Atmospheric Pollution Research 16 (2025) 102459

Contents lists available at ScienceDirect

Atmospheric Pollution Research

journal homepage: www.elsevier.com/locate/apr

Household $PM_{2.5}$ in a South African urban and rural setting: A comparative analysis using low-cost sensors

Matthew Benyon [a,b,c,*], Ngwako Kwatala [d], Tracey Laban [e], Thandi Kapwata [e,f], Chiara Batini [g,h], Samuel Cai [b], Lisa K. Micklesfield [i], Rikesh Panchal [a,b], Siyathemba Kunene [d], Sizwe B. Zondo [d], Brigitte Language [j], Bianca Wernecke [k], Scott Hazelhurst [d,l], F. Xavier Gómez-Olivé [m], Joshua Vande Hey [a,b,c], Caradee Y. Wright [e,k]

[a] School of Physics and Astronomy, University of Leicester, Leicester, LE1 7RH, United Kingdom
[b] Centre for Environmental Health and Sustainability, University of Leicester, Leicester, LE1 7RH, United Kingdom
[c] Institute for Environmental Futures, University of Leicester, Leicester, LE1 7RH, United Kingdom
[d] Sydney Brenner Institute for Molecular Bioscience, University of the Witwatersrand, Private Bag 3, 9 Jubilee Road, 2050, Gauteng, South Africa
[e] Climate Change and Health Research Programme, Environment and Health Research Unit, South African Medical Research Council, Private Bag x385, 1 Soutpansberg Road, Pretoria, 0001, South Africa
[f] Department of Environmental Health, University of Johannesburg, Johannesburg, South Africa
[g] Genetic Epidemiology Group, Department of Population Health Sciences, University of Leicester, Leicester, LE1 7RH, United Kingdom
[h] University Hospitals of Leicester NHS Trust, Groby Road, Leicester, Groby Rd, Leicester, LE3 9QP, United Kingdom
[i] SAMRC/Wits Developmental Pathways for Health Research Unit, Department of Paediatrics, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Private Bag X3, Wits, Johannesburg, 2050, South Africa
[j] Unit for Environmental Science and Management, North-West University, Potchefstroom, South Africa
[k] Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, 0001, South Africa
[l] School of Electrical and Information Engineering, University of the Witwatersrand, 1 Jan Smuts Avenue, Braamfontein, Johannesburg, 2000, South Africa
[m] MRC/Wits Rural Public Health and Health Transitions Research Unit, School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, 1 Jan Smuts Avenue, Braamfontein, Johannesburg, 2000, South Africa

ARTICLE INFO

Keywords:
Air quality
Environmental health
Particulate matter
Low-middle-income-countries
Indoor air pollution

ABSTRACT

Household air pollution (HAP) is responsible for millions of premature deaths each year. Exposure to household air pollutants as a risk factor for poor health has not been adequately quantified in many parts of the world, especially Sub-Saharan Africa. We aimed to assess HAP, specifically $PM_{2.5}$, and its associations with dwelling and household characteristics in urban (Soweto) and rural (Agincourt) settings in South Africa. We monitored indoor $PM_{2.5}$ concentrations in 40 unique households using low-cost sensors, across two study sites and seasons. Low-cost sensors were calibrated by collocation, and associations between dwelling and household characteristics with indoor $PM_{2.5}$ concentrations were assessed using a log-linear regression model. $PM_{2.5}$ concentrations were greater in urban households in the summer (50 $\mu g/m^3$ (95% CI: 41–63) and in the winter (82 $\mu g/m^3$ (95% CI: 62–109)) compared to rural households (summer: 19 $\mu g/m^3$ (95%: CI 14–26) and winter: 48 $\mu g/m^3$ (95% CI: 44–53)). The log-linear model (n = 39) explained 74% of the variance in leave-one-out cross validation. Significant associations with household $PM_{2.5}$ were observed with the following: the season, study setting, presence of tobacco smoking, presence of incense burning inside the dwelling, and the use of heating. This study found significant variations in HAP concentrations within and across the urban and rural communities, likely influenced by differences in ambient outdoor concentrations and individual behaviours such as incense burning. It is crucial to enhance community and policy maker awareness regarding the dangers of indoor smoking and the harmful effects of burning incense indoors.

* Corresponding author. Space Park Leicester, 92 Corporation Road, Leicester, LE4 5SP, United Kingdom.
E-mail address: mb932@leicester.ac.uk (M. Benyon).

https://doi.org/10.1016/j.apr.2025.102459
Received 22 August 2024; Received in revised form 20 January 2025; Accepted 11 February 2025
Available online 13 February 2025

1. Introduction

Ambient and household air pollution (HAP) pose one of the greatest threats to the health and wellbeing of people around the world, causing a combined 6.7 million deaths in 2021 (Fuller et al., 2022). Fine particulate matter ($PM_{2.5}$; particulate matter smaller than 2.5 μm in diameter) is a component on HAP and is particularly harmful due to its ability to penetrate deeply into the lungs. Over two billion people worldwide rely on solid fuels—such as wood, coal, animal dung, and crop waste—for cooking and heating, which can produce hazardous levels of $PM_{2.5}$ (World Health Organization, 2023). Recent studies indicate that indoor environments could be at least twice as polluted as outdoor environments, and infiltration by outdoor sources of air pollution can also influence indoor concentrations, especially in homes that are located near major roads with heavy traffic or industrialised areas (González-Martín et al., 2021; European Commission, 2003). A study in the United Kingdom showed that indoor NO_2 levels, particularly in the kitchen area, surpassed outdoor levels (Singh et al., 2024). In 2020, more than three million people died likely as a result of exposure to HAP (World Health Organization, 2023).

South Africa is considered to be an upper-middle income country (World Bank, no date), though internal hardships such as high rates of unemployment and wealth inequality mean that it is considered an economically developing country by the United Nations (United Nations, 2024). In 2012, about 97% of South Africa's population were exposed to ambient $PM_{2.5}$ (particulate matter with a diameter $\leq$2.5 μm) concentrations above the 2005 World Health Organization (WHO) annual guideline of 10 $\mu m/m^3$; and the highest number of attributable deaths due to ambient (outdoor) $PM_{2.5}$ was found to be in the Gauteng province of South Africa (Roomaney, Cairncross, et al., 2022a). Air pollution in Gauteng is mainly derived from industrial emissions, dust, vehicles, and domestic combustion of fuels such as coal, wood, and biomass (Adeyemi et al., 2021; Muyemeki et al., 2021).

In 2012, an estimated 17.6% of the South African population were exposed to HAP, with almost 1.7% of all deaths attributed to HAP exposure (Roomaney, Wright, et al., 2022b). In contrast to urban communities, South Africa's low-income (rural) communities rely heavily on sources other than electricity for heating and cooking. In one rural South African province, around half of the population rely on solid fuels such as wood, coal, and animal dung for cooking and space heating activities (Kapwata et al., 2018). These fuel types can be significant sources of indoor-originated PM (Salvi and Barnes, 2010; Paulin et al., 2013; Semmens et al., 2015). Adesina et al. (2020) found that indoor PM concentrations were higher than outdoor levels during mornings and evenings, coinciding with the fuel-burning pattern of households where coal was the dominant fuel type. Studies in low-income communities living on the Mpumalanga Highveld – an area known for severe air pollution – have shown that domestic fuel use habits vary (Nkosi et al., 2017; Wernecke et al., 2024). While most rural households rely on coal for heating and cooking, households move between various sources to meet their energy needs. The impact of these energy choices on HAP remains underexplored, and further studies are needed to examine the differences in HAP between urban and rural households (Roomaney et al., 2022a,b).

Low-cost air quality monitoring instruments (low-cost sensors: LCS) can be used to measure HAP to bridge the significant data gap in low- and middle-income countries (LMICs) due to affordability, small size and portability (Raysoni et al., 2023). The performance of LCS has been shown to be reasonably effective, especially when considering the cost-to-performance ratio (Kang et al., 2022). The development of LCS provides a pathway for needed epidemiological research to determine the strength of associations between the HAP exposure and health. However, the installation and use of low-cost sensors in indoor environments require non-trivial efforts, including careful calibration through co-location with reference instruments (Wernecke et al., 2021).

Given the concerns about the impacts of HAP exposure on human health, studies are needed in both urban and rural areas – where there are differences in dwelling or household characteristics – in LMICs to better understand HAP concentrations and the factors influencing these concentrations. Here, we assessed $PM_{2.5}$ concentrations in dwellings in both urban and rural areas in South Africa to investigate their associations with dwelling and household characteristics. The following section provides the methods we applied to achieve the study aim.

2. Methods

2.1. Data collection

2.1.1. Study sites

This study took place in two locations: Soweto (urban), situated in the Vaal Air-Shed Priority Area and Agincourt (rural), in the Highveld Priority Area. These priority areas are defined by their poor air quality and are located in prominent mining and industrial zones situated in the north-east region of South Africa. The urban study site was in Soweto in the Gauteng province of South Africa (Fig. 1). Soweto is around 200 km^2 and in 2011 had a population of 1.3 million (Statistics South Africa, no date). The urban winter and summer campaigns occurred in the Soweto suburbs of Jabavu and Diepkloof, respectively. These locations were chosen because they were less than 10 km apart and since the research team had worked there before, it helped to facilitate community entry. Moreover, Jabavu and Diepkloof are similar in many aspects including housing type and air pollution sources, namely, vehicles, household solid fuel burning and local industry. Interaction terms were explored to account for interactions between the theoretically related variables primary cooking fuel the 17th of October and 2nd of November 2022, while the urban winter campaign was from the 16th to the 29th May 2023. We highlight that we did not intend to capture data representative of an entire season (which was limited due to budget and number of instruments) but rather to gather data that were illustrative of air quality during a two-week period in a specific season.

The rural study site was located around the villages surrounding the South African Medical Research Council (MRC)/Wits Rural Public Health and Health Transitions Research Unit in Agincourt in the Bushbuckridge Municipality, Mpumalanga Province. The population of the MRC/Wits Agincourt Research study site is approximately 120,000 people distributed in 31 villages in an area of 450 Km^2. The area is typical of rural South Africa with high unemployment rates, high labour migration and limited infrastructure (Kahn et al., 2012). The study site is located in between the dry savannah of Kruger National Park to the east, and the Drakensberg escarpment to the West. The rural summer campaign took place between the 15th of February and 2nd of March 2023, while the rural winter campaign was from 5th to the 18th of July 2023.

