala Govender

EMAIL:
rgovender@iziko.org.za

DATES:
Received: 27 Aug. 2020
Revised: 15 Jan. 2021
Accepted: 15 Jan. 2021
Published: 29 Mar. 2021

HOW TO CITE:
Govender R. Shark–seal interaction off South Africa's west coast during the early Pliocene (5 Ma). S Afr J Sci. 2021;117(3/4), Art. #8837. https://doi.org/10.17159/sajs.2021/8837

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITOR:
Margaret Avery

KEYWORDS:
white shark, phocid seal, *Homiphoca, Carcharodon carcharias*, Langebaanweg, bite mark

FUNDING:
NRF/African Origins Platform grant (grant numbers UID 98834 and UID117782)

Shark–seal interaction off South Africa's west coast during the early Pliocene (5 Ma)

A study of the Mio-Pliocene marine palaeoenvironment on South Africa's west coast revealed aspects of the biology and behaviour of fossil marine mammals. Close examination showed that seals from Langebaanweg suffered from pathologies and bore marks of marine carnivore activity. This study adds to our knowledge of shark feeding behaviour in the geological past and is one of a few studies of sharks feeding on seals in the fossil record. Two incomplete seal humeri with shark tooth marks are the first documented evidence from South Africa's Mio-Pliocene of such behaviour. These injuries show no healing, which suggests that the animals were most likely scavenged.

Significance:

- Fossil rich deposits at Langebaanweg contribute to the knowledge of South African Mio-Pliocene fossils by placing them in a global context.
- This study is one of five globally that have documented marine carnivores feeding on seals.
- This is the first description of white sharks feeding on seals from South Africa's geological past.
- This study shows this behaviour was in place on South Africa's coast as early as 5 million years ago.
- The injuries show no signs of healing, suggesting the most parsimonious explanation is that white sharks were scavenging seal carcasses.

Introduction

Along South Africa's south and west coasts today, there are offshore breeding colonies of South African fur seals (*Arctocephalus pusillus pusillus*). White shark (*Carcharodon carcharias*) predation on these colonies is well documented.[1-3] It therefore follows that there should be evidence of similar behaviour in the geological past. Diedrich[4] showed that the correlation between the presence of seal fossils and abundance of shark teeth in the Proto North Sea may indicate the earliest specialisation for seal hunting in the fossil record in the northern hemisphere. This indicates that a specialisation by white sharks to hunt seals was already in place 50 Ma.[4] Collareta et al.[5] described bites from a megatooth shark (*Carcharocles megalodon*) on the scapula of a large seal from Peru's late Miocene, while *Squalus* (dogfish shark) tooth marks were identified on an *Allodesmus* skull (a large extinct seal-like pinniped known from the Northern Pacific Ocean) from the early Late Miocene of southwestern Washington State.[6] Purdy et al.[7] described a long unserrated bite mark on the distal humerus of *Callophoca obscura* (an extinct earless seal) left by *Isurus xiphodon* (an extinct mako shark), eliciting the question as to whether this shark preferred warm-water seals.

Part of this study was to understand the interactions occurring between animals living on or visiting the west coast. An analysis of the marine mammal fossils has produced evidence of white and mako sharks feeding on cetaceans off the coast 5 Ma, like they do today.[8-10] It was therefore expected to find similar evidence of sharks feeding on seals as they do today off South Africa's south and west coasts. This study documents the first evidence of sharks feeding on seals from South Africa's Mio-Pliocene west coast.

Geological setting

Langebaanweg is located on the stable platform of South Africa's west coast; 13 km inland from the current coastline and less than 2° north of the southern tip of Africa (Figure 1a).[11-14] The Varswater Formation at 'E' Quarry spans the middle Miocene (Langhian) to the early Pliocene (Zanclean) and is associated with a number of changes in sea level.[12-16] The bonebed is at the base of the Muishond Fontein Pelletal Phosphorite Member (MPPM) within a channel fill in the Langeberg Quartzose Sand Member (LQSM)[13] (Figure 1b). The Langebaanweg fossils were deposited during the 30-m high sea level when 'E' Quarry was less than a kilometre from the river mouth. During the earliest Pliocene, the ambient and sea surface temperatures cooled, and the Benguela Upwelling System intensified, providing a rich feeding ground.[11-13,17]

'E' Quarry sedimentary facies include a river channel deposit near the river mouth, shallow marine, estuarine, marsh and fluvial systems.[14,18] At Langebaanweg during the earliest Pliocene (Zanclean) there was a protected lagoon open to the ocean as well as an estuary.[19] Marine mammal skeletal elements were deposited by storm surges flooding the neighbouring land surface[11] and by terrestrial carnivore activity. Areas of high relief formed islands off the west coast during the early Pliocene (Zanclean)[11,18,20] and provided ideal conditions for seals to breed.

Materials and methods

There is a total of 3131 complete, incomplete, and fragmentary cranial and postcranial elements in the phocid seal (*Homiphoca*) assemblage from Langebaanweg.[21] Two isolated right humeri (SAM-PQL-34631 and SAM-PQL-60698) have shark tooth marks and root etching on the distal two-thirds (Table 1; Figures 2 and 3).

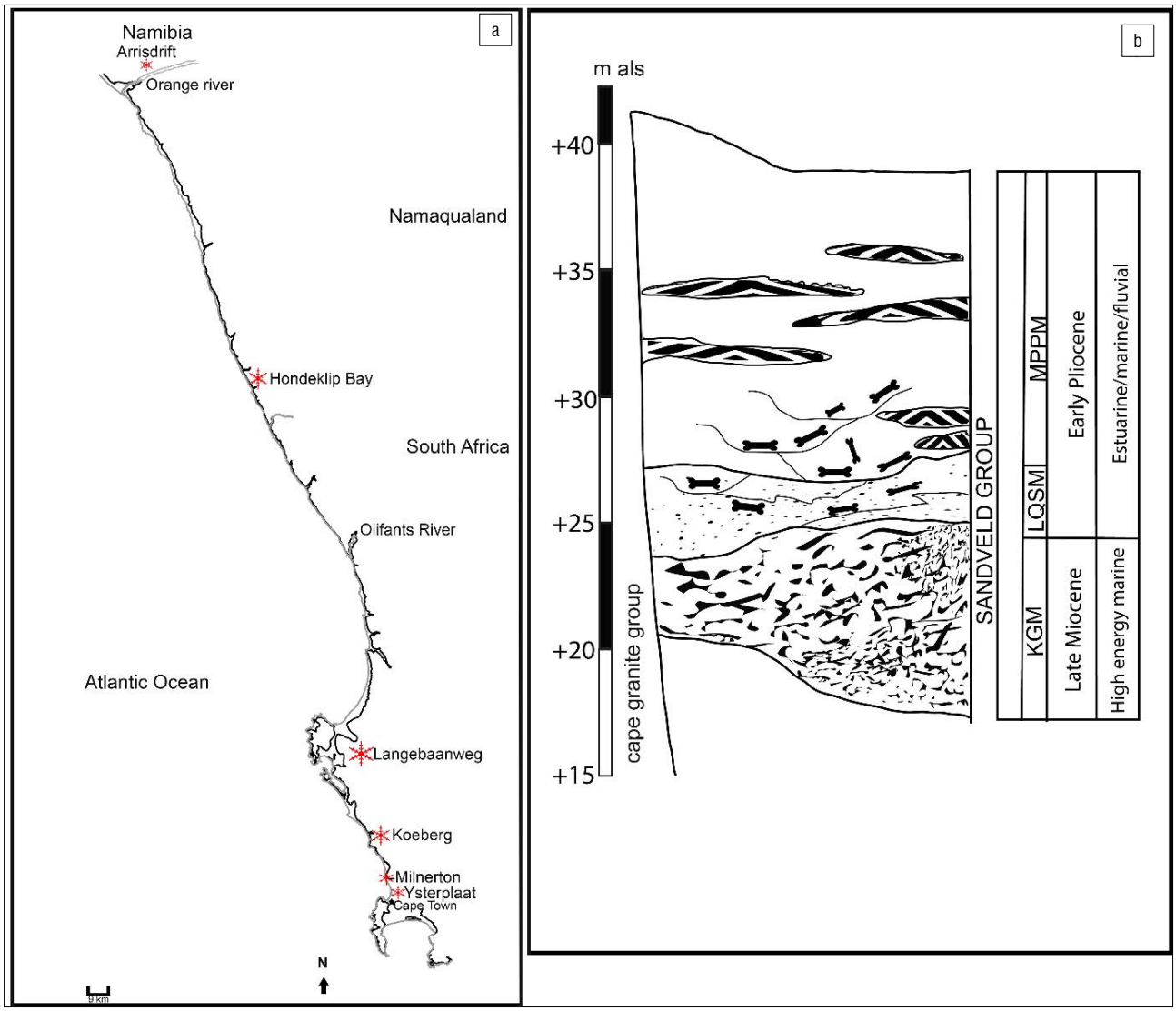


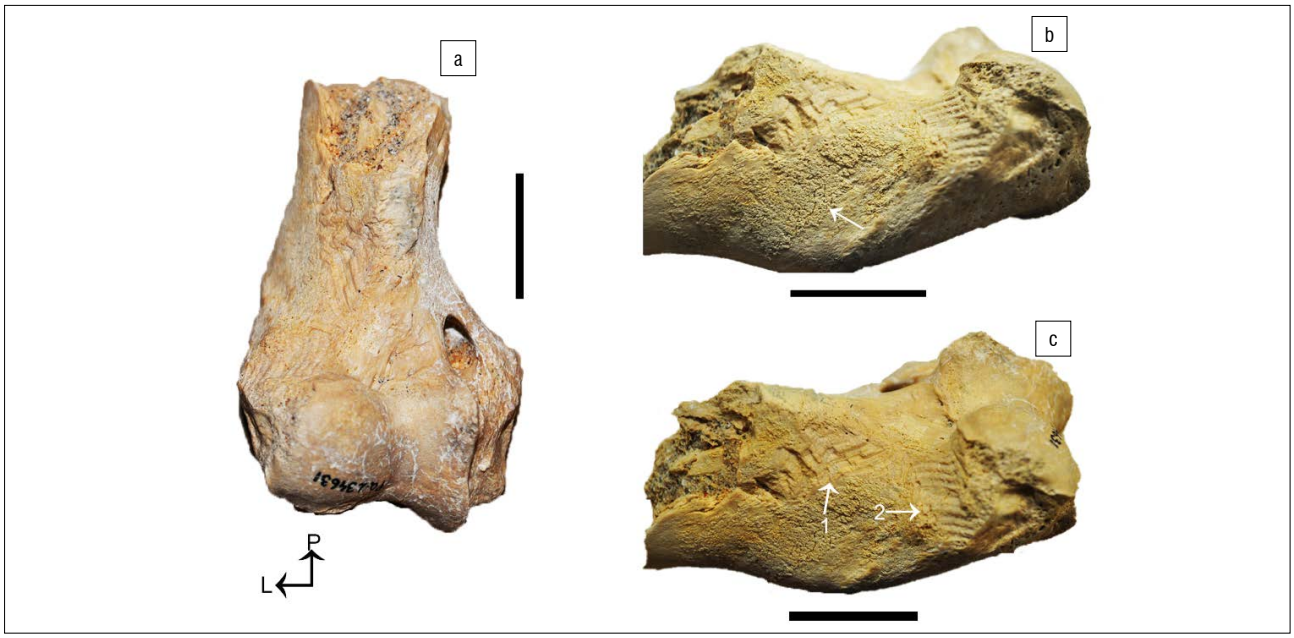
Figure 1: (a) Locality map showing the location of Langebaanweg during the 30-m high (the modern coastline is shown in grey). (b) Abbreviated stratigraphic section of Langebaanweg [Konings Vlei Gravel (KGM), Langeberg Quartz Sand (LQSM) and Muishond Fontein Pelletal Phosphorite (MPPM) Members] showing the fossil finds concentrated between 26 m and 30 m above sea level (m asl) (after Roberts et al.[13]; Govender and Chinsamy[8]; Govender et al.[28]).

