Estimating Particulate Matter (PM) concentrations from a meteorological index for data-scarce regions: A pilot study

Anzel de Lange, *[a] Rebecca M. Garland, [a,b] Liesl L. Dyson [a]

[a] *Department of Geography, Geoinformatics and Meteorology, University of Pretoria (UP), Hatfield, 0028, South Africa*

[b] *Council for Scientific and Industrial Research (CSIR), Climate Studies, Modelling and Environmental Health Research Group, Pretoria, 0001, South Africa*

*Corresponding author

E-mail: anzel.delange@up.ac.za

Tel.: +27 82 779 8493

Abstract

In regions where air quality data are scarce or access thereto is limited, a comprehensive understanding of air pollution is hindered by a lack of emission data and ambient air pollution measurements. Therefore, in this pilot study, we assess the feasibility of estimating particulate matter (PM) mass concentrations from a meteorological index. Measured PM concentrations from air quality monitoring stations (2013–2016) situated in and around South African air pollution priority areas were analysed. Simulated meteorological parameters were used to calculate the newly-developed Air Dispersion Potential (ADP) index, which describes the meteorological potential for pollution dispersion in the atmosphere. For most conditions, there exists weak (r=0.1–0.29) to moderate (r=0.30–0.49) correlations between the ADP index and PM classes. At the three stations with adequate data availability, it was found that the ADP index was relatively successful in predicting conditions of high PM concentrations. An investigation of the effect of meteorological conditions on the diurnal variation of PM concentrations led to both the quantification of this effect, and the realization that at these diverse sites, up to 29% of variation in hourly PM concentrations can be explained by variations in meteorology. The application of the index in this way can play an important role in air quality management by quantifying the impacts of meteorological drivers on PM peaks.

Keywords: Air pollution; South Africa; Particulate Matter; pollution dispersion; Air Dispersion Potential (ADP); meteorological parameters

1. Introduction

Particulate matter (PM) pollution is an important global issue due to its effects on health and the environment. Owing to these negative impacts, countries need to set standards for and attempt to lower PM concentrations. Dangerously high concentrations of PM are especially problematic for rapidly growing and developing countries throughout the world (Cohen *et al.*, 2005; Mannucci and Franchini, 2017; Panyacosit, 2000). South Africa (SA) is also affected by high levels of ambient PM; this becomes especially clear when considering the World Health Organization (WHO) Global Urban Ambient Air Pollution Database (World Health Organization (WHO), 2016), wherein several South African sites were studied. In this analysis, the annual South African National Ambient Air Quality Standard (SA NAAQIS) for PM_{10} (40 μgm^{-3}) was exceeded by over 60% of the sites, and the annual SA NAAQIS for $PM_{2.5}$ (20 μgm^{-3}) was exceeded at almost 70% of the sites. In all cases, for all sites, the WHO annual recommendations (PM_{10} = 20 μgm^{-3}, $PM_{2.5}$ = 10μgm^{-3}), which are stricter than the SA NAAQIS, were exceeded. In many developing countries, as in SA, exposure to high concentrations of air pollution is worse in regions where population and, consequently, anthropogenic emissions are concentrated (Hersey *et al.*, 2015).

A comprehensive understanding of air pollution in SA, and other developing countries, is hindered by the lack of emission and measured air pollution concentration data. It is improbable for air quality to improve in countries where these types of data are scarce (Fajersztajn *et al.* 2014). In SA, there is no default government emission inventory, emission factors and activity data are difficult to attain, and available datasets often contain large uncertainties (Garland *et al.,* 2017; Naidoo *et al.,* 2014). Only recently has the government developed a system for the reporting and tracking of annual emissions from regulated industrial sources (*i.e.* National Atmospheric Emissions Inventory System). Air quality modelling in SA is further complicated by the diverse and numerous emission sources contributing to the air pollution problem. These include emissions from sources as diverse as industry, vehicles, biomass burning, biogenic, domestic fuel use, waste burning, and wind-blown dust.

While there are some measurements of air quality in South Africa, (e.g. Alade, 2010; Belelie *et al.,* 2019; Hersey *et al.,* 2015; Josipovic *et al.,* 2009; Witi, 2005; Wright *et al.,* 2011) in general, monitoring is limited. There is a countrywide network of air quality monitoring stations, however, making use of ground-based measured air pollution concentration data from compliance monitoring stations for research purposes in SA is challenging. This is due in part to issues such as limited spatial and temporal coverage, and varying degrees of data quality. The use of remote sensing data, e.g. estimating PM ground-level concentrations from satellite-derived Aerosol Optical Depth (AOD), provides great opportunity to estimate PM concentrations in areas with sparse measurements, at good spatial coverage (Chu et al., 2016). Although this opportunity exists, translating column AOD to surface level PM concentration for SA is not a trivial process and more research is needed (Hersey *et al.,* 2015). Many developing countries face large uncertainties in emissions and a lack of observational data, but still require an understanding of air quality in order to combat their pollution problems. Meteorological science is an important tool in the field of air pollution research. This is especially true when access to good quality air pollutant concentration and emission data are limited.

Daily PM concentrations and their diurnal cycles respond to variations in atmospheric stability, mixing depth, and local and meso-scale winds (Tyson and Preston-Whyte, 2000). Southern Africa is situated in the subtropics. General circulation over the region is dominated by a semi-permanent, subtropical high-pressure cell. The mean circulation over southern Africa is anti-cyclonic throughout the year (except at the surface) and associated with divergence and subsidence, which results in clear skies and mostly rain-free conditions (Tyson and Preston-Whyte, 2000). This circulation exists for prolonged periods and causes increased stability over the region. Consequently, very persistent absolutely stable layers, which can occur over SA throughout the year, form at distinct levels throughout the troposphere and inhibit vertical mixing (Cosijn & Tyson, 1996; Garstang et al., 1996).

The meteorology of an area, together with the characteristics of an emission source, are the two most important factors determining the way in which pollutants disperse in the atmosphere (Kanevce & Kanevce, 2006). Not considering the amount of pollutant emitted, its resultant concentration depends on vertical dispersion, horizontal movement of air, and rate of deposition (Holzworth, 1971). Since pollution concentrations in the atmosphere are heavily dependent on meteorological variables, a way to investigate and predict air pollution without the use of the emission data has led researchers to quantify the influence of atmospheric variables and meteorology on pollutant concentrations.

Numerous international studies have investigated the use of individual meteorological parameter to describe atmospheric pollutant concentrations (e.g. Grundström *et al.*, 2015; Kim *et al.*, 2005; Li *et al.*, 2017a; Alvarez *et al.*, 2018; Zhang *et al.*, 2015). Individual parameters such as temperature, vertical temperature gradient, wind speed (WS), relative humidity, surface pressure, and weather type are relevant to the air pollutant transport process, but an individual parameter is not solely responsible for the spread and dispersion of pollutants. Therefore, by making use of multiple meteorological variables, as in the index described in this paper, a more comprehensive representation of the air pollution climate is likely.

This pilot study investigates the possibility of using an index, calculated from simulated meteorological data, to estimate PM concentrations in SA, a region where emission data are not freely available and ambient PM data are limited. Defining a relationship between meteorological parameters and PM concentrations in SA could lead to the possibility of using these simulated meteorological parameters as a proxy for pollutant concentrations, or as a method to characterize variation in PM concentrations. The development of an index based on only simulated meteorological parameters has many benefits for countries lacking the capacity to forecast air pollution or the infrastructure to measure pollutant concentrations. Pollutant concentration forecasts can be vital in health and early warning systems for high-risk groups. Thus, a multi-parameter index is proposed and its performance in air quality hotspots in SA is assessed as a case study.

2. Air Dispersion Potential

The combined characterization of the ability of the atmosphere to adequately dilute and disperse any admixture is often referred to as Air Pollution Potential (APP) and was proposed by Niemeyer (1960). The criteria for APP is related to the simultaneous occurrence of conditions associated with slowly moving anticyclones (low wind speed and stable atmospheric conditions), which are forecasted to continue for at least 36 hours. In South Africa, these conditions occur mostly in winter, and are often disrupted by synoptic scale weather systems (e.g. cold fronts and cut off lows). In summer, these conditions generally do not prevail. Therefore, for the study region, APP is restrictive as the aforementioned synoptic conditions are seasonal and do not occur regularly.

In this paper, we present Air Dispersion Potential (ADP), a comprehensive and contemporary representation of the characteristics of air pollution dispersion. ADP is a joint probability distribution that considers the combined effect of relevant dynamic, thermodynamic, and turbulence processes that determine the conditions of air pollutant dispersion in the atmosphere. The ADP calculation is used to determine the potential for air to disperse pollutants based on three meteorological parameters; atmospheric stability in the form of Monin-Obukhov Length (MOL) and Mixing Height (MH), both measured in metres (m), as well as WS, which is measured in metre per second (ms^{-1}). The ADP index is calculated per hour.

The ADP index uses input information required to run a contemporary air pollution model, which includes the following dynamic factors; wind velocity ($|\vec{V}|$), appropriate atmospheric stability information (MOL (L)), height of the Planetary Boundary Layer (PBL) (H) and/or inversion height ($H_{inversion}$) (Swart, 2016). The ADP index is based on the conditional probability distribution of these parameters, which allows for constructing the relevant probability tree. Probability of the comprehensive ADP index (Eq. 1) is the multiple of probabilities for WS [$P(|\vec{V}|)$], MH [$P(H)$], and stability [$P(L)$].

Therefore, the probability tree for individual realization of ADP will be:

$$P(ADP)=P(|\vec{V}|)P(H)P(L) \qquad (1)$$

The unit of the ADP is m^3s^{-1} and its value describes how many cubic meters per second (ventilation rate) are passing through a certain point and thus, what the conditions are for the pollutants to diffuse.