2.1.2. Participant recruitment

Participants were recruited from the Africa Wits-INDEPTH partnership for Genomic Studies (AWI-Gen) (Ramsay et al., 2016; Ali et al., 2018) led by the University of the Witwatersrand. For this small-scale feasibility study, we used a convenience sample, as the primary objective was to assess the feasibility of the proposed methods and interventions, and the secondary objective was to understand HAP concentrations. Given the study's exploratory nature, representativeness was not part of the main objectives. All AWI-Gen participants' street addresses were obtained and those participants living closest to government-run ambient (outdoor) air quality monitoring stations were invited to participate in the study. A fieldworker visited the household and requested to speak with a member of the household older than 18 years of age. The fieldworker explained the purpose of the study after which the participant provided informed consent. A total of 44 unique households were recruited for the study, 11 per site per season. This was determined by the number of low-cost sensors that were available for use. Research ethics clearance was provided by the University of

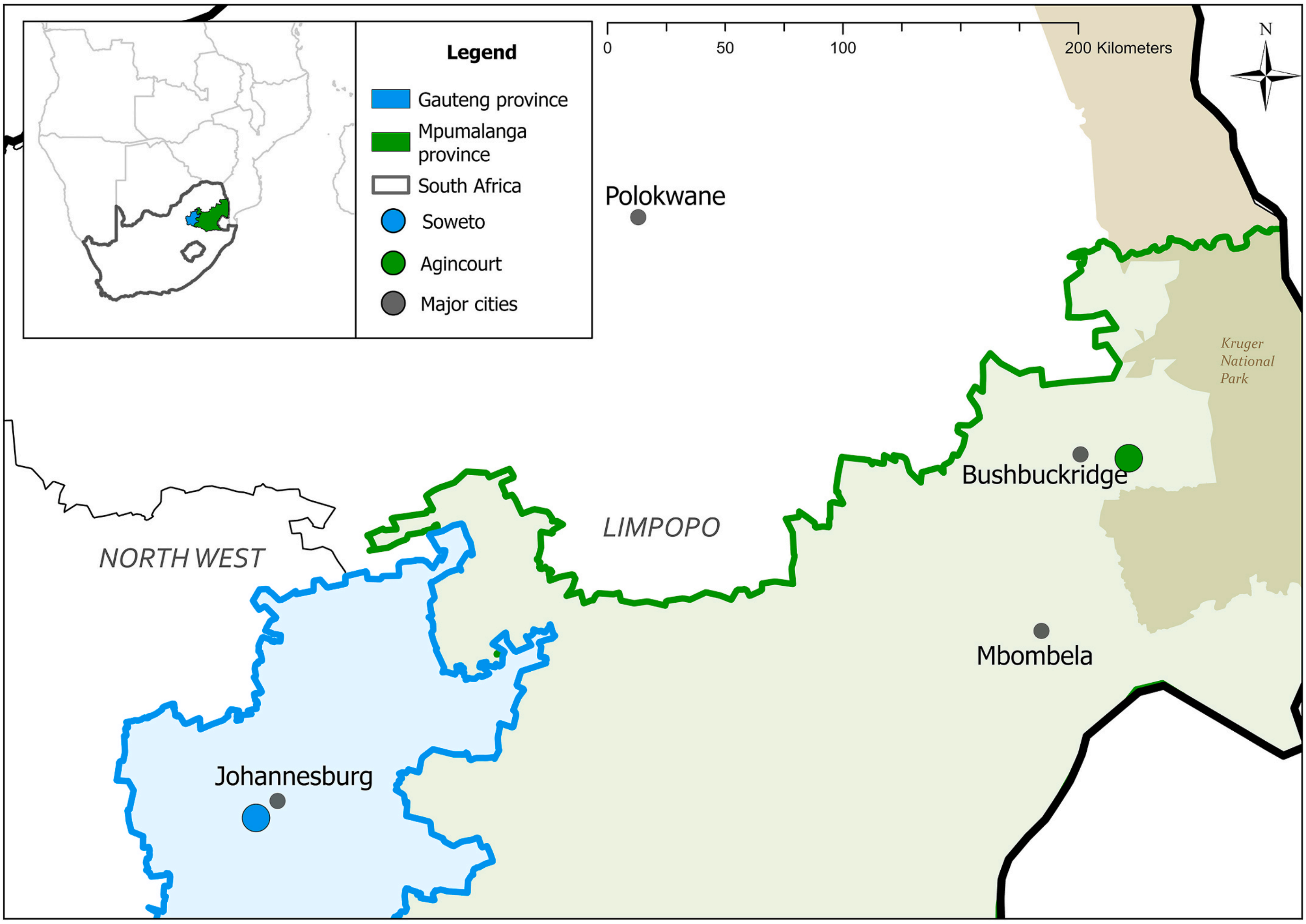

Fig. 1. Map showing the location of the study sites within the Gauteng- and Mpumalanga Provinces in South Africa. Map lines delineate study areas and do not necessarily depict accepted national boundaries.

Witwatersrand Research Ethics Committee (Protocol M210670; 10 September 2021).

2.1.3. Questionnaire and dwelling observation checklist

In this study, we collected data on several elements of the participants' living arrangements, including both dwelling and household characteristics. A dwelling refers to the main building or unit that may be occupied by a person or a group of people, while a household consists of the individuals who share occupancy of the dwelling and its facilities, such as cook stoves (Office for National Statistics (ONS), 2014).

To determine various household characteristics (and some dwelling characteristics) a questionnaire was administered by a fieldworker to the participant. The questionnaire was derived from the International Study of Asthma and Allergies in Childhood (ISAAC) questionnaire, an international tool used to understand asthma symptoms and air pollution exposure (Asher et al., 1995). Participants were asked about the following items: type of dwelling; number of people living in the household; number of rooms in the household; primary cooking/heating fuel of the household; whether cooking happens indoors or outdoors; whether anyone smokes indoors; whether windows or other forms of ventilation are used during cooking; whether incense is burned indoors; and whether there is any exhaust ventilation (e.g., chimney). In addition, two questions were asked about how long people spend cooking when and where; however, these data were deemed unreliable and are not presented here. Questions were answered by the participants, and the fieldworkers entered the answer into REDCap (Harris et al., 2009) software on a tablet. No personal identifying information was captured, and a unique identifier was used to link data to the participant.

A dwelling observation checklist was also completed by the field worker to determine dwelling characteristics. The checklist items covered: dwelling characteristics (e.g., roofing material, ceiling material, wall material, number of windows, flooring material, etc.), whether the front door faced the road (and if the road was paved or not) and if the house was next to tree shade. For both the questionnaire and the dwelling observation checklist, pre-determined answers with check boxes were created with the option to enter text manually if required. This was not required for any of the responses thus there was no ambiguity during the data analysis phase.

2.1.4. Air quality instruments, monitoring and data pre-processing

We used purpose-designed $PM_{2.5}$, temperature, and relative humidity (RH) low-cost monitors, hereafter referred to as "Bonolos" (Imatech, 2022). These monitors housed passive-sampling, light scattering $PM_{2.5}$ sensors and were powered by an external battery with on-board data-logging capabilities and a clock for the data to be time-stamped. They have several advantages over high-grade air quality monitoring equipment in that they are relatively low-cost and portable. The on-board data logger logged time-stamped data approximately every 1 min and stored these on a memory card for later download to a computer (timestamps outputted by the Bonolo would inconsistently deviate from a minute, so data were time averaged to 3 min). The sensors were programmed to measure the temperature, RH and $PM_{2.5}$ continuously for approximately 2 weeks in each dwelling. The range for the Bonolos, as specified by the manufacturer, was 0–500 μg/m^3 for $PM_{2.5}$, −40 to 125 °C for temperature and 0–100% for relative humidity. A total of 166 data points were removed from the field data across three sensors, due to falling outside this range (in the 1-min time resolution).

The Bonolos were placed inside the household in a location which was convenient for the participant and on a flat, stable surface. Since the sensors were powered by external batteries, there was flexibility in the location where they could be installed due to not being limited by proximity to a power socket. They were not installed in direct sunlight or placed on top of or near heat-generating appliances, such as refrigerators or microwaves, as these could create unique temperature microenvironments that might interfere with the sensor readings. The selection criteria of where to install the Bonolos included: rooms where the people spent most of their time (kitchen, living room, bedroom, etc.); breathing height places either standing or sitting; and in places where people would not be disturbed by the Bonolos.

We also measured ambient (outdoor) $PM_{2.5}$ concentrations using Earthsense Zephyr air quality sensors (Earthsense, no date) at both sites. For the rural winter campaign, the Zephyr was installed at a centrally located high school, however for the rural summer campaign, we encountered logistical constraints impeding the installation of Zephyr monitors and therefore we did not collect outdoor pollutant measurements during this period. For the urban site, the Zephyr was installed at the South African Ambient Air Quality Information System (SAAQIS) Jabavu-NAQI (27°52′19.6″E 26°15′09.4″S) and Diepkloof-NAQI (27°57′23.1″E, 26°15′02.6″S) monitoring stations for the summer and winter campaigns, respectively. The calibration of the Zephyr instruments was conducted through collocation with reference instruments situated at the Automatic Urban and Rural Network (AURN) air quality monitoring station located at the University of Leicester in the United Kingdom, operated by the Department for Environment, Food and Rural Affairs (DEFRA). Zephyr calibration methodologies were implemented by Earthsense, with data provided by the researchers, in accordance with procedures outlined in prior studies (Peterson et al., 2017; Panchal et al., 2022). Proprietary quality control procedures are applied to the data during calibration by Earthsense. The Zephyr $PM_{2.5}$ sensor is an optical particle counter characterised by a specified accuracy of ±5 μg/m^3.

2.1.5. Low-cost sensor collocation and calibration

The sensors were calibrated by collocating them alongside a TSI DustTrak® II aerosol monitor (model 8530) in four, 2-week long, intermittent campaigns which ran between the field campaigns. These two-week sampling campaigns were conducted during the summer and winter months to capture and compare potential peak air pollution levels across distinct seasonal intervals. The sensors were installed indoors in one of the researchers' dwellings where conditions are similar to those they would measure in the field-the interquartile range (IQR) of indoor temperatures measured by the Bonolos in colocation was between 21 °C and 25 °C, while in the field data it was between 21 °C and 28 °C. The IQR of relative humidity was between 22% and 49% in colocation and 33% and 54% for colocation. These ranges show that the distribution of these variables was reasonably consistent between collocation and field campaigns. The DustTrak monitor is an active sampling, light scattering instrument and was factory calibrated using Arizona road test dust (ISO 12103-1) before the campaign had begun. The specified range for the DustTrak is 1–400,000 μg/m^3 (TSI, 2023). Calibration coefficients were calculated for individual sensors by fitting a multiple linear regression (MLR) to DustTrak data. Multiple linear regression was chosen for the calibration models to control for temperature and relative humidity components, as these have been shown to have a significant effect on low-cost sensor performance (Malings et al., 2020). Outliers were removed based on data that were three median absolute deviations from the mean, as has been done in other low-cost sensor studies (Lu et al., 2021). Colocation data was split, such that the models were trained on 80% of the data, then subsequently tested on the remaining 20%. The resulting calibration models were based on Equation (1) and data for the calibration coefficients and metrics of the regression models are presented in the supplementary materials (Table S1), where $\boldsymbol{\beta}$ are the model coefficients.

$$DustTrak\ PM_{2.5} = \beta_0 + \beta_1\ (Bonolo\ PM_{2.5}) + \beta_2\ (Bonolo\ Temperature) + \beta_3\ (Bonolo\ RH\%) \tag{1}$$

The calibration models were then used to convert $PM_{2.5}$ data collected from sensors placed in dwellings to adjusted $PM_{2.5}$ concentrations. Fig. 2 shows the underestimation of Bonolo $PM_{2.5}$ compared to the DustTrak $PM_{2.5}$ during colocation. Assessment of the calibration models was conducted using guidance from the United States Environmental Protection Agency (US EPA) (National Exposure Research Laboratory, 2018). The calibration model's prediction errors were relatively low, with a mean cross validation RMSE of 7.9 μg/m^3 (range: 5.4–9.7) across all calibration models. A mean cross validation R^2 of 0.91 (range: 0.87–0.95) indicates that the calibration models explained a high proportion of the variability. Mean normalized bias (MNB) indicated that our models may tend to overestimate, given that the MNB averaged across all calibration models was +14% (range: +10% to +19%). However, these fit within the range described by the US EPA air sensor guideline for measuring "personal exposure" (Williams et al., 2014). Overall, the performance metrics indicate that the models perform well in making accurate predictions and are comparable to those of previous studies (Hong et al., 2021; McFarlane et al., 2021).