Table 1: List of seal specimens with the number and type of bites identified

Element	Specimen number	Locality	Number of bites	CF1	CF2	CF3	CF4
Incomplete right humerus	SAM-PQL-34631	MPPM (BCWW T2 area RS)	2	–	–	x	–
Incomplete right humerus	SAM-PQL-60698	MPPM (W Wall IWRP 1976/2)	19	x	x	–	–

Bone fragments are embedded in the bone cortex due to the force of a bite or the pressure causing the fracture (Figure 2a). The bites are preserved on the shaft and distal part of the bone. These specimens were collected in 1976 from MPPM as part of an excavation. The bites described below follow Govender and Chinsamy[8]:

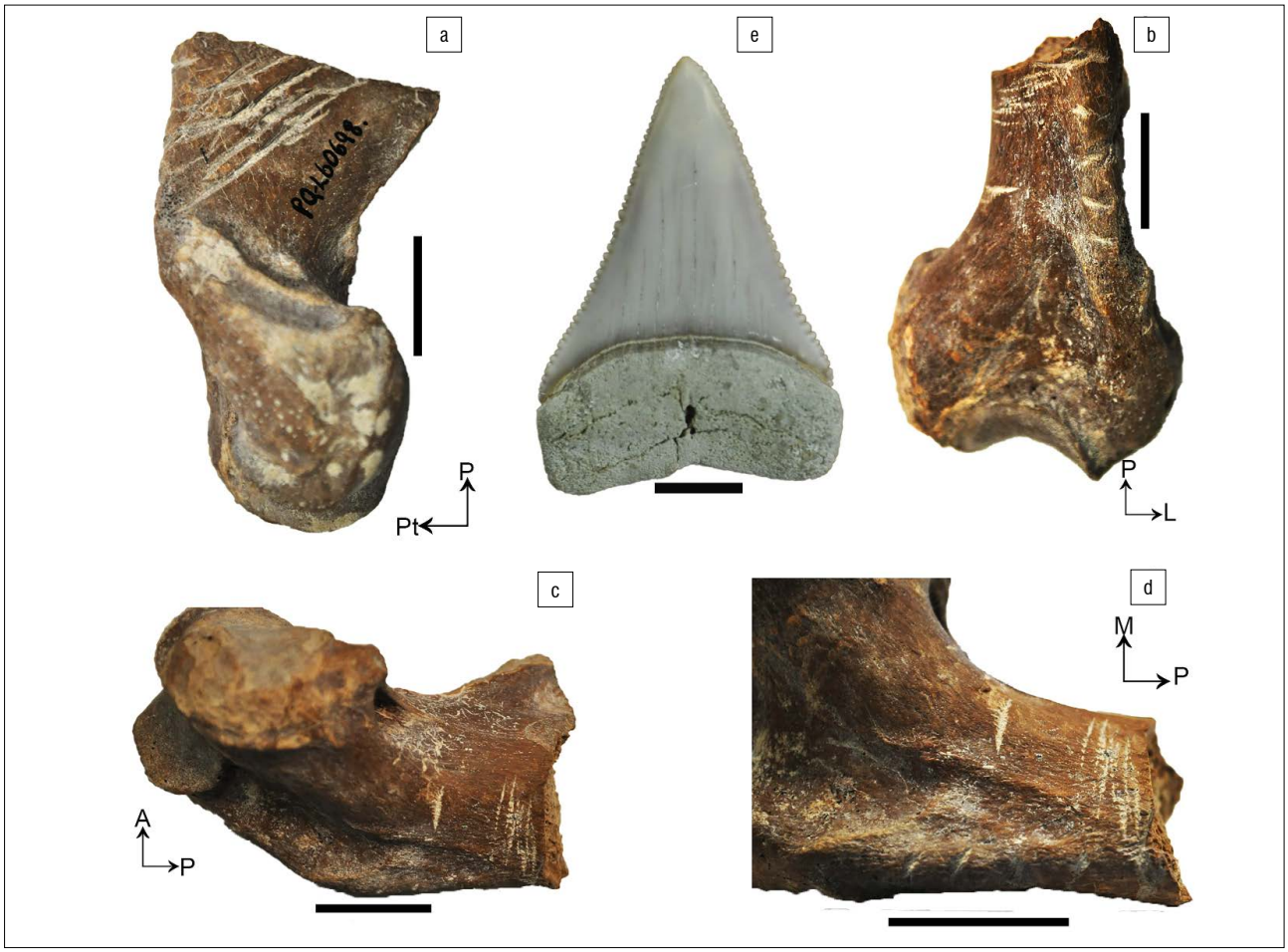
1. CF1 damage can be subdivided into two types (herein designated serrated Types CF1a and CF1b); CF1a is a simple, superficial groove with dotted markings left by serrations, whereas CF1b is a deeper groove with ridges and grooves caused by tooth serrations.
2. CF2 is a simple groove with tapered end and no trace of serrations.
3. CF3 damage has subparallel ridges and grooves corresponding with the tooth's serrated edge. There is no cut groove.
4. CF4 damage suggests cutting and/or scraping action with rotating movement. This action leaves curvilinear markings caused by the rotation of the tooth.



Scale = 2 cm

L, lateral; P, proximal

Figure 2: Right humerus SAM-PQL-34631. (a) Anterior view showing the proximal break. (b) Lateral view showing the periosteal reaction indicated by the arrow. (c) Arrow 1 showing the undulating CF3 and arrow 2 showing the CF3.



Scale = 2 cm (a–d); scale = 1 cm (e)

A, anterior; L, lateral; M, medial; P, proximal; Pt, posterior

Figure 3: Right humerus SAM-PQL-60698. (a) CF2 on the lateral surface showing the tooth moving over the surface a few times. (b) Posterior view showing laterally placed CF2. (c) CF1a and CF1b medially located on the posterior surface. (d) Close-up of CF1a and CF1b. (e) Fossil white shark tooth [SAM-PQSS-225 (SAM-PQY-24A)] from Saldanha Steel – a similar aged site west of Langebaanweg.

Description of bite marks

SAM-PQL-34631 shows periosteal reaction like that seen in SAM-PQL-30080[21], suggesting that this seal was suffering from an infection (Figure 2b arrow). The shark tooth marks are on the lateral surface of the bone; one is directed at an angle down the long axis of the bone and is proximally incomplete while the other is directed across the bone towards the lateral margin. There are groups of subparallel ridges and grooves with a wavy appearance (Figure 2c arrow 1). Another set of ridges and grooves (Figure 2c arrow 2) run to the lateral margin just above the distal articulation. This damage suggests that the serrated edge of the tooth scraped across the surface of the bone (CF3).

SAM-PQL-60698 has at least four bites along the lateral margin. The shark's tooth moved across the bone and left behind a simple groove with tapered ends (CF2; Figure 3a); the tooth moved over the same area a few times. Laterally, on the posterior surface, there are four CF2 bites, two bites leading to a bite on the lateral surface. From the middle of the bone to the medial border below the broken edge of the shaft, there are six bites caused by the shark biting down but not penetrating the bone causing only surface damage (CF1a; Figure 3b). Two bites – one towards the lateral surface and the other towards the medial surface – show evidence of the shark biting down and leaving grooves with ridges and grooves inside (CF1b, Figure 3c, d). The damage suggests that the tooth moved across the bite area multiple times (see Govender and Chinsamy's fig. 4E[8]).

Discussion

Extant sharks, including white (*Carcharodon carcharias*), tiger (*Galeocerdo cuvier*), Zambezi (bull) (*Carcharhinus leucas*), whale (*Rhincodon typus*), ragged tooth (sandtiger, *Carcharias tarsus*) and shortfin mako (*Isurus oxyrinchus*) sharks, inhabit varying habitats along South Africa's coast.[8] Sharks from 'E' Quarry include *Carcharias tarsus*, *Carcharodon carcharias* and *Isurus* sp.[22] A re-examination of the shark teeth identified probable *Carcharhinus* cf. *leucas* from Langebaanweg (personal observation 2013). Today, studies document extant adult white shark (*Carcharodon carcharias*) predation on South African fur seals on islands off the coast.[1-3] In the vicinity of Langebaanweg, 'E' Quarry, offshore islands and a protected lagoon that was open to the ocean were present 5 Ma.[11,18] The islands were surrounded by shallow water[23], making this ideal for haul out and rookeries. Shark tooth marks have been identified on two fragmentary humeri, which limits the discussion to hypothetical scenarios extrapolated from the damage to the bones.

Two potential scenarios are hypothesised from the damage and the lack of healing around the bites: an active hunted seal or a scavenged seal carcass. The islands and the lagoon supplied an environment that was ideal for carcasses to beach and refloat or float in quiet waters. It is also interesting that of the 3131 seal specimens, only 2 have any evidence of shark tooth marks (0.06%). This low incidence could be related to the fact that the coast was also frequented by cetaceans which were more likely the focus of shark feeding[8,9] or that there are still specimens to be found in the collection and through excavation.

The environment in the Langebaanweg area allowed carcasses to remain afloat and when beached may have been refloated by wave action[8]; these carcasses leaching fluid would have attracted sharks into the area to feed, as evidenced by the presence of white shark and mako shark damage on the cetacean bones.[8-10] This put the sharks in contact with the seals living on the islands. The lagoon and surrounding islands were linked like 'Shark Alley' today[24] where seal carcasses may have remained afloat. It could also be interpreted that sharks did not frequent the area until there was a cetacean carcass to feed on and were opportunistically scavenging seal carcasses when excluded from the cetacean carcasses.

White sharks ambush seals off Seal Island, biting prey obliquely using their anterolateral teeth in a lateral snap[25], and attack seals at the water surface using a steep vertical attack[26]. The bites on the humeri vary from superficial scrapes to deep bites penetrating the bone, suggesting that the shark and seal (carcass) were in motion. The presence of scrapes on SAM-PQL-34631 suggests that the shark tooth/teeth slid across the bone as the shark was biting. This type of damage could have resulted from the carcass or shark or both being in motion and it not being able to get a firm grip on the bone. There is evidence that this animal was also suffering from an infection which would potentially affect its ability to evade a predator or it may have succumbed to its illness.

SAM-PQL-60698 shows the presence of CF1a, CF1b and CF2 bites. The shark's teeth slid across the surface of the bone without penetrating the periosteum (CF1a) while it fed, as a result of the shark or the carcass or both being in motion.[8] This could be a result of the shark propelling itself forward to bite and then straight reverse.[27] The bites on SAM-PQL-60698 suggest that the shark bite was from the right side behind the seal's front flipper. The bite went through the muscles of the humerus, leaving evidence on the bone itself, or the carcass was partially skeletonised.

Sharks with serrated teeth left the bites seen on these seal bones (Figure 3e), which resembles the damage described in Govender and Chinsamy[8] on cetaceans from Langebaanweg caused by the white shark (*Carcharodon carcharias)*. The bite marks on the bone suggests that some skeletonisation had occurred, further supporting the hypothesis that white sharks were scavenging seal carcasses. Deposition at 'E' Quarry occurred as a result of terrestrial carnivore activity or storm surges.

Conclusion

This study is the first description of sharks interacting with seals in South Africa's geological past. At least one individual was suffering from an infection at the time of death. These carcasses remained afloat in the channels connecting the surrounding islands with the lagoon at Langebaanweg, making them available for sharks to feed on while they were also feeding on nearby whale carcasses. Based on the bone damage, the seals were carcasses that were scavenged by white sharks. Shark tooth marks are preserved on only two seal humeri. Simulations/impact models and statistical analyses are currently beyond the scope of this study but may prove beneficial in future if more specimens are collected during future excavations and/or identified from the collection.

Acknowledgements

I thank the Cenozoic collections staff, Sarena Govender (Assistant Collections Manager: Cenozoic Palaeontology), Angus Rayners (Technical Assistant: Cenozoic Palaeontology) and Mark de Benedictis (Technical Assistant: Cenozoic Palaeontology) at Iziko South African Museum, Cape Town, for their assistance. I also thank Dr Graham Avery for his encouragement, and the two anonymous reviewers whose comments greatly enhanced the manuscript. This research was funded by an NRF/African Origins Platform grant (grant numbers UID 98834 and UID117782).

Competing interests

There are no competing interests to declare.

References

1. De Vos A, O'Riain MJ, Meyer MA, Kotze PG, Kock AA. Behavior of Cape fur seals (*Arctocephalus pusillus pusillus*) in relation to temporal variation in predation risk by white sharks (*Carcharodon carcharias*) around a seal rookery in False Bay, South Africa. Mar Mamm Sci. 2015;31(3):1118–1131. https://doi.org/10.1111/mms.12208
2. Johnson RL, Keswick T, Bester MN, Oosthuizen WH. Encounters between white sharks and Cape fur seals in a shallow channel. Mar Biodiv Rec. 2009;1:2:1–5. https://doi.org/10.1017/s1755267209000682
3. Hammerschlag N, Martin RA, Fallows C. Effects of environmental conditions on predator–prey interactions between white sharks (*Carcharodon carcharias*) and Cape fur seals (*Arctocephalus pusillus pusillus*) at Seal Island, South Africa. Environ Biol Fish. 2006;76:341–35. https://doi.org/10.1007/s10641-006-9038-z
4. Diedrich CG. Evolution of white and megatooth sharks, and evidence for early predation on seals, sirenians, and whales. Natural Sci. 2013;5(11):1203–1218. http://dx.doi.org/10.4236/ns.2013.511148

5. Collareta A, Lambert O, Landini W, Di Celma C, Malinverno E, Varas-Malca R, et al. Did the giant extinct shark *Carcharocles megalodon* target small prey? Bite marks on marine mammal remains from the late Miocene of Peru. Palaeogeogr Palaeoclimatol Palaeoecol. 2017;469:84–91. http://dx.doi.org/10.1016/j.palaeo.2017.01.001

6. Bigelow PK. Occurrence of a squaloid shark (Chondrichthyes: Squaliformes) with the pinniped *Allodemus* from the Upper Miocene of Washington. J Paleont. 1994;68(3):680–684. https://doi.org/10.1017/s0022336000026032

7. Purdy R, Schneider VP, Applegate SP, McLellan JH, Meyer RL, Slaughter BH. The Neogene sharks, rays, and bony fishes from Lee Creek Mine, Aurora, North Carolina. In: Ray CE, Bohaska DJ, editors. Geology and paleontology of the Lee Creek Mine, North Carolina. Vol. III. Smithsonian Contribution to Paleobiology 90. Washington DC: Smithsonian Institution Press; 2001. p. 71–202.

8. Govender R, Chinsamy A. Early Pliocene (5 Ma) shark–cetacean trophic interaction from Langebaanweg, western coast of South Africa. Palaios. 2013;28(5):270–277. http://dx.doi.org/10.2110/palo.2012.p12-058r

9. Govender R. Shark-cetacean trophic interaction, Duinefontein, Koeberg, (5 Ma), South Africa. S Afr J Sci. 2015;111(11–12), Art. #2014-0453. http://dx.doi.org/10.17159/sajs.2015/20140453

10. Govender R. Early Pliocene fossil cetaceans from Hondeklip Bay, Namaqualand, South Africa. Hist Biol. 2019;15:1–20. https://doi.org/10.1080/08912963.2019.1650273

11. Hendey QB. Palaeoecology of the Late Tertiary fossil occurrences in 'E' Quarry, Langebaanweg, South Africa, and a reinterpretation of their geological context. Ann S Afr Mus. 1981;84(1):1–104.

12. Smith RMH, Haarhoff P. Sedimentology and taphonomy of an early Pliocene Sivathere bonebed at Langebaanweg, Western Cape Province, South Africa. Afr Nat Hist. 2006;2:197.

13. Roberts DL, Matthews T, Herries AI, Boulter C, Scott L, Dondo C, et al. Regional and global context of the Late Cenozoic Langebaanweg (LBW) palaeontological site: West Coast of South Africa. Earth-Sci Rev. 2011;106(3–4):191–214. https://doi.org/10.1016/j.earscirev.2011.02.002

14. Hendey QB. Langebaanweg: A record of past life. Cape Town: South African Museum; 1989.

15. Tankard AJ, Rogers J. Late Cenozoic palaeoenvironments on the west coast of southern Africa. J Biogeogr. 1978;1:319–337. https://doi.org/10.2307/3038026

16. Siesser WG, Dingle RV. Tertiary sea-level movements around Southern Africa. J Geol. 1981;89(4):523–536.

17. Siesser WG. Late Miocene origin of the Benguela upswelling system off Northern Namibia. Science. 1980;208(4441):283–285. https://doi.org/10.1126/science.208.4441.283

18. Kensley BF. Pliocene marine invertebrates from Langebaanweg, Cape Province. Ann S Afr Mus. 1972:173–190.

19. Erasmus L. Virtual reconstruction of stratigraphy and past landscapes in the West Coast Fossil Park Region [MSc thesis]. Stellenbosch: Stellenbosch University; 2005.

20. Hendey QB, Dingle RV. Onshore sedimentary phosphate deposits in south-western Africa. In: Notholt AJG, Sheldon RP, Davidson DF, editors. Phosphate deposits of the world. Vol. 2: Phosphate rock resources. Cambridge: Cambridge University Press; 1989. p. 200–206.

21. Govender R, Avery G, Chinsamy A. Pathologies in the early Pliocene phocid seals from Langebaanweg, South Africa. S Afr J Sci. 2011;107:72–77. https://doi.org/10.4102/sajs.v107i1/2.230

22. Tulu Y, Chinsamy-Turan A. Langebaanweg Quarry, Western Cape, South Africa: The Elasmobranch fauna and comparisons to faunas of PCS (Lee Creek) phosphate mine and Sharktooth Hill. J Vert Paleontol. 2011;31:207.

23. Rogers J. First report on the Cenozoic sediments between Cape Town and Elands Bay. Geological Survey of South Africa. 165:1–64.

24. Towner AV, Leos-Barajas V, Langrock R, Schick RS, Smale MJ, Kaschke T, et al. Sex-specific and individual preferences for hunting strategies in white sharks. Funct Ecol. 2016;30(8):1397–1407. https://doi.org/10.1111/1365-2435.12613

25. Martin RA, Hammerschlag N, Ralph OP, Collier S, Fallows C. Predatory behaviour of white sharks (*Carcharodon carcharias*) at Seal Island, South Africa. J Mar Biol Ass UK. 2005;85:1121–1135. https://doi.org/10.1017/s002531540501218x

26. Martin RA, Rossmo DK, Hammerschlag N. Hunting patterns and geographic profiling of white shark predation. J Zool. 2009;279:111–118. https://doi.org/10.1111/j.1469-998.2009.00586.x

27. Tricas TC. Feeding ethology of the white shark, *Carcharodon carcharias*. Mem Calif Acad Sci. 1985;9:81–91.

28. Govender R, Bisconti M, Chinsamy A. A late Miocene–early Pliocene baleen whale assemblage from Langebaanweg, west coast of South Africa (Mammalia, Cetacea, Mysticeti). Alcheringa. 2016;40(4):542–555.

AUTHORS:
Thomas N. Huffman[1]
Stephan Woodborne[1,2]

AFFILIATIONS:
[1]School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand, Johannesburg, South Africa
[2]iThemba LABS, Johannesburg, South Africa

CORRESPONDENCE TO:
Thomas Huffman

EMAIL:
thomas.huffman@wits.ac.za

DATES:
Received: 07 Oct. 2020
Revised: 09 Nov. 2020
Accepted: 09 Nov. 2020
Published: 29 Mar. 2021

HOW TO CITE:
Huffman TN, Woodborne S. New AMS dates for the Middle Iron Age in the Mapungubwe landscape. S Afr J Sci. 2021;117(3/4), Art. #8980. https://doi.org/10.17159/sajs.2021/8980

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

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

EDITORS:
Margaret Avery

KEYWORDS:
abandonment, droughts, ethnic interaction, Mapungubwe landscape, population dynamics

FUNDING:
Archaeological Resources Management, University of the Witwatersrand Research Incentive Scheme (HUFF013), South African National Research Foundation (85983)

New AMS dates for the Middle Iron Age in the Mapungubwe landscape

Research in the Limpopo Valley has documented over 500 Middle Iron Age sites (AD 900–1320) relevant to the origins of Mapungubwe – the capital of the first indigenous state in southern Africa. Fifteen new accelerator mass spectrometry (AMS) dates from 11 of these archaeological sites establish the boundaries of the ceramic facies that form the culture-history framework for such diverse topics as land use, ethnic stratification, population dynamics and rainfall fluctuations. Mapungubwe was abandoned at about AD 1320.

Significance:

- Because Mapungubwe developed relatively recently (circa AD 1200), it can clarify the origins of older states.
- Environmental factors such as droughts, along with agriculture and trade, played a role in the abandonment of Mapungubwe.

Mapungubwe was the capital of the first indigenous state in southern Africa, laying the foundations for Great Zimbabwe[1] (Figure 1). As with states elsewhere, external trade wealth and intensive agriculture were critical agents of change: they helped to transform a ranked-based society with hereditary leadership at the capital K2 into a class-based society with sacred leadership at Mapungubwe.

Archaeologically, Mapungubwe belongs to the Iron Age, a 1500-year long era dominated by Bantu-speaking farmers.[2] By convention, archaeologists divide this era into three arbitrary periods: the Early Iron Age (AD 300–900), the Middle Iron Age (AD 900–1300) and the Late Iron Age (AD 1300–1840). Characteristic ceramic facies form the basis of the culture-history sequence. Although problematic in terms of real cultural groups, it is another convention to apply the facies name to people who produced the style: thus, Mapungubwe people produced the *Mapungubwe* style.

For the origins of Mapungubwe, the most important period is the Middle Iron Age. Stratigraphic relationships for this period have helped to produce a definitive ceramic sequence (Figure 2): it includes the facies known as *Zhizo*, *Leokwe*, *K2*, *Transitional K2* (*TK2*) and *Mapungubwe*.

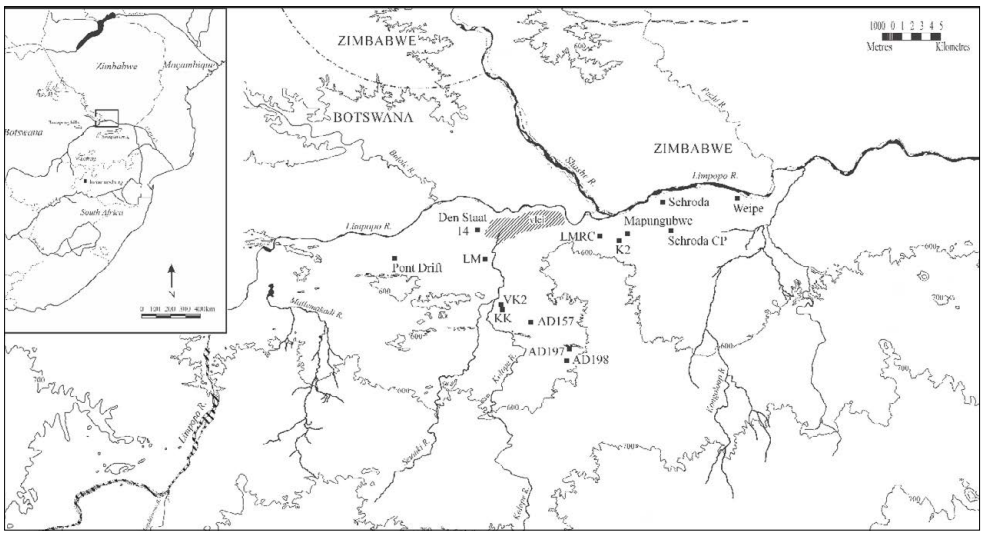


Figure 1: New accelerator mass spectrometry (AMS)-dated sites within the Mapungubwe landscape.