The intervals for WS, MH, and MOL are quantified by proxy for very unfavourable, unfavourable, moderate, favourable, and very favourable for pollution dispersion (Table 1). The thresholds for the relevant meteorological parameters are based on a combination of existing classifications. WS was classed according to the Beaufort Wind Scale, MH classes for ADP calculation were adapted from definitions of stable and unstable PBL conditions, as defined in Seibert *et al.* (2000), and MOL was classed as in Gryning *et al.* (2007), Peña *et al.* (2010), and Sathe *et al.* (2013).

Table 1. Wind Speed (WS), Mixing Height (MH) and Monin-Obukhov Length (MOL) intervals, as well as resultant ADP value ranges.

Meteorological parameter	Very unfavourable	Unfavourable	Moderate	Favourable	Very favourable
WS	0 to 0.2 ms^{-1}	0.3 to 1.5 ms^{-1}	1.6 to 3.3 ms^{-1}	3.4 to 5.4 ms^{-1}	> 5.4 ms^{-1}
MH	0 to 200 m	> 200 to 400 m	> 400 to 500 m	> 500 to 800 m	> 800 m
MOL	10 to 200 m	200 to 500 m	> 500 or < -500 m	-200 to -500 m	-50 to -200 m
ADP	20	> 20 to 40	> 40 to 60	> 60 to < 80	>= 80

Depending on the classification of each variable, it is assigned a coefficient from which the ADP value is calculated. The coefficients for the components of the ADP index are chosen in such a way that the resultant ADP values range from 20 to 100. A value of 20 is very unfavourable, 50 is moderate and 100 represents absolutely favourable air dispersion conditions. Table 1 provides the resultant ADP values, which are then classed as; Class 1 (very unfavourable), Class 2 (unfavourable), Class 3 (moderate), Class 4 (favourable) or Class 5 (very favourable).

ADP Class 1 differs from the rest of the ADP classes in that it is a single value of 20, and not a range. This is attributed to the coefficients assigned to very unfavourable contributions of parameters in Class 1. The remaining classes (2 to 5) have ranges of ADP values so that a combination of classes for WS, MH and MOL may result in these classifications.

3. Data and methodology

3.1 Study region

All the air quality monitoring stations investigated (Table 2) are located in South Africa's industrialized regions. Their location in, or in close proximity to residential and rural areas, mines and power stations, make for elevated pollution levels. Three South African pollution Priority Areas have been declared (Fig. 1), namely the Vaal Triangle Airshed Priority Area (VTAPA), the Highveld Priority Area (HPA), and the Waterberg-Bojanala Priority Area. The declarations are based on the fact that ambient air quality standards are exceeded in these areas, or a situation exists that is causing or may cause a significant negative impact on air quality in these areas (SAAQIS, 2018). These areas require specific air quality management action, in order to rectify the situation.

Table 2. Sites for which hourly PM pollution concentration data were obtained from SAAQIS.

Site name	Latitude	Longitude	Elevation (m)	Classification	Data period	Pollutants
Xanadu	-25.75º	27.92º	1192	Residential background	2014 - 2016	PM_{10}, $PM_{2.5}$
Lephalale	-23.68º	27.72º	834	Rural, residential	2013 - 2016	PM_{10}, $PM_{2.5}$
Camden	-26.62º	30.11º	1637	Industrial background	2013 - 2016	PM_{10}, $PM_{2.5}$ (only 2016)
Zamdela	-26.84º	27.86º	1486	Industrial and residential (low-income)	2013 - 2016	PM_{10}, $PM_{2.5}$
Witbank	-25.88º	29.19º	1629	Urban, industrial, rural (low-income)	2013 - 2016	PM_{10}, $PM_{2.5}$, PM_1

Five air quality monitoring stations (Table 2) were chosen for this study based on their location (near or in the declared priority areas), site classification, types of pollutants monitored, and availability of data. Site classifications are based on the location of the monitoring station and sources of pollution that affect the station.

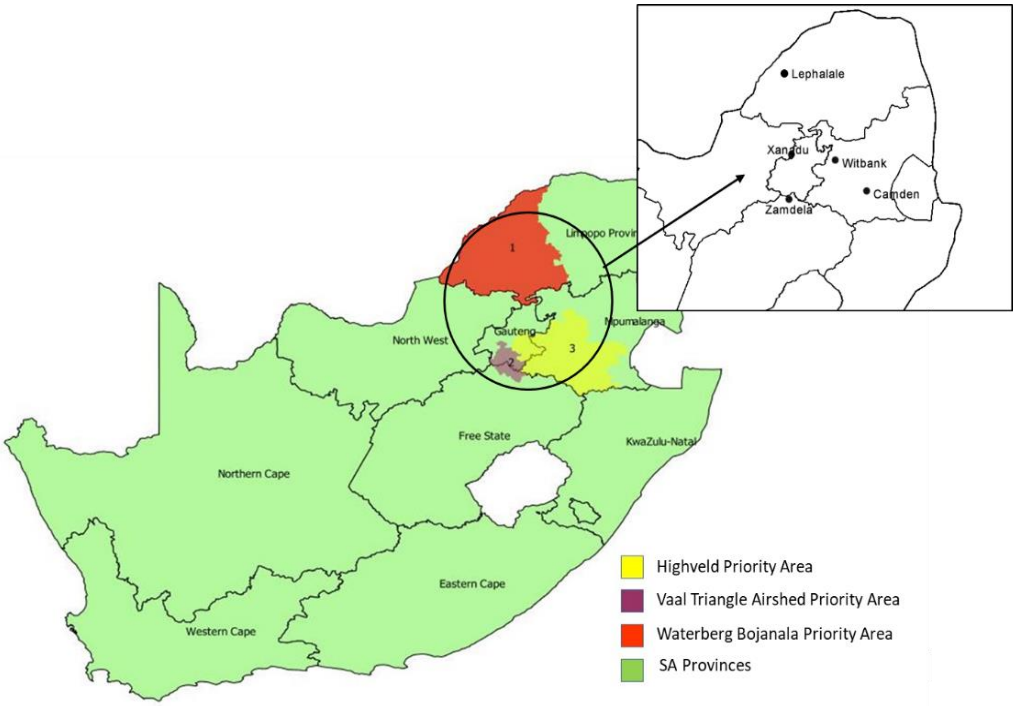

Figure 1. South African Air Quality Priority Areas; the insert shows the locations of five sites chosen for this study (Adapted from SAAQIS, 2018).

3.2 Measured hourly PM concentrations

The South African Air Quality Information System (SAAQIS) website provides researchers and stakeholders with air quality and some meteorological data at locations throughout the country. Hourly PM concentration and wind-speed data for five ambient air pollution measurement stations (Table 3) were obtained for the period 2013 to 2016 from SAAQIS. The hourly WS data were used for the verification of simulated WS produced by the meteorological model, while PM concentrations are used to investigate its relationship with the ADP index. Data Quality Control (QC) includes the removal of unrealistic values, negative values, and, repeated values; repeated values were identified as periods where more than two hourly values in a row were exactly the same (US EPA, 2017; Zahumensky, 2004). Table 3 summarises the availability of hourly PM concentration data prior to and after the QC procedure.

Table 3. Combined availability of PM_{10}, $PM_{2.5}$ and PM_1 (Witbank only) hourly data prior to, and after QC, for the entire period.

Station name	PM concentration data availability	
	Prior to QC	After QC
Xanadu	39.90%	38.77%
Lephalale	95.11%	94.03%
Camden	66.84%	65.76%
Zamdela	76.75%	66.34%
Witbank	83.56%	82.73%

Hourly WS, MOL and MH were simulated and used to calculate the ADP index. The performance of the meteorological model in the prediction of WS was assessed using the following statistical parameters: mean (M), standard deviation (SD), mean absolute error (MAE), mean squared error (MSE), and Pearson product-moment correlation coefficient (R).

Simulated MOL and MH data were not verified in this manner because these parameters are not measured at the monitoring stations. MH values are expected to be low during the night when conditions are stable, and can increase to as much as 3000 m in the noon and afternoon when convective conditions dominate (Quan *et al.,* 2013; Seibert *et al.,* 2000). MOL values are based on the same principle as PBL height, whereas unstable conditions with values less than -50 m, occur predominantly in the noon and afternoon when solar radiation peaks. MOL length values that represent stable conditions (>10 m) occur mostly during the night when the PBL height is low and atmospheric conditions are stable (Ashrafi & Hoshyaripour, 2008).

3.3 The Air Pollution Model (TAPM)

The Air Pollution Model (TAPM), developed by the Australian CSIRO Atmospheric Research Division, is a prognostic air pollution model that also simulates meteorology. Air pollution models typically require observed data from a surface meteorological station or a diagnostic wind-field model. TAPM differs from these models as its meteorological component solves the momentum equations, the incompressible continuity equation, scalar equations, the Exner pressure function, and the Poisson equation for the prediction of meteorology (Hurley, 2008a). Predicted meteorological parameters include; temperature, WS, wind direction, relative humidity, radiation, heat flux, MH, MOL and rainfall rate.

TAPM predicts the flows that are important to local-scale air pollution transport (e.g., sea breezes, terrain-induced flow, etc.) against a background of larger scale meteorology provided by synoptic analyses. The model includes parameterizations for the micro-physical processes associated with cloud and precipitation, turbulence, fluxes, and soil processes. The databases provided with TAPM include gridded global terrain height, vegetation and soil type, Leaf Area Index (LAI), sea-surface temperature, and synoptic-scale meteorology. From a summary of international verification studies presented by Hurley (2008b), it was found that the meteorological component of TAPM performs well in coastal, inland, and complex terrain, in subtropical to mid-latitude conditions, for both short periods, i.e. case studies and year-long simulations.

The meteorological component of TAPM was used to simulate hourly meteorological parameters, for the period 2013 to 2016, as required for the calculation of the ADP index. The model simulations were set up to run with four nested grids of 27 km, 9 km, 3 km and 1 km resolution. The grids were chosen in this way so that a resolution of 1 km could be achieved for the inner grid. Meteorological parameters were simulated for specific station locations in Table 2 separately. Therefore, the model output was at the location of the observational stations, and it was not necessary to adopt any interpolation scheme (other than those used by TAPM) to extract the simulated meteorology.