2.2. Statistical methods

2.2.1. Household air pollution

We conducted exploratory data analysis, investigating time series and diurnal patterns of indoor $PM_{2.5}$, and made comparisons to outdoor concentrations. Mann-Whitney U tests were employed to assess differences in measured outdoor $PM_{2.5}$ concentrations between the study settings. Comparison to the South African National Ambient Air Quality Standards (NAAQS) and the World Health Organization Air Quality Guideline Values (WHO AQG) were made using boxplots and considering the proportion of days in the study period where these thresholds were exceeded. The most recent 24-h NAAQS for $PM_{2.5}$ is 40 μg/m^3 (South African Department of Environmental Affairs (DEA), 2012) and

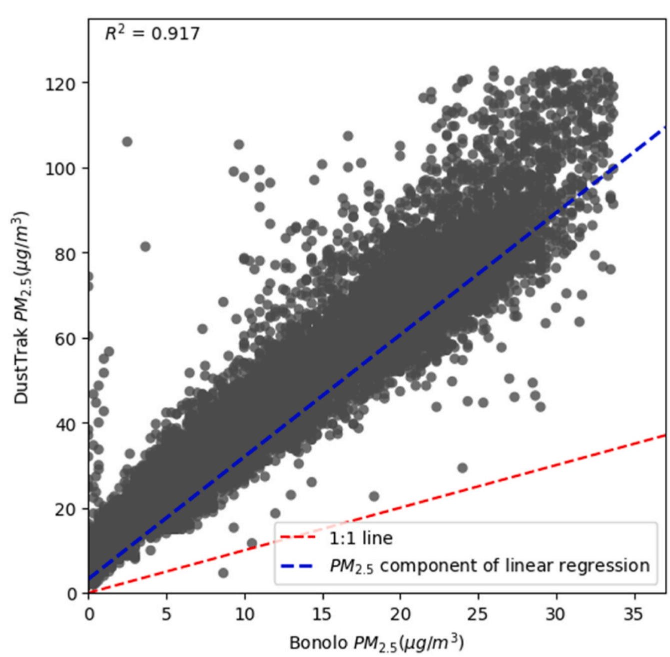

Fig. 2. Scatterplot showing the $PM_{2.5}$ component of one of the calibration models developed (sensor 5; Table S1).

the 24-h WHO AQG is 15 μg/m^3, based on a maximum of 3–4 exceedance day per year (World Health Organization, 2021). To compare the hourly and daily $PM_{2.5}$ concentrations across different households, one-way ANOVA analysis was conducted. The dependent variable was the hourly/daily $PM_{2.5}$ concentration, and the independent variable was the household ID. This analysis allowed us to assess the proportion of variability in $PM_{2.5}$ concentrations that was attributable to differences between households versus variability within households. $PM_{2.5}$ concentration means presented are based on the geometric mean to account for positively skewed data and 95% confidence interval estimates are reported. For the estimates, 95% confidence intervals were computed using the t-distribution and sample statistics to indicate the likely range of the true population parameters.

2.2.2. Dwelling and household characteristics in relation to HAP

To determine the associations of dwelling and household characteristics with 24-h indoor $PM_{2.5}$ concentrations, we conducted regression analysis by developing an MLR model. Households with missing $PM_{2.5}$ data were excluded listwise and imputation methods were not considered due to the possibility that $PM_{2.5}$ sensor data was Missing Not At Random (MNAR), for example due the sensor being damaged from very high exposure to $PM_{2.5}$. To manage the observed non-normality in the dependent variable, we applied a log transformation, in line with previous related studies that have highlighted the benefits of log transformations for HAP data (Shezi et al., 2018; Shupler et al., 2022). As a result of this transformation, the model coefficients were interpreted as proportional changes to the dependent variable rather than an absolute change (Fandiño-Del-Rio et al., 2020). Several variables with multiple responses were redefined as binary variables to avoid multicollinearity and increase the utility of the analysis, since imbalances in responses and a small sample size can lead to a less robust model (Kuhn and Johnson, 2013). The responses to the Primary Heating Fuel variable were converted to a binary variable, "Used any form of heating". The sensor location variable was also redefined as a binary variable where the sensor was placed in the kitchen or not and the presence of ceiling was derived from the ceiling material variable. Furthermore, a variable was created that combined responses of "outdoor" and "both" from the cooking location variable, as there was potential that these variables shared similar information. By combining the variables, the statistical power of the model could potentially be increased due to fewer variables with small numbers per group thus leading to a more robust model. The variable "presence of tree shade" was not included in the initial model, despite meeting criteria, due to being an almost identical proxy for the study setting variable.

The process of feature selection for the MLR model was based on a systematic approach combining exploratory data analysis techniques, including the use of box plots and count plots, alongside fitting univariate models with the log-transformed 24-h indoor $PM_{2.5}$ as the dependent variable. Independent variables demonstrating statistical significance, determined by a *p*-value (F-test) below 0.2, were then used to form the foundation of the initial multivariable model, in line with methods used in previous research (Jafta et al., 2017; Fandiño-Del-Rio et al., 2020). Subsequent refinement of the multivariable model was conducted through an exhaustive regression subset selection process. This method, facilitated by the R package Leaps (v3.1), prioritised features optimising for Bayesian Information Criterion (BIC) and Mallow's Cp. The presence of multicollinearity within the model was assessed using variance inflation factors (VIF). Consistent with established guidelines (Akinwande, Dikko and Samson, 2015), independent variables with VIF values exceeding 10 were excluded from the final model. Interaction terms were explored to account for interactions between the theoretically related variables primary cooking fuel and cooking location; however, these showed signs of high multicollinearity (VIF: >10) so were excluded. Additional measures were also used to assess multicollinearity (Cramér's V and Theil's U). Our final model was evaluated through leave-one-out, 5-fold, and 10-fold cross-validation methodologies, employing root mean squared error (RMSE) and R^2 as performance metrics.

A combination of the statsmodels (v0.14.0) package (Skipper and Perktold, 2010) in Python and the Car (v3.2-2) (Fox and Weisberg, 2019) and Caret (v6.0-94) (Kuhn, 2008) packages in R were used to conduct the regression analysis. Significance testing, visualisations and models were developed in python, then models were recreated in R to conduct the regression subset selection, cross validation, and the assessment of VIF. To ensure the models were identical between programmes, we ensured that the coefficients, standard errors, F-statistics and R^2 were consistent.

3. Results

Questionnaires were administered in all 44 recruited households. Three households had missing HAP data (H1SW, H9AW & H11AW[1]) and one withdrew early resulting in only two days of data (H4AS), so our analysis is drawn from 40 households. To inform our regression model, we excluded one more household (H10SS) due to incomplete household characteristic information, resulting in a sample size of 39 households for the modelling process.

3.1. Ambient (outdoor) air pollution

During winter, the average outdoor temperatures measured by the Zephyrs were higher in the rural setting (mean: 20 °C, range: 12–25) than in the urban setting (mean: 16 °C, range: 13–19). Since we could not collect outdoor measurements for the rural setting in the summer, we cannot compare measured outdoor temperatures for this campaign. However, temperatures measured by the Bonolos showed that indoor temperatures were higher in the rural setting than in the urban setting during the summer months (Table S2). The mean daily outdoor $PM_{2.5}$ concentrations for the urban campaigns were 12 μg/m^3 (range: 10–14) for summer and 27 μg/m^3 (range: 14–69) for winter. For the rural winter campaign, the concentration was 10 μg/m^3 (range: 2–31) which was significantly lower than the outdoor concentrations of $PM_{2.5}$ in the urban winter campaign ($p = 0.001$) (Fig. 3).

Possible sources of air pollution in the urban study areas were from vehicles, local industry, household burning of solid fuels, mine dust and dust from unpaved roads. In the rural study area, air pollution sources included dust from unpaved roads, household solid fuel burning, some (but to a lesser extent) vehicle emissions and veld fires, for example.

3.2. Household air pollution

Across all households, the 24-h average NAAQS for $PM_{2.5}$ was exceeded on 52% of days measured across all households, 31% in rural households and 70% in urban households (Fig. 4). It was also exceeded on 71% of winter and 34% of summer days measured (across both communities). Likewise, the 24-h average WHO 2021 AQG for $PM_{2.5}$ was exceeded on 86% of days measured across all households, 72% in rural households and 97% in urban households. It was also exceeded on 99% of winter and 74% of summer days measured (across both communities). One-way ANOVA analysis showed that 86% of the variance for hourly averages was within households while 14% was between households, with statistically significant differences between the households (($F(39, 11735) = 49.6$, $p < 0.001$)). For 24-h averaged concentrations, 47% of the variance was within and 53% between households (($F(39, 506) = 14.3$, $p < 0.001$)).

During winter, the average outdoor temperatures measured by the Zephyrs were higher in the rural setting (mean: 20 °C, range: 12–25 °C) than in the urban setting (mean: 16 °C, range: 13–19 °C). Since we could not collect outdoor measurements for the rural setting in the summer,

[1] Household number, Agincourt/Soweto, Summer/Winter.

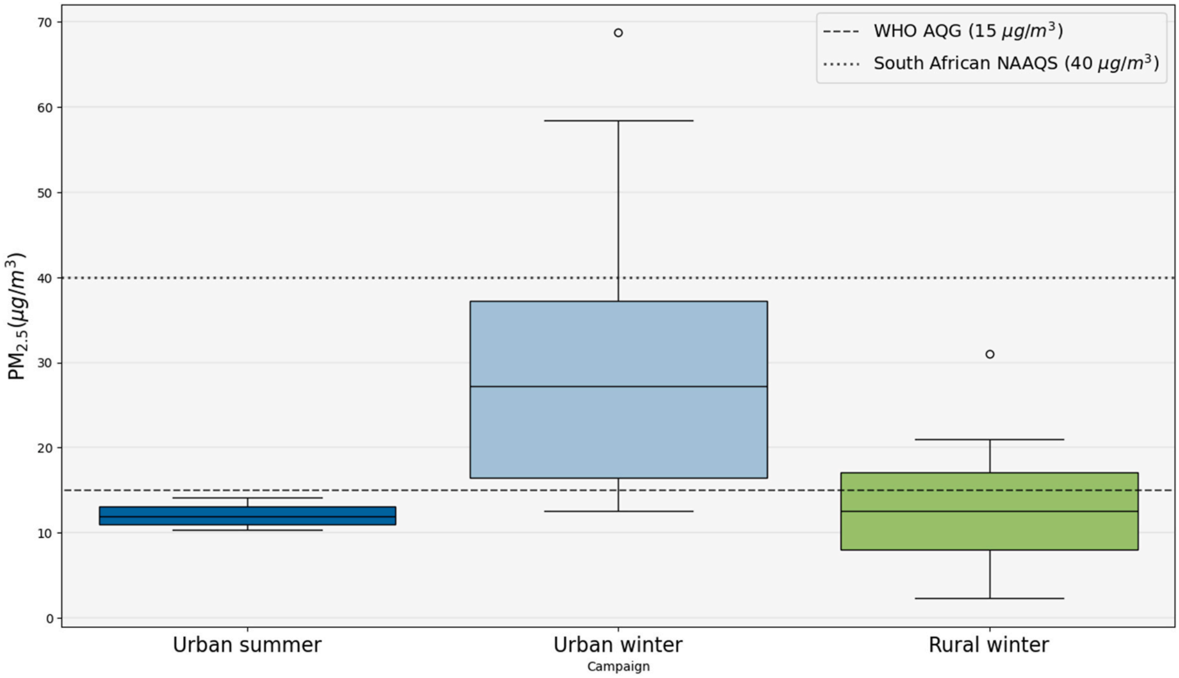

Fig. 3. Boxplots showing the 24-h ambient concentration distributions for each campaign as measured by the Zephyr's. Urban summer only recorded for the last 6 days of the campaign. No data was available for the rural summer campaign.