Origins of Mapungubwe Project

Since 1999, foot surveys in the Mapungubwe National Park and surrounding Buffer Zone have recorded some 1150 Iron Age sites. This large number has helped to clarify different land uses, ethnic stratification, population dynamics and droughts. As part of our project, we have processed 15 new accelerator mass spectrometry (AMS) dates from 11 Middle Iron Age sites and other researchers have produced a few more[3-5] (Table 1). We report them here by ceramic facies and research topic.

For Table 1, we first calibrated the BP (Before Present) dates using Calib 8.10 and the Southern Hemisphere data set (SHCal20) using Stuiver and Reimer[6] and Hogg et al.[7] This calibration programme includes the median age for the radiocarbon date, but this often falls outside the 1-sigma range.

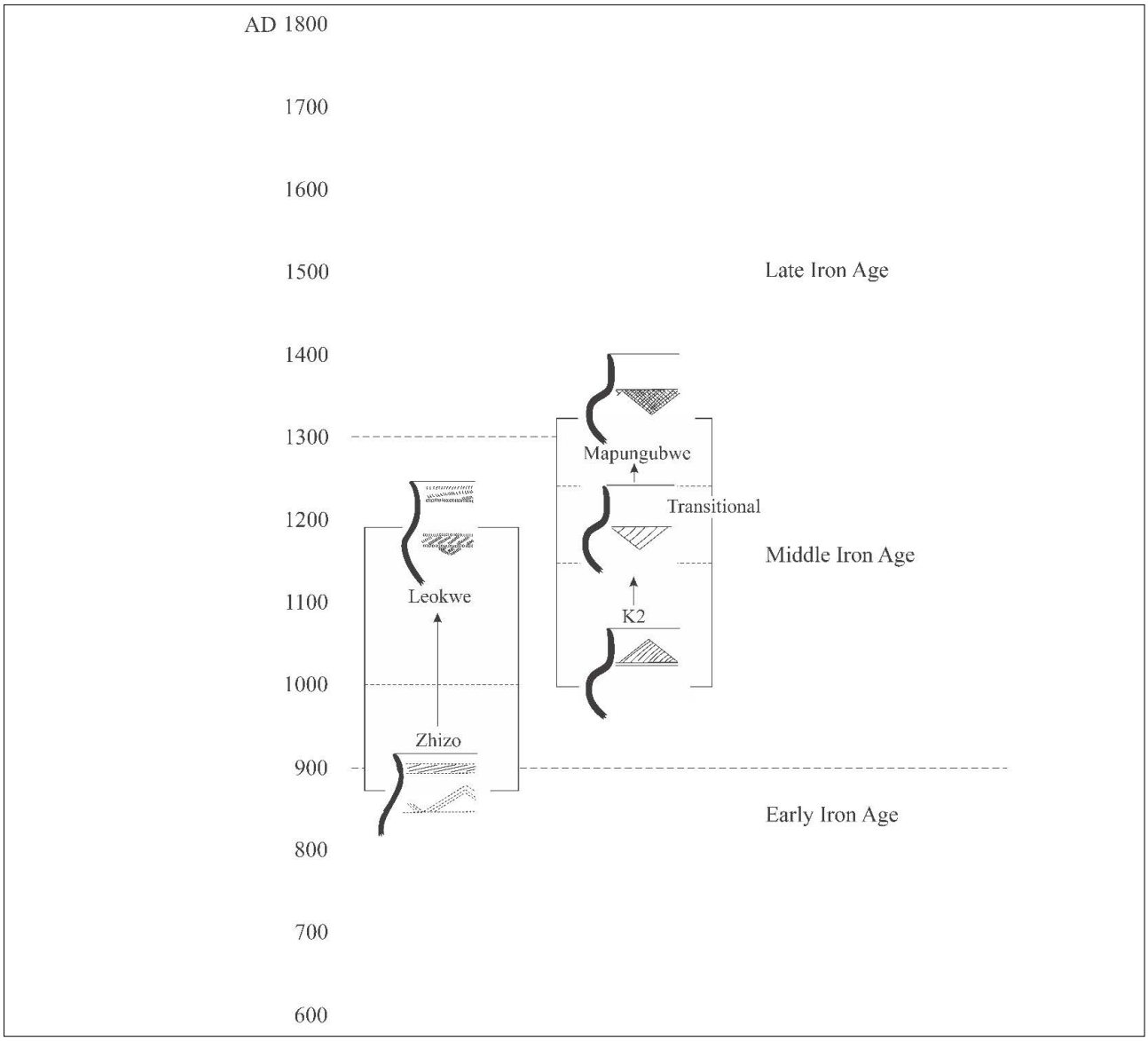


Figure 2: Ceramic sequence for the Middle Iron Age in the Mapungubwe landscape.

Furthermore, the radiocarbon curve fluctuates markedly during the Middle Iron Age, so that one radiocarbon result may have two or more possible calendar dates. To help choose between different calibration spans, we consider the midpoints of the 1-sigma ranges along with the known stratigraphic sequence and then order the possibilities. Thus, a hypothetical date of BP 1000±1 calibrates to AD 1033–1048 for sites with *K2* ceramics but to AD 1120–1137 for the *TK2* facies.

The Middle Iron Age

According to isotopic analysis, when Zhizo people moved into the Limpopo Valley from southwest Zimbabwe at about AD 900, the climate was similar to that today.[8] This means that Zhizo people would have found farming difficult, and some other reason probably accounts for their presence. Ivory artefacts and imported glass beads at Schroda[9,10] and the locations of other Zhizo sites[1] indicate that these people may have purposefully moved into the basin to hunt elephants for the ivory trade.

Land use

At about AD 1000, or slightly later, Leopard's Kopje people established their capital at the site K2 near the Shashe-Limpopo confluence in South Africa.[11] In contrast to the earlier Zhizo phase, Leopard's Kopje people began to cultivate the margins of the large vlei there (Figure 1). Models of vlei and riverbed cultivation in Zimbabwe[12] suggest that they planted sorghums in the rich loams along the wet edges and millets on the sandy fluvial terraces.[13,14] In typical farming homesteads, many grainbin foundations encircle a central cattle kraal (e.g. Liz 197: IT-C-2042; Edmondsberg 157: IT-C-2047). Besides these homesteads, some settlements were cattle posts located on spurs near springs on the escarpment, or otherwise well away from agricultural land (e.g. on Schroda: IT-C-2041). In addition, field camps were located near agricultural land but on small hills and rises in situations unsuitable for settled villages: they have granaries, small stock kraals and middens, but lack permanent housing and cattle kraals. Rainmaking hills are a fourth kind of site.[5,15]

Ethnic interaction

When K2 people took over the valley, many Zhizo people went west to Botswana to become the Toutswe group.[16] Some Zhizo people, however, stayed behind to live within the K2 interaction sphere.[17,18] Because their ceramic style has changed somewhat, it is called *Leokwe* after the hill where it was first recognised. A Leokwe site in the Venetia Reserve, KK110, upstream of the vlei, has been AMS dated (IT-C-2038) to the 10th century, somewhat earlier than most dates. Antonites[3] added three new dates for the Leokwe levels at Schroda, one of which is also early.

Table 1: New accelerator mass spectrometry (AMS) dates for the Mapungubwe landscape

Site	Lab. no.	$\delta^{13}C$	BP ± 1σ	SHCal20 1σ	Means
Mapungubwe facies					
Weipe508 VII/C$_1$/4	IT-C-785 charcoal	-24.0	660±52	1302–1362, ~~1380–1399~~	1332
Transitional K2 facies					
VK2	IT-C-730 post	-23.5	740±68	1235–1242, ~~1270–1322, 1350–1388~~	1239
–	IT-C-1498 post	-23.5	990±44	~~1032–1053, 1060–1070~~,1080–1150	1135
Den Staat 14B	IT-C-2040 charcoal	-23.0	1050±34	~~993–1009, 1014–1046, 1088–1105~~, 1123–1130	1127
Den Staat 14C	IT-C-671 wild seeds	-21.5	950±52	~~1046–1087~~, ~~1106–1122~~, 1131–1186, 1196–1209	1159, 1203
Liz 198 small kraal	IT-C-2037 dung	-13.8	930±27	~~1055–1058, 1072–1079,~~1152–1189, 1192–1211	1171, 1202
-IX/gb	IT-C-1500 charcoal	-26.4	960±42.6	~~1046–1088~~, ~~1105–1123,~~ 1130–1181	1155
-IV/midden	IT-C-2034 sorghum	-10.0	970±30	~~1045–1089,~~ ~~1104–1124~~, 1129–1157	1143
-IX/S/3	IT-C-2033 charcoal	-24.9	1010±28	~~1027–1048,~~ 1085–1138	1112
Little Muck 138	IT-C-2039 post	-23.7	960±27	~~1047–1086, 1109–1120,~~1134–1161, 1168–1180	1148, 1174
K2 facies					
Liz 197	IT-C-2042 dung	-12.7	1020±27	1024–1028, 1085–1112, ~~1117–1137~~	1036, 1099
VAD13/157	IT-C-2047 dung	-13.3	1090±27	991–1021	1016
Schroda CP	IT-C-2041 dung	-11.3	940±27	1051–1063, 1067–1082,~~1148–1186,~~ ~~1198–1208~~	1057, 1075
Leokwe facies					
LMRC II/M/4-5	IT-C-733 charcoal	-26.7	960±50	1045–1089, 1097–1100, ~~1102–1124~~,- ~~1129–1182~~	1067, 1099
KK II/H/3	IT-C-2038 charcoal	-24.4	1130±28	~~900–925~~, ~~967–993~~,-1008–1014	1011

Note: Crossed out dates were eliminated for stratigraphic or other archaeological reasons.

These dates show that *Zhizo* ceramics began to change into *Leokwe* when Leopard's Kopje people first moved into the valley. This contact represents the first 'ethnic stratification'[19] during the Iron Age in southern Africa. The new dates and ceramic analyses show that this relative status started at the beginning of contact, contra some interpretations.[17]

Although under the political authority of K2, Leokwe people maintained their own material-culture signature for several decades. It is common in such situations for earlier people to assume ritual roles: this gives them respect but not political power.[20] Among other tasks, Leokwe people probably supervised the initiation school[21] at Schroda.

Besides ritual specialists, Leokwe people appear to have herded cattle for K2 elite, as several Leokwe settlements have 'extra' kraals.[22] A large Leokwe complex (2229AB223/224) inside the National Park yielded a mid- to late-11th century date from the main midden (IT-C-733), placing it in the mid-K2 phase.

Universally, states tend to subsume ethnic differences in favour of a national identity. In this regard, a few *Leokwe* vessels occur in K2 and TK2 sites (presumably through marriage alliances), but not in Mapungubwe. A national identity thus appears to have replaced ethnic differences by the early-13th century when large-scale centralised authority was established, but before sacred leadership had fully materialised.

Population dynamics

As the state grew, so did populations. For population estimates, we need accurate spans for each facies. Until now, the boundary between *K2* and *TK2* has been unclear. Carbonised seeds from Den Staat 14C (IT-C-671), along with dates from Liz198 (IT-C-1500, IT-C-2033, IT-C-2034 and IT-C-2037), a burnt hut on Little Muck (IT-C-2039) and a burnt granary at VK2 (IT-C-1498) together show that *K2* ceramics transformed into *TK2* around AD 1150. *TK2* in turn became *Mapungubwe* about 120 years later, while *Mapungubwe* pottery lasted for about 50 years.

Using these new time spans, we assign 50 people (half of them adults) to each homestead, based on the Middle Iron Age burials at Kgaswe[16] in Botswana. We then assign 50 years duration to each homestead and divide the time span of each facies, and population, by the number of 50-year units. Thus, if 7650 K2 people (153 x 50) lived in the valley, then 2550 people (7650 ÷ 3) lived there at any one time. This formula determines general populations in relative terms rather than as an absolute census. In addition to ordinary people, total populations need to include the capitals (Schroda, K2 and Mapungubwe). It is likely, however, that K2 started as a small capital before reaching its maximum extent. We thus present the first 25 years of its lifespan as one half of a 50-year unit (2550 ÷ 2 = 1275 people) and then calculate the remaining population (7650 – 1275 = 6375 people) and duration minus that amount (150 – 25 = 125 years). Whatever formula is used, the K2 population was larger than Zhizo and over 10 000 people lived in the valley during the Mapungubwe phase (Table 2).

Table 2: Population dynamics for the South African portion of the Shashe-Limpopo valley

Phase	Homesteads	Time span	General population	Capital	Total
Mapungubwe AD 1270–1320	114	50	5700	5000	10 700
Transitional AD 1150–1270	143	120	2979	2500	5479
K2 AD 1000–1150	153	125 25	2550 1275	1500 300	4050 1575
Leokwe AD 1000–1200	63	200	787	none	787
Zhizo AD 900–1000	22	100	550	300	850

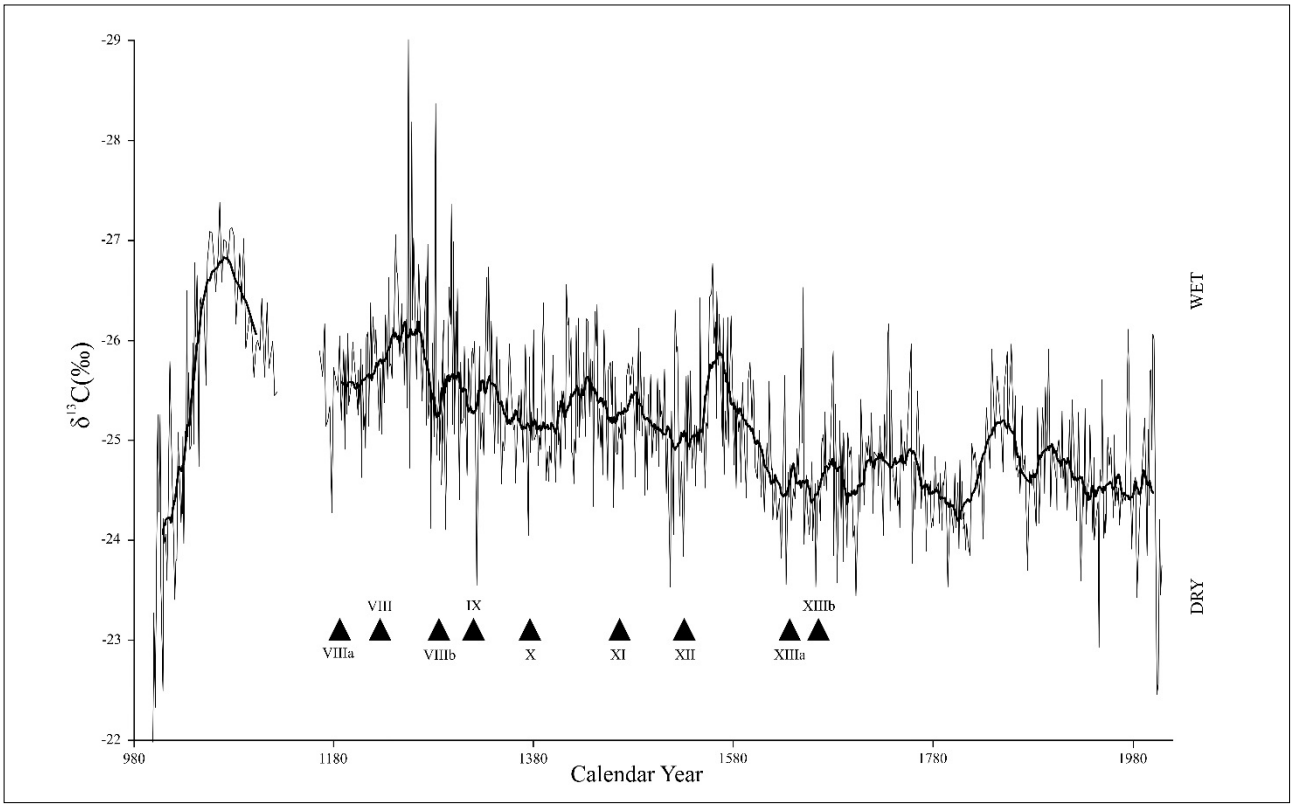
Number of sites ÷ 50 years x 50 people

Droughts and abandonment

Rainfall affected political stability as well as agricultural production. We know that some farmers burnt their grainbins as a ritual of cleansing related to severe drought (3–5 years in a row).[23] We first used the traditional radiocarbon method to date the burnings and droughts but fluctuations in the calibration curve confounded the results. We later added detailed baobab data (based on the isotopic component of successive growth rings)[24,25] that eliminate the multiple choices in the calibration curves. These data reveal a few droughts not previously noted (Table 3). One drought (Group IX) in particular contributed to the abandonment of Mapungubwe. The baobab sequence dates this important episode somewhat later than expected, to about AD 1310±5.

Following the principles of sacred leadership, the leader's right to rule would have been questioned as a result of this drought.[12] With Mapungubwe leadership in turmoil, Great Zimbabwe was able to seize control of the gold belt, the most important source of trade wealth, and Great Zimbabwe became the new centre of power. Thus, the 14th-century drought was an indirect cause of Mapungubwe's abandonment.

Table 3: Baobab climatic sequence and severe droughts recognised in the archaeological record



Group	Cal AD	Baobab cal AD
XIIIb		1660±5
XIIIa	1650	1635±5
XII	1530	1530±10
XI	1440–1450	1465±5
X	1350–1400	1390±10
IX	1300	1310±5
VIIIb		1285±5
VIII	1200–1250	1208, 1226, 1256
VIIIa		1185±10
VII	1020–1070	
VI	900–1000 (Two episodes)	

Dates from Weipe508 (IT-C-785, BP 650±52; and Pta-9549, BP 630±70) show that Mapungubwe people remained in the valley until about AD 1320 – the same date as the drought.

Acknowledgements

We thank John Calabrese, Kerry Fatherley, Myra Mashimbye, McEdward Murimbika, Justine du Piesanie, Alex Schoeman, Jeannette Smith, Peter Venter, Jaco van der Walt and Marilee Wood for their help with the surveys, excavations and specialist studies. The University of the Witwatersrand Honours classes of 2000 to 2009 also participated in the research. Wendy Voorvelt prepared the illustrations and Moshabi Silidi and Stephan Winkler processed the AMS dates. Fieldwork was sponsored by Archaeological Resources Management and the University of the Witwatersrand Research Incentive Scheme (HUFF013) and the South African National Research Foundation (85983).

Competing interests

We have no competing interests to declare.

Authors' contributions

T.N.H. directed the field work and S.W. the AMS dating. Both authors prepared the manuscript.

References

1. Huffman TN. Mapungubwe and the origins of the Zimbabwe culture. In Leslie M, Maggs T, editors. African Naissance: The Limpopo Valley 1 000 years ago. S Afr Archaeol Soc Goodwin Series 8. 2000. p. 14–29. https://doi.org/10.2307/3858043
2. Huffman TN. Handbook to the Iron Age: The archaeology of pre-colonial farming societies in southern Africa. Pietermaritzburg: University of KwaZulu-Natal Press; 2007.
3. Antonites AR. A revised chronology for the Zhizo and Leokwe horizons at Schroda. South Afr Humanities. 2018;31:223–246.
4. Manyanga M. Resilient landscapes: Socio-environmental dynamics in the Shashi-Limpopo Basin, Southern Zimbabwe c. AD 800 to the present. Studies in Global Archaeology 11. Uppsala: Department of Archaeology and Ancient History; 2007.
5. Schoeman MH. Excavating the 'waterpits in the mountain': The archaeology of Shashe-Limpopo Confluence Area rain-hill rock tanks. South Afr Humanities. 2009;21:275–298.
6. Stuiver M, Reimer PJ. Extended ^{14}C radiocarbon data base and revised CALIB 3.0 ^{14}C age calibration program. Radiocarbon. 1993;35(1):215–230. https://doi.org/10.1017/S0033822200013904
7. Hogg AC, Hua Q, Palmer JG, Turney CSM, Southon J, Bayliss A, et al. SHCal20 Southern hemisphere calibration, 0–55,000 years cal BP. Radiocarbon. 2020;62(4):759–778. https://doi.org/10.1017/RDC.2020.59
8. Smith J, Lee-Thorp J, Hall S. Climate change and agropastoralist settlement in the Shashe-Limpopo River Basin, southern Africa: AD 880 to 1700. S Afr Archaeol Bull. 2007;62:115–125.
9. Voigt EA. Mapungubwe: An archaeozoological interpretation of an Iron Age community. Museum Monograph 1. Pretoria: Transvaal Museum; 1983.
10. Wood M. A glass bead sequence for southern Africa from the 8th to the 16th century AD. J Afr Archaeol. 2011;9(1):67–84. https://doi.org/10.3213/1612-1651-10184
11. Meyer A. The archaeological sites of Greefswald: Stratigraphy and chronology of the sites and a history of investigations. Pretoria: University of Pretoria; 1998.
12. Murimbika MT. Sacred powers and rituals of transformation: An ethnoarchaeological study of rainmaking rituals and agricultural productivity during the evolution of the Mapungubwe state, AD 1000 to AD 1300 [dissertation]. Johannesburg: University of the Witwatersrand; 2006.
13. Huffman TN. Test excavations at Liz 197: A K2-period homestead in the Limpopo Valley. South Afr Humanities. 2016;28:135–152.
14. Huffman TN. Origins of Mapungubwe Project: Test excavations at Den Staat 14B and 14C. South Afr Humanities. 2017;30:185–194.
15. Schoeman MH. Imagining rain-places: rain-control and changing ritual landscapes in the Shashe-Limpopo Confluence Area, South Africa. S Afr Archaeol Bull. 2006;61:152–165. https://doi.org/10.2307/20474923
16. Denbow JR. A new look at the later prehistory of the Kalahari. J Afr Hist. 1986;27:3–29. https://doi.org/10.1017/S0021853700029170
17. Calabrese JA. The emergence of social and political complexity in the Shashi-Limpopo Valley of southern Africa, AD 900 to 1300: Ethnicity, class, and polity. BAR International Series 1617. Cambridge Monographs in African Archaeology 69. Oxford: Archaeopress; 2007. https://doi.org/10.30861/9781407300290
18. Huffman TN. Leokwe and K2: Ethnic stratification during the Middle Iron Age in Southern Africa. J Afr Archaeol. 2007;5(2):3–27. https://doi.org/10.3213/1612-1651-10091
19. Hammond-Tooke WD. Ethnicity and ethnic group in Iron Age southern Africa. S Afr J Sci. 2000;96(8):421–422.
20. Kopytoff I. The African frontier: The reproduction of traditional African societies. Bloomington, IN: Indiana University Press; 1989.
21. Van Schalkwyk JF, Hanisch EOM, editors. Sculptured in clay: Iron Age figurines from Schroda, Limpopo Province, South Africa. Pretoria: National Culture History Museum; 2002. p. 21–39.
22. Huffman TN. Salvage excavations on Greefswald: Leokwe commoners and K2 cattle. South Afr Humanities. 2014;26:101–128.
23. Huffman TN. A cultural proxy for drought: Ritual burning in the Iron Age of southern Africa. J Archaeol Sci. 2009;36:991–1005. https://doi.org/10.1016/j.jas.2008.11.026
24. Woodborne S, Hall G, Robertson I, Patrut A, Rouault M, Loader N, et al. 1000-year carbon isotope rainfall proxy record from South Africa baobab trees (*Adansonia diagata* L). PLoS ONE. 2015;10(5), e0124202. https://doi.org/10.1371/journal.pone.0124202
25. Huffman TN, Woodborne S. Archaeology, baobabs and drought: Cultural proxies and environmental data from the Mapungubwe landscape. Holocene. 2016;26(3):464–470. https://doi.org/10.1177/0959683615609753

AUTHORS:
Ivan Turok[1,2]
Justin Visagie[3]

AFFILIATIONS:
[1]Human Sciences Research Council, Cape Town, South Africa
[2]DSI/NRF Research Chair in City-Region Economies, Department of Economics and Finance, University of the Free State, Bloemfontein, South Africa
[3]Human Sciences Research Council, Durban, South Africa

CORRESPONDENCE TO:
Ivan Turok

EMAIL:
iturok@hsrc.ac.za

DATES:
Received: 24 Sep. 2020
Revised: 08 Dec. 2020
Accepted: 09 Dec. 2020
Published: 29 Mar. 2021

HOW TO CITE:
Turok I, Visagie J. COVID-19 amplifies urban inequalities. S Afr J Sci. 2021;117(3/4), Art. #8939. https://doi.org/10.17159/sajs.2021/8939

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:
coronavirus pandemic, spatial inequalities, informal settlements, social grants, place-based policies

FUNDING:
NIDS-Cram

COVID-19 amplifies urban inequalities

COVID-19 has had asymmetrical spatial impacts across South Africa. New evidence from the National Income Dynamics Study: Coronavirus Rapid Mobile (NIDS-CRAM) survey shows that the pandemic and lockdown reflex have magnified pre-existing divisions within cities. Although COVID-19 has severely impacted the whole country, townships and informal settlements have proved more vulnerable than suburbs. As South Africa was already one of the most unevenly developed countries in the world, COVID-19 has widened the gap between places, which face very different levels of risk and resilience.

Significance:

- We present original evidence that COVID-19 has affected poor urban communities more than it has suburbs in South Africa. This is apparent in terms of employment and hunger. The effect has been to magnify territorial divisions and exacerbate social discontent. Premature withdrawal of government relief will aggravate the hardships facing poor communities that rely on these resources following the slump in jobs.

Introduction

South Africa introduced one of the earliest and strictest lockdowns in the world in an effort to contain the coronavirus pandemic and to prepare the health-care system for the anticipated upsurge in patients needing treatment. This approach was driven by fear that the population was particularly susceptible to the disease and a desire to minimise the loss of life. However, the lockdown reflex shuttered much of the economy (production, consumption and distribution), with unintended socio-economic consequences.

While many other governments introduced exceptional support programmes for businesses and households to mitigate the damage caused by the restrictions on activity, South Africa's response was limited by the poor state of public finances.[1] The devastating effects are gradually becoming apparent. Surveys suggest that between 2 million and 3 million people lost their jobs between February and April 2020.[2] Other data show a 16.4% contraction in GDP between April and June 2020.[3] In September, the OECD announced that it expected South Africa's economy to shrink 11.5% in 2020.[1] This decline is the biggest amongst the 19 countries that feature in the Organisation for Economic Co-operation and Development's (OECD) analysis.

The geography of South Africa's twin public health and economic crises has received little attention to date, given the focus of the pandemic analysis and response at the national and provincial levels. The COVID-19 lockdown regulations and special support programmes have been uniform and 'place-blind' in the interests of simplicity and fairness. Yet South Africa is one of the most unevenly developed countries in the world, with stark contrasts in the risks and resilience of communities living in different places.[4,5] Casual observation suggests that some urban communities have been buffeted more than others. Here we present original evidence of the unequal impact of COVID-19 on livelihoods and well-being across different types of urban area.

Data and methods

The evidence comes from Waves 1 and 2 of the ongoing National Income Dynamics Study: Coronavirus Rapid Mobile Survey (NIDS-CRAM). The NIDS-CRAM was designed as a 'barometer' to assess the socio-economic impact of COVID-19.[2,6] It is based on a sample of adults who were previously surveyed in Wave 5 of the NIDS in 2017. Hence, the NIDS-CRAM provides another two rounds of socio-economic data for a subsample of 7073 adults from NIDS Wave 5 who were re-interviewed in May/June 2020 (NIDS-CRAM: Wave 1) and 5676 adults who were interviewed again in July/August (NIDS-CRAM: Wave 2). The first wave allows an assessment of the initial shock of COVID-19 and the second wave provides insights into the subsequent trajectory, including any signs of recovery.

The locational typology used for the analysis focuses on the differences within cities between four different kinds of neighbourhood. The classification is intuitive and draws on urban residents' own perceptions to distinguish between suburbs, townships, shack dwellers (informal settlements and backyarders) and peri-urban areas (which include smallholdings, farms or tribal land on the urban fringe). NIDS-CRAM is a telephonic survey in which respondents are asked about the kind of neighbourhood in which they live. In addition, information on the quality of housing was used to identify backyarders. The sample was restricted to individuals living in urban areas. The rationale is that one would expect the suburbs to cope better with the lockdown and social distancing protocols because residents tend to have secure jobs, more savings to rely on, and find it easier to work from home. In contrast, people living in townships, informal settlements and peri-urban areas are likely to have more precarious livelihoods, fewer resources to withstand shocks, and their neighbourhoods are likely to have weaker social infrastructure and safety nets.

We describe the impact of COVID-19 on these different areas. Doing so is important to improve understanding of the distinctive challenges facing different places and for more targeted, place-based responses, including consultation with local communities. We reveal that a blanket, nationwide approach that treats places equally does not diminish the gaps between them. The analysis is novel because initial studies of the impact of the crisis have focused on the attributes of individuals (race, gender, education, occupation, earnings, etc.)[2] and paid little attention to spatial considerations. The analysis does not control for the possibility that some people moved between locations – either

temporarily or permanently – as part of their livelihood strategies. A study based on Wave 1 of the NIDS-CRAM[7] estimated that about 15% of adults moved to a different household during the first few months of the lockdown, some of whom would have been in a different location. It is not technically possible to pinpoint the origins and destinations of these movers. The findings presented here are based on where everyone in the survey was living during Wave 2.

The analysis focuses on three particular impacts of COVID-19: the labour market, household incomes and food security. These are clearly linked, with causation running from the labour market to household incomes and onto hunger. Changes in employment (job loss) result in a loss of earnings for households, which affects whether people go hungry. The severity of the shock has been moderated by social grants from the government. A special COVID-19 social relief of distress grant of ZAR350 per month was introduced in June 2020, along with various top-ups to existing grants and temporary relief for workers made unemployed.

Asymmetrical urban impacts

Labour market effects

The first and most important finding from the survey is that the pandemic has magnified pre-existing economic disparities between suburbs, townships and informal settlements within cities. The suburbs proved more resilient to the lockdown than other types of urban settlement. Suburbs started out in February in a much stronger position with 58% of adults in paid employment, compared with 51% in the townships, 45% in peri-urban areas and 59% among shack dwellers (Figure 1). The latter reflects the high level of informal enterprise among shack dwellers.

After the level 5 lockdown was imposed, the suburbs lost one in seven jobs (14%) by April, compared with one in four in the townships (24%) and peri-urban areas (23%) and more than a third of jobs (36%) in shack areas. Shack dwellers were extremely vulnerable to the restrictions on informal trading during the shutdown, demonstrating the precarious nature of their livelihoods.

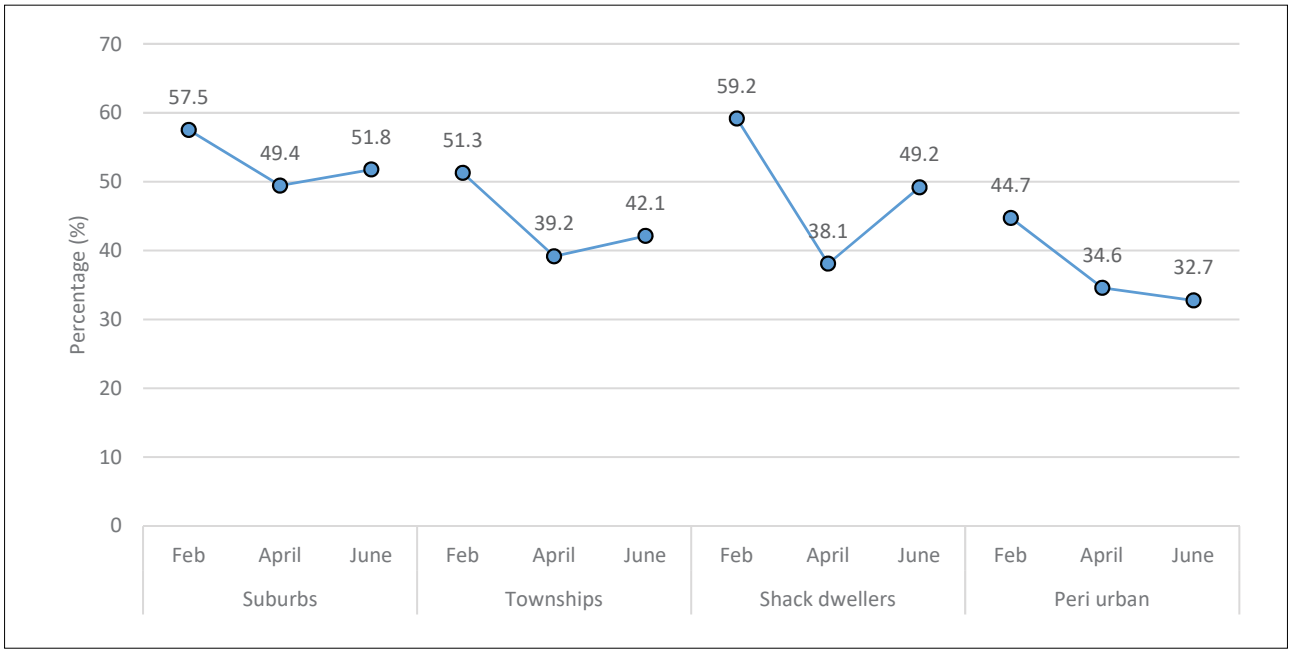
The hard lockdown was eased to level 3 in May. Between April and June 2020, the suburbs showed slight signs of recovery, with approximately 5% of workers going back to paid employment. There was a similar bounce-back in the townships, but no sign of recovery in the peri-urban areas. Meanwhile, approximately half of the shack dwellers who lost their livelihoods were also able to resume their activities, presumably because the costs of restarting were limited.

The net result was that by June 2020, the economic slump had hit poor urban communities harder than it had the suburbs. This is most apparent from the divergent rates of unemployment between the neighbourhood types (Figure 2). The 2017 NIDS Wave 5 survey provides a useful baseline for purposes of comparison. There was only 12 percentage points difference in the unemployment rate between the different locations in 2017. However, by April 2020, the gap had widened to 20 percentage points. Every location suffered a sharp rise in unemployment, but particularly the peri-urban and shack areas. By June 2020, the unemployment differential had widened further to 27 percentage points. The suburbs and shack areas showed signs of bouncing back, but the townships and peri-urban areas did not. Consequently, the positions of the three poorer neighbourhood types were far worse in June 2020 relative to the suburbs than they were before COVID-19 struck.

Social relief

The second finding relates to the provision of social support. Peri-urban communities have been much bigger beneficiaries of government grants than have suburban residents. More than half of peri-urban respondents (54%) lived in households that received social grants in June 2020, compared with less than half of township residents (45%), two in five shack dwellers (40%) and one in four suburban residents (26%). This was because peri-urban residents were far less likely to be in paid employment.

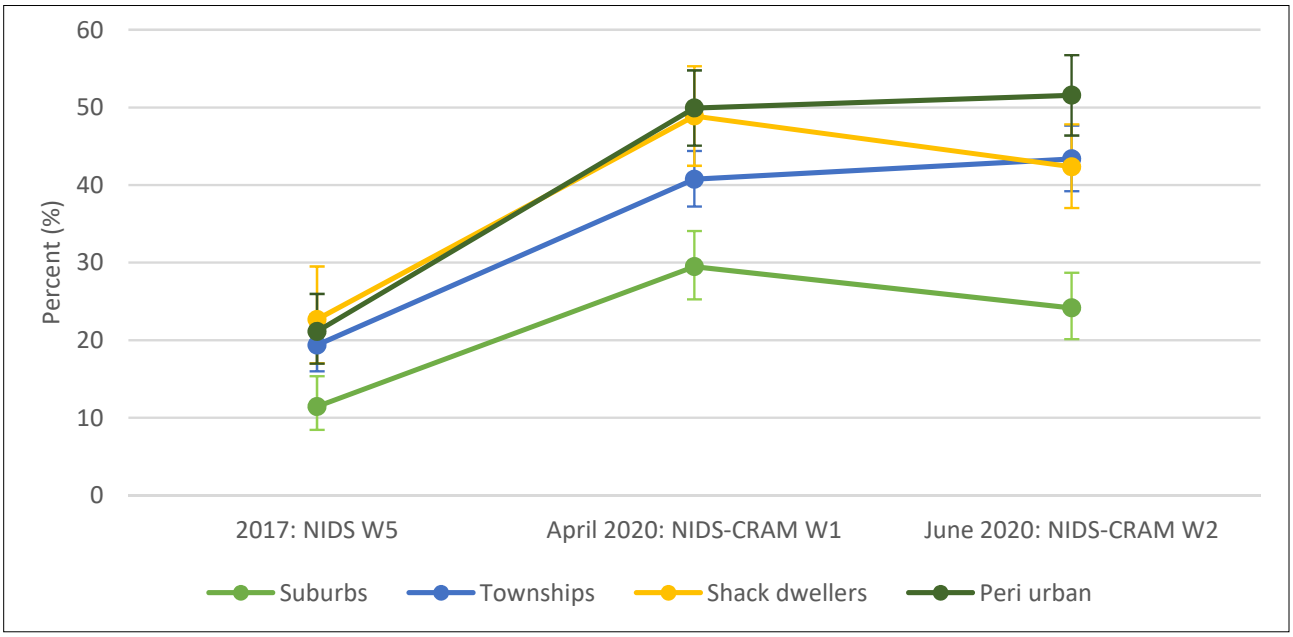
In terms of special relief from the crisis, less than one in three peri-urban residents (29%) said that someone in their household had received the COVID-19 grant, compared with 27% in townships, 18% of shack dwellers and 16% in suburban areas. These differences are smaller than for other grants, suggesting that the COVID-19 grant is benefiting people who did not qualify for government support before, such as unemployed men. The proportion of shack dwellers receiving these and other grants is surprisingly low considering their levels of poverty and distress. Further research is required to explain this finding.



Sources: NIDS-CRAM W1[2] and W2[6]

Notes: The sample comprises adults aged 18 years and older. Workers who had a job but reported zero earnings were not counted as employed. Self-identified neighbourhood type is defined in W2. The data are weighted.

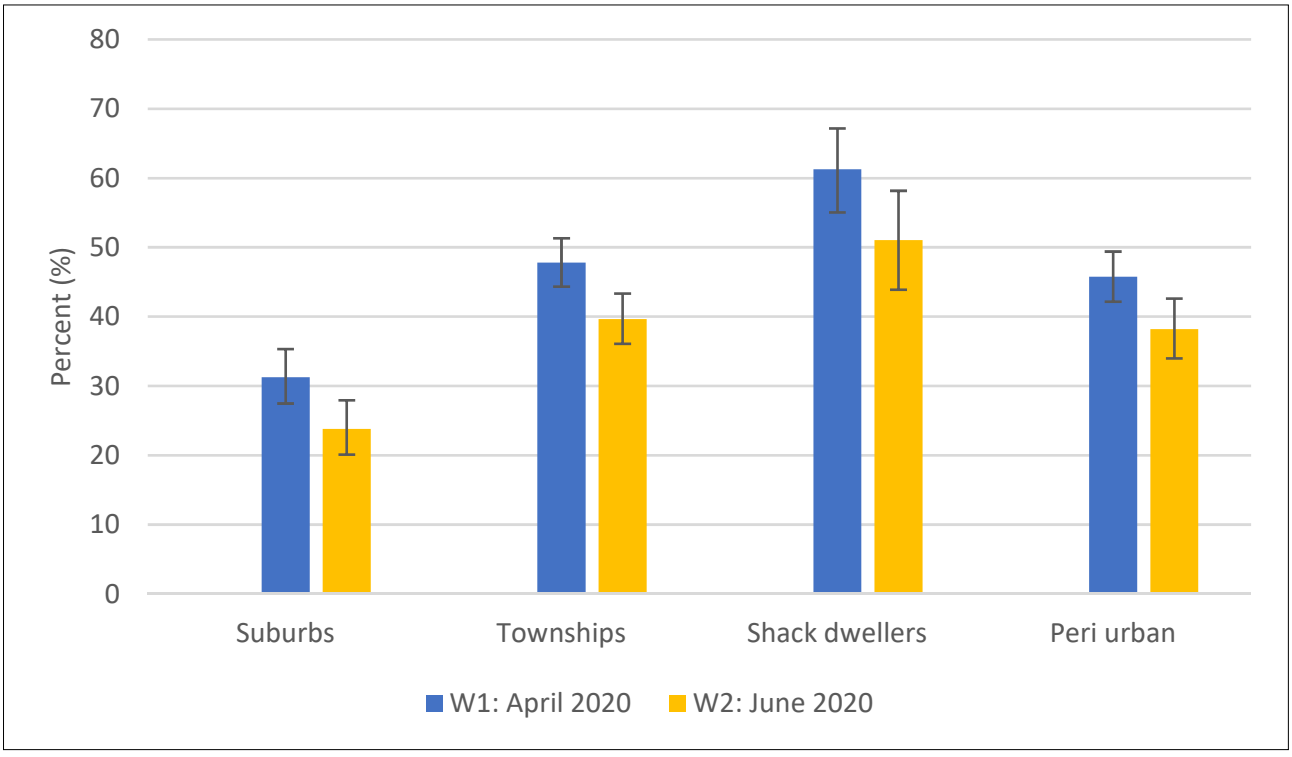
Figure 1: Percentage of respondents (adults 18 years and older) in paid employment, 2020.



Sources: NIDS W5[8], NIDS-CRAM W1[2] and W2[6].

Notes: Expanded rate of unemployment (i.e. includes the non-searching unemployed). The sample comprises adults aged 18 years and older. Error bars are 90% confidence intervals. The data are weighted.

Figure 2: Rate of unemployment, 2017 to June 2020.



Sources: NIDS-CRAM W1[2] and W2[6].

Notes: Error bars are 90% confidence intervals. The data are weighted.

Figure 3: Percentage of adults who reported that their household ran out of money to buy food.

Government grants have clearly helped to protect livelihoods in poor communities and compensate for high unemployment rates. However, there is a corresponding risk to living standards if the temporary relief is withdrawn before the labour market has recovered. Such a scenario would aggravate human suffering and misery.

Food security

The third finding relates to hunger. The proportion of respondents who said their household had run out of money to buy food in April 2020 was 31% in the suburbs, 48% in the townships and 61% in the shack areas (Figure 3). Shack-dwellers were noticeably worse off than other urban respondents. This adds to the concern noted above that far fewer shack-dwellers receive government grants. By June 2020, these proportions had come down to 24% in the suburbs, 40% in the townships and 50% in the shack areas. In other words, hunger had declined everywhere, although the gap between the shack-dwellers and other groups was still large.

The proportion of respondents who said that someone in their household had gone hungry in the last seven days (in May/June 2020) was 11% in the suburbs, 22% in the townships and 32% in the shack areas. By July/August 2020, these proportions had come down to 7% in the suburbs, 16% in the townships and 22% in the shack areas. Conditions clearly improved, but the differences between urban neighbourhoods remained very large.

Summing up, government social grants helped to offset the large economic gaps between places, but the incidence of hunger was still much higher in informal settlements, townships and peri-urban areas than in the suburbs.

Conclusion

Further research is required to substantiate and extend this initial evidence of the asymmetric impact of COVID-19 across South Africa. The situation is dynamic and evolving as the restrictions are relaxed and different economic sectors show different levels of recovery. It is vital to recognise that different parts of the country face different challenges and risks of further setbacks, depending on how the pandemic unfolds and evolves. Poor communities have borne more of the burden of the lockdown than have suburban communities. Treating unequal places in the same way will not narrow the gap between them. Blanket national measures have not been sensitive to these variations, with the unintended consequence of amplifying inequalities. National programmes need complementary efforts to boost jobs and livelihoods in and around vulnerable communities. This means targeting particular kinds of places as well as specific groups of people in tackling poverty and unemployment. It also means working closely with affected communities in formulating and implementing appropriate responses.

Acknowledgements

We thank NIDS-Cram for funding the research report on which this paper is based.

Competing interests

We have no competing interests to declare.

Authors' contributions

I.T.: Conceptualisation, oversight, writing. J.V.: Conceptualisation, methodology, data analysis.

References

1. Organisation for Economic Co-operation and Development (OECD). OECD Economic outlook: Interim report September 2020. Paris: OECD; 2020. https://doi.org/10.1787/34ffc900-en
2. Spaull N, Ardington C, Bassier I, Bhorat H, Bridgman G, Brophy T, et al. NIDS-CRAM Wave 1 synthesis report: Overview and findings [document on the Internet]. c2020 [cited 2020 Sep 22]. Available from: https://cramsurvey.org/reports/#wave-1
3. Statistics South Africa. Key findings: P0441 – Gross Domestic Product, second quarter 2020 [document on the Internet]. c2020 [cited 2020 Sep 22]. Available from: http://www.statssa.gov.za/publications/P0441/P04411stQuarter2020.pdf
4. Turok I. Worlds apart: Spatial inequalities in South Africa. In: Smith MN, editor. Confronting inequality: The South African crisis. Johannesburg: Jacana Media; 2018. p.129–151.
5. Todes A, Turok I. Spatial inequalities and policies in South Africa: Place-based or people-centred? Progr Plann. 2018;123:1–32.
6. Spaull N, Oyenubi A, Kerr A, Maughan-Brown B, Ardington C, Christian C, et al. NIDS-CRAM Wave 2 synthesis report: Overview and findings [document on the Internet]. c2020 [cited 2020 Sep 22]. Available from: https://cramsurvey.org/reports/#wave-2
7. Posel D, Casale D. Who moves during times of crisis? Mobility, living arrangements and COVID-19 in South Africa. NIDS-CRAM Wave 1 Working Paper [document on the Internet]. c2020 [cited 2020 Sep 22]. Available from: https://cramsurvey.org/reports/#wave-1
8. National Income Dynamics Survey Wave 5 [data set on the Internet]. c2020 [cited 2020 Sep 22]. Available from: http://www.nids.uct.ac.za/nids-data/data-access

Corrigendum

[Original article] Tembe D, Mukaratirwa S. Forensic entomology research and application in southern Africa: A scoping review. S Afr J Sci. 2020;116(5/6), Art. #6065, 8 pages. https://doi.org/10.17159/sajs.2020/6065

HOW TO CITE:
Corrigendum: Forensic entomology research and application in southern Africa: A scoping review [S Afr J Sci. 2020;116(5/6), Art. #6065, 8 pages]. S Afr J Sci. 2021;117(3/4), Art. #6065C. https://doi.org/10.17159/sajs.2021/6065C

Errors that appear in the Review Article by Tembe and Mukaratirwa are corrected here. Prof. Martin Villet (Director: Southern African Forensic Entomology Research Laboratory and Department of Zoology & Entomology, Rhodes University) is acknowledged for drawing the authors' attention to these errors.

Page 2, right column, section 'Search strategy and selection of the literature', paragraph 1, lines 13–17:

"Two exclusion criteria were also identified: (1) no focus on forensic entomology research, such as articles that dealt with identification or distribution of arthropods in southern African countries, but not undertaken in the context of forensic entomology; (2) no information points that contributed to answering the scoping question."

SHOULD BE REPLACED WITH:

"The exclusion criteria for this review were developed based on the three inclusion criteria which were peer-reviewed research articles from southern Africa explicitly reporting on forensic entomology research in a country or countries from southern Africa between 1932 and 2017 that included: (1) colonisation and succession pattern of arthropods during different stages of decomposition; (2) variation spectrum of carrion-feeding insects; and (3) diversity and/or abundance of arthropods colonising a carcass during different seasons. Any study that may have been referenced under forensic entomology, but did not report on one of the above-mentioned inclusion criteria, was not included. Furthermore, any study that had no information to contribute to answering the scoping questions was also excluded."

Page 4, Table 2, caption and columns 7 and 8 subheading AND page 6, left column, subheading for paragraph 2 AND page 6, right column, subheading:

Carrion-feeding SHOULD BE REPLACED WITH Carrion-associated

Page 4, Table 2, column 6, line 13:

"*Histeridae*" SHOULD BE DELETED

Page 4, Table 2, column 6, line 14:

"Fabricius" SHOULD BE: *"Thanatophilus (Chalcosilpha) micans* (Fabricius)"

Page 4, Table 2, column 6, line 49:

"*Dermestes maculatus"* SHOULD BE DELETED

Page 5, Table 2, column 6, line 89:

"*Muscidae*" SHOULD BE REPLACED WITH: "*Atherigona aberrans* (Mallocch), *A. naqvii* (Steyskal), *A. steeleae* (Emden), *Atherigona* spp. indet."

Page 5, Table 2, columns 6 and 7, line 99:

"*Lucilia* sp. –" SHOULD BE: "*Lucilia* sp. 47"

Page 6, left column, paragraph 2, lines 6–8:

"In both studies, *D. maculatus and Lucilia* spp. were found on decomposing carcasses during the dry season only."

SHOULD BE REPLACED WITH:

"In both studies, *Lucilia* spp. were found on decomposing carcasses during the dry season only. *Dermestes maculatus* was found only during the dry season in the study of Ellison[38], and was not found in either season in the study of Braack[37]."

Relevant species presented by Braack[37] that were omitted have been added to Table 2. These are: *Chrysomia marginalis* (Wd.), *Chrysomia putoria* (Wd.), *Rhinia apicalis* (Wd.), *Rhyncomya forcipata, Sarcophaga haemorrhoidalis* (Fallen), *S. hirtipes* (Wd.), *S. nodosa* (Engel), *Brachyponera sennaarensis* (Mayr), *Pheidole crassinoda* (Em.), *P. liengmei* (For.), *Axestotrigona togoensis* (Cockerell).

Relevant results from Kelly et al.[36] that were omitted have been added to Table 2.

The corrected Table 2 appears below.

Page 6, Table 3, caption and column 5:

Results by Williams et al.[39] obtained from rats and chicken liver that were omitted have been added to Table 3.

The corrected Table 3 appears below.

Page 8, reference 39:

Three of the authors were omitted. The corrected reference is:

Williams KA, Wallman JF, Lessard BD, Kavazos CRJ, Mazungula DN, Villet MH. Nocturnal oviposition behavior of blowflies (Diptera: Calliphoridae) in the southern hemisphere (South Africa and Australia) and its forensic implications. Forensic Sci Med Pathol. 2017;13(2):123–134. https://doi.org/10.1007/s12024-017-9861-x

Table 2: Summary of studies (1934–2017) on the diversity and abundance of carrion-associated arthropods collected during different seasons in southern Africa

Study	Country of study	Location of study	Objectives of study	Host animal	Outcome of study		
					Order/family/species	Average number of carrion-associated arthropods	
						Dry season	Rainy season
Braack[37]	South Africa	Kruger National Park	To collect and identify the species found on the large mammal carcasses during both summer and winter	Impala	*Anisolabis* sp.	–	<10
					Bormansia meridionalis Burr	–	<10
					Euborellia annulipes (Lucas)	–	<10
					Fusius rubricosus (Stal)	–	<10
					Lisarda rhodesiensis Miller	–	<10
					Rhinocoris albopunctatus (Stal)	–	<10
					R. violentus (Germar)	–	<10
					Xylocoris (Proxylocoris) afer Reuter	–	±60
					Solenostethium liligerum	–	<10
					Metagonum sp.	–	<10
					Platymetopus curtulus (Peringuey)	–	<10
					Xenodochus melanarius (Boheman)	–	<10
					Thanatophilus (Chalcosilpha) micans (Fabricius)	–	265
					Staphylinidae	–	625
					Trogidae	–	1422
					Allogymnopleurus thalassinus (Klug)	–	<30
					Anachalcos convexus (Boheman)	–	164
					Aphodius sp.	–	<100
					Caccobius convexifrons (Roth)	–	<30
					C. nigritulus (Klug)	–	<30
					Catharsius philus (Kolbe)	–	<30
					Copris amyntor (Klug)	–	<30
					C. elphenor (Klug)	–	<30
					C. evanidus (Klug)	–	<30
					C. mesacanthus (Harold)	–	<30
					Garreta nitens (Olivier)	–	<30
					Gymnopleurus virens (Erichson)	–	<30
					Metacatharsius opacus (Waterhouse)	–	<30
					Milichus sp. *probably apicalis* (Fahraeus)	–	<30
					Onitis fulgidus (Klug)	–	<30
					O. granulisetosus (Ferreira)	–	<30
					O. inversidens (van Lansberge)	–	<30
					O. obenbergeri (Balthasar)	–	<30
					O. picticollis (Fabricius)	–	<30
					Onthophagus (Proagoderus) dives (Klug)	–	5670
					Pedaria sp.	–	<30
					Phaeochrous madagascariensis (Westwood)	–	4486
					Phalops ardea (Klug)	–	<30
					Sarophorus costatus (Fahraeus)	–	2304
					Scarabaeus ebenus (Klug)	–	<30
					Sisyphus calcaratus (Klug)	–	<30
					S. goryi (Harold)	–	<30
					S. impressipennis (van Lansberge)	–	<30
					S. injuscatus (Klug)	–	<30
					S. seminulum (Gerstaecker)	–	<30
					Sybax distortus (Schaum)	–	<30
					Tiniocellus spinipes (Peringuey)	–	<30
					Necrobia rufipes (De Geer)	–	2572
					Phloeocopus sp.	–	1
					Carpophilus nr. *quadrisignatus* Er.	–	<10
					Carpophilus sp.	–	<10
					Bactria sp.	–	<10
					Euscelidia rapax (Westwood)	–	<10
					Hoplistomerus nobilis (Loew)	–	<10
					Neolophonotus (Lophopeltis) sp.	–	<10

Table 2 continued

Study	Country of study	Location of study	Objectives of study	Host animal	Outcome of study		
					Order/family/species	Average number of carrion-associated arthropods	
						Dry season	Rainy season
					Ommatius sp.	–	<10
					Stichopogon caffer (Hermann)	–	<10
					S. punctus (Loew)	–	<10
					Hypocerides spinulicosta (Beyer)	–	<10
					Megaselia curtineura	–	<10
					Megaselia sp. n. *pauculitincta*	–	<10
					Plethysmochaeta sp.	–	<10
					Australosepsis niveipennis (Becker)	–	<50
					Paratoxopoda depilis (Walker)	–	97
					Xenosepsis sp.	–	<50
					Cestrotus n. sp.	–	<10
					Homoneura (Keisomyia) n. sp.	–	<10
					Curtonotum cuthbertsoni (Duda)	–	<10
					Atherigona aberrans (Mallocch), *A. naqvii* (Steyskal), *A. steeleae* (Emden), *Atherigona* spp. indet.	–	1289
					Fannia leucosticta (Meigen)	–	1
					Graphomya leucomelas (Wiedemann)	–	1
					Gymnodia mervinia (Walker)	–	5
					Gymnodia tonitrui (Wiedemann)	–	3
					Haematobosca latifrons (Malloch)	–	1
					H. spinigera (Malloch)	–	6
					H. thirouxi ssp. *potans* (Bezzi)	–	7
					Morellia nilotica (Loew)	–	3
					Lucilia sp.	47	–
					Chrysomya marginalis (Wd.)	–	>991
					Chrysomya putoria (Wd.)	–	<10
					Rhinia apicalis (Wd.)	<10	–
					Rhyncomya forcipata	<10	–
					Sarcophaga haemorrhoidalis (Fallen)	<10	–
					S. hirtipes (Wd.)	<10	–
					S. nodosa (Engel)	<10	–
					Nasonia vitripennis (Walker)	–	<40
					Trichopria lewisi (Nixon)	–	>35
					Brachyponera sennaarensis (Mayr)	–	<5000
					Pheidole crassinoda (Em.)	–	<5000
					P. liengmei (For.)	–	>5000
					Axestotrigona togoensis (Cockerell)	–	30–50
					Lardoglyphus sp.	–	<100
					Macrocheles muscaedomesticae (Scopoli)	–	<100
					Pygmephorus sp.	–	<100
Ellison[38]	South Africa	Klaserie Private Nature Reserve	The effect of scavenger mutilation on the subsequent rate of decomposition and insect colonisation of such carcasses	Impala	*Saprinus* spp.	1.3	–
					Necrobia rufipes	6.6	–
					Dermestes maculatus	9.2	–
					Aleochara spp.	<1	–
					Thanatophilus spp.	<1	–
					Mycetophagidae spp.	<1	–
					Onthophagus spp.	<1	–
					Piophila spp.	36.5	–
					Ophyra capensis	3.4	–
					Musca spp.	10.9	–
					Chrysomya albiceps	3.4	–
					Chrysomya chloropyga	<1	–
					Chrysomya marginalis	4	–
					Chrysomya putoria	<1	–
					Tricyclea spp.	9.7	–
					Lucilia spp.	11	–
					Sarcophaga spp.	0.75	–
					Auchmeromyia luteola	0.25	–
					Ceratophaga vastella	<1	–
					Brachynieria spp.	<1	–
					Acrididae spp.	<1	–

Table 2 continued

Study	Country of study	Location of study	Objectives of study	Host animal	Outcome of study		
					Order/family/species	Average number of carrion-associated arthropods	
						Dry season	Rainy season
†Kelly et al.[36]	South Africa	University of the Free State, Bloemfontein	(1) Study the effect of various wound types on the detection and selection of the carcasses by Diptera (2) Study the early dipteran colonisation and overall arthropod succession patterns on wounded and non-wounded carcasses (3) Compare unclothed and clothed carcasses decomposition and arthropod succession during all seasons	Pig (*Sus scrofa* L.)	*Chrysomya albiceps* *Chrysomya chloropyga* *Chrysomya marginalis* *Calliphora vicina* *Sarcophagidae* *Dermestes maculatus* *Necrobia rufipes* *Musca* spp. *Thanatophilus micans*	– 2 – 1 4 1 2 – 3	2 4 1 – 3 1 2 – –

†Note: 1, most abundant (highest counts); 2, second most abundant; 3, few individuals; 4, few individuals (least counts)
–, None present or identified

Table 3: Summary of the diurnal and nocturnal oviposition by forensically important arthropods on pork chops, chicken liver and rat carcasses in southern Africa

Study	Country of study	Location of study	Objective of study	Host animal	Outcome of study		
					Species identified	Day	Night
Williams et al.[39]	South Africa	Grahamstown and Durban	To determine the nocturnal oviposition behaviour of blowflies in the southern hemisphere	Pig (pork chops)	*Chrysomya megacephala*	1	0
					Lucilia sericata	8	1
					Chrysomya putoria	7	1
					Chrysomya chloropyga	2	1
				Rat (carcasses)	*Lucilia sericata*	0	1
					Sarcophaga sp.	+	0
				Chicken (liver)	*Sarcophaga* sp.	+	0
					Chrysomya megacephala	0	+
					Lucilia cuprina	+	+