3.4 Evaluation of relationships and ADP index performance

In order to statistically investigate the relationship between PM concentrations and the ADP index, both variables are classed according to predetermined intervals. Categories used for PM concentrations are site specific and based on 20th percentiles of the actual measured pollutant-concentration data after quality correction. Very favourable ADP and very low PM concentrations were allocated classifications of 5, while very unfavourable ADP and very high concentrations of PM were classed as 1 (Table 4). This process assists in establishing a relationship between the two variables.

Table 4. Description of all PM concentration and ADP classifications. Class combination used to assess ADP's performance as predictor are underlined.

ADP classes		PM concentration classes 1	2	3	4	5
	1	<u>ADP very unfavourable, PM concentrations very high (CLASS 1)</u>	ADP very unfavourable, PM concentrations high	ADP very unfavourable, PM concentrations moderate	ADP very unfavourable, PM concentrations low	ADP very unfavourable, PM concentrations very low
	2	ADP unfavourable, PM concentrations very high	<u>ADP unfavourable, PM concentrations high (CLASS 2)</u>	ADP unfavourable, PM concentrations moderate	ADP unfavourable, PM concentrations low	ADP unfavourable, PM concentrations very low
	3	ADP moderate, PM concentrations very high	ADP moderate, PM concentrations high	<u>ADP moderate, PM concentrations moderate (CLASS 3)</u>	ADP moderate, PM concentrations low	ADP moderate, PM concentrations very low
	4	ADP favourable, PM concentrations very high	ADP favourable, PM concentrations high	ADP favourable, PM concentrations moderate	<u>ADP favourable, PM concentrations low (CLASS 4)</u>	ADP favourable, PM concentrations very low
	5	ADP very favourable, PM concentrations very high	ADP very favourable, PM concentrations high	ADP very favourable, PM concentrations moderate	ADP very favourable, PM concentrations low	<u>ADP very favourable, PM concentrations very low (CLASS 5)</u>

Visually examining the bivariate data with the use of a scatterplot (not shown) was the first step to find a relationship. Thereafter, the classed data were ranked and the Spearman rank correlation coefficient (r) was calculated in order to statistically measure the degree of relationship between the two variables

(Wilks, 2011). The strength of the absolute value of the Spearman rank correlation coefficient is classified as follows: Small/weak: r=0.1 to 0.29; medium/moderate: r=0.30 to 0.49; and large/strong: r=0.50 to 1.0 (Cohen, 1988).

Contingency tables were used to verify categorical forecasts and to show the joint distribution of forecast and observations in various categories (Jolliffe and Stephenson, 2012). These tables assisted in identifying under which ADP conditions PM is predicted successfully. In order to calculate statistics from contingency tables, all data was classified as hits (event forecasted and observed); misses (event not forecasted, but observed); false alarms (event forecasted, but not observed); and correct negatives (event forecasted not to occur and was not observed). When evaluating the ADP index performance, we considered statistics calculated from a multi-category contingency table as described in Table 5.

Table 5. Description, equation and ranges for the chosen categorical statistics, which include; AOF (Accuracy of Forecast), POD (Probability of Detection), SR (Success Ratio), and FAR (False Alarm Ratio). These statistics use data classified as H (Hits), M (Misses), FA (False Alarms), and CN (Correct Negatives) to evaluate forecast performance (Adapted from CAWCR, 2015; Done *et al.,* 2004).

Statistic	Description	Equation	Range	Perfect score
AOF	AOF gives the fraction of correct forecasts in each category.	AOF=(H+CN)/Total	0 to 1	1
POD	POD (hit rate) gives an indication of the observed events in classes that were correctly forecasted. POD does not consider false alarms.	POD=H/(H + M)	0 to 1	1
SR	SR gives an indication of the forecasted events in classes that were correctly observed. SR does not consider misses.	SR=H/(H+FA)	0 to 1	1
FAR	FAR is an indication of the predicted events that did not occur.	FAR=FA/(FA+CN)	0 to 1	0

3.5 Quantifying the effect of meteorology on diurnal PM variation

We consider known patterns in emissions and meteorology (diurnal and weekday/weekday variation) to attempt to estimate the effects of emissions and meteorological conditions on PM concentration levels. It is known that air quality monitoring stations over the interior of SA generally measure higher $PM_{2.5}$ and PM_{10} concentrations between Monday and Friday, and a decrease on Saturday and Sunday, because of differences in emissions between weekdays and weekends (Feig *et al.*, 2016).

Considering diurnal variation, aerosol concentrations typically have two peaks, one in the morning and another in the evening, as well as a concentration minimum during the day, typically between 12:00 and 14:00 Local Time (LT). The occurrence of the two peaks results from a combination of diurnal variations of emissions and meteorological factors, which includes the PBL (Tie *et al.*, 2007). The morning decrease in PM concentrations corresponds with the break-up of the PBL, the formation of the mixed layer (unstable conditions), and the rapid increase in MH. The afternoon increase in PM concentrations corresponds with meteorological conditions becoming unfavourable for pollution dispersion, which includes with the formation of the stable nocturnal BL and a decrease in MH. Along with PBL height and stability conditions, changes in emissions also affect PM concentrations on a diurnal scale (Quan *et al.,* 2013). This can especially be seen at urban and/or low-income sites, where morning and afternoon peaks correspond to increased emissions from sources like domestic burning (Hersey *et al.*, 2015; Mdluli, 2008). Since there are many emission sources (e.g. vehicular, domestic burning, industrial, etc.) contributing to the diurnal variation in PM concentrations, it is difficult to quantify the relative impacts of changing emissions and the changing PBL on ambient concentrations from measurements alone.

We attempt to quantify the impact of meteorological drivers, like the changing PBL, on PM peaks. Diurnal variations of PM concentrations were examined in order to identify the optimum period of study at each site. Thereafter, the Coefficient of Determination (R^2) was calculated during the relevant hours. Wilks (2011) defines R^2 as the "proportion of variation in the predictand that is described or accounted for by the regression". Based on the preceding information, the effect of change in meteorological conditions (represented by the ADP index) on PM concentrations is quantified as a percentage (%) deduced from the R^2 statistic.

4. Results

4.1 Verification of simulated WS

Hourly observed WS, as simulated by TAPM, were compared against WS from the SAAQIS observational stations in Table 6. For the period considered, all sites have mean (M) simulated WS within 0.4 ms^{-1} of observed mean WS, except for Witbank, where the mean observed (1.9 ms^{-1}) and mean simulated (3.2 ms^{-1}) WS differ considerably. Standard deviation (SD) values for simulated WS are relatively close to the mean, with values varying between 1 ms^{-1} and 1.6 ms^{-1}. SD values for observed WS have a similar range, except for Camden, where the SD is 2.4 ms^{-1}. All correlations between simulated and observed WS range from 0.4 to 0.6 and, according to a Student's t-test, the relationship between the simulated and observed WS data for all sites is statistically significant at a 95% confidence interval.

Table 6. Statistics for assessment of simulated and observed wind speed (WS) measured in ms^{-1} for each site. Statistics used to assess model performance include; mean (M), standard deviation (SD), mean absolute error (MAE), mean squared error (MSE), and Pearson product-moment correlation coefficient (R).

		Xanadu	Lephalale	Camden	Zamdela	Witbank
Observed WS	**M**	2.5	1.6	3.1	2.1	1.9
	SD	1.5	1.1	2.4	1.6	1.4
Simulated WS	**M**	2.9	1.6	3.0	2.5	3.2
	SD	1.3	1.0	1.6	1.4	1.5
Observed &	**MAE**	1.3	0.7	1.7	1.1	1.5
Simulated WS	**MSE**	2.6	1.0	4.9	2.1	3.3
	R	0.4	0.5	0.5	0.6	0.6

MAE of simulated WS ranges from smallest at Lephalale (0.7 ms^{-1}) to largest at Camden (1.7 ms^{-1}). This means that the sum of the absolute differences between observed and model simulations are below 1.7 ms^{-1} for all sites. MSE values closer to zero indicate better model performance. MSE for Lephalale is 1 ms^{-1}, this is the lowest for all sites. Camden has the largest MSE (4.9 ms^{-1}); this can be attributed to some very high hourly measured WS in the observed data. All simulated wind speeds are considered acceptable for the task of calculating the ADP index.

MOL and MH could not be verified in the same way as WS as there are no measurements at these sites. Their diurnal variations do coincide with the expected diurnal variation of these variables found in previous studies (not shown). Simulated average MH values are at a minimum throughout the night (<500m), starts to increase at 08:00 LT, and reaches a peak at 15:00/16:00 LT. These night-time lows and daily peaks coincide with the findings of Liu and Liang (2010), who considered soundings collected in 14 major field campaigns around the world. Seasonally, simulated MH also produce the expected cycle with MH being higher during summer and spring, and lower in winter and autumn (El-Shazly *et al.,* 2012). MOL is an indication of atmospheric stability; its classes are classified as unstable, neutral, or stable. Simulated MOL values exhibit mostly stable conditions throughout the night, while unstable conditions prevail throughout the day (approximately 09:00 to 17:00 LT); from literature, this is what is expected (Ashrafi & Hoshyaripour, 2008). Overall, stable atmospheric conditions dominate more than 50% of the time in the simulated MOL data. As for the seasonality of atmospheric stability, stable conditions dominate during summer and winter. Summer months experience a higher frequency of unstable hours, while stable conditions are more frequent during the winter. These seasonal patterns are reproduced in the simulated MOL data.