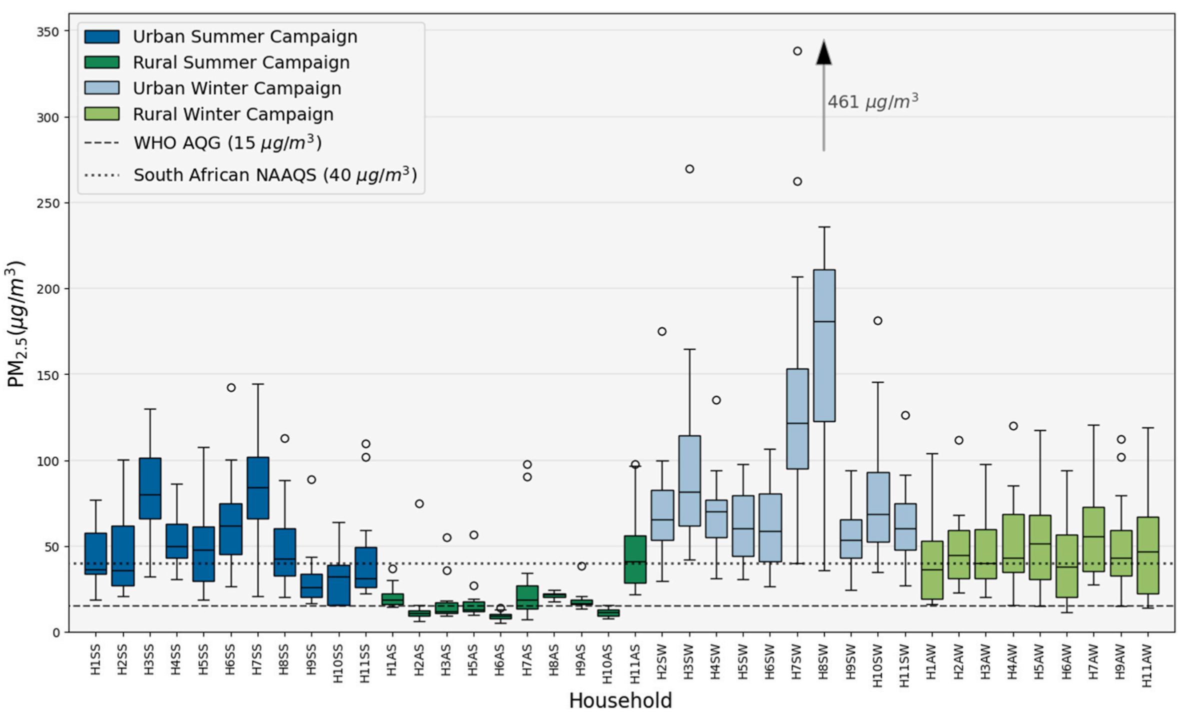

Fig. 4. Boxplots showing the 24-h concentration distributions for each dwelling compared to the South African NAAQS and the WHO AQG. Arrow shows an additional outlier not shown on the plot. Please refer to Table 2 below for the details of statistical significance by study location and season.

we cannot compare measured outdoor temperatures for this campaign. However, temperatures measured by the Bonolos showed that indoor temperatures were higher in the rural setting than in the urban setting during the summer months (Table S2).

Fig. 5 displays the diurnal trends in household $PM_{2.5}$ concentrations during the study periods at both sites. A bimodal peak is evident, especially in the urban winter campaign. Evidence of a bimodal peak is less evident for both seasons in the rural setting, gradually increasing to a peak in the evening with no clear peak in the morning.

3.3. Comparison of dwelling and household characteristics

Of the 44 households, the arithmetic mean household size was 4.0 occupants (s.d. = 2.5, median = 3.5) and the arithmetic mean number of rooms in the main dwelling was 4.4 (s.d. = 2.8, median = 4) (Table 1). Households had more members in the rural setting (median = 5) than in

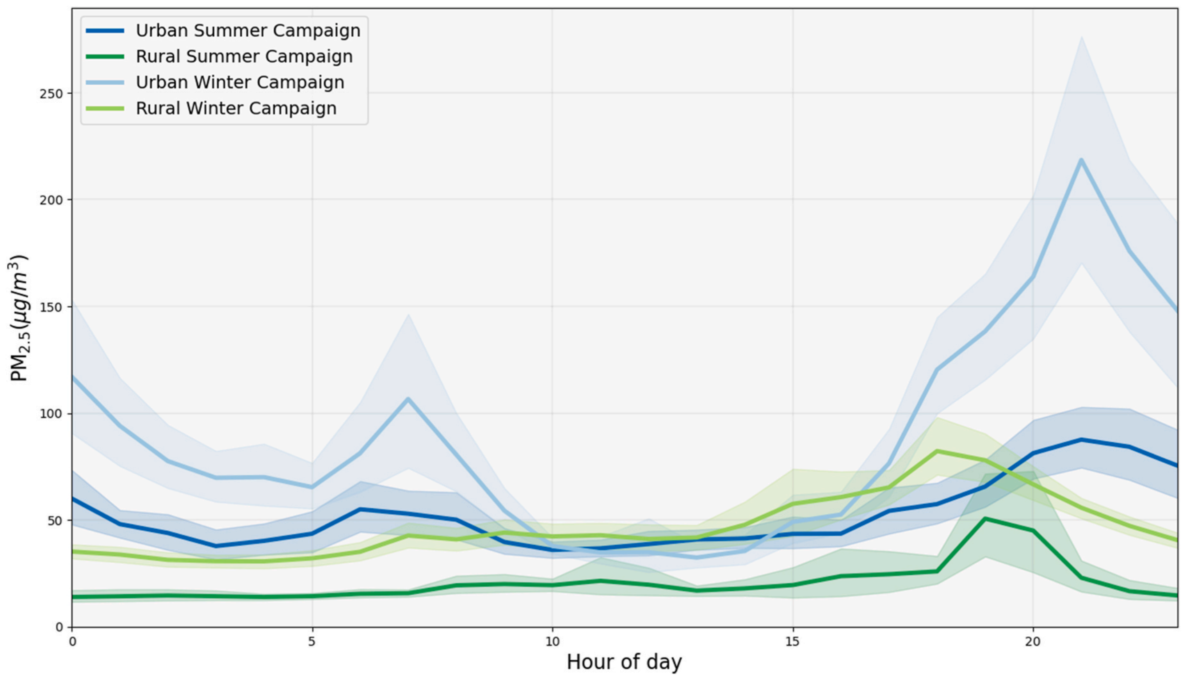

Fig. 5. Diurnal variation in indoor $PM_{2.5}$ at the two study sites in summer and winter. Shaded regions indicate 95% confidence intervals of the mean estimate. A bimodal peak is evident, especially in the urban winter (light blue line) campaign.

the urban setting (median = 2.5). All but one household was described as "dwelling/house or brick/concrete block structure", with the remaining household being described as "informal".

Table 1 shows the proportion of dwelling and household characteristics with 24-h $PM_{2.5}$ geometric means and 95% CI, stratified by each community. Environmental tobacco smoking (ETS) was observed in three (16%) rural households and six (29%) urban households. In these nine households, the $PM_{2.5}$ concentrations were approximately 14 μg/m^3 higher than all non-smoking households from both the urban and rural sample. In our sample, incense burning was only practiced in urban households. $PM_{2.5}$ concentrations were approximately 20 μg/m^3 higher in the 12 urban households that did burn incense, compared to the nine urban households that did not. Among all households in the sample, only 4 had instances of both incense burning and ETS: the average concentration in these households was 90 μg/m^3. The average of the households with both no ETS and no incense burning (N = 23) was 34 μg/m^3.

Table 2 provides the results of the univariate regression models (initial variable screening) considering log $PM_{2.5}$ in relation to dwelling and household characteristics among the rural and urban sample (n = 39). Variables in bold formed the basis of the initial multivariable model, due to the *p*-value being less than 0.2. The performance metrics for the initial model are given in Table S3.

Table 3 shows which features were selected for the final model by our optimisation process. There were five variables, namely, used any form of heating, presence of ETS and presence of incense burning, the season and location setting. Our final model was able to explain 79% of the variance in 5-fold, and 74% in both 10-fold and leave-one-out cross validation (Table S3).

4. Discussion

This study showed that household concentrations of $PM_{2.5}$ were significantly higher in the urban community compared to the rural community, underscoring a substantial difference in potential exposure between these areas. Industry and mining operations are more common in urban areas, which has a direct influence on indoor air pollution. Additionally, prior research indicates that there is a higher concentration of residential wood combustion heaters, diesel exhaust, crystal dust, and transportation in metropolitan locations (Hime et al., 2018). Daily PM_{10} and NO_2 concentrations were found to be associated with hospital admissions for respiratory diseases in Cape Town, South Africa, with NO_2 most likely coming from traffic pollution (Adebayo-Ojo et al., 2022).

ANOVA analysis indicated that hourly $PM_{2.5}$ levels are more variable within households, suggesting that indoor sources are the primary drivers of short-term peaks in HAP. Strong associations with household $PM_{2.5}$ were observed for ETS and burning incense indoors (the latter observed exclusively in the urban community). Additionally, we found strong associations with the season variables and the use of heating.

The results were consistent with existing literature in the region. Kapwata et al. (2018) found similar concentrations of PM_4 (particulate matter smaller than 4 μm in diameter) in a rural community in Limpopo province, South Africa. In terms of seasons, we found that $PM_{2.5}$ concentrations were higher in the winter than in the summer in the rural community, which was also observed in Limpopo (Kapwata et al., 2018). Gravimetric sampling, using a DustTrak, was conducted in their study, focusing on PM_4 concentrations in kitchens which helped to focus on the implications of cooking with dirty fuels; this was not the focus of our study which aimed to do an intercomparison between urban and rural sites. Another study conducted on the Highveld in South Africa that found similar concentrations of PM_4 in non-solid fuel burning household and a solid fuel burning household (Adesina et al., 2020) also observed lower concentrations in the summer than in the winter. We also observed higher concentrations of $PM_{2.5}$ in the winter compared to the summer in the urban community. In the urban setting, diurnal patterns showed the typical double peaks characteristic of urban morning and evening higher air pollution concentrations in both seasons. In the rural setting, there was only a single peak in the evening observed in both seasons which is in contrast to other studies in rural South Africa that observed a double peak in $PM_{2.5}$ concentrations (Kapwata et al., 2018; Adesina et al., 2020). This is potentially due to a much smaller volume of traffic in the total community compared to the urban community since morning rush traffic is a primary driver of $PM_{2.5}$ (Jeong et al., 2019).

