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

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

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

CORRESPONDENCE TO:
Caradee Wright

EMAIL:
caradee.wright@mrc.ac.za

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

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

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

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

PUBLISHED:
31 August 2023

Significance:

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

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

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

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

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

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

Air quality monitoring in South Africa

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

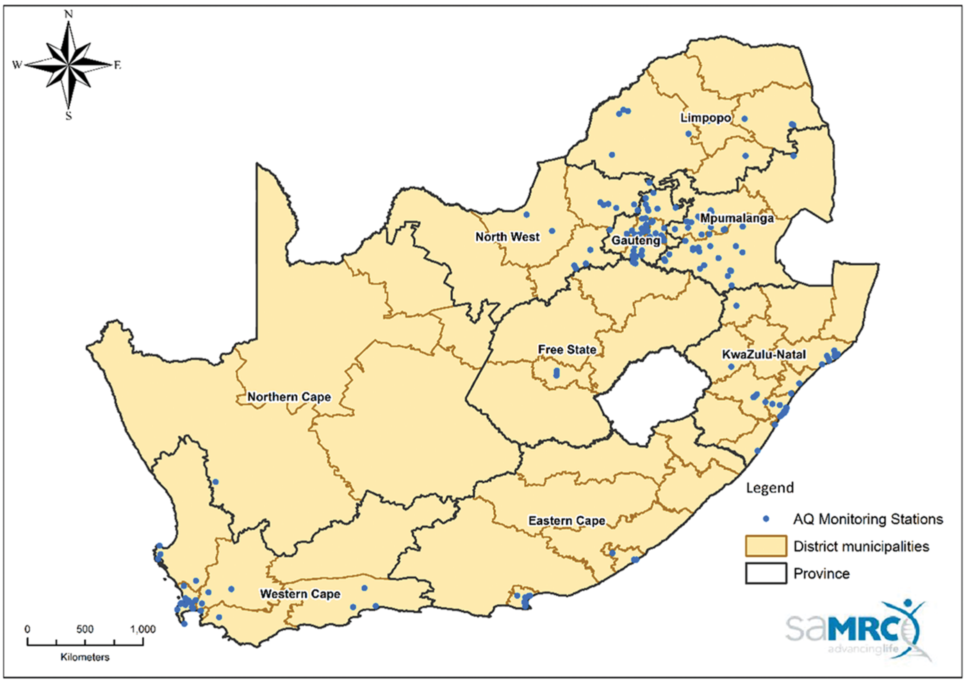

Figure 1: Location of the air quality monitoring stations in South Africa that report to the South African Air Quality Information System (SAAQIS) as of January 2023.

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

Load shedding impacts air quality monitoring data collection

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

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

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

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

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

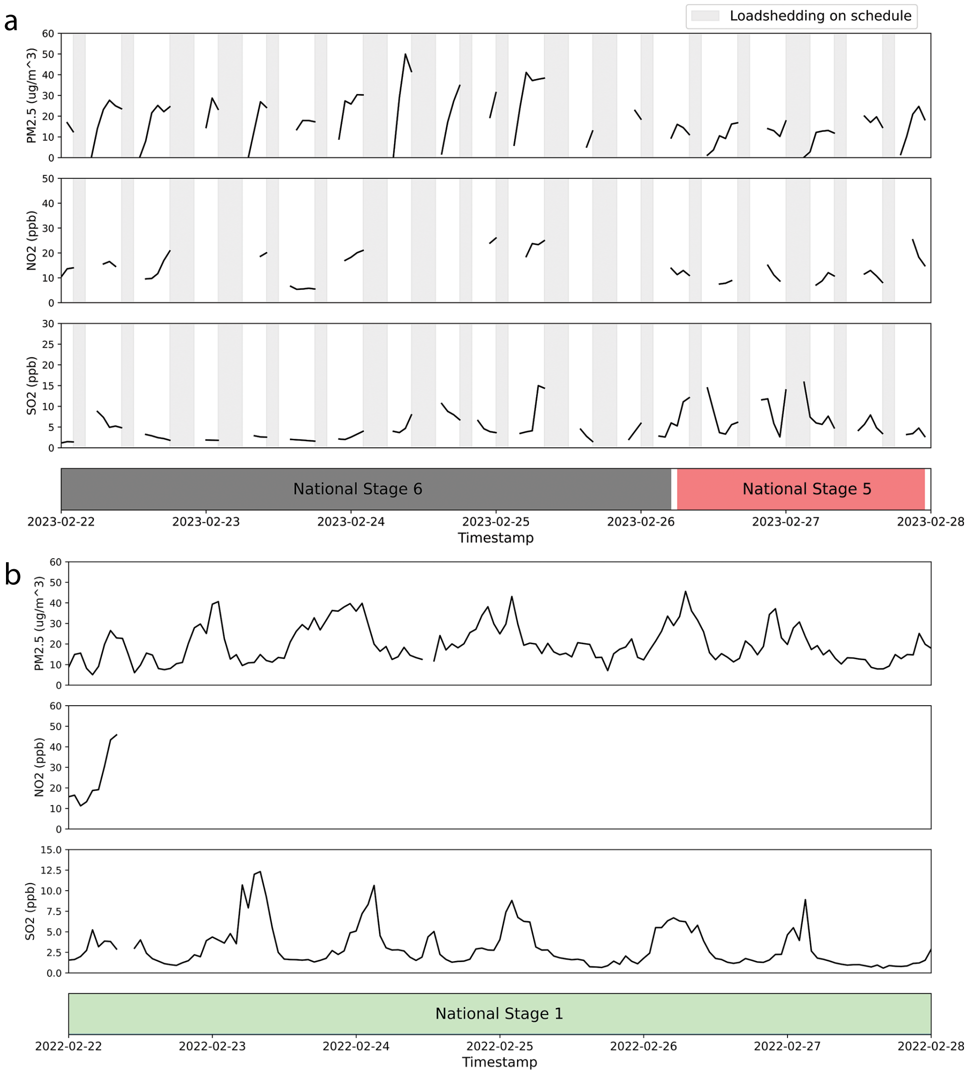

Figure 2: Concentrations of $PM_{2.5}$, NO_2 and SO_2 at the Diepkloof monitoring station between 22 and 27 February 2023 during: (a) load shedding Stages 6 and 5 (the width of the grey column is indicative of the time when load shedding was occurring) and (b) no load shedding (there were no data for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

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

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

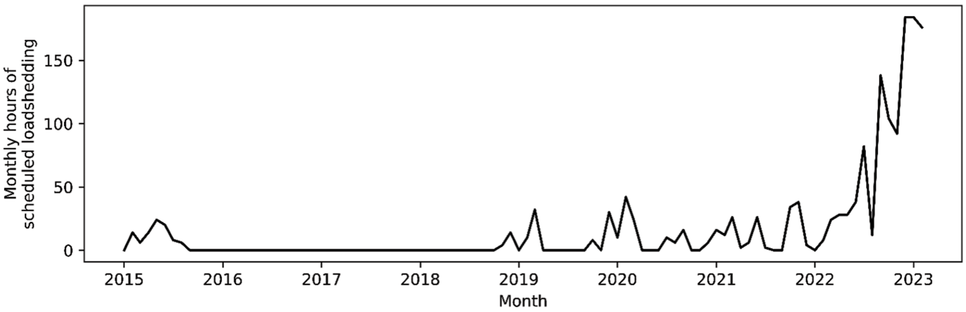

Figure 3: Monthly hours of scheduled load shedding between 2015 and 2023.

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

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

Challenges, opportunities and way forward

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

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

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

Conclusions

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

Acknowledgements

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

Competing interests

We have no competing interests to declare.

References

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

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

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

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

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

6. Eskom. Interpreting load-shedding stages in South Africa [webpage on the Internet]. No date [cited 2023 Jun 05]. Available from: https://loadshedding.eskom.co.za/LoadShedding/ScheduleInterpretation

7. Mbomvu L, Hlongwane IT, Nxazonke NP, Qayi Z, Bruwer JP. Load shedding and its influence on South African small, medium and micro enterprise profitability, liquidity, efficiency and solvency (2021 April 21). Business Re-Solution Working paper BRS/2021/001. SSRN [cited 2023 Apr 18]. Available from: https://doi.org/10.2139/ssrn.3831513

8. Gehringer C, Rode H, Schomaker M. The effect of electrical load shedding on pediatric hospital admissions in South Africa. Epidemiology. 2018;29(6):841–847. https://doi.org/10.1097/EDE.0000000000000905

9. Government of South Africa. Government terminates National State of Disaster on electricity supply constraints [media release on the Internet]. c2023. [cited 2023 Apr 18]. Available from: https://www.gov.za/speeches/government-terminates-national-state-disaster-%C2%A0electricity-supply-constraints-5-apr-2023

10. South Africa warned of load shedding 'beyond stage 6' as winter approaches. BusinessTech. 2023 April 12. Available from: https://businesstech.co.za/news/energy/679511/south-africa-warned-of-load-shedding-beyond-stage-6-as-winter-approaches/

11. Arowosegbe OO, Röösli M, Künzli N, Saucy A, Adebayo-Ojo TC, Schwartz J, et al. Ensemble averaging using remote sensing data to model spatiotemporal PM_{10} concentrations in sparsely monitored South Africa. Environ Pollut. 2022;310, Art. #119883. https://doi.org/10.1016/j.envpol.2022.119883

12. World Health Organization. Ambient (outdoor) air pollution fact sheet [webpage on the Internet]. c2022 [cited 2023 May 24]. Available from: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health

13. South African Air Quality Information System [homepage on the Internet]. No date [cited 2023 Jun 05]. Available from: https://saaqis.environment.gov.za/

14. World Health Organization. Global air quality guidelines: Particulate matter ($PM_{2.5}$ and PM_{10}), ozone, nitrogen dioxide, sulphur dioxide and carbon. Geneva: World Health Organization; 2021. https://apps.who.int/iris/handle/10665/345329

15. Parliamentary Monitoring Group. DFFE & SAWS Q1 & 2 2022/23 Performance with Minister [webpage on the Internet]. c2023 [cited 2023 Apr 18]. Available from: https://pmg.org.za/committee-meeting/36426/

16. ESP [load shedding application]. c2023 [cited 2023 Apr 18]. Available from: https://esp.info/

17. Gwaze P, Mashele SH. South African Air Quality Information System (SAAQIS) mobile application tool: Bringing real time state of air quality to South Africans. Clean Air J. 2018;28(1):3. https://doi.org/10.17159/2410-972X/2018/v28n1a1

18. Altieri KE, Keen SL. Public health benefits of reducing exposure to ambient fine particulate matter in South Africa. Sci Total Environ. 2019;684:610–620. https://doi.org/10.1016/j.scitotenv.2019.05.355

19. Hersey SP, Garland RM, Crosbie E, Shingler T, Sorooshian A, Piketh S, et al. An overview of regional and local characteristics of aerosols in South Africa using satellite, ground, and modeling data. Atmos Chem Phys. 2015:15:4259–4278. https://doi.org/10.5194/acp-15-4259-2015

20. Matandirotya NR, Moletsane SD, Matandirotya E, Burger RP. State of ambient air quality in a low-income urban settlement of South Africa. Sci Afr. 2022;16:e01201. https://doi.org/10.1016/j.sciaf.2022.e01201

21. South African Department of Forestry, Fisheries, and the Environment (DFFE). Update on the air quality monitoring strategy and AAQI stations project [document on the Internet]. c2017 [cited 2023 Jun 05]. Available from: https://saaqis.environment.gov.za/Lekgotla%20Proceedings/2017/2017_2.2-update-on-aqm_strategy---naqi-station-project.pdf

22. South African Department of Forestry, Fisheries, and the Environment (DFFE). 2022 State of Air Report and AQM highlights [document on the Internet]. c2022 [cited 2023 Jun 05]. Available from: https://www.dffe.gov.za/sites/default/files/docs/2022airqualitylekgotlapresentations_stateofair.pdf

23. De Jongh D, Ghoorah D, Makina A. South African renewable energy investment barriers: An investor perspective. J Energy South Afr. 2014;25(2):15–27. https://doi.org/10.17159/2413-3051/2014/v25i2a2665

24. Murombo T. Legal and policy barriers to renewable and sustainable energy sources in South Africa. J World Energy Law Bus. 2016;9:142–165. https://doi.org/10.1093/jwelb/jww001