4.2 ADP and PM concentrations

All sites show a similar pattern when it comes to the diurnal variation of ADP, with values peaking during the day, between 07:00 and 18:00 Local Time, and reaching a maximum at 14:00/15:00 LT (Fig. 2). Values stay mostly constant during the night, ranging between 20 and 40, and dropping to a minimum at 20:00 LT. The ADP calculation is based on parameters such as MOL and MH that are influenced by solar radiation, and it is expected that ADP values will be more favourable (larger) during the day, when solar radiation is at its maximum. ADP is more favourable for pollution dispersion during summer months (December, January and February (DJF)). Due to the prevalence of low mixing heights and dominant stable atmospheric conditions during the winter in this region, ADP is less favourable during winter months (June, July and August (JJA)) (not shown).

While ADP peaks during the day, PM pollution concentrations peak in the morning and again in the evening at all of the sites except Camden. The PM peak in the morning and afternoon may be attributed to increased emissions from sources, such as traffic and domestic fuel burning, as well as unfavourable meteorological conditions for pollution dispersion. Increased atmospheric mixing, due to the break-up of the PBL, causes the decrease in PM concentration throughout the day. Although the diurnal PM variations for most of the sites are similar, with a peak in the morning and evening (Fig. 2), PM mass concentration levels between sites vary greatly. On average, $PM_{2.5}$ values are between 40% and 50% of average hourly PM_{10} values.

PM concentrations at Camden differ from the typical pattern, only showing a slight peak in the afternoon between 18:00 and 19:00 LT. The Camden monitoring station is located near various mining and power-generation activities, including a coal-fired power station. All of the other stations are located within residential areas, while Camden is not.

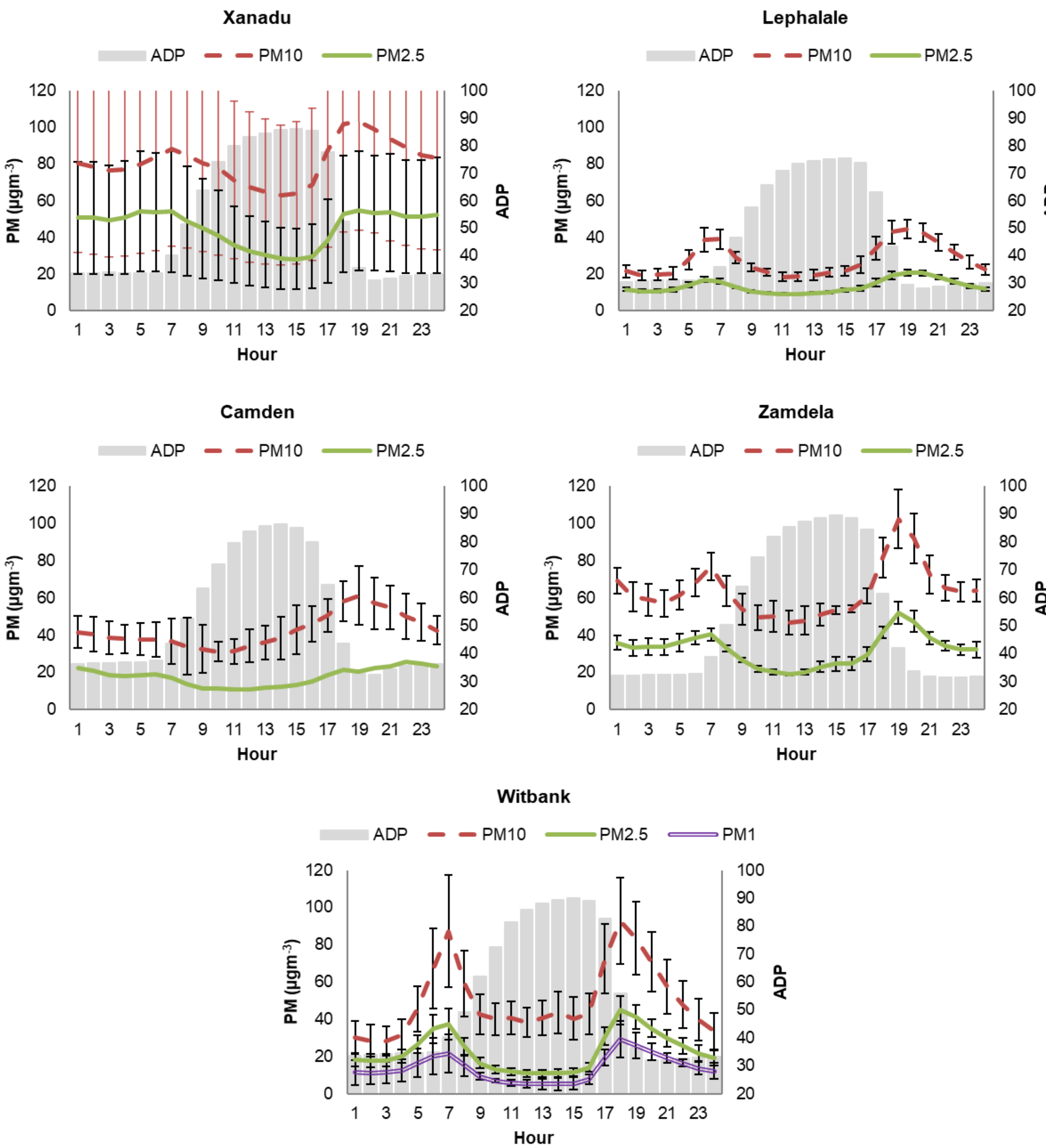

Figure 2. Average hourly PM concentrations in units of µgm^{-3} (lines) and ADP values (bars), demonstrating diurnal variation, for all sites for the entire period. Error bars indicate ± 1 standard deviation (SD) of hourly PM concentrations.

4.3 Identifying relationships

The distribution of PM_{10}, $PM_{2.5}$ and PM_1 per ADP class was investigated. The dominant ADP class for all observations is Class 2, this may be interpreted as meaning that most of the hours with PM observations experienced unfavourable conditions for pollution dispersion. The dominant class combination is PM Class 1 (high pollutant concentration) and ADP Class 2 (unfavourable conditions). Very unfavourable ADP conditions (Class 1) occur the least.

In order to investigate the strength of the relationship between ADP and PM classes, correlations between the two datasets were calculated (Table 7). The entire period, summer (DJF), and winter (JJA) are investigated to identify the conditions under which ADP and PM classes show the strongest correlation. Summer and winter seasons are investigated separately due to the significantly different meteorological conditions that dominate in each period over the study region. During summer, PM concentrations are generally lower, due to a decreased need for heating (by domestic burning), unstable atmospheric conditions, and an increase in wet deposition (all sites are within the summer rainfall region). PM concentrations are higher during the cold winter months due to increased emissions from domestic burning (e.g. coal and wood) for heating. In addition, meteorological conditions are typically not conducive to pollution dispersion; in the winter at these sites, atmospheric conditions tend to be more stable, MH can be extremely low, and therefore PM is confined near the earth's surface (Hersey et al., 2015).

Table 7. Correlations (r) between ADP and PM classes for the entire period, summer and winter. The underlined values show correlations of moderate strength (0.30–0.49).

Station name	Period			Summer (DJF)			Winter (JJA)		
	PM_{10}	$PM_{2.5}$	PM_1	PM_{10}	$PM_{2.5}$	PM_1	PM_{10}	$PM_{2.5}$	PM_1
Xanadu	0.13	0.22		0.07	0.14		<u>0.31</u>	<u>0.48</u>	
Lephalale	0.22	0.22		0.08	0.08		<u>0.34</u>	<u>0.38</u>	
Camden	0.18	<u>0.31</u>		0.20	0.24		0.11	<u>0.46</u>	
Zamdela	0.22	0.24		0.21	0.20		0.25	<u>0.32</u>	
Witbank	0.10	<u>0.36</u>	<u>0.46</u>	-0.02	<u>0.30</u>	<u>0.41</u>	0.07	<u>0.41</u>	<u>0.48</u>

Although the correlations between ADP and PM classes vary only from weak (0.1–0.29) to moderate (0.30–0.49) strength, most are at least positive. This means that higher classes (more favourable) ADP are related to higher classes (lower concentrations) of PM. All correlation coefficients in Table 7 have calculated p-values less than 0.05; this means that the observed differences between PM concentration and ADP index classes are unlikely to be due to chance. Xanadu and Lephalale have relatively good correlations in the winter for both PM_{10} and $PM_{2.5}$, and very weak correlations in the summer. At Zamdela, correlations are between 0.20 and 0.32 for all cases. For all sites, except PM_{10} at Camden , the correlation between ADP and PM classes is stronger in the winter. In most of the cases considered, coarse particulates show a weaker relationship between their PM classes and ADP, and fine particulate show more promising results with medium-strength relationships, especially during winter. It is well documented that higher concentrations of $PM_{2.5}$ and PM_{10} are associated with stable atmospheric conditions, low wind speeds, and the occurrence of inversion layers (Czernecki et al., 2017; Di Virgilio et al., 2018; Perrino et al., 2008; Xu et al., 2018). The results suggest that, because $PM_{2.5}$ presents stronger correlations with the ADP index than PM_{10}, it is influenced by these meteorological variables to a larger extent. Additionally, at higher wind speeds, PM_{10} concentrations in some areas may increase due increased wind-blown dust. This influences the relationship between PM_{10} and WS classes, because larger WS could lead to higher PM_{10} concentrations.

Wet deposition plays an important role in the removal of PM, especially fine particulates, from the atmosphere (Wu *et al.*, 2018). The process of wet deposition might negatively affect the correlations between the ADP index and PM concentration classes as rainfall is not considered in the ADP index. The study region is situated in the austral summer rainfall region of SA (Tyson and Preston-Whyte, 2000). Therefore, the process of wet deposition may explain some of the weaker correlations in Table 7 during summer.

Since the diurnal cycles of PM concentrations and ADP index values differ remarkably (Fig. 2), we consider the average daily PM and ADP classes to investigate the effect of meteorological conditions on daily average PM levels. Hourly ADP and PM classes were used to calculate these daily averages.

Table 8. Correlations (r) between hourly and daily average ADP and PM classes for the entire period. The underlined values indicate moderate (0.30–0.49) and strong (0.50 to 1.0) correlations.

Station name	Hourly			Daily average		
	PM_{10}	$PM_{2.5}$	PM_1	PM_{10}	$PM_{2.5}$	PM_1
Xanadu	0.13	0.22		<u>0.48</u>	<u>0.45</u>	
Lephalale	0.22	0.22		<u>0.58</u>	<u>0.53</u>	
Camden	0.18	<u>0.31</u>		<u>0.35</u>	<u>0.38</u>	
Zamdela	0.22	0.24		<u>0.35</u>	<u>0.34</u>	
Witbank	0.10	<u>0.36</u>	<u>0.46</u>	<u>0.62</u>	<u>0.66</u>	<u>0.67</u>

Table 8 correlations show that relationships between average daily ADP and PM classes are significantly stronger than between hourly ADP and PM classes. From daily averages, the effect of meteorological condition are more noticeable because emissions tend to be quite similar from day-to-day. It should be noted that the thresholds for the variable classes contained in the ADP calculation (Table 1) were developed for hourly, and not daily, values. Therefore, although the daily values appear to produce stronger relationships, a proper comparison and conclusion would require the calculation of daily average ADP with its own thresholds. In addition, the averaging of ADP values from hourly to daily dilutes the data to a point where certain classes no longer occur. Before the averaging of hourly data, all ADP classes (1 to 5) were represented in the data for every site. After averaging to daily values, only ADP classes 2 to 4 are present in the data. Consequently, while these daily average correlations indicate stronger relationships, hourly values were used in the following analyses.

4.4 Individual meteorological parameters

Individual parameters were examined to assess their performance compared to the ADP index. The variables contained in the ADP calculation, namely MH, MOL and WS, were classed as in Table 1 and examined in the same manner as ADP. The association between MH, MOL, WS and PM classes was investigated, in order to identify their influence on ADP as a predictor, as well as their individual performance to predict PM concentrations.

On average, the individual variables have stronger relationships with PM classes in winter and for finer particulates (i.e. $PM_{2.5}$ and PM_1). MOL, MH and WS correlations for all cases vary from weak to strong. PM concentrations in the winter are very much dependant on variables influenced by the diurnal cycle of solar radiation such as MOL and MH, and less by WS. In the summer, PM concentrations are overall more reliant on MOL and WS, and MH was never the variable with the strongest correlation (Table 9). The sites investigated are situated in the summer rainfall region of South Africa. Winters in this region are characterized by stable atmospheric conditions and cloud-free skies, whereas summers experience mainly conditions of atmospheric instability, as well as frequent thunderstorms in the afternoon and early evenings (Tyson and Preston-Whyte, 2000). Of the individual variables in the ADP index, MOL is the best performer producing strongest relationships with PM classes for 13 out of the 22 cases.

Table 9. Variable in the ADP index (MOL, Mixing Height (MH) or Wind Speed (WS)), which produced the strongest correlation (r) with hourly PM classes at each site during summer and winter. Strength of the correlations are indicated in brackets. The underlined values indicate moderate (0.30–0.49) and strong (0.50 to 1.0) correlations.

Station name	Summer (DJF)			Winter (JJA)		
	PM_{10}	$PM_{2.5}$	PM_1	PM_{10}	$PM_{2.5}$	PM_1
Xanadu	WS (0.07)	MOL (0.13)		<u>MH (0.39)</u>	<u>MH (0.53)</u>	
Lephalale	WS (0.21)	WS (0.20)		<u>MH (0.37)</u>	<u>MH (0.38)</u>	
Camden	MOL (0.26)	MOL (0.28)		MOL (0.14)	<u>MOL (0.45)</u>	
Zamdela	MOL (0.23)	MOL (0.22)		MOL (0.25)	<u>MOL (0.32)</u>	
Witbank	WS (0.10)	<u>MOL (0.31)</u>	<u>MOL (0.41)</u>	WS (0.22)	<u>MOL (0.41)</u>	<u>MOL (0.48)</u>

Since MOL is the individual variable with the strongest relationship with PM, we compare correlations between PM and ADP (Table 7) with correlations between PM and MOL (not shown). Although differences in correlation strengths between the above mentioned are in some cases quite small, ADP outperforms MOL 55% of the time for the cases considered in Table 9.

Data from all stations were considered for the previous analyses (Section 4.2 to 4.4) because the calculation of correlations consider data hour by hour. However, the following results (Section 4.5 and 4.6) consider the number of cases present in various class combinations, and are based on the diurnal variation of ADP and PM, respectively. As such, Xanadu was not considered for any further analysis due to data availability being less than 40% after QC (as shown in Table 3), and Camden was left out because of insufficient $PM_{2.5}$ data.

4.5 Forecast performance

Data contained in multi-category contingency tables are used to assess the performance of ADP as a predictor in the different classes. Plotted in Fig. 3 are POD, SR and FAR for Zamdela, Lephalale and Witbank.

POD and SR are plotted together with the FAR, in order to assess the performance of ADP as a forecasting tool. Lephalale in winter (Fig. 3) produces a high SR score in Class 1. This means that a relatively large fraction of Class 1 PM events forecasted by ADP were observed correctly. However, this score may be misleading because of the small fraction (<1%) of events forecasted in Class 1. FAR is also at its lowest level (0.5) in Class 1. In contrast to Class 1, Class 2 has a high POD score, which translates to Class 2 having many PM-observed events that were correctly forecasted by ADP, for both summer and winter.

In both summer and winter, Zamdela has a relatively high FAR for all classes, never dropping below 0.6. The best performing class with respect to POD is Class 2, whereas Class 5 has the highest score for SR. POD for the forecasts in Class 1, in both seasons, are close to zero. The fact that POD scores in Class 1 are close to zero means that there were very little forecasts for very unfavourable ADP for Zamdela. Witbank has relatively high SR for Class 1 in winter, and POD is the highest in Class 2 for both summer and winter.

Accuracy of Forecast (AOF) is the level of agreement between the forecast and measurements. All cases considered have an AOF of less than 0.30. This means that the fraction of forecasts in the correct category is less than 30% for each case considered. Based on POD and SR scores, the ADP index forecasts best for high (Class 2), but not highest (Class 1) PM concentration events. The relatively high POD scores (>0.5) for Class 2 indicate that unfavourable ADP conditions correctly forecast high PM concentrations at least 50% of the time. Class 2 is by far the most commonly occurring ADP class; this undoubtedly has a favourable influence on its performance scores.

The poor performance of ADP forecasts in Class 1 are due to very few hours (less than 1%) being classed as having very unfavourable conditions for ADP. In future research, it would be worthwhile to reconsider the MOL, MH and WS intervals used in the ADP calculation in order to improve ADP forecasts, and decrease false alarms, in Class 1.

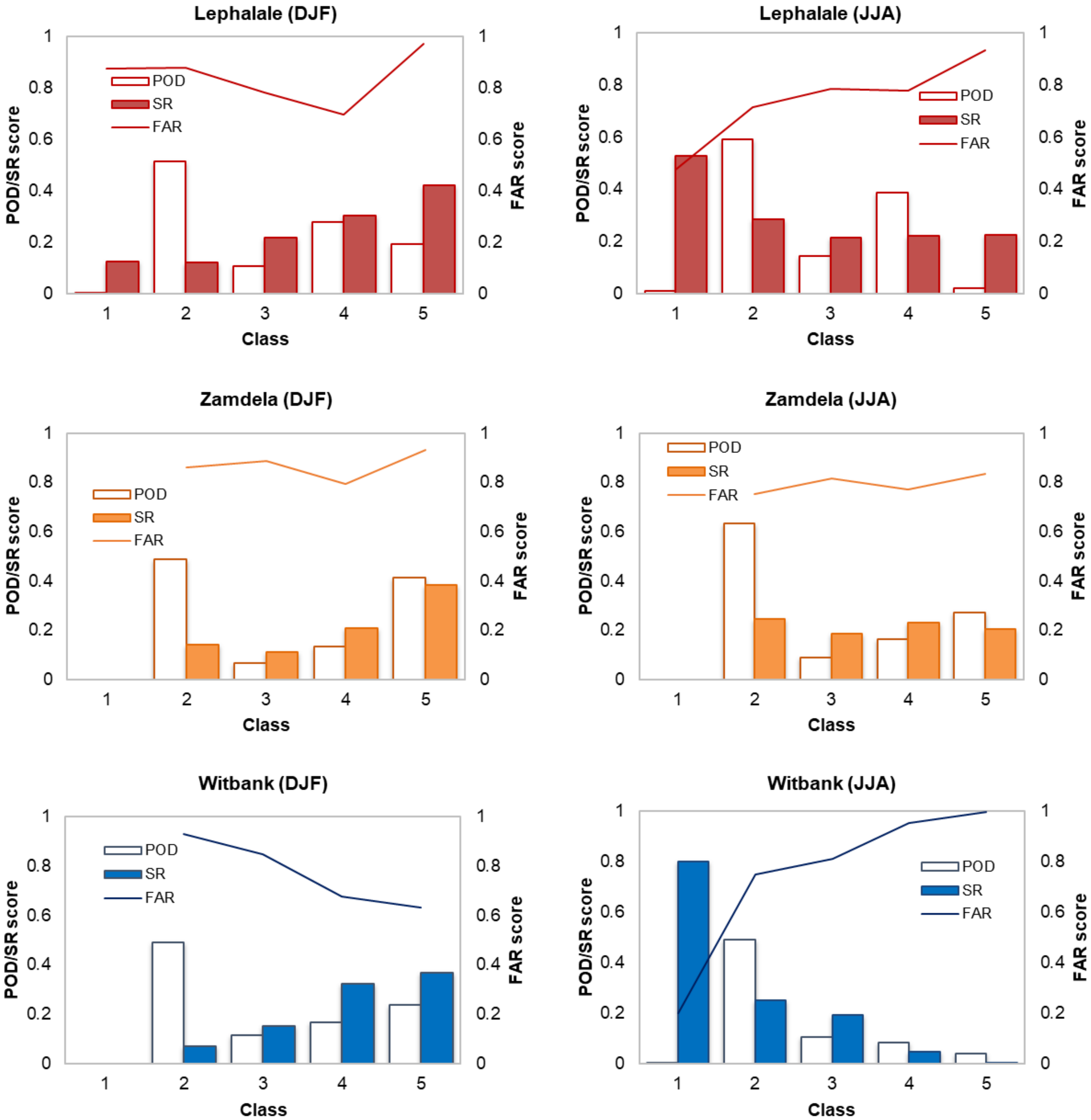

Figure 3. Probability of detection (POD), success ratio (SR), and False Alarm Ratio (FAR) scores for Zamdela, Lephalale and Witbank in both the winter (JJA) and the summer (DJF) for PM_{10}. High POD and SR scores (close to 1) indicate better performance, whereas lower FAR scores (close to 0) indicate a better forecast with less false alarms. Classes 1 to 5 on the x-axis refer to the class combinations, as described in Table 4. $PM_{2.5}$ graphs (not shown) exhibit very similar patterns for all sites.