ANOVA analysis indicated that hourly $PM_{2.5}$ levels are more variable within households, suggesting that short-term peaks in HAP may be primarily driven by indoor sources such as combustion-related

Table 1
Proportion of dwelling and household characteristics with 24-h $PM_{2.5}$ geometric means and 95% CI, stratified by each study setting.

Variable	Response	Both communities (n = 40)		Urban (n = 21)		Rural (n = 19)	
		n (%)	$PM_{2.5}$ (μg/m^3)	n (%)	$PM_{2.5}$ (μg/m^3)	n (%)	$PM_{2.5}$ (μg/m^3)
Season	Summer	21 (52)	32 (24–42)	11 (52)	50 (41–63)	10 (53)	19 (14–26)
	Winter	19 (48)	64 (53–77)	10 (48)	82 (62–109)	9 (47)	48 (44–53)
	Total	40 (100)	44 (36–54)	21 (100)	64 (52–77)	19 (100)	29 (22–39)
Sensor location[a]	Dining room	17 (42)	35 (25–48)	6 (29)	58 (49–68)	11 (58)	26 (18–39)
	Kitchen	11 (28)	64 (48–86)	8 (38)	71 (48–106)	3 (16)	49 (45–54)
	Bedroom	8 (20)	41 (23–75)	4 (19)	62 (30–126)	4 (21)	28 (9–88)
	Sitting room	4 (10)	49 (15–155)	3 (14)	58	1 (5)	29
	Missing	0					
Primary cooking fuel[a]	Electricity	29 (74)	48 (38–61)	18 (90)	65 (53–81)	11 (58)	29 (20–43)
	Gas	3 (8)	65	2 (10)	71	1 (5)	54
	Wood	7 (18)	27 (15–48)	0 (0)	N/a	7 (37)	27 (15–48)
	Missing	1					
Primary heating fuel	Electricity	18 (45)	57 (45–72)	16 (76)	62 (50–77)	2 (11)	28
	Wood	6 (15)	50 (47–54)	0 (0)	N/a	6 (32)	50 (47–54)
	Gas	3 (8)	58	2 (10)	61	1 (5)	54
	Coal	1 (3)	83	1 (5)	83	0 (0)	N/a
	No heating	12 (30)	25 (16–39)	2 (10)	67	10 (53)	20 (14–29)
	Any form of heating[b]	28 (70)	56 (48–66)	19 (90)	63 (53–76)	9 (47)	44 (43–57)
	Missing	0					
Cooking location[a]	Indoor	19 (49)	61 (47–78)	18 (90)	65 (53–81)	1 (5)	18
	Outdoor	6 (15)	35 (18–68)	1 (5)	82	5 (26)	29 (15–57)
	Both	14 (36)	32 (23–46)	1 (5)	62	13 (68)	31 (21–45)
	Any cooking outdoors[c]	20 (30)	33 (25–44)	2 (10)	71	18 (95)	30 (23–41)
	Missing	0					
Exhaust ventilation (e.g., chimney)	Yes	8 (20)	47 (27–83)	6 (29)	65 (41–100)	17 (89)	31 (23–42)
	No	32 (80)	43 (34–54)	15 (71)	63 (49–81)	2 (11)	19
	Missing	0					
Ventilation during cooking (e.g., opening windows, doors)	Yes	31 (79)	44 (34–57)	18 (90)	67 (53–83)	13 (68)	25 (18–36)
	No	8 (21)	45 (32–63)	2 (10)	56	6 (32)	42 (27–66)
	Missing	1					
ETS and Incense burning	ETS & Incense burning	4 (10)	90 (40–203)	4 (19)	90 (40–203)	0 (0)	N/a
	ETS or Incense burning	17 (42)	63 (47–83)	14 (67)	74 (58–94)	3 (16)	29
	Incense burning	12 (30)	73 (55–96)	12 (57)	73 (55–96)	0 (0)	N/a
	No incense burning	28 (70)	36 (28–45)	9 (43)	53 (40–71)	19 (100)	29 (22–39)
	ETS	9 (22)	61 (36–104)	6 (29)	88 (57–136)	3 (16)	29
	No ETS	31 (78)	47 (32–50)	15 (71)	56 (45–69)	16 (84)	30 (21–41)
	Neither	23 (57)	34 (27–43)	7 (33)	46 (36–61)	16 (84)	30 (21–41)
	Missing	0					
Number of occupants	1	6 (15)	52 (38–73)	5 (24)	53 (35–82)	1 (5)	48
	2–5	23 (57)	51 (39–67)	13 (62)	69 (52–93)	10 (53)	34 (23–51)
	6+	11 (28)	30 (19–46)	3 (14)	59 (43–80)	8 (42)	23 (14–38)
	Missing	0					
Rooms	1	8 (20)	66 (43–102)	6 (29)	74 (41–133)	2 (11)	48 (38–60)
	2–5	23 (57)	40 (30–54)	14 (67)	61 (49–76)	9 (47)	21 (13–32)
	6+	9 (22)	39 (29–54)	1 (5)	44	8 (42)	39 (27–56)
	Missing	0					
Floor material	Cement	13 (32)	40 (24–65)	4 (19)	94 (45–194)	9 (47)	27 (17–44)
	Tiles	27 (68)	47 (38–57)	17 (81)	58 (48–71)	10 (53)	32 (21–48)
	Missing	0					
Roof material	Metal sheeting	18 (45)	45 (32–62)	9 (43)	67 (46–96)	9 (47)	30 (19–48)
	Clay tiles	9 (22)	28 (17–44)	0 (0)	N/a	9 (47)	28 (17–44)
	Asbestos	8 (20)	51 (40–65)	8 (38)	51 (40–64)	0 (0)	N/a
	Concrete	5 (12)	78 (46–134)	4 (19)	89 (50–160)	1 (5)	47
	Missing	0					
Ceiling material	No ceiling	21 (52)	41 (29–57)	8 (38)	80 (58–110)	13 (68)	27 (18–39)
	Ceiling boards	17 (42)	47 (37–59)	11 (52)	54 (41–72)	6 (32)	36 (22–58)
	Wood	1 (2)	49	1 (5)	49	0 (0)	N/a
	Asbestos	1 (2)	83	1 (5)	83	0 (0)	N/a
	Any ceiling present[d]	19 (48)	48 (39–60)	13 (62)	55 (43–71)	6 (32)	36 (22–58)
	Missing	0					
Exterior wall material	Plastered brick	36 (90)	44 (36–53)	19 (90)	60 (50–72)	17 (89)	30 (23–40)
	Brick	2 (5)	23	0 (0)	N/a	2 (11)	23
	Metal sheeting	2 (5)	108	2 (10)	108	0 (0)	N/a
	Missing	0					
Interior wall material	Plastered brick	35 (18)	43 (35–53)	19 (90)	60 (50–72)	16 (84)	29 (22–39)
	Brick	3 (8)	30	0 (0)	N/a	3 (16)	30
	Missing	0					
Window number	0–2	9 (22)	63 (40–101)	8 (38)	81 (49–108)	1 (5)	20
	3–4	13 (32)	47 (31–71)	9 (43)	62 (43–88)	4 (21)	26 (8–81)
	5–8	9 (22)	40 (29–56)	4 (19)	52 (40–68)	5 (26)	33 (18–60)
	9–16	9 (22)	31 (19–50)	0 (0)	N/a	9 (47)	31 (19–50)
	Missing	0					
Front door faces unpaved road	Yes	5 (12)	79 (41–150)	5 (24)	79 (41–151)	0 (0)	N/a

(continued on next page)

Table 1 (*continued*)

Variable	Response	Both communities (n = 40)		Urban (n = 21)		Rural (n = 19)	
		n (%)	$PM_{2.5}$ (μg/m^3)	n (%)	$PM_{2.5}$ (μg/m^3)	n (%)	$PM_{2.5}$ (μg/m^3)
	No	35 (88)	41 (33–50)	16 (76)	59 (48–73)	19 (100)	29 (22–39)
	Missing	0					
Presence of tree shade	Yes	18 (45)	29 (21–38)	0 (0)	N/a	18 (95)	29 (21–38)
	No	22 (55)	63 (52–76)	21 (100)	64 (52–77)	1 (5)	47
	Missing	0					

Note. Where no 95% confidence interval is reported, this is due to the extremely small sample size (1–3) as it is not a reliable measure of the true variability in this context. Readers should interpret the mean with caution, recognizing the high degree of uncertainty due to small samples within each stratum.
[a] n = 39 due to missing data from one urban household, resulting in 20 urban samples.
[b] Electricity, wood, gas & coal combined.
[c] Outdoor & both combined.
[d] Ceiling boards, wood & asbestos combined.

Table 2
Results of the univariate regression models (*p*-values) considering log $PM_{2.5}$ in relation to dwelling and household characteristics among the rural and urban sample (n = 39).

Household/dwelling characteristics	*p*-value
Season (e.g., winter month, summer month)	**< 0.001**
Study setting (e.g., urban, rural)	**< 0.001**
Sensor location	0.10
Sensor in kitchen	**0.022**
Primary cooking fuel	**0.061**
Primary heating fuel	0.002
Used any form of heating	**< 0.001**
Exhaust ventilation (e.g., chimney)	0.75
Ventilation during cooking (e.g., opening windows, doors)	0.95
Cooking location	0.008
Any cooking outdoors	**0.002**
Incense burning	**< 0.001**
ETS	**0.093**
Occupants	**0.041**
Rooms	0.067
Floor material	0.42
Roof material	**0.018**
Ceiling material	0.67
Any ceiling present	0.35
Exterior wall material	**0.048**
Window number	**0.064**
Front door faces unpaved road	**0.031**
Temperature	**< 0.001** (r = −0.66, R^2 = 0.44)

Note. *P*-values are obtained from univariate linear regression analysis (F-test) against the log-transformed mean 24-h household $PM_{2.5}$ concentrations. Bold text indicates that the variable met the criteria of $p < 0.2$ and was included as an independent variable in the initial multivariable model. Italic text indicates that a binary variable was derived from the related variable and tested separately.

Table 3
Coefficients of the multivariable regression model for 24-h $PM_{2.5}$.

Variable	β (95% CI)	P>|t|	e^{β} (95% CI)
Constant	2.85 (2.66, 3.04)	<0.001	17.3 (14.3, 20.9)
Used any form of heating | Yes	0.274 (0.0204, 0.528)	0.035	1.31 (1.02, 1.70)
ETS | Yes	0.389 (0.163, 0.615)	0.001	1.47 (1.18, 1.84)
Incense burning | Yes	0.386 (0.163, 0.615)	0.006	1.47 (1.18, 1.84)
Season | Winter	0.713 (0.515, 0.911)	<0.001	2.04 (1.67, 2.49)
Study setting | Urban	0.367 (0.0916, 0.642)	0.011	1.44 (1.10, 1.90)

Note. Variable | Reference group. e^{β} = exponentiated coefficient (ratio).Exponentiated coefficients are interpreted as proportional changes to the dependent variable rather than an absolute change. For example, consider an independent variable X with a coefficient and a dependent variable Y, which is modelled as ln (Y) = βX. For this example, a one unit change in X corresponds to an eˆβ change in Y. If, for instance, β = 0.2, this translates to a ratio of eˆ0.2 ≈ 1.22, indicating that the effect of X on Y is a 1.22-fold increase for every one-unit change in X, which is equivalent to a 22% increase.

activities, which vary significantly over time. In contrast, the variance in 24-h averaged concentrations was more evenly distributed between and within households reflecting the influence of both household-specific factors (e.g., ventilation and source activities) and structural or environmental differences across households, aligning with the proportions of variance observed in hourly concentrations reported by (Jafta et al., 2017). These findings underscore the importance of targeted interventions at the household level to address HAP variability effectively.

Our regression model was able to explain a high proportion of the variance in 24-h indoor $PM_{2.5}$ concentration (with an R^2 ranging from 0.74 to 0.79 depending on the method of cross validation used). This is considerably higher than previous $PM_{2.5}$ modelling studies in South Africa; Shezi et al. (2018) (N = 300, sampling period: 24-h) achieved an R^2 of 0.50, later improving this number to 0.54 by considering a mixed-effect model and repeated seasonal measurements (Shezi et al., 2020) (N = 30, sampling period: 24-h). In addition, Jafta et al. (2017) were able to develop a model for $PM_{10,}$ which could explain 41% of the model variance (N = 105, sampling period: 24-h). Outside of South Africa, other studies have employed similar methodology, with moderate proportions of variance explained: Clark et al. (2010) developed a model which was able to explain more than 50% of the variance (N = 54, sampling period: 8-h) while Shupler et al. (2022) were able to explain 54% (N = 2384, sampling period: 48-h). Of these modelling studies, ours is the only to monitor for an extended period of 2 weeks; an increased proportion of variance explained by our model could be a result of obtaining a more representative sample of the true 24-h average concentrations. Despite reporting higher R-squared values compared to previous studies, we acknowledge the variability of R-squared estimates across different cross-validation methods. For instance, Shezi et al. (2018) reported an R^2 of 0.55 for leave-one-out cross validation, whereas for 10-fold cross validation they reported 0.86 which is greater than the R^2 of our model using the same method (0.74). Our initial model (before selecting optimising using the LEAPS package in R) performed similarly to reported values, with R^2 ranging from 0.38 to 0.55. By optimising the model, we negated multicollinearity issues, which are a hindrance to interpretable linear regression models (Vatcheva et al., 2016). When interpreting our results, it is important to emphasize that, while we do not see evidence of overfitting, the number of households measured was smaller than other modelling studies so the results should be interpreted with a degree of caution.

The seasonal component of our model had the largest effect, which was expected due to observing large seasonal differences in $PM_{2.5}$ concentrations in both the rural and urban locations. Outdoor concentrations were significantly higher during winter compared to summer in the urban community. Unfortunately, comparison between seasons was not possible for the rural community due to lack of available measured data. Thus, we cannot determine either the impact of outdoor $PM_{2.5}$ levels on HAP in Agincourt or whether any change in fuel use during summer had a significant impact on outdoor $PM_{2.5}$ levels. This highlights the need for more comprehensive seasonal data collection in future research. In addition, there is a growing need for in-depth analysis to better delineate the complex relationships between outdoor and indoor pollutant levels, which may enhance our understanding of the impacts on both environments specifically by using source apportionment and particle characterization. These seasonal effects could be due to changes in fuel usage throughout the community, leading to elevated household and outdoor concentrations (Language et al., 2016). Study setting emerged as a significant factor in our model, suggesting that there were significant differences between the communities which were not captured by the other included variables, for example socioeconomic status or the proximity of households. Previous work in South Africa has shown that population density and the number of households using solid fuels in an area can increase levels of outdoor $PM_{2.5}$ in that area (Lindeque et al., 2018; Simon D. et al., 2021). One notable observation is the widespread practice of incense burning within the urban community. Within the urban community, we discovered that concentrations were notably higher in households where incense was burned compared to those where it was not. Across both communities, higher concentrations were also observed where ETS was present. This issue was further exacerbated in households where both smoking and incense burning occurred; a 3-fold increase in $PM_{2.5}$ concentrations compared to households that had neither. Our model showed that household $PM_{2.5}$ was associated with the presence of burning incense (44% increase) and environmental tobacco smoke (also a 44% increase). Similar findings have been observed in analogous studies conducted in South Africa (Shezi et al., 2018, 2020). Shezi et al. (2020) reported an ETS coefficient of 0.56, corresponding to an exponentiated coefficient ratio of 1.75 which, while greater than what we report (1.44), is within the range of our confidence intervals.

We found that the most common primary fuel used for both cooking and heating was electricity. While electricity supply in South Africa is presently facing major challenges resulting in regular and prolonged power outages (Wiese and van der Westhuizen, 2024), people living in low-income communities are provided with an electricity subsidy (Republic of South Africa, 2024), which may be the reason for the relatively high electricity use in our study sample. Alternatively, another reason could be that while electricity was the most commonly reported primary fuel type for cooking and heating, households might not use this fuel consistently ("energy switching") or might use secondary fuels ("energy stacking"). Despite significant differences of HAP concentrations between primary cooking fuel types, it was not included in our optimised model. This is despite the use of solid fuels as being identified as a source of HAP among existing literature (Smith, Mehta and Maeusezahl-Feuz, 2004; Bonjour et al., 2013). Adesina et al. (2020) found higher concentrations of PM_4 in a solid fuel burning household (including coal and wood) compared to a non-solid fuel burning household on the South Africa Highveld. In our sample, cooking outdoors was uncommon in the urban communities, while in the rural communities, it was common for households to cook both indoors and outdoors. Many households in South Africa and Sub-Saharan Africa that do use wood for cooking often do so outdoors or in external structures (Masekela and Semenya, 2021; Shupler et al., 2024), and this may help mitigate infiltration of $PM_{2.5}$ to the main dwelling. Future work should aim to install a sensor, both inside the main dwelling and outside to investigate indoor/outdoor ratios.

We were unable to assess the association with the types of heating fuels used, since there was a low occurrence of some types of heating. To counter this, we considered households that used heating or not. We observed an association with the use of heating which is in line with existing literature, for example a multinational HAP modelling study of peri-urban communities in LMIC's which found the primary heating fuel, particularly wood, as one of the main predictors of both kitchen concentrations and personal exposure to $PM_{2.5}$ (Shupler et al., 2022). A significantly higher proportion of urban households reported using electricity for heating (76%) compared to rural households, where wood was the predominant choice (six out of nine households reported they heated their homes using wood). Among urban households, there was no substantial difference in $PM_{2.5}$ concentrations between those using heating ($n = 19$, 63 $\mu g/m^3$) and those that did not ($n = 2$, 67 $\mu g/m^3$). In contrast, rural households using heating had double the concentration of $PM_{2.5}$ ($n = 9$, 44 $\mu g/m^3$) compared to those that used no heating ($n = 10$, 20 $\mu g/m^3$). While it is unlikely that wood was burned indoors for heating, no data were collected to confirm this. If wood burning occurred outdoors, it is plausible that smoke infiltrated the dwellings. The observed association between heating fuel type and $PM_{2.5}$ levels—absent for cooking fuel—may reflect the closer proximity of heating fires to the dwellings compared to cooking fires.

During our feature selection process (univariate regression model), we noted significant differences between sensors placed in kitchens and those placed in other rooms. However, our optimisation process eliminated this variable, suggesting that these differences were adequately accounted for when controlling for other included variables (Table 3). We identified a slight negative correlation between 24-h $PM_{2.5}$ concentrations and temperature. However, it was eliminated by our optimisation process and not incorporated into our final model potentially due to redundancy with factors like heating usage and seasonal variations. These seasonal variations include phenomena like temperature inversion layers when warm air is trapped at the surface beneath cooler air, often in association with the occurrence of domestic solid fuel burning for indoor heating that is released into the air at the surface creating a temperature inversion, especially when the community is located in a valley or bowl (Matandirotya et al., 2022). Indoor heating during cold weather also contributes to HAP when heating is done using 'dirty' fuels and ventilation is reduced or removed entirely to keep temperatures indoors as warm as possible (Wang, Wang and Norbäck, 2022).

4.1. Study strengths and limitations

Our study had several strengths. We addressed a particular research gap by applying low-cost sensor for a reasonably extended indoor monitoring campaign in an under-researched setting. We measured air quality in each household for two weeks during one month of two different seasons and in two distinct contexts using the same methodology using low-cost sensors, which potentially can be applied at scale in similar settings. Our data identified seasonal differences and particularly interesting diurnal trends in the rural community which, to the best of our knowledge, have not been observed as of yet in South Africa. We conducted regression analysis using data collection from a questionnaire and observation sheet. The statistical model developed was able to explain a higher portion of the variance compared to existing HAP models in South Africa. Within the urban community, we observed that concentrations were higher in dwellings where incense was burned compared to those where it was not. Across both urban and rural communities, higher concentrations were also observed where ETS was present. This issue was further exacerbated in dwellings where both smoking and incense burning occurred with a 3-fold increase in $PM_{2.5}$ concentrations compared to households that had neither. We believe that this is one of the first air quality studies to report on incense burning as well as the presence of ETS.

There were some study limitations. Simultaneous gravimetric samples were not acquired during the colocation campaign therefore the DustTrak may not have been appropriately calibrated for the environmental context and indoor $PM_{2.5}$ composition. This meant that we did not effectively account for systematic overestimation by the DustTrak (Tasić et al., 2012; Kapwata et al., 2018; Javed and Guo, 2021). Comparisons were made with the NAAQS and WHO AQG, showing that many households exceeded these values. However, this may portray an overly **pessimistic** scenario as these guidelines are intended to be used with reference instrumentation. Additionally, the use of Arizona road test dust for factory calibration may have led to increased uncertainty in the measurements from both the DustTrak and Bonolos, given the former was employed as a reference instrument for correcting data from the latter. Our calibration models exhibited a positive mean normalized bias which, although the magnitudes align with existing EPA guidelines (National Exposure Research Laboratory, 2018), represents an additional potential source of overestimation in the corrected LCS field data. The upper range of $PM_{2.5}$ concentrations observed in our collocation data was lower than that measured in the households (Figs. 2 and 4). Consequently, extrapolation was required when applying the calibration models, introducing additional uncertainty and a potential overestimation of $PM_{2.5}$ concentrations, as a non-linear relationship may exist beyond the calibration range.