4.6 Effect of meteorology on diurnal PM variation

In this section, we consider known patterns in emissions and meteorology (diurnal and weekday/weekday variation) to estimate the effects of emissions and meteorological conditions on PM concentration levels. There is no significant difference in the strengths of weekday and weekend correlations between ADP and PM (not shown). Therefore, we can conclude that the weekday/weekend variances in emissions are not driving the relationship between ADP and PM for the sites considered here.

The next aspect to consider the diurnal variation of PM concentrations. Meteorological conditions (represented by the ADP index) have an effect on the morning decrease and afternoon increase of PM concentrations. This effect was calculated for Lephalale, Zamdela and Witbank, in both summer and winter periods. Lephalele results are presented in detail in Fig 4, while all sites are summarized in Table 10.

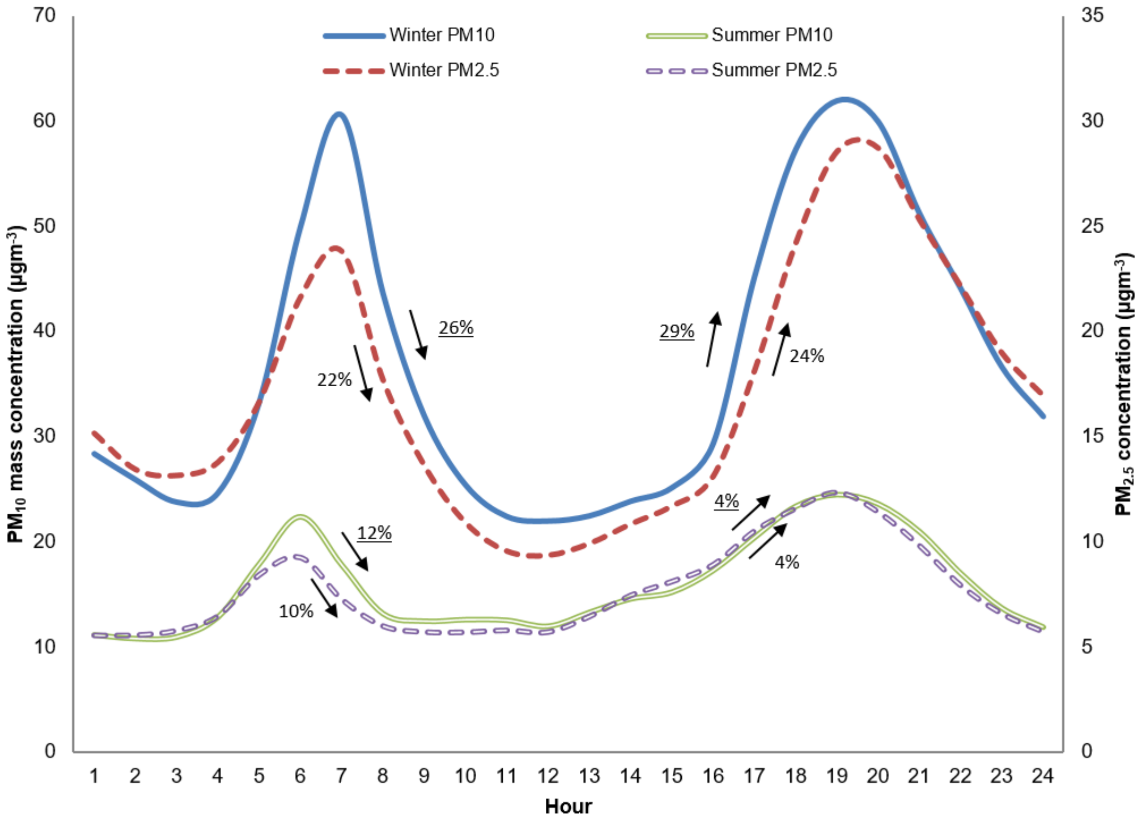

Figure 4. Diurnal variation of PM_{10} and $PM_{2.5}$ concentrations (μgm^{-3}) at Lephalale for both winter and summer. Arrows and percentages (PM_{10} underlined) indicate the direction and amount of change in PM concentrations attributed to the change in meteorological conditions (i.e. ADP).

PM concentration levels at Lephalale (Fig. 4) follow the typical diurnal variation, with two PM peaks (morning and evening)and a minimum at noon, as expected (Mdluli, 2008; Hersey *et al.*, 2015). The percentages displayed on Fig. 4 are based on the R^2 statistic (Coefficient of Determination) and display the proportion of variation in PM that is accounted for by ADP. The afternoon increase in PM concentrations is largely affected by the onset of unfavourable conditions for pollution dispersion in the evening, especially during winter (29% PM_{10}, 24% $PM_{2.5}$).

Zamdela PM concentrations have a morning and afternoon peak in winter and summer. At this site, afternoon PM increases are influenced more by the change in meteorological conditions, than PM decreases in the mornings (Table 10). Witbank PM_{10} concentrations are less dependent on changes in meteorological conditions than $PM_{2.5}$. Considering PM_{10} in winter, only 10% of the morning decrease and 13% of the afternoon increase in concentrations is attributed to change in ADP, compared to 27% and 26% for $PM_{2.5}$. This is also true during the summer months, when $PM_{2.5}$ is more dependent on change in meteorological conditions than PM_{10}.

Table 10. Percentage change in PM concentrations attributed to the change in meteorological conditions for all sites investigated.

	PM_{10}				**$PM_{2.5}$**			
	Winter		**Summer**		**Winter**		**Summer**	
	PM decrease	**PM increase**	**PM decrease**	**PM increase**	**PM decrease**	**PM increase**	**PM decrease**	**PM increase**
Zamdela	6%	24%	4%	13%	12%	22%	6%	11%
Lephalale	26%	29%	12%	4%	22%	24%	10%	4%
Witbank	10%	13%	0.3%	1%	27%	26%	9%	9%

PM concentrations levels are affected by meteorological conditions studied to a different degree at each site, but there are some common threads. In general, meteorological conditions have a stronger influence on the diurnal variation of PM concentrations in the winter as compared to the summer, and the afternoon increase in PM concentrations is influenced more by a change in the meteorological conditions than the morning decrease.

The onset of "bad" meteorology (unfavourable conditions) in the afternoon has a significant effect on PM concentrations at most sites, especially in the winter. "Good" meteorology, associated with the mixed layer, influences the decrease in PM concentrations, but to a lesser extent than unfavourable conditions, in most cases. Diurnal variation in PM concentrations, at the sites investigated, is dominated by variability of emissions rather than meteorological conditions contained in the ADP index.

Here we investigated the PM concentration slopes that should be highly influenced by the variables contained in the ADP index. The remaining two slopes (PM increase in the morning and PM decrease in the evening) should not be affected by these meteorological parameters to the same extent. Therefore, when R^2 for the remaining slopes were calculated, all results were less than those presented in Table 10. On average across all sites, the influence of change in ADP on morning PM increase was 2.7%, and on evening PM decrease, only 1.3%.

5. Discussion and Conclusions

The ADP index is based on meteorological variables instructive to the dispersion of air pollution in the atmosphere on an hourly timescale. ADP was developed and presented for the first time here, and the estimation of PM concentrations at sites in SA, using this new index, was investigated. ADP varies diurnally, with a clear peak in values between 14:00 and 15:00 LT, and remains mostly constant during the night, with values between 20 and 40. PM pollution concentrations peak in the morning and again in the evening at most of the sites investigated.

There exists a relationship between PM classes and ADP; PM concentration classes are positively correlated with ADP classes under all circumstances investigated. Due to the inverse relationship between meteorological parameters such as WS, MH, MOL and pollutants in the atmosphere, this result is expected. PM_{10} has a weaker relationship with ADP, whereas $PM_{2.5}$ performs best in most cases. Correlation coefficients vary from weak to moderate. On average, the degree of relationships in summer is weaker than in winter. A possible explanation for this phenomenon is the fact that all sites investigated are situated in the summer rainfall region of SA. During precipitation days, wet deposition and the mechanism of scavenging of PM has a lowering effect on PM concentration and, at the same time, weaken the effect of other meteorological variables (like those contained in the ADP index) on its concentration (Holst *et al.,* 2008). Additionally, there are other meteorological influences, not considered here, that most likely influenced the strength of relationships found. Strong wind speeds can lead to increases in wind-blown dust instead of facilitating dispersion, transformation of particles by photochemical processes intensified by high air temperature, and high air temperature aiding in vertical dispersion of pollutants (Li et al., 2017a). Future studies could investigate the impact of these variables on ADP performance.

In the presented research, the correlation strengths vary significantly between sites, seasons and particulate size. Correlation strengths range from; a weak negative correlation found between PM_{10} and ADP at Witbank (summer), to moderate strength relationships for $PM_{2.5}$ and PM_{10} at Lephalale (winter), and a correlation as strong as a 0.48 for PM_1 at Witbank (winter). In other studies where the relationship between meteorological parameters and pollutant concentrations were investigated, the correlations strengths were similar. For instance, Zhang et al. (2015) studied the relationships between meteorological parameters and ambient air pollutant concentrations in three megacities in China. Correlations between pollutants and meteorological parameters differed considerably and varied from weak to moderate in strength, with some strong correlations (> ~0.70) also found. Li et al. (2017a) investigated $PM_{2.5}$ and PM_{10} and their relationships with different meteorological parameters for 11 monitoring stations in Shenyang, China. Correlations strengths and directions differed significantly between season and PM sizes, and all were weak to moderate in strength (no correlation larger than ~0.52). Alvarez et al. (2018) assessed the impact of temperature, relative humidity, and WS on the variability of $PM_{2.5}$ at three sites in El Paso, Texas, USA. Again, correlations varied greatly in strength between seasons and sites, and for most sites, correlations either changed from positive to negative or dissolved depending on the season. Most of the correlations calculated were weak (< ~0.30), with some moderate strength cases (between ~0.30 and ~0.50).

Noticing only weak- to moderate-strength relationships between hourly ADP and PM classes in this study led to the consideration of the meteorological variables contained in ADP, separately. Previous studies by Li et al. (2017b), Nath & Patil (2006), Yin et al. (2016), as well as Ziomas et al. (1995) have shown the possibility of predicting pollutant concentrations based on relationships between pollutants and individual meteorological parameters (relative humidity, MH, temperature, etc.). It was found that PM concentrations in the winter are more dependent on variables influenced by solar radiation such as MOL and MH and less by WS. In the summer, PM concentrations are overall more reliant on MOL and WS; in this instance, MH has the least influence. Overall relationships between hourly WS and PM classes are the weakest, while MOL is the best-performing individual variable. In terms of relationship strength with PM classes, ADP outperforms MOL 55% of the time for the cases considered.

The forecast performance was assessed (using SR, POD and FAR); all sites produced very similar results for PM_{10} and $PM_{2.5}$. A combination of POD and SR scores indicated that, for all cases, ADP predictions work best in the frequently occurring Class 2. Class 2 represents unfavourable ADP and high PM concentrations. Since there are serious health risks associated with high concentrations of particulate pollution, this is a significant result. Non-meteorological influences contribute to weak correlations, POD scores, and SR scores.

The ADP index, as used in this pilot study, is not sufficient to estimate PM concentration at the sites considered because of relatively weak correlations, and low SR and POD scores. The ADP index may be applied elsewhere, but further research and optimization of the index for South African conditions is recommended first. MOL, MH and WS intervals need to be re-evaluated in order to find the ideal thresholds for use in SA. In future research, precipitation could be included in the index to account for the effect that rainfall has on ADP index performance. Although this might add value to the index, accurately forecasting hourly precipitation is not a trivial task, especially due to the spatial heterogeneity of rainfall over the study region. The classification of the monitoring stations must be taken into account when considering the results obtained. The sites studied here are located either in rural low-income areas near mines, power stations and industries or close to residential areas. The pollutant concentrations measured, and fluctuations observed at these stations are heavily influenced by residential (including domestic burning), industrial, and traffic pollution, and not purely by meteorological factors. Due to a lack of data, no proper background sites, located away from industry, residential areas or roads, could be considered. Further research could therefore include the index being tested at appropriate background sites. The forecast capabilities of the ADP index may also be significantly improved by a complete dataset of good quality PM concentration data from a well-maintained monitoring station.

The ADP index was designed for variables that vary on an hourly scale. While the strength of the relationship between ADP and PM did improve using daily averaged classes, this averaging decreased the range of the data (e.g. decreased the number of classes with data). Future studies could investigate the optimal time-resolution of the variables driving and contained in ADP on PM concentrations. Since acute exposure to PM is regulated through ambient standards using 24-hour averages, a re-designed index based on a daily timescale may be something to consider.

Even though the relationship between ADP index and PM concentrations is weak overall, there are cases and periods where ADP correlates well with PM concentrations. The typical morning decrease in PM concentrations (coincides with the formation of the mixed layer), and afternoon increase of PM concentrations (coincides with the formation of the stable boundary layer) is influenced, to a degree, by the meteorological variables contained in the ADP calculation. Using the R^2 statistic to describe this influence has led to the quantification of the influence of “good” (favourable) and “bad” (unfavourable) meteorological conditions on PM concentrations. These finding allowed us to quantify the impact that these meteorological variables (i.e. ADP) have on the diurnal cycle of the PM. The application of this index in this way can play an important role in air quality management when quantifying the impacts of drivers of PM peak concentrations.

It was found that each site is affected differently by meteorological factors, but in general, ADP has a stronger effect on the diurnal variation of PM concentrations in the winter. The afternoon increase in PM concentrations is also influenced more by meteorology than the morning decrease. For both PM_{10} and $PM_{2.5}$, ADP accounted for more than 20% of the afternoon increase in PM levels at 2 of the 3 sites studied in the winter. This is an important result for air quality management in the area, as it quantifies, for the first time to our knowledge, the role that meteorology plays in this diurnal cycle across these sites, and highlights the large role that emissions play in the diurnal cycle.

Acknowledgements

The authors acknowledge South African Department of Environmental Affairs, ESKOM and The South African Weather Service (SAWS) for the use of their PM and meteorological data, and SAWS for providing pollutant concentration and some meteorological data for all investigated sites through SAAQIS. This work was supported by SASOL through the Laboratory for Atmospheric Studies (LAS) at the University of Pretoria. The authors extend appreciation to the late Professor George Djolov without whom this research would not have been possible.

References

Alade, O.L., 2010. Characteristics of Particulate Matter over the South African industrialised Highveld. M.Sc. Dissertation, University of the Witwatersrand, Johannesburg, South Africa, 122 pages. Retrieved from: http://hdl.handle.net/10539/9101.

Alvarez, H.A.O., Myers, O., Weigel, M., Armijos, R., 2018. The value of using seasonality and meteorological variables to model intra-urban PM2.5 variation. Atmos. Environ. 182, 1-8. https://doi.org/10.1016/j.atmosenv.2018.03.007.

Ashrafi, K., Hoshyaripour, A.G., 2008. A model to determine atmospheric stability and its correlation with CO concentration. World Academy of Science, Engineering and Technology. Int. J. Environ. Chem. Ecol. Geol. Geophys. Eng. 2, 143-148.

Belelie, M.D., Piketh, S.J., Burger, R.P., Venter, A.D., Naidoo, M., 2019. Characterisation of ambient Total Gaseous Mercury concentrations over the South African Highveld. Atmos. Pollut. Res. 10, 12-23. https://doi.org/10.1016/j.apr.2018.06.001.

Chu, Y., Liu, Y., Li, X., Liu, Z., Lu, H., Lu, Y., Mao, Z., Chen, X., Li, N., Ren, M., Liu, F., Tian, L., Zhu, Z., Xiang, H.A., 2016. Review on Predicting Ground PM2.5 Concentration Using Satellite Aerosol Optical Depth. Atmos. 7, 129. https://doi.org/10.3390/atmos7100129.

Cohen, J., 1988. Statistical power analysis for the Behavioral Sciences, second ed., Lawrence Erlbaum Associates, New Jersey.

Cohen, A.J., Anderson, H.R., Ostro, B., Pandey, K.D., Krzyzanowski, M., Künzli, N., Gutschmidt, K., Pope, A., Romieu, I., Samet, J.M., Smith, K., 2005. The global burden of disease due to outdoor air pollution. J. Toxicol. Environ. Health. 68, 1301-1307. https://doi.org/10.1080/15287390590936166.

Collaboration for Australian Weather and Climate Research (CAWCR), 2015. http://www.cawcr.gov.au/projects/verification/ Accessed in June 2018.

Cosijn, C., Tyson, P.D., 1996. Stable discontinuities in the atmosphere over South Africa. S. Afr. J. Sci. 92, 381-386. https://hdl.handle.net/10520/AJA00382353_7756.

Czernecki, B., Półrolniczak, M., Kolendowicz, L. Marosz, M., Kendzierski, S., Pilguj, N., 2017. Influence of the atmospheric conditions on PM10 concentrations in Poznań, Poland. J. Atmos. Chem. 74, 115-139. https://doi.org/10.1007/s10874-016-9345-5.

Di Virgilio, G., Hart, M.A., Jiang, N., 2018. Meteorological controls on atmospheric particulate pollution during hazard reduction burns. Atmos. Chem. Phys., 18, 6585-6599. https://doi.org/10.5194/acp-18-6585-2018.

Done, J., Davis, C., Weisman, M. 2004. The next generation of NWP: Explicit forecasts of convection using the weather research and forecasting (WRF) model. Atmos. Sci. Lett. 5, 110-117. https://doi.org/10.1002/asl.72.

El-Shazly, S.M., Kassem, K.O., Hassan, A.A., Hala., E. A., 2012. Assessment of Mixing Height at Qena/Upper Egypt Based on Radiosonde Data. Resour. Environ. 2, 275-280. https://doi.org/10.5923/j.re.20120206.05.

Fajersztajn, L., Veras, M., Barrozo, L., Saldiva, P., 2014. Air monitoring coverage in low-income countries: An observational study. Lancet. 384, S14. https://doi.org/10.1016/S0140-6736(14)61877-8.

Feig, G., Naidoo, S., Ncgukana, N., 2016. Assessment of ambient air pollution in the Waterberg Priority Area 2012-2015. Clean Air J. 26, 21-28. http://dx.doi.org/10.17159/2410-972X/2016/v26n1a9.

Garland, R.M., Naidoo M., Sibiya, B., Oosthuizen, R., 2017. Air quality indicators from the Environmental Performance Index: Potential use and limitations in South Africa. Clean Air J. 27, 33-41. https://doi.org/10.17159/2410-972X/2017/v27n1a8.

Garstang, M., Tyson, P.D., Swap, R., Edwards, M., Kållberg, P., Lindesay, J.A., 1996. Horizontal and vertical transport of air over southern Africa. J. Geophys. Res. 101, 23721-23736. https://doi.org/10.1029/95JD00844.

Grundström, M., Hak, C., Chen, D., Hallquist, M., Pleijel, H., 2015. Variation and co-variation of PM_{10}, particle number concentration, NO_x and NO_2 in the urban air - Relationships with wind speed, vertical temperature gradient and weather type. Atmos. Environ. 120, 317-327. https://doi.org/10.1016/j.atmosenv.2015.08.057.

Gryning, S.E., Batchvarova, E., Brummer, B., Jørgensen, H., Larsen, S., 2007. On the extension of the wind profile over homogeneous terrain beyond the surface layer. Bound.-Layer Meteorol. 124, 251-268. https://doi.org/10.1007/s10546-007-9166-9.

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

Holst, J., Mayer, H., Holst, T., 2008. Effect of meteorological exchange conditions on PM10 concentration. Meteorol. Z. 17, 273-282. https://doi.org/10.1127/0941-2948/2008/0283.

Holzworth, G., 1971. Air pollution climatology. American Institute of Chemical Engineers, 64th Annual Meeting. San Francisco, California.

Hurley, P., 2008a. TAPM V4. Part 1: Technical description. CSIRO Marine and Atmospheric Research Paper No. 25. ISBN: 978-1-921424-71-7.

Hurley, P., 2008b. TAPM V4. User Manual. CSIRO Marine and Atmospheric Research Internal Report No. 5. ISBN: 978-1-921424-73-1.

Jolliffe, I., Stephenson, D., 2012. Forecast verification: A practitioner's guide in atmospheric science, second ed. John Wiley & Sons, New Jersey. ISBN: 978-0-470-66071-3.

Josipovic, M., Annegarn, H., Kneen, M., Pienaar, J., Piketh, S., 2009. Concentrations, distributions and critical level exceedance assessment of SO2, NO2 and O3 in South Africa. Environ. Monit. Assess. 171, 181-196. https://doi.org/10.1007/s10661-009-1270-5.

Kanevce, G., Kanevce, L., 2006. Dispersion modelling for regulatory applications. Therm. Sci. 10, 141-154. https://doi.org/10.2298/TSCI0602141K.

Kim, K., Kim, M., Hong, S., Youn, Y., Hwang, S., 2005. The effects of wind speed on the relative relationships between different sized-fractions of airborne particles. Chemosphere. 59, 929-937. https://doi.org/10.1016/j.chemosphere.2004.11.042.

Li, X., Ma, Y., Wang, Y., Liu, N., Hong, Y., 2017a. Temporal and spatial analyses of particulate matter (PM_{10} and $PM_{2.5}$) and its relationship with meteorological parameters over an urban city in northeast China. Atmos. Res. 198, 185-193. https://doi.org/10.1016/j.atmosres.2017.08.023.

Li, L., Wu, A., Cheng, I., Chen, J., Wu, J., 2017b. Spatiotemporal estimation of historical $PM_{2.5}$ concentrations using PM_{10}, meteorological variables, and spatial effect. Atmos. Environ. 166, 182-191. https://doi.org/10.1016/j.atmosenv.2017.07.023.

Liu, S., Liang, X., 2010. Observed Diurnal Cycle Climatology of Planetary Boundary Layer Height. J. Climate, 23, 5790-5809. https://doi.org/10.1175/2010JCLI3552.1.

Mannucci, P.M., Franchini, M., 2017. Health effects of ambient air pollution in developing countries. Int. J. Environ. Res. Public Health. 14, 1048. https://doi.org/10.3390/ijerph14091048.

Mdluli, T., 2008. The societal dimensions of domestic coal combustion: People's perceptions and indoor aerosol monitoring. Ph.D. Thesis. University of the Witwatersrand, Johannesburg, South Africa, 170 pages. Retrieved from: http://hdl.handle.net/10539/5814.

Naidoo, S., Piketh, S., Curtis, C., 2014. Quantification of emissions generated from domestic burning activities from townships in Johannesburg. Clean Air J. 2, 34-41. Retrieved from: http://hdl.handle.net/10520/EJC154554.

Nath, S., Patil, R., 2006. Prediction of air pollution concentration using an in-situ real time mixing height model. Atmos Environ. 40, 3816-3822. https://doi.org/10.1016/j.atmosenv.2006.02.034.

Niemeyer, L.E., 1960. Forecasting air pollution potential. Mon. Weather Rev. 88, 88-96. https://doi.org/10.1175/1520-0493(1960)088<0088:FAPP>2.0.CO;2.

Panyacosit, L., 2000. A review of Particulate Matter and health: Focus on developing countries. International Institute for Applied Systems Analysis IR-00-005. https://doi.org/10.2139/ssrn.235099.

Peña, A., Gryning, S., Mann, J., 2010. On the length-scale of the wind profile. Q. J. R. Meteorol. Soc. 136, 2119-2131. https://doi.org/10.1002/qj.714.

Perrino, C., Catrambone, M., Pietrodangelo, A., 2008. Influence of atmospheric stability on the mass concentration and chemical composition of atmospheric particles: A case study in Rome, Italy. Environ. Int. 34, 621-628. https://doi.org/10.1016/j.envint.2007.12.006

Quan, J., Gao, Y., Zhang, Q., Tie, X., Cao, J., Han, S., Meng, J., Chen, P., Zhao, D., 2013. Evolution of planetary boundary layer under different weather conditions, and its impact on aerosol concentrations. Particuology. 11, 34-40. https://doi.org/10.1016/j.partic.2012.04.005.

Sathe, A., Mann, J., Barlas, T., Bierbooms, W., Van Bussel, G., 2013. Influence of atmospheric stability on wind turbine loads. Wind Energy. 16, 1013-1032. https://doi.org/10.1002/we.1528.

Seibert, P., Beyrich, F., Gryning, S., Joffre, S., Rasmussen, A., Tercier, P., 2000. Review and intercomparison of operational methods for the determination of the mixing height. Atmos. Environ. 34, 1001-1027. https://doi.org/10.1016/S1352-2310(99)00349-0.

South African Air Quality Information System (SAAQIS), 2018. http://www.saaqis.org.za/Default.aspx Accessed in February 2018.

Swart, A., 2016. Assessment of the baseline meteorological and air quality conditions over Uubvlei, Oranjemund, Namibia. M.Sc. Dissertation, University of Pretoria, Pretoria, South Africa, 94 pages. Retrieved from: http://hdl.handle.net/2263/60862.

Tie, X., Madronich, S., Li G., Ying, Z., Zhang, R., Garcia, A.R., Lee-Taylor, J., Liu, Y., 2007. Characterizations of chemical oxidants in Mexico City: A regional chemical dynamical model (WRF-Chem) study. Atmos. Environ. 41, 1989-2008. https://doi.org/10.1016/j.atmosenv.2006.10.053.

Tyson, P.D., Preston-Whyte, R.A., 2000. The weather and climate of Southern Africa. Oxford University Press Southern Africa, Cape Town, South Africa. ISBN: 978-0-195-71806-5.

US EPA (United States Environmental Protection Agency), 2017. QA Handbook for air pollution measurement systems: Volume II: Ambient air quality monitoring program. EPA-454/B-17-001.

Wilks, D., 2011. Empirical distributions and exploratory data analysis. Statistical Methods in the Atmospheric Sciences, third ed. International Geophysics Series, Volume 100, Academic Press. https://doi.org/10.1016/B978-0-12-385022-5.00003-8.

Witi, J., 2005. Report on ambient PM10 and PM2.5 estimates from monitoring stations data. Cape Peninsula University of Technology. Cape Town, South Africa. Retrieved from: http://www.ehrn.co.za/publications/download/95.pdf.

World Health Organization (WHO), 2016. WHO Global Urban Ambient Air Pollution Database, http://www.who.int/phe/health_topics/outdoorair/databases/cities/en/ Accessed in April 2018.

Wright, C., Oosthuizen, M., Mostert, J., Van Niekerk, L., 2011. Investigating air quality and air-related complaints in the City of Tshwane, South Africa. Clean Air J. 20, 3-12. Retrieved from: http://hdl.handle.net/10204/5997.

Wu, Y., Liu, J., Zhai, J., Cong, L., Wang, Y., Ma, W., 2018. Comparison of dry and wet deposition of particulate matter in near-surface waters during summer. PLOS ONE 13, e0199241. https://doi.org/10.1371/journal.pone.0199241.

Xu, Y., Xue, W., Lei, Y., Yang, Z., Cheng, S., Ren, Z., Huang, Q., 2018. Impact of Meteorological Conditions on PM2.5 Pollution in China during Winter. Atmosphere. 9. 429. https://doi.org/10.3390/atmos9110429.

Yin, Q., Wang, J., Hu, M., Wong, H., 2016. Estimation of daily $PM_{2.5}$ concentration and its relationship with meteorological conditions in Beijing. J. Environ. Sci. 48, 161-168. https://doi.org/10.1016/j.jes.2016.03.024.

Zahumensky, I., 2004. Guidelines on quality control procedures for data from automatic weather stations. WMO-No. 955. Geneva, Switzerland.

Zhang, H., Wang, Y., Hu, J., Ying, Q., Hu, X., 2015. Relationships between meteorological parameters and criteria air pollutants in three megacities in China. Environ. Res. 140, 242-254. https://doi.org/10.1016/j.envres.2015.04.004.

Ziomas, I., Melas, D., Zerefos, C., Bais, A., Paliatsos, A., 1995. Forecasting peak pollutant levels from meteorological variables. Atmos. Environ. 29, 3703-3711. https://doi.org/10.1016/1352-2310(95)00131-H.