We were unable to validate our Zephyr ambient outdoor measurements due to missing data at the Diepkloof-NAQI and Jabavu-NAQI stations. A higher prevalence of missing data from the South African Ambient Air Quality Information Service has been attributed to the increase in loadshedding faced by the country (Wright et al., 2023). We were also unable to provide ambient outdoor measurements for the rural summer campaign due to logistical issues. In the urban setting, the summer ambient outdoor monitor was only installed for six days. The Zephyr sensor, like the DustTrak, has a limitation as it is a light scattering sensor, and thus benefits from field calibration (Wang et al., 2016). However, calibrating Zephyr sensors requires colocation at an ambient outdoor reference site and providing the data to the manufacturer for calibration, which was not feasible in this study. Consequently, while outdoor measurements in this study help provide context for the HAP measurements, they may not accurately represent the ambient air pollution in the study locations as the colocation was completed in the UK under different environmental conditions where factors such as relative humidity, temperature and $PM_{2.5}$ composition differ. Additionally, we were not able to assess the true outdoor contribution to the indoor $PM_{2.5}$ levels via indoor/outdoor ratios due to the calibration methodology for the Zephyr instrument. Comparisons between the Bonolo and Zephyr sensors would have been too uncertain to draw any reasonable conclusions.

While efforts were made to ensure a large sample within the constraints of the study context, the sample size may still be considered relatively small compared to the broader population and may not be fully representative. The sample size was also reduced since we did not impute missing values. By removing values, the statistical power of the analysis was reduced but may have been more representative of the true sample. Future studies should aim to have more sensors to increase the households included in the study. This study demonstrated the feasibility and potential utility of using low-cost sensors for HAP monitoring in South Africa. The small sample size also limited aspects of our analysis; a larger sample would allow for a greater range of variables to be considered when creating a statistical model. Our model needed to be simplified to retain robustness under the constraints of a small sample. A larger sample may have resulted in additional characteristics showing significant associations with HAP. Additionally, our small sample size and any potential biases in our questionnaire and dwelling checklist may limit the generalisability of our regression model.

We did not find an association with the primary cooking fuel type, despite the use of solid fuels being a common source of HAP. This is potentially due to several reasons. Firstly, many households used electricity as their primary fuel but may have used a solid fuel as a secondary fuel. Our questionnaire did not adequately capture the potential for "energy stacking." For example, we lacked information on whether households used solid fuels during certain parts of the cooking process while primarily relying on electricity. Consequently, variations in $PM_{2.5}$ levels arising from secondary fuel use could not be accounted for in our regression model. This limitation also reduced the statistical power of the primary fuel variable, as the unmeasured contributions of secondary fuels introduced additional variability, potentially obscuring the true relationship between primary fuel use and indoor air pollution levels. Secondly, households that did cook with solid fuels, for example wood, likely cooked outdoors or in an external structure and concentrations inside the dwelling may not have increased. Due to our low sample size, we were unable to investigate the interaction between cooking location and primary cooking fuel in our regression analysis. Incorporating personal exposure measurements would have enhanced our comprehension of the impact of solid fuel usage on HAP exposure. Our measurements were confined to the primary living space, which may not necessarily coincide with the area of highest exposure due to cooking activities.

5. Conclusions

This study aimed to evaluate HAP in two distinct contexts measured in summer and winter in South Africa. To the best of our knowledge, this is the first time low-cost $PM_{2.5}$ sensors have been used for HAP monitoring on this scale and in two distinct communities (i.e., urban and rural) in the same study in South Africa. We found that HAP concentrations in many dwellings were above the South African National Ambient Air Quality Standards. Exposure to air pollution is a risk factor for many diseases and HAP is a particular concern, especially in LMICs. In our study of unique urban/rural settings in South Africa, HAP was found to be higher in urban dwellings, with the winter period being particularly problematic when ambient outdoor concentrations are higher and the need for heating increases due to lower temperatures. Our study demonstrated that low-income communities, with limited access to cleaner fuels are disproportionately affected by HAP. Awareness of the health impacts of incense burning and tobacco smoking within dwellings may be limited in these communities. Emphasizing education of the health benefits of reducing these practices is crucial, as they represent relatively straightforward interventions compared to more difficult to implement alternatives. This study identifies two major contributors to HAP that are relatively easy to prevent-it is imperative that community and policy maker awareness is raised about the need to avoid smoking and burning incense indoors, and to encourage the use of clean fuels for heating, especially in winter when meteorological conditions lead to stagnation of cold air worsening the air pollution effects.

CRediT authorship contribution statement

Matthew Benyon: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Ngwako Kwatala:** Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation. **Tracey Laban:** Writing – original draft, Formal analysis. **Thandi Kapwata:** Writing – review & editing, Methodology, Conceptualization. **Chiara Batini:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **Samuel Cai:** Writing – review & editing, Methodology, Conceptualization. **Lisa K. Micklesfield:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **Rikesh Panchal:** Writing – review & editing, Resources, Methodology, Conceptualization. **Siyathemba Kunene:** Writing – review & editing, Methodology, Investigation. **Sizwe B. Zondo:** Writing – review & editing, Methodology, Investigation. **Brigitte Language:** Writing – review & editing, Methodology, Conceptualization. **Bianca Wernecke:** Writing – review & editing, Methodology, Conceptualization. **Scott Hazelhurst:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **F. Xavier Gómez-Olivé:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **Joshua Vande Hey:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Caradee Y. Wright:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Conceptualization.

Data availability

The data presented in this study is under use for additional research under the LEAP-EPI project. Data may be made available to researchers upon a reasonable request.

Funding information

This work was supported in part by Royal Academy of Engineering GCRF grant FF\1920\1\52. J. We acknowledge support from the National Institute for Health Research (NIHR) Health Protection Research Unit (HPRU) in Environmental Exposures and Health at the University of Leicester development award, a partnership between the UK Health Security Agency, the Health and Safety Executive and the University of Leicester. SH is partially supported by the SA National Research Foundation (RA201216579883). The views expressed are those of the author (s) and not necessarily those of the funding agencies. The funders had no role in the study design, analysis or interpretation of data, decision to publish, or preparation of the manuscript.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Joshua Vande Hey reports financial support was provided by Royal Academy of Engineering. Joshua Vande Hey reports financial support was provided by National Institute for Health Research. Scott Hazelhurst reports financial support was provided by SA National Research Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to express gratitude to Natasha Naidoo for their invaluable assistance in the editing and proofreading of this manuscript at the revisions stage.

For the purpose of open access, the author has applied a Creative Commons Attribution license (CC BY) to any Author Accepted Manuscript version arising from this submission.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apr.2025.102459.

References

Adebayo-Ojo, T.C., et al., 2022. Short-term joint effects of PM_{10}, NO_2 and SO_2 on cardio-respiratory disease hospital admissions in Cape Town, South Africa. Int. J. Environ. Res. Publ. Health 19, 495. https://doi.org/10.3390/ijerph19010495. Available at:

Adesina, J.A., et al., 2020. Contrasting indoor and ambient particulate matter concentrations and thermal comfort in coal and non-coal burning households at South Africa Highveld. Sci. Total Environ. 699, 134403. https://doi.org/10.1016/j.scitotenv.2019.134403. Available at:

Adeyemi, A., et al., 2021. Source apportionment of fine atmospheric particles using positive matrix factorization in Pretoria, South Africa. Environ. Monit. Assess. 193 (11), 716. https://doi.org/10.1007/s10661-021-09483-3. Available at:

Akinwande, M.O., Dikko, H.G., Samson, A., 2015. Variance inflation factor: as a condition for the inclusion of suppressor variable(s) in regression analysis. Open J. Stat. 5 (7), 754–767. https://doi.org/10.4236/ojs.2015.57075. Available at:

Ali, S.A., et al., 2018. Genomic and environmental risk factors for cardiometabolic diseases in Africa: methods used for Phase 1 of the AWI-Gen population cross-sectional study. Glob. Health Action 11 (Suppl. 2), 1507133. https://doi.org/10.1080/16549716.2018.1507133. Available at:

Asher, M., et al., 1995. International study of asthma and Allergies in childhood (ISAAC): rationale and methods. Eur. Respir. J. 8 (3), 483–491. https://doi.org/10.1183/09031936.95.08030483. Available at:

Bonjour, S., et al., 2013. Solid fuel use for household cooking: country and regional estimates for 1980–2010. Environ. Health Perspect. 121 (7), 784–790. https://doi.org/10.1289/ehp.1205987. Available at:

Clark, M.L., et al., 2010. Indoor air pollution, cookstove quality, and housing characteristics in two Honduran communities. Environ. Res. 110 (1), 12–18. https://doi.org/10.1016/j.envres.2009.10.008. Available at:

Earthsense(no date) Zephyr® Air Quality Monitor. Available at: https://www.earthsense.co.uk/zephyr (Accessed: 8 July 2024).

European Commission, 2003. Indoor air pollution: new EU research reveals higher risks than previously thought. https://ec.europa.eu/commission/presscorner/detail/en/ip_03_1278. (Accessed 8 January 2025).

Fandiño-Del-Rio, M., et al., 2020. Household air pollution exposure and associations with household characteristics among biomass cookstove users in Puno, Peru. Environ. Res. 191, 110028. https://doi.org/10.1016/j.envres.2020.110028. Available at:

Fox, J., Weisberg, S., 2019. An R Companion to Applied Regression, Third. Sage, Thousand Oaks (CA). Available at: https://socialsciences.mcmaster.ca/jfox/Books/Companion/. (Accessed 25 June 2024).

Fuller, R., et al., 2022. Pollution and health: a progress update. Lancet Planet. Health 6 (6), e535–e547. https://doi.org/10.1016/S2542-5196(22)00090-0. Available at:

González-Martín, J., et al., 2021. A state–of–the-art review on indoor air pollution and strategies for indoor air pollution control. Chemosphere 262, 128376. https://doi.org/10.1016/j.chemosphere.2020.128376. Available at:

Hime, N.J., Marks, G.B., Cowie, C.T., 2018. A comparison of the health effects of ambient particulate matter air pollution from five emission sources. Int. J. Environ. Res. Publ. Health 15, 1206. https://doi.org/10.3390/ijerph15061206. Available at:

Harris, P.A., et al., 2009. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J. Biomed. Inf. 42 (2), 377–381.

Hong, G.-H., et al., 2021. Long-term evaluation and calibration of three types of low-cost PM2.5 sensors at different air quality monitoring stations. J. Aerosol Sci. 157, 105829. https://doi.org/10.1016/j.jaerosci.2021.105829. Available at:

Imatech, 2022. Bonolo air quality sensor. Available at: https://iot.circor.co.za/manuals/BonoloAQS%20V4%20-%20Manual.pdf. (Accessed 1 July 2024).

Jafta, N., et al., 2017. Indoor air quality of low and middle income urban households in Durban, South Africa. Environ. Res. 156, 47–56. https://doi.org/10.1016/j.envres.2017.03.008. Available at:

Javed, W., Guo, B., 2021. Performance evaluation of real-time DustTrak monitors for outdoor particulate mass measurements in a desert environment. Aerosol Air Qual. Res. 21 (6), 200631. https://doi.org/10.4209/aaqr.200631. Available at:

Jeong, C.-H., et al., 2019. Temporal and spatial variability of traffic-related PM2.5 sources: comparison of exhaust and non-exhaust emissions. Atmos. Environ. 198, 55–69. https://doi.org/10.1016/j.atmosenv.2018.10.038. Available at:

Kahn, K., et al., 2012. Profile: Agincourt health and socio-demographic surveillance System. Int. J. Epidemiol. 41 (4), 988–1001. https://doi.org/10.1093/ije/dys115. Available at:

Kang, Y., et al., 2022. Performance evaluation of low-cost air quality sensors: a review. Sci. Total Environ. 818, 151769. https://doi.org/10.1016/j.scitotenv.2021.151769. Available at:

Kapwata, T., et al., 2018. Variation of indoor particulate matter concentrations and association with indoor/outdoor temperature: a case study in rural Limpopo, South Africa. Atmosphere 9 (4), 124. https://doi.org/10.3390/atmos9040124. Available at:

Kuhn, M., 2008. Building predictive models in *R* using the **caret** package. J. Stat. Software 28 (5). https://doi.org/10.18637/jss.v028.i05. Available at:

Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer New York, New York, NY. https://doi.org/10.1007/978-1-4614-6849-3. Available at:

Language, B., et al., 2016. Household Air Pollution in South African Low-Income Settlements: a Case Study, pp. 227–236. https://doi.org/10.2495/AIR160211. Available at:

Lindeque, F., et al., 2018. Land use regression as a tool to quantify $PM_{2.5}$ concentrations and identify suitable offset sites on the Highveld. In: Conference of the National Association for Clean Air. Vanderbijlpark, South Africa, 30 October – 1 November 2018. Available at: https://www.naca.org.za/uploads/NACA_Programme_2018.pdf. (Accessed 8 July 2024).

Lu, Y., Giuliano, G., Habre, R., 2021. Estimating hourly PM2.5 concentrations at the neighborhood scale using a low-cost air sensor network: A Los Angeles case study. Environ. Res. 195, 110653. https://doi.org/10.1016/j.envres.2020.110653. Available at:

Malings, C., et al., 2020. Fine particle mass monitoring with low-cost sensors: corrections and long-term performance evaluation. Aerosol. Sci. Technol. 54 (2), 160–174. https://doi.org/10.1080/02786826.2019.1623863. Available at:

Matandirotya, N.R., et al., 2022. State of ambient air quality in a low-income urban settlement of South Africa. Science Africa 16, e01201. https://doi.org/10.1016/j.sciaf.2022.e01201. Available at:

Masekela, M.E., Semenya, K., 2021. Factors influencing the use of firewood post-electrification in rural South Africa: the case of Ga-Malahlela village. J. Energy South Afr. 32 (3). https://doi.org/10.17159/2413-3051/2021/v32i3a7781. Available at:

McFarlane, C., et al., 2021. First measurements of ambient PM2.5 in kinshasa, Democratic Republic of Congo and Brazzaville, Republic of Congo using field-calibrated low-cost sensors. Aerosol Air Qual. Res. 21 (7), 200619. https://doi.org/10.4209/aaqr.200619. Available at:

Muyemeki, L., et al., 2021. Source apportionment of ambient PM10−2.5 and PM2.5 for the vaal triangle, South Africa. South Afr. J. Sci. 117 (5/6). https://doi.org/10.17159/sajs.2021/8617. Available at:

National Exposure Research Laboratory, 2018. How to Evaluate Low-Cost Sensors by Collocation with Federal Reference Method Monitors. United States Environmental Protection Agency. Available at: https://www.epa.gov/sites/default/files/2018-01/documents/collocation_instruction_guide.pdf. (Accessed 8 July 2024).

Nkosi, C.N., et al., 2017. Variability of domestic burning habits in the South African Highveld: a case study in the KwaDela Township (April 2017). In: 2017 International Conference on the Domestic Use of Energy (DUE). IEEE, pp. 23–29. https://doi.org/10.23919/DUE.2017.7931820. Available at:

Office for National Statistics (ONS), 2014. Families and households statistics explained. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/articles/familiesandhouseholdsstatisticsexplained/2021-03-02. (Accessed 6 August 2024).

Panchal, R., et al., 2022. Personal air pollution exposure during morning commute car and active transport journeys. J. Transport Health 26, 101365. https://doi.org/10.1016/j.jth.2022.101365. Available at:

Paulin, L.M., et al., 2013. Indoor air quality in central appalachia homes impacted by wood and coal use. J. Environ. Protect. 4 (1), 67–71. https://doi.org/10.4236/jep.2013.41007. Available at:

Peterson, P., et al., 2017. Practical use of metal oxide semiconductor gas sensors for measuring nitrogen dioxide and ozone in urban environments. Sensors 17 (7), 1653. https://doi.org/10.3390/s17071653. Available at:

Ramsay, M., et al., 2016. H3Africa AWI-Gen Collaborative Centre: a resource to study the interplay between genomic and environmental risk factors for cardiometabolic diseases in four sub-Saharan African countries. Global Health, Epidemiology and Genomics 1, e20. https://doi.org/10.1017/gheg.2016.17. Available at:

Raysoni, A.U., et al., 2023. A review of literature on the usage of low-cost sensors to measure particulate matter. Earth 4 (1), 168–186. https://doi.org/10.3390/earth4010009. Available at:

Republic of South Africa, 2024. How do I access free basic municipal services? Available at: https://www.gov.za/faq/government-services/how-do-i-access-free-basic-municipal-services. (Accessed 9 April 2024).

Roomaney, R.A., Cairncross, E., et al., 2022a. Estimating the burden of disease attributable to ambient air pollution (ambient PM2.5 and ambient ozone) in South Africa for 2000, 2006 and 2012. S. Afr. Med. J. 705–717. https://doi.org/10.7196/SAMJ.2022.v112i8b.16483. Available at:

Roomaney, R.A., Wright, C.Y., et al., 2022b. Estimating the burden of disease attributable to household air pollution from cooking with solid fuels in South Africa for 2000, 2006 and 2012. S. Afr. Med. J. 718–728. https://doi.org/10.7196/SAMJ.2022.v112i8b.16474. Available at:

Salvi, S., Barnes, P.J., 2010. Is exposure to biomass smoke the biggest risk factor for COPD globally? Chest 138 (1), 3–6. https://doi.org/10.1378/chest.10-0645. Available at:

Semmens, E.O., et al., 2015. Indoor particulate matter in rural, wood stove heated homes. Environ. Res. 138, 93–100. https://doi.org/10.1016/j.envres.2015.02.005. Available at:

Shezi, B., et al., 2018. Developing a predictive model for fine particulate matter concentrations in low socio-economic households in Durban, South Africa. Indoor Air 28 (2), 228–237. https://doi.org/10.1111/ina.12432. Available at:

Shezi, B., et al., 2020. Predictors of urban household variability of indoor PM $_{2.5}$ in low socio-economic communities. Environ. Sci. J. Integr. Environ. Res.: Process. Impacts 22 (6), 1423–1433. https://doi.org/10.1039/D0EM00035C. Available at:

Shupler, M., et al., 2022. Multinational prediction of household and personal exposure to fine particulate matter ($PM_{2.5}$) in the PURE cohort study. Environ. Int. 159, 107021. https://doi.org/10.1016/j.envint.2021.107021. Available at:

Shupler, M., et al., 2024. Household concentrations and female and child exposures to air pollution in peri-urban sub-Saharan Africa: measurements from the CLEAN-Air (Africa) study. Lancet Planet. Health 8 (2), e95–e107. https://doi.org/10.1016/S2542-5196(23)00272-3. Available at:

Simon D, M., et al., 2021. Intra-urban variability of PM2. 5 in a dense, low-income settlement on the South African Highveld. Clean Air J. 31 (1), 1–9.

Singh, A., et al., 2024. Impacts of daily household activities on indoor particulate and NO_2 concentrations; a case study from oxford UK. Heliyon 10 (15), e34210. https://doi.org/10.1016/j.heliyon.2024.e34210. Available at:

Skipper, S., Perktold, J., 2010. statsmodels: econometric and statistical modeling with python. In: Proceedings of the 9th Python in Science Conference.

Smith, R.S., Mehta, S., Maeusezahl-Feuz, M., 2004. Indoor air pollution from household use of solid fuels. In: Ezzati, M., et al. (Eds.), Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors. World Health Organisation, Geneva, pp. 1435–1494.

South African Department of Environmental Affairs (DEA), 2012. National Environmental Management: air Quality Act: national ambient air quality standard for particulate matter with aerodynamic diameter less than 2.5 micron metres (PM 2.5). Available at: https://www.gov.za/sites/default/files/gcis_document/201409/35463gon486.pdf. (Accessed 9 April 2024).

Statistics South Africa (no date). [Information about Soweto in the city of Johannesburg]. Available at: https://www.statssa.gov.za/?page_id=4286&id=11317. (Accessed 5 July 2024).

Tasić, V., et al., 2012. Comparative assessment of a real-time particle monitor against the reference gravimetric method for PM_{10} and $PM_{2.5}$ in indoor air. Atmos. Environ. 54, 358–364. https://doi.org/10.1016/j.atmosenv.2012.02.030. Available at:

TSI, 2023. Dustrak II aerosol monitor model 8530/8530EP/8532: operation and Service manual. Available at: https://tsi.com/getmedia/7c608b93-b6d1-459a-a6a8-2b0e2a55ba91/8530-8531-8532-DustTrak_II-6001893-web?ext=.pdf. (Accessed 1 July 2024).

United Nations, 2024. World Economic Situation and Prospects as of Mid-2024. E/2024/56. New York.

Vatcheva, K.P., et al., 2016. Multicollinearity in regression analyses conducted in epidemiologic studies. Epidemiology 6 (2), 227. https://doi.org/10.4172/2161-1165.1000227. Available at:

Wang, C., Wang, J., Norbäck, D., 2022. A systematic review of associations between energy use, fuel poverty, energy efficiency improvements and health. International Journal of Environmental Research and Public Health, 16 19 (12), 7393. https://doi.org/10.3390/ijerph19127393. Available at:

Wang, Z., et al., 2016. Comparison of real-time instruments and gravimetric method when measuring particulate matter in a residential building. J. Air Waste Manag. Assoc. 66 (11), 1109–1120. https://doi.org/10.1080/10962247.2016.1201022. Available at:

Wernecke, B., et al., 2021. Opportunities for the application of low-cost sensors in epidemiological studies to advance evidence of air pollution impacts on human health. Clean Air J. 31 (1), 1–4. https://doi.org/10.17159/caj/2021/31/1.11219. Available at:

Wernecke, B., et al., 2024. Fuel switching and energy stacking in low-income households in South Africa: a review with recommendations for household air pollution exposure research. Energy Res. Social Sci. 109, 103415. https://doi.org/10.1016/j.erss.2024.103415. Available at:

Wiese, M., van der Westhuizen, L.-M., 2024. Impact of planned power outages (load shedding) on consumers in developing countries: evidence from South Africa. Energy Policy 187, 114033. https://doi.org/10.1016/j.enpol.2024.114033. Available at:

Williams, R., Kilaru, V., Snyder, E., Kaufman, A., Dye, T., Rutter, A., Russell, A., Hafner, H., 2014. Air Sensor Guidebook. US Environmental Protection Agency. EPA/600/R-14/159 (NTIS PB2015-100610).

World Bank (no date)World Bank Country and Lending Groups. Available at: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups (Accessed: 8 July 2024).

WHO, 2021. Global Air Quality Guidelines: Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide. World Health Organization, Geneva. Licence: CC BY-NC-SA 3.0 IGO.

World Health Organization, 2023. Household air pollution. Available at: https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health. (Accessed 22 February 2024).

Wright, C.Y., et al., 2023. Data gaps will leave scientists “in the dark”: how load shedding is obscuring our understanding of air quality. South Afr. J. Sci. 119 (9/10). https://doi.org/10.17159/sajs.2023/16009. Available at: