https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

Measurement report: Statistical modelling of long-term atmospheric inorganic gaseous species trends within proximity of the pollution hotspot in South Africa

Jan-Stefan Swartz[1], Pieter G. Van Zyl[1]*, Johan P. Beukes[1], Corinne Galy-Lacaux[2], Avishkar Ramandh[3], Jacobus J. Pienaar[1]

[1] Unit for Environmental Sciences and Management, North-West University, Potchefstroom Campus, Potchefstroom, 2520, South Africa

[2] Laboratoire d'Aerologie, UMR 5560, Universit´e Paul-Sabatier (UPS) and CNRS, Toulouse, France

[3] Sasol Technology R&D (Pty) Limited, Sasolburg, South Africa

*Corresponding author: P.G. van Zyl (pieter.vanzyl@nwu.ac.za)

Abstract

South Africa is considered an important source region of atmospheric pollutants, which is compounded by high population- and industrial growth. However, this region is understudied, especially with regard to evaluating long-term trends of atmospheric pollutants. The aim of this study was to perform statistical modelling of SO_2, NO_2 and O_3 long-term trends based on 21-, 19- and 16-year passive sampling datasets available for three South African INDAAF (International network to study Atmospheric Chemistry and Deposition in Africa) sites located within proximity of the pollution hotspot in the industrialised north-eastern interior in South Africa. The interdependencies between local, regional and global parameters on variances in SO_2, NO_2 and O_3 levels were investigated in the model. Long-term temporal trends indicated seasonal and inter-annual variability at all three sites, which could be ascribed to changes in meteorological conditions and/or variances in source contribution. Local, regional and global parameters contributed to SO_2 variability, with total solar irradiation (TSI) being the most significant factor at the regional background site, Louis Trichardt (LT). Temperature (T) was the most important factor at Skukuza (SK), located in the Kruger National Park, while population growth (P) made the most substantial contribution at the industrially impacted Amersfoort (AF) site. Air masses passing over the source region also contributed to SO_2 levels at SK and LT. Local and regional factors made more substantial contributions to modelled NO_2 levels, with P being the most significant factor explaining NO_2 variability at all three sites, while relative humidity (RH) was the most important local and regional meteorological factor. The important contribution of P on modelled SO_2 and NO_2 concentrations was indicative of

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

the impact of increased anthropogenic activities and energy demand in the north-eastern interior of South Africa. Higher SO_2 concentrations, associated with lower temperatures, as well as the negative correlation of NO_2 levels to RH, reflected the influence of pollution build-up and increased household combustion during winter. ENSO made a significant contribution to modelled O_3 levels at all three sites, while the influence of local and regional meteorological factors was also evident. Trend lines for SO_2 and NO_2 at AF indicated an increase in SO_2 and NO_2 concentrations over the 19-year sampling period, while an upward trend in NO_2 levels at SK signified the influence of growing rural communities. Marginal trends were observed for SO_2 at SK, as well as SO_2 and NO_2 at LT, while O_3 remained relatively constant at all three sites. SO_2 and NO_2 concentrations were higher at AF, while the regional O_3 problem was evident at all three sites.

Keywords: passive sampling; sulphur dioxide; nitrogen dioxide; ozone; DEBITS; multiple-linear regression

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

1. Introduction

Although Africa is regarded as one of the most sensitive continents with regard to air pollution and climate change, it is the least studied (Laakso et al., 2012). South Africa is considered an important source region of atmospheric pollutants within the African continent, which is attributed to its highly industrialised economy with the most significant industrial activities including mining-, metallurgical- and petrochemical activities, as well as large-scale coal-fired electricity generation (Rorich and Galpin, 1998; Tiitta et al., 2014). Atmospheric pollution associated with South Africa is compounded by high population growth that, in turn, drives further economic and industrial growth leading to an ever-increasing energy demand (Tiitta et al., 2014). The extent of air pollution in South Africa is illustrated by the well-known NO_2 pollution hotspot revealed by satellite data over the Mpumalanga Highveld, where 11 coal-fired power stations are located (Lourens et al., 2011), which was also recently indicated by the newly launched European Space Agency Sentinel 5P satellite (Meth, 2018).

The importance of long-term atmospheric chemical measurements has been indicated by numerous studies on atmosphere-biosphere interactions (Fowler et al., 2009) and air quality (Monks et al., 2009). These long-term assessments are crucial in identifying relevant policy requirements on local and global scales, as well as the most topical atmospheric chemistry research questions (Vet et al., 2014; IPCC, 2014). In 1990, the International Global Atmospheric Chemistry (IGAC) programme, in collaboration with the Global Atmosphere Watch (GAW) network of the World Meteorological Organisation (WMO) initiated the Deposition of Biogeochemically Important Trace Species (DEBITS) project with the aim to conduct long-term assessments of atmospheric biogeochemical species in the tropics – a region for which limited data existed (Lacaux et al., 2003). The programme is currently operated within the framework of the third phase of IGAC and within the context of the International Nitrogen Initiative (INI) programme. The African component of this initiative was historically referred to as IGAC DEBITS Africa (IDAF), which was relabelled in 2015/2016 under the International Network to study Atmospheric Chemistry and Deposition in Africa (INDAAF) programme. The INDAAF long-term network currently consists of 13 monitoring sites, strategically positioned in southern-, western- and central Africa, which are representative of the most important African ecosystems (http://indaaf.obs-mip.fr). Typical measurements at the INDAAF sites include wet-only rain collection, aerosol composition and inorganic gaseous concentrations, determined with passive samplers.

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

Long-term measurements have been conducted at three dry-savannah southern African INDAAF sites, which include Amersfoort (AF), Louis Trichardt (LT) and Skukuza (SK) located within proximity of the pollution hotspot in the north-eastern interior of South Africa. Measurement of inorganic gaseous pollutant species i.e. sulphur dioxide (SO_2), nitrogen dioxide (NO_2) and ozone (O_3), have been conducted since 1995 at LT, 1997 at AF and 2000 at SK utilising passive samplers. These gaseous species are generally associated with the above-mentioned major sources of atmospheric pollutants in South Africa (Connell, 2005). Moreover, a large number of these sources are located within the north-eastern interior of South Africa, and include the Mpumalanga Highveld, the Johannesburg-Pretoria conurbation and the Vaal Triangle. Laban et al. (2018), for instance, recently indicated high O_3 levels in this north-eastern interior of South Africa, while it was also indicated that O_3 formation in this region can be considered NO_x-limited due to high NO_2 concentrations. Therefore, the South African INDAAF sites were strategically positioned to be representative of the South African interior, with AF an industrially influenced site, LT a rural background site and SK a background site located in the Kruger National Park, as indicated in Fig. 1.

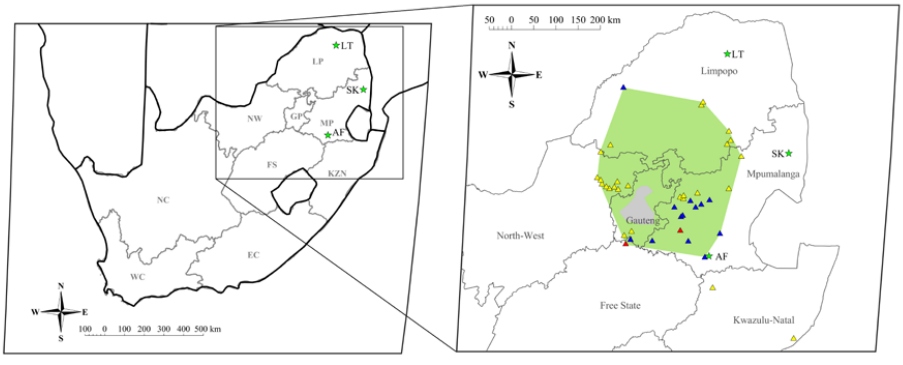


Figure 1: Regional map of South Africa indicating the measurement sites at Amersfoort (AF), Louis Trichardt (LT) and Skukuza (SK) with green stars. A zoomed-in map indicates the defined source region, the Johannesburg-Pretoria Megacity (grey polygon) and large point sources, i.e. power stations (blue triangles), petrochemical plants (red triangles) and pyrometallurgical smelters (yellow triangles)

A number of studies have been reported on measurements conducted within the INDAAF network (Martins et al., 2007; Adon et al., 2010; Josipovic et al., 2011; Adon et al., 2013),

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

104 presenting inorganic gaseous concentrations at southern-, as well as western- and central
105 African sites, respectively. Conradie et al. (2016) recently reported on precipitation chemistry
106 at the South African INDAAF sites, while Maritz et al. (2019) conducted an assessment of
107 particulate organic- and elemental carbon at these sites. However, in-depth analysis of long-
108 term trends of atmospheric pollutants at the INDAAF sites has not been conducted due to the
109 non-availability of long-term data. Therefore, the aim of this study was to perform statistical
110 modelling of SO_2, NO_2 and O_3 long-term trends based on 21-, 19- and 16-year datasets
111 available for LT, AF and SK, respectively. The influences of sources together with local,
112 regional and global meteorological patterns on the atmospheric concentrations of SO_2, NO_2
113 and O_3 were considered in the model.

114

115 **2 Measurement site and experimental methods**

116 **2.1 Site description**

117 Detailed site descriptions have been presented in literature, e.g. Mphepya et al. (2004),
118 Mphepya et al. (2006) and Conradie et al. (2016). AF (1 628 m amsl) and LT (1 300 m amsl)
119 are located within the South African Highveld, while SK is situated in the South African
120 Lowveld. As indicated in Fig.1, AF is in close proximity to the major industrial activities in
121 the Mpumalanga Highveld (~50 to 100 km north-west) and ~200 km east of the Johannesburg-
122 Pretoria conurbation. LT is located in a rural region mainly associated with agricultural activity,
123 while SK (267 m amsl) is situated in the Kruger National Park, i.e. natural bushveld in a
124 protected area.

125 A summary of the regional meteorology of the South African interior, especially relating to the
126 north-eastern part, was presented by Laakso et al. (2012) and Conradie et al. (2016).
127 Meteorology in the South African interior exhibits strong seasonal variability. This region is
128 characterised by anticyclonic air mass circulation, which is especially predominant during
129 winter, resulting in pronounced inversion layers trapping pollutants near the surface (Tyson et
130 al., 1996; Garstang et al., 1996). In addition, the north-eastern interior (as most parts of the
131 South African interior) is also characterised by distinct wet and dry seasons, with the wet season
132 occurring typically from mid-spring up to autumn (mid-October to mid-May) (Hewitson and
133 Crane, 2006; Conradie et al., 2016).

134 In Fig. 2, the air mass history for LT, AF and SK for the entire sampling periods at each site is
135 presented by means of overlaid back trajectories. 96-hour back trajectories arriving hourly at

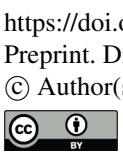
https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

each site at a height of 100 m were calculated with the Hybrid Single-Particle Langrangian Integrated Trajectory (HYSPLIT) model (version 4.8), developed by the National Oceanic and Atmospheric Administration (NOAA) Air Resources Laboratory (ARL) (Draxler and Hess, 2014).

Meteorological data was obtained from the GDAS archive of the National Centre for Environmental Prediction (NCEP) of the United States National Weather Service. Back trajectories were overlaid with fit-for-purpose programming software on a map area divided into grid cells of 0.2° x 0.2°. A colour scale presents the frequency of back trajectories passing over each grid cell, with dark blue indicating the lowest and dark red the highest percentage. The predominant anticyclonic air mass circulation over the interior of South Africa is reflected by the overlay back trajectories at each site, while it also indicates that AF is frequently impacted by air masses passing over the major sources in the north-eastern interior. In addition, it is also evident that the rural background sites (LT and SK) are also impacted by the regional circulation of air masses passing over the major sources.

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020



Atmospheric Chemistry and Physics Discussions

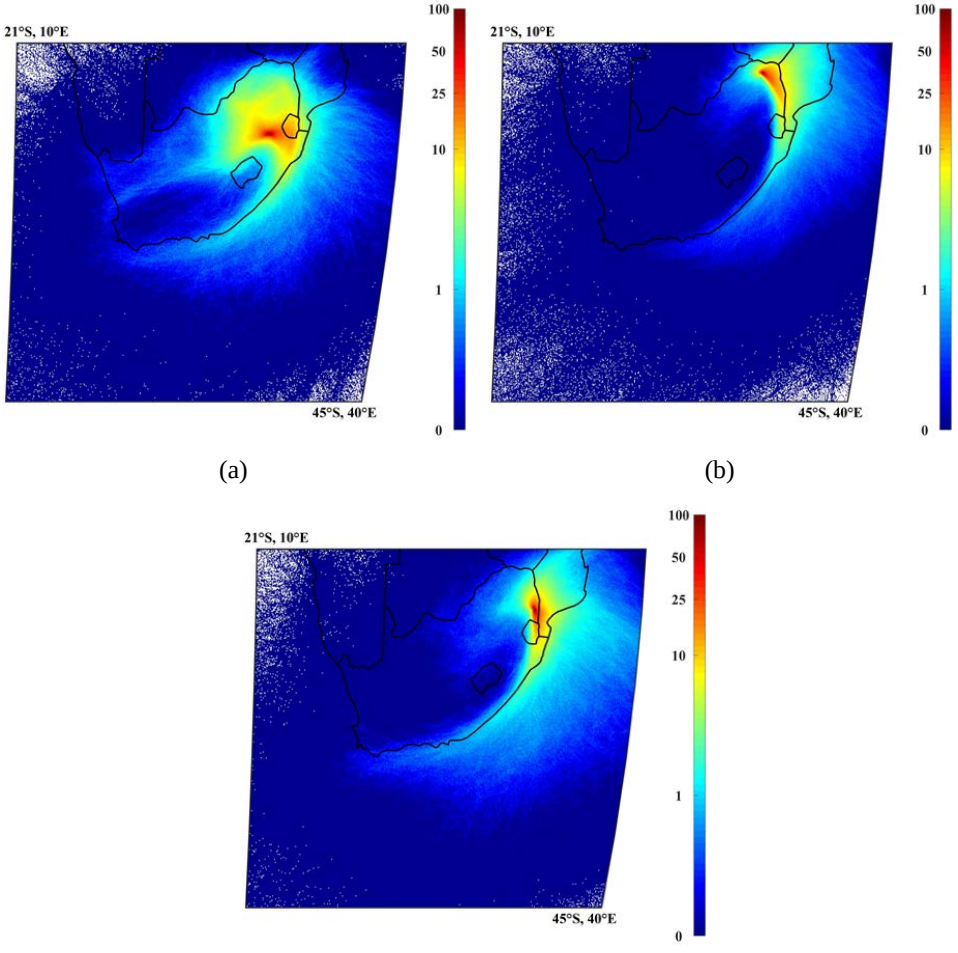


Figure 2: Overlaid hourly arriving 96-hour back-trajectories for air masses arriving at (a) AF from 1997 to 2015, (b) LT from 1995 to 2015 and (c) SK from 2000-2015

2.2 Sampling, analysis and data quality

Passively derived SO_2, NO_2 and O_3 concentrations were available from 1995 to 2015, 1997 to 2015 and 2000 to 2015 for LT, AF and SK, respectively. Gaseous SO_2, NO_2 and O_3 concentrations were measured utilising passive samplers manufactured at the North-West University, which are based on the Ferm (1991) passive sampler. Detailed descriptions on the theory and functioning of these passive samplers, which are based on laminar diffusion and chemical reaction of the atmospheric pollutant of interest, have been presented in literature

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

(Ferm, 1991; Dhammapala, 1996; Martins et al., 2007; Adon et al., 2010). In addition, the passive samplers utilised in this study have been substantiated through a number of inter-comparison studies (Martins et al., 2007; He and Bala, 2008).

Samplers were exposed in duplicate sets for each gaseous species at each measurement site (1.5 m above ground level) for a period of approximately one month and returned to the laboratory for analysis. Blank samples were kept sealed in the containers for each set of exposed samplers. Prior to 2008, SO_2 and O_3 passive samples were analysed with a Dionex 100 Ion Chromatograph (IC), while NO_2 samples were analysed with a Cary 50 uv/vis spectrometer up until 2012. SO_2 and O_3 samples collected after 2008, and NO_2 samples collected after 2012, were analysed with a Dionex ICS-3000 system. Data quality of the analytical facilities is ensured through participation in the World Meteorological Organisation (WMO) bi-annual Laboratory Inter-Comparison Study (LIS). The results of the 50th LIS study in 2014 indicated that the recovery of each ion in standard samples was between 95 and 105% (Conradie et al., 2016). Analysed data was also subjected to the Q-test, with a 95% confidence threshold to identify, evaluate and reject outliers in the datasets.

2.3 Multiple linear regression model

Similar to the approach employed by Swartz et al. (2019) for the Cape Point GAW station, a multiple linear regression (MLR) model was utilised to statistically evaluate the influence of sources and meteorology on the concentrations of SO_2, NO_2 and O_3 at AF, LT and SK. This model was also utilised by Toihir et al. (2018) and Bencherif et al. (2006) for trend estimates of O_3 and temperature, respectively. MLR analysis models the relationship between two or more independent variables and a dependant variable by fitting a linear equation to the observed data, which can be utilised to calculate values for the dependent variable. In this study, concentrations of inorganic gaseous species (SO_2, NO_2 and O_3) were considered the dependent variable (C(t)), while local, regional and global factors were considered independent variables to yield the following general equation:

$$C(t) = \sum_{k=1}^{p} a(k) \times f(t,k) + R'(t) \quad (1)$$

where f(t,k) describes the specific factor k at time t; a(k) is the coefficient calculated by the model for the factor k that minimises the root mean square error (RMSE); and R′(t) is the

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

residual term that accounts for factors that may have an influence on the model, which are not considered in the MLR model. The RMSE compares the calculated values with the measured values as follows;

$$\chi^2 = [\sum_t C(t) - \sum_k a(k) \times f(t,k)]^2 \quad (2)$$

The trend was parameterised as linear: Trend (t) = $\alpha_0 + \alpha_1.t$, where t denotes the time range, α_0 is a constant, α_1 is the slope of Trend(t) line that estimates the trend over the time scale.

The significance of each of the independent variables on the calculated C(t) was evaluated by the relative importance weights (RIW) approach, which examines the relative contribution that each independent variable makes to the dependent variable and ranks independent variables in order of significance (Nathans et al., 2012; Kleynhans et al., 2017). The RIW approach was applied with IBM® SPSS® Statistics Version 23, together with program syntaxes and scripts adapted from Kraha et al. (2012) and Lorenzo-Seva et al. (2010).

2.4 Input data

Global meteorological factors considered in the model included Total Solar Irradiation (TSI), the El-Niño Southern Oscillation (ENSO), the Indian Ocean Dipole (IOD), the Quasi-Biennial Oscillation (QBO) and the Southern Annular Mode (SAM). Data for the ENSO and QBO cycles was obtained from the National Oceanic and Atmospheric Administration (NOAA) (NOAA, 2015a; NOAA, 2015b), while TSI and IOD data was obtained from the Royal Netherlands Meteorological Institute (*"Koninklijk Nederlands Meteorologisch Instituut"*) (KMNI, 2016a; KMNI, 2016b). SAM data was obtained from the National Environmental Research Council's British Antarctic Survey (Marshall, 2018). The initial input parameters for the model only included the global force factors in order to assess the importance of individual global predictors on measured gaseous concentrations.

Local and regional meteorological parameters included in the model were rain depth (RF), relative humidity (RH) and ambient temperature (T), as well as monthly averaged wind direction (Wd) and -speed (Ws). Since meteorological parameters were not measured at the three sites during the entire sampling period, meteorological data was obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis-interim archive (ERA). Although meteorological measurements were conducted by the South African Weather Service within relative proximity of the locations of the three sites, the data coverage for all

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

the meteorological parameters for the entire sampling period was relatively low (<50%). Planetary boundary layer (PBL) heights were obtained from the global weather forecast model operated by the ECMWF (Korhonen et al., 2014). Population data (P) from three separate national censuses was obtained from local municipalities and was also included in the model.

Daily fire distribution data from 2000 to 2015 was derived from the National Aeronautics and Space Administration's (NASA) Moderate Resolution Imaging Spectrometer (MODIS) satellite retrievals. MODIS is mounted on the polar-orbiting Earth Observation System's (EOS) Terra spacecraft and globally measures, among others, burn scars, fire and smoke distributions. This dataset was retrieved from the NASA Distributed Active Archive Centres (DAAC) (Kaufman et al., 2003). Fire events were separated into local fire events (LFE), occurring within a 100 km radius from a respective site, and regional fire events (DFE), taking place between 100 km and 1 000 km from each site.

Hourly arriving back trajectories (as discussed above) were also used to calculate the percentage time that air masses spent over a predefined source region (Fig. 1) before arriving at each of the sites for each month, which was also a parameter (SR) included in the statistical model. The source region is a combination of source regions defined in previous studies, e.g. Jaars et al. (2014) and Booyens et al. (2019), which comprised the Mpumalanga Highveld, Vaal Triangle, the Johannesburg-Pretoria conurbation, the western- and the eastern Bushveld Igneous Complex, as well as a region of anticyclonic recirculation (Fig. 1).

Since data was not available for certain local and regional factors considered in the model for the entire sampling periods at AF, LT and SK, and, in an effort to include the optimum number of local and regional factors available for each site, modelled concentrations could not be calculated for the entire sampling periods when global, regional and local factors were included in the MLR model.

3 Results

Fig. A1, A2 and A3 present the time series of monthly average SO_2, NO_2 and O_3 concentrations measured at AF (1997 - 2015), LT (1995 - 2015) and SK (2000 - 2015). Seasonal and inter-annual variability associated with changes in the prevailing meteorology and source

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

contributions will be evaluated and statistically assessed using multiple linear regression model in subsequent sections.

3.1 Seasonal and inter-annual variability

In Fig. 3, 4 and 5, the monthly SO_2, NO_2 and O_3 concentrations, respectively at AF, LT and SK, determined for the entire sampling periods, are presented. Monthly variability in concentrations of these species at these three sites is expected. The north-eastern interior of South Africa, where these sites are located, is generally characterised by increased concentrations in pollutant species during the dry winter months (June to September) due to the prevailing meteorological conditions (Conradie et al., 2016). More pronounced inversion layers trap pollutants near the surface, which, in conjunction with increased anticyclonic recirculation and decreased wet deposition, leads to the build-up pollutant levels (Conradie et al., 2016; Laban et al., 2018). In addition, increased household combustion for space heating during winter also contributes to higher levels of atmospheric pollutants, while open biomass burning (wild fires) is also a significant source of atmospheric species in late winter and spring (August to November). Species typically associated with biomass burning (open or household) include particulate matter (PM), CO and NO_2, while household combustion can also contribute to SO_2 emissions depending on the type of fuel consumed. CO and NO_2 are also important precursors of tropospheric O_3, which also lead to increased surface O_3 concentrations, especially with increased photochemical activity in spring (Laban et al., 2018). From Fig. 3, it is evident that SO_2 concentrations peaked in winter months at LT and SK, while SO_2 levels did not reveal significant monthly variability at AF throughout the year. NO_2 and O_3 concentrations at all three sites are higher during August to November, coinciding with open biomass burning. NO_2 and O_3 levels at AF do not reflect the influence of pollutant build-up in winter, although the whiskers in July do indicate more instances of higher NO_2 concentrations. SK did indicate higher NO_2 and O_3 concentrations during June and July, while LT also had relatively higher O_3 concentrations during July.

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

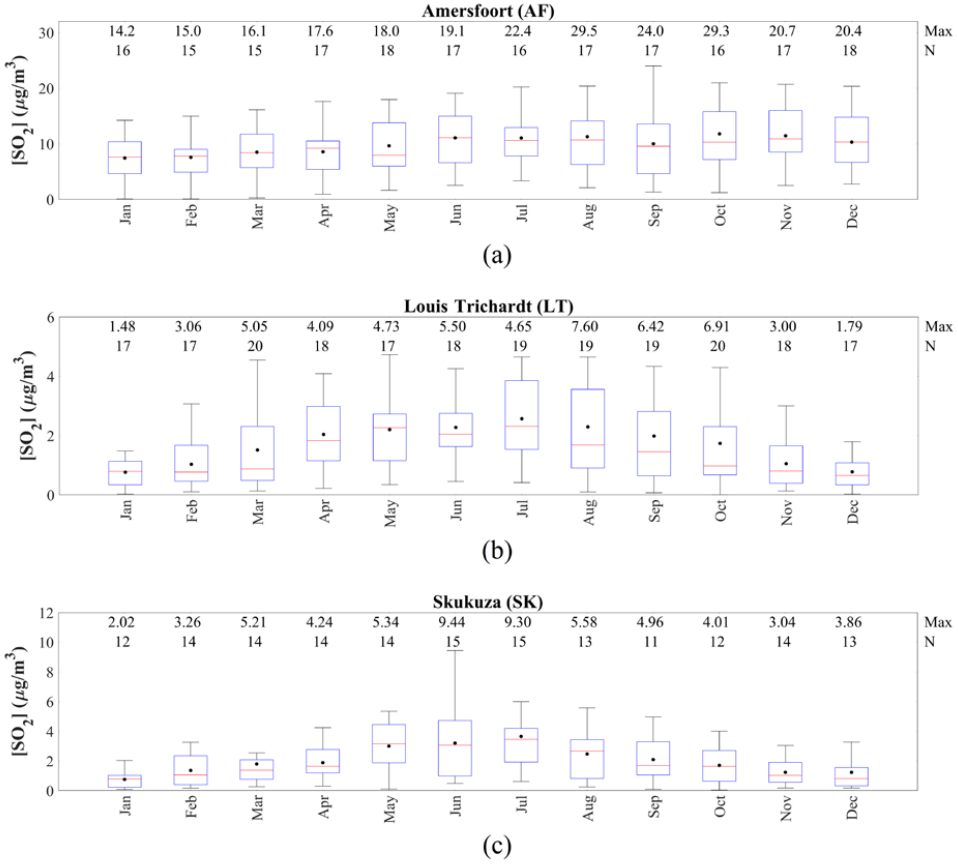
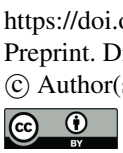


Figure 3: Monthly SO_2 concentrations measured at (a) AF from 1997 to 2015, (b) LT from 1995 to 2015 and (c) SK from 2000 to 2015. The red line of each box represents the median, the top and bottom edges of the box the 25th and 75th percentiles, respectively, the whiskers ±2.7σ (99.3% coverage if the data has a normal distribution) and the black dots the averages. The maximum concentrations and the number of measurements (N) are presented at the top

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020



Atmospheric Chemistry and Physics Discussions

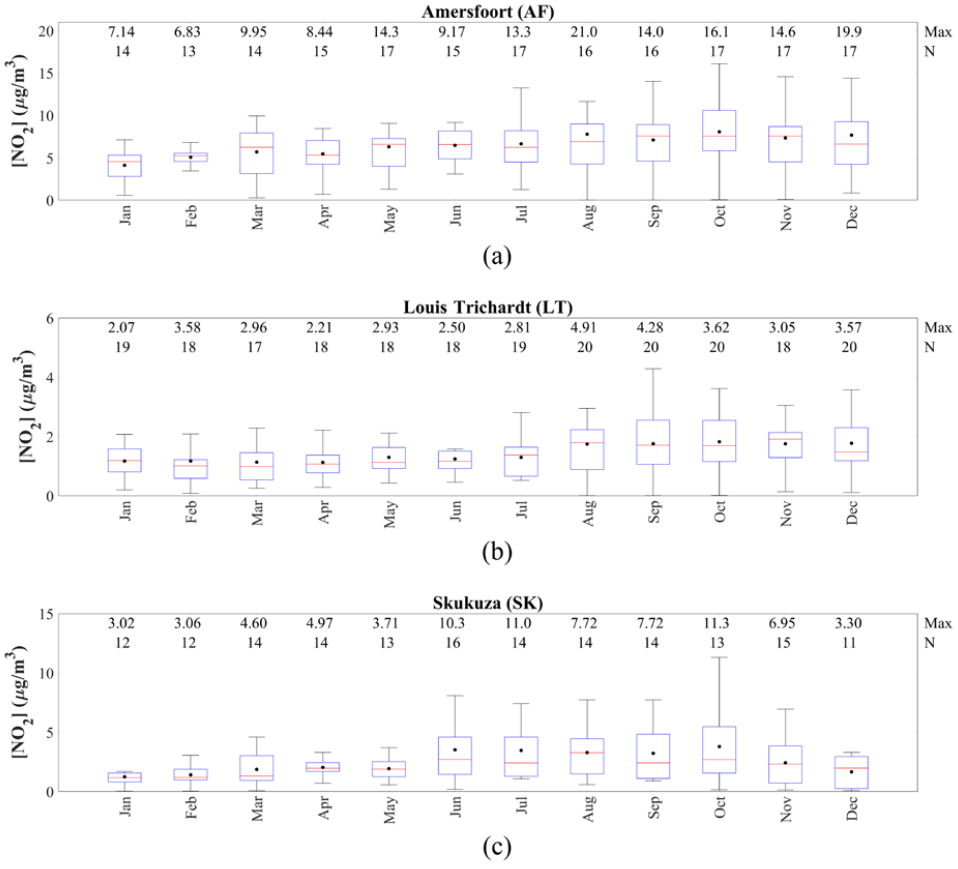


Figure 4: Monthly NO_2 concentrations measured at (a) AF from 1997 to 2015, (b) LT from 1995 to 2015 and at (c) SK from 2000 to 2015. The red line of each box represents the median, the top and bottom edges of the box the 25th and 75th percentiles, respectively, the whiskers ±2.7σ (99.3% coverage if the data has a normal distribution) and the black dots the averages. The maximum concentrations and the number of measurements (N) are presented at the top

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

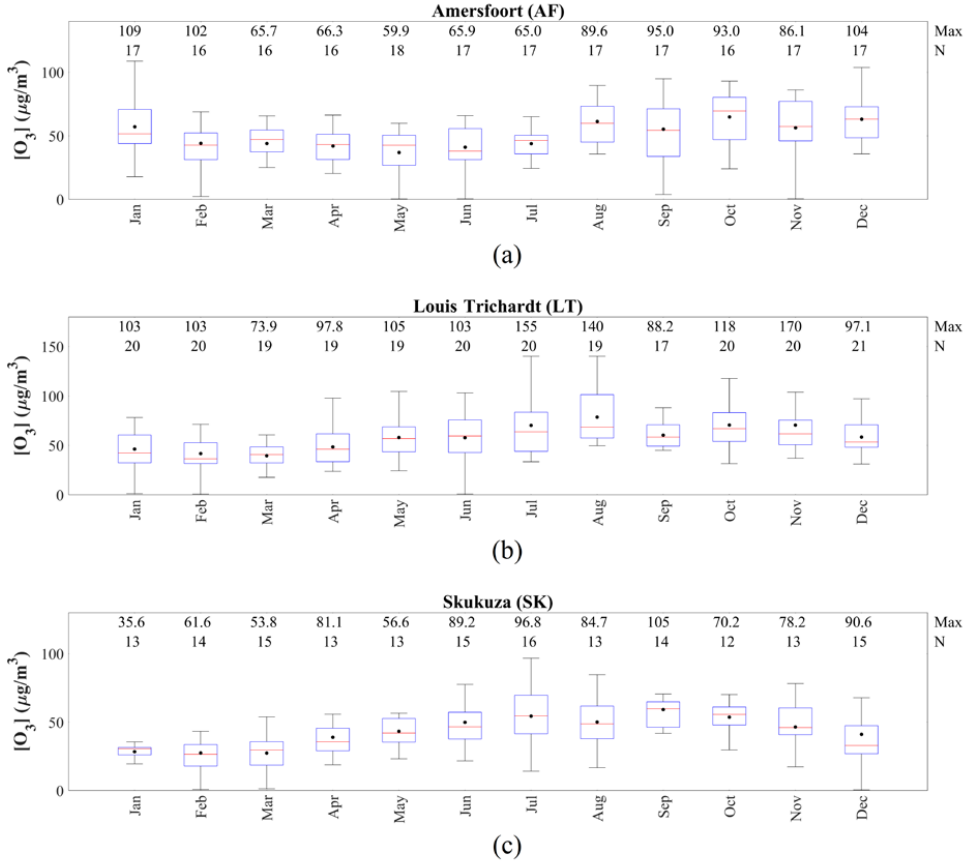


Figure 5: Monthly O_3 concentrations measured at (a) AF from 1997 to 2015, (b) LT from 1995 to 2015 and (c) SK from 2000 to 2015. The red line of each box represents the median, the top and bottom edges of the box the 25th and 75th percentiles, respectively, the whiskers ±2.7σ (99.3% coverage if the data has a normal distribution) and the black dots the averages. The maximum concentrations and the number of measurements (N) are presented at the top

The inter-annual variability of SO_2, NO_2 and O_3 levels is presented in Fig. 6, 7 and 8, respectively for AF, LT and SK. Noticeable from the SO_2 and NO_2 inter-annual fluctuations at all three sites is that the annual average SO_2 and NO_2 concentrations decreased up until 2003/2004 and 2002, respectively, which is followed by a period during which levels of SO_2 and NO_2 increased up until 2009 and 2007, respectively. After 2009, annual average SO_2 concentrations remained relatively constant, while NO_2 showed relatively large inter-annual variability, with annual NO_2 concentrations reaching a maximum in 2011 and 2012. These

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

observed periods of decreased and increased SO_2 and NO_2 levels are also indicated by the three-year moving averages of the annual mean SO_2 and NO_2 concentrations at all three sites. Since these trends are observed at all three sites, located several kilometres apart in the north-eastern interior, these inter-annual trends seem real and not merely a localised artefact. Furthermore, monthly SO_2 and NO_2 measurements conducted at the Cape Point Global Atmosphere Watch station on the west coast of South Africa also indicate similar periods of increase and decrease in SO_2 and NO_2 levels (Swartz et al., 2019). Although annual O_3 concentrations indicate inter-annual variances, annual average O_3 concentrations remained relatively constant at all three sites, with the exception of a decreasing trend observed from 1995 to 2001 at LT corresponding to the period during which SO_2 and NO_2 decreased. Similar to seasonal variances, inter-annual fluctuations can also be ascribed to changes in meteorological conditions and/or variances in source contribution. Conradie et al. (2016), for example, indicated that rain samples collected from 2009 to 2014 at these three sites had higher SO_4^{2-} and NO_3^- concentrations compared to rain samples collected in 1986 to 1999 and 1999 to 2002, which is attributed to increased energy demand and a larger vehicular fleet associated with economic- and population growth.

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

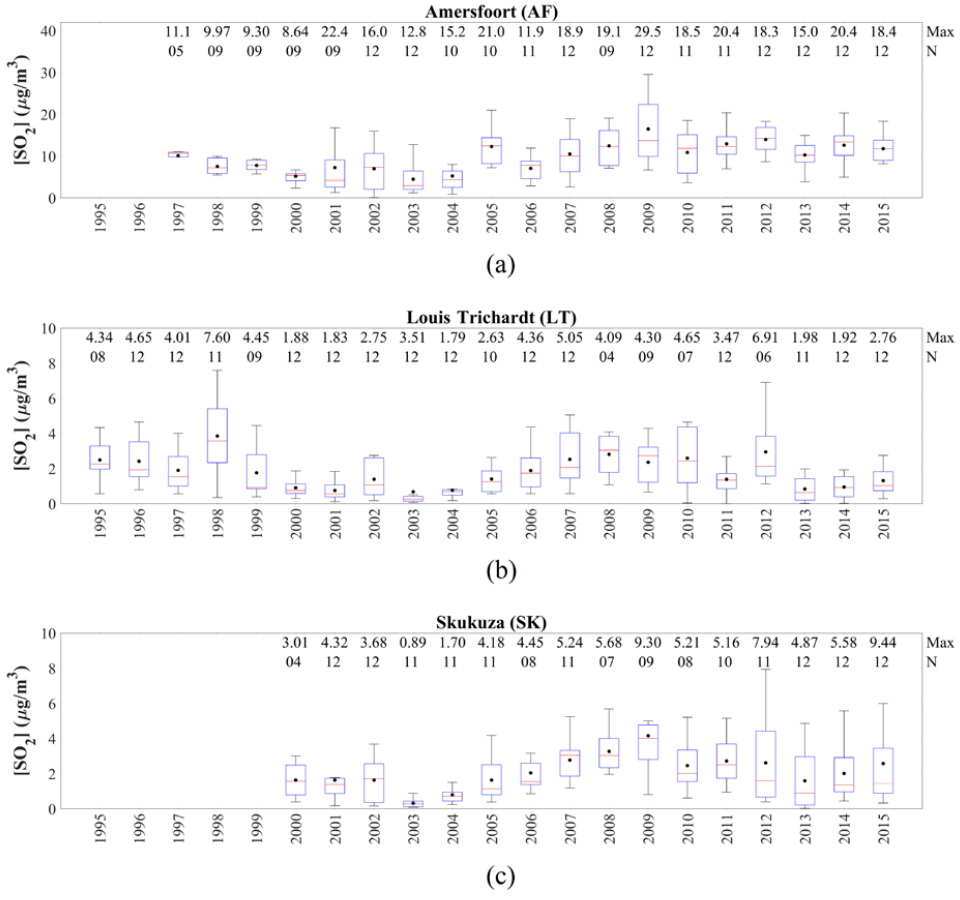


Figure 6: Annual SO_2 concentrations at (a) AF, (b) LT and (c) SK. The red line of each box represents the median, the top and bottom edges of the box the 25th and 75th percentiles, respectively, the whiskers ±2.7σ (99.3% coverage if the data has a normal distribution) and the black dots the averages. The maximum concentrations and the number of measurements (N) are presented at the top

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

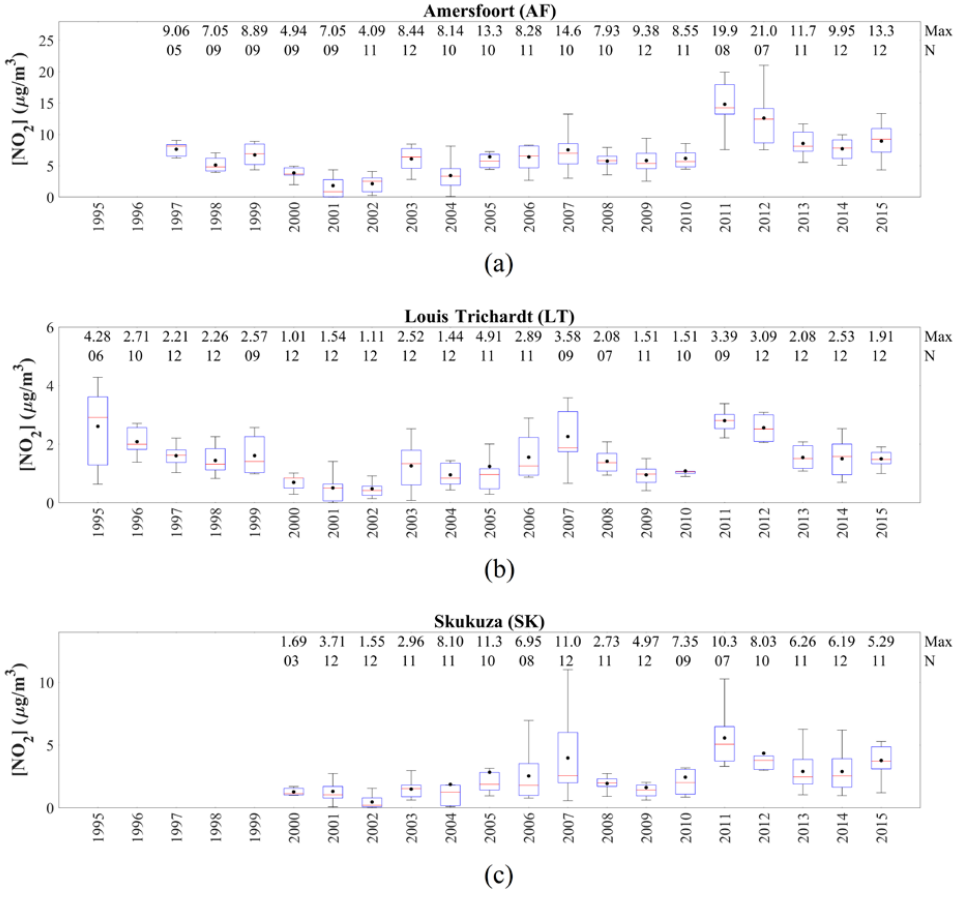


Figure 7: Annual NO_2 concentrations at (a) AF, (b) LT and (c) SK. The red line of each box represents the median, the top and bottom edges of the box the 25th and 75th percentiles, respectively, the whiskers ±2.7σ (99.3% coverage if the data has a normal distribution) and the black dots the averages. The maximum concentrations and the number of measurements (N) are presented at the top

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

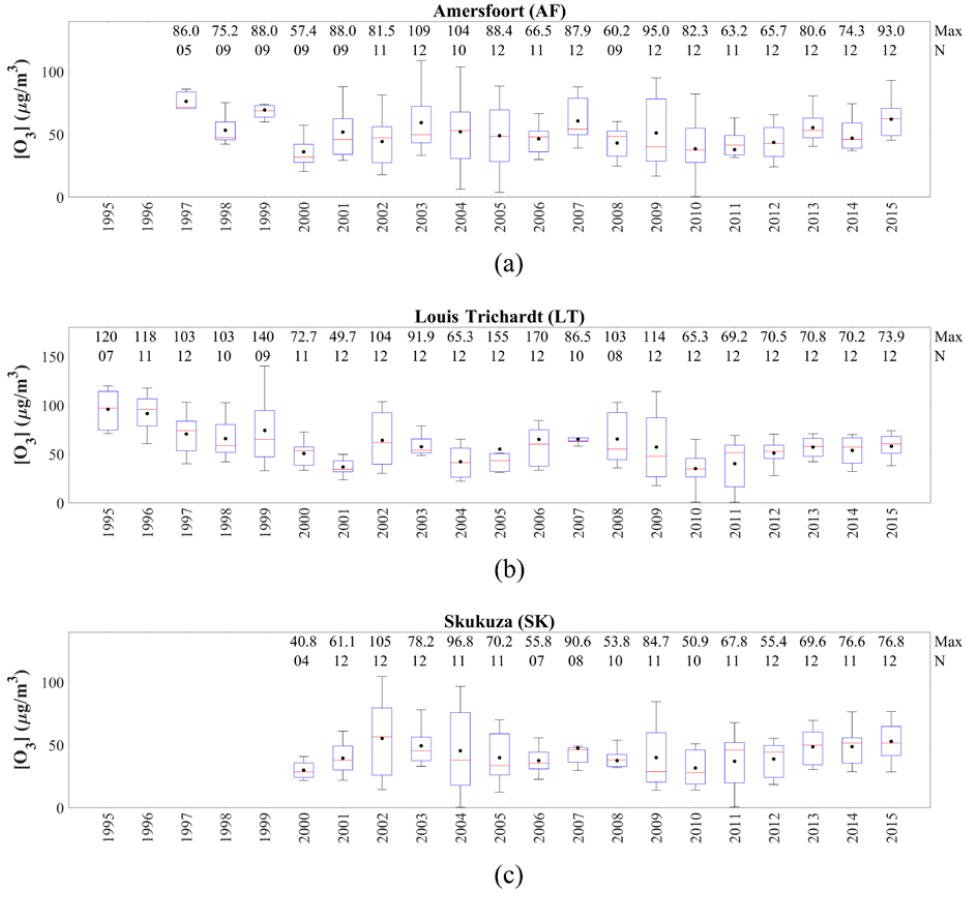


Figure 8: Annual O_3 concentrations at (a) AF, (b) LT and (c) SK. The red line of each box represents the median, the top and bottom edges of the box the 25th and 75th percentiles, respectively, the whiskers ±2.7σ (99.3% coverage if the data has a normal distribution) and the black dots the averages. The maximum concentrations and the number of measurements (N) are presented at the top

3.2 Statistical modelling of variability

3.2.1 Sulphur dioxide (SO_2)

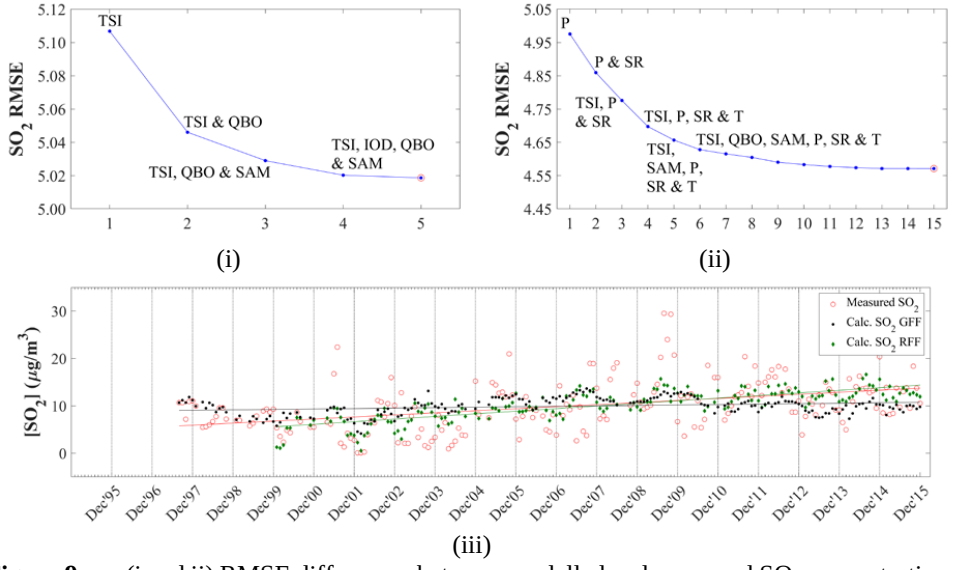
The SO_2 concentrations calculated with the MLR model are compared to measured SO_2 levels in Fig. 9 for AF (Fig. 9a), LT (Fig. 9b) and SK (Fig. 9c). In each sub-figure, the RMSE differences between measured and modelled SO_2 concentrations are presented as a function of the number of independent variables included in the model (i and ii), while the differences between modelled and measured SO_2 levels for each sample are also indicated (iii). As

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

354 indicated above, in the initial run of the model, only global factors were included (i and iii),
355 after which all factors (local, regional and global) were incorporated in the model (ii and iii).
356 In Table 1, the coefficients and RIW% of each of the independent variables are included in the
357 optimum MLR equation containing all global factors, as well as in the optimum MLR equation
358 when all local, regional and global factors are included. It is evident from Fig. 9 (iii) that the
359 correlations between measured and modelled SO_2 levels are significantly improved when all
360 factors are considered in the MLR model compared to only including global factors at all three
361 sites. The R^2 values are improved from 0.122 to 0.330, 0.078 to 0.257, and 0.100 to 0.389 at
362 AF, LT and SK, respectively. Although relatively weak correlations are observed between
363 modelled and measured SO_2 levels, the general trend of the measured SO_2 concentrations is
364 mimicked by the modelled values, even when only global factors are included in the MLR
365 model. In addition, the R^2 values at AF and SK when all factors are considered (0.330 and
366 0.389) can be considered moderate correlations (Kleynhans et al., 2017). It also seems that very
367 high and low SO_2 levels are underestimated by the model. Swartz et al. (2019) attributed
368 differences between monthly concentrations of species measured with passive samplers at CPT
369 GAW and modelled levels to the limitations associated with the use of passive samplers.

370



371 **Figure 9a:** (i and ii) RMSE differences between modelled and measured SO_2 concentrations
372 as a function of the number of independent variables included in the model, as
373 well as comparison between modelled and measured SO_2 levels (iii) for global
374 force factors only (GFF), and for global, regional and local factors (RFF)
375 determined for AF

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

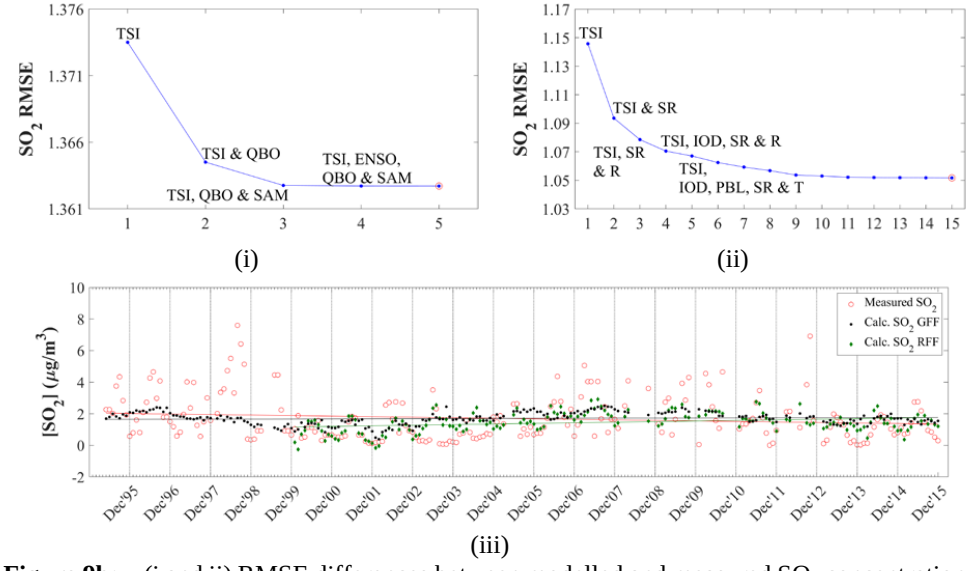


Figure 9b: (i and ii) RMSE differences between modelled and measured SO_2 concentrations as a function of the number of independent variables included in the model, as well as comparison between modelled and measured SO_2 levels (iii) for global force factors only (GFF), and for global, regional and local factors (RFF) determined for LT

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

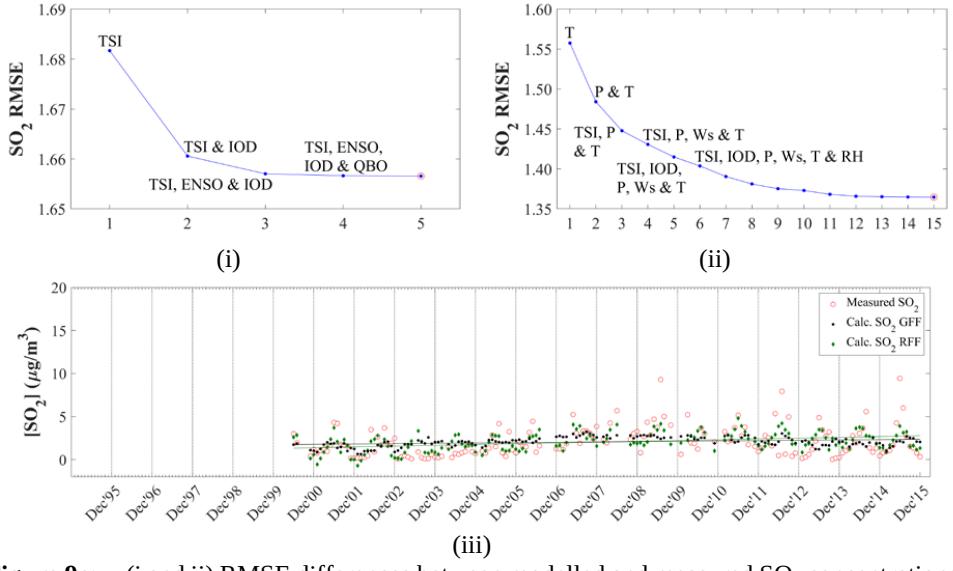


Figure 9c: (i and ii) RMSE differences between modelled and measured SO_2 concentrations as a function of the number of independent variables included in the model, as well as comparison between modelled and measured SO_2 levels (iii) for global force factors only (GFF), and for global, regional and local factors (RFF) determined for SK

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

Table 1: Regression coefficients (b) and relative important weight percentage (RIW%) of each independent variable included in the MLR model to calculate SO_2 concentrations at AF, LT and SK

	AF			**LT**			**SK**	
	b	***RIW%***		***b***	***RIW%***		***b***	***RIW%***
i) Global forcing factors								
TSI	-3.563	66.2	TSI	-0.875	80.2	TSI	-0.988	61.6
QBO	-0.057	21.2	QBO	-0.011	15.2	IOD	1.183	33.8
IOD	0.818	5.5	SAM	-0.042	3.9	ENSO	-0.158	3.7
SAM	-0.209	5.0	IOD	-0.011	0.5	QBO	-2.500×10^{-3}	0.7
ENSO	0.170	2.0	ENSO	-0.012	0.2	SAM	-0.010	0.3
ii) Global, regional and local factors								
P	1.927×10^{-3}	54.5	TSI	-0.827	34.7	T	-0.281	15.9
TSI	-2.373	14.6	SR	0.069	11.3	TSI	-0.820	12.0
SR	0.189	6.2	T	-0.109	9.9	SR	0.076	9.9
T	-0.588	4.5	IOD	0.588	8.0	P	5.610×10^{-6}	9.1
QBO	-0.034	4.4	R	6.448×10^{-4}	6.7	Ws	-1.357	9.1
RH	0.043	3.9	RH	-0.014	6.2	PBL	3.134×10^{-3}	8.4
PBL	6.396×10^{-3}	2.8	Ws	-0.404	5.1	R	9.233×10^{-4}	7.4
SAM	-0.406	2.6	PBL	1.520×10^{-3}	4.9	RH	-0.024	7.0
R	-1.104×10^{-3}	1.8	Wd	2.746×10^{-3}	3.1	IOD	1.011	6.7
Ws	0.076	1.5	P	-1.035×10^{-6}	2.7	Wd	-4.034×10^{-4}	5.6
IOD	-0.674	0.9	SAM	-0.049	2.4	LFE	5.827×10^{-5}	4.5
LFE	1.114×10^{-4}	0.9	DFE	-2.892×10^{-7}	2.0	DFE	-3.355×10^{-6}	2.2
Wd	-3.502×10^{-3}	0.6	QBO	-6.471×10^{-3}	1.6	ENSO	-0.260	1.7
DFE	-1.319×10^{-5}	0.5	LFE	-8.706×10^{-5}	0.8	SAM	-0.078	0.5
ENSO	-0.310	0.3	ENSO	-0.034	0.6	QBO	-2.726×10^{-3}	0.2

The interdependencies between TSI and QBO at AF and LT, as well as TSI and IOD at SK yielded the largest decreases in RMSE when only global parameters were considered. The RIW% calculated for these parameters in the optimum MLR equation containing all global factors also indicates that these factors are the most significant. When all factors (local, regional and global) were considered in the model, the combinations between P, TSI, SR and T at AF, TSI, SR, IOD and R at LT, and T, TSI, P and Ws contributed to the most significant decrease in RMSE for each of the sites. According to the RIW% calculated for each parameter in the optimum MLR equation containing all factors P (54.5%) and TSI (14.6%) at AF, TSI (34.7%), SR (11.3%), T (9.9%) and IOD (8.0%) at LT, and T (15.9%), TSI (12.0%), SR (9.9%), P (9.1%) and Ws (9.1%) at SK were the most important factors contributing to variances. From the MLR

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

model, it is evident that global meteorological factors contribute to SO_2 variability at each of these sites located in the north-eastern interior of South Africa. The model also indicates that the influence of global factors is more significant at the rural background site LT, where TSI made the largest contribution to the modelled value, while IOD also made a relatively important contribution. Although TSI was the second most significant factor at AF and SK, local and regional parameters were more important to variances in modelled SO_2 levels at these sites.

Population growth had the most substantial contribution to the dependent variable at the industrially influenced AF, which is indicative of the impacts of increased anthropogenic activities and energy demand in this region. Therefore, it is most-likely that the observed inter-annual variability observed at AF, i.e. periods of decreased and increased SO_2 levels, can mainly be attributed to changes in source contribution. The decrease in SO_2 concentrations up until 2003/2004 is associated with a period post-1994 (when the new democracy was established) during which many companies obtained environmental accreditation (ISO 14000 series, ISO survey (2015)) and implemented mitigation technologies in order to comply with international trade requirements, e.g. certain large metallurgical smelters applied desulphurisation technologies (e.g. Westcott et al., 2007). The period was characterised by an increased awareness of air pollution and its impacts in South Africa. However, it seems that these improvements made with regard to air pollution were offset from 2003/2004 due to rapid economic growth associated with increased industrial activities, e.g. increased production by pyrometallurgical industries (ICDA, 2012), as well as the increase in population growth accompanied by higher energy demand (Vet et al., 2014). Electricity consumption is a good indicator of increased anthropogenic activities, with Inglesi-Lotz and Blignaut (2011) indicating that electricity consumption in South Africa increased by 131 024 GWh from 1993 to 2006. In 2007/2008, the global financial crisis occurred, which forced numerous South African commodity-based producers (e.g. platinum group metal, base metal, ferrochromium, ferromanganese, ferrovanadium and steel smelters) to completely discontinue production. Ferrochromium production in South Africa, for instance, decreased by approximately 35% from 2007 to 2009 (ICDA, 2013), while energy consumption in the manufacturing sector dropped by approximately 34% from 2007 to 2008 (Statistics South Africa, 2012). Furthermore, these variances in source contribution associated with anthropogenic activities are also observed at LT and SK distant from the major sources due to these sites also being impacted by the regional circulation of air masses passing over major sources, as indicated in Fig. 2. In addition, the RIW% associated with P (9.1%) in the optimum MLR equation

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

containing all factors at SK is also indicative of not only the influence of population growth within the source region (Fig. 1), but also the increased populations of rural communities on the border of the Kruger National Park. Maritz et al. (2019) attributed higher organic- and elemental carbon concentrations measured at SK to increased biomass burning by these rural communities.

Temperature had the largest contribution to the variances of the modelled SO_2 at SK, while it was also an important parameter at LT. In addition, the source region (SR) factor made significant contributions to the dependent variable at SK and LT, while it also made a relative contribution at AF. These two factors are indicative of the influence of changes in local and regional meteorological conditions on SO_2 concentrations, as well as the important influence of air mass movement over the source region. The contribution of SR at all the sites indicated that months and/or years coinciding with these sites being more frequently impacted by air masses passing over the defined source region (Fig. 1) corresponded to increased SO_2 concentrations, while it also substantiates the afore-mentioned deduction that increased anthropogenic activities in the source region also influenced LT and SK. As indicated in section 3.1, SK and LT revealed the expected higher SO_2 levels during winter, while AF had a less distinct seasonal pattern. Therefore, the strong negative correlation between temperature and modelled SO_2 concentrations at SK and LT, i.e. higher SO_2 levels associated with lower temperature, reflects the influence of local and regional meteorology on monthly SO_2 variability, i.e. build-up of pollutant concentrations during winter. At SK, the influence of local meteorology is also indicated by the relative strong negative correlation to Ws, i.e. more stable conditions in winter coinciding with higher SO_2 concentrations. Furthermore, the influence of the rural communities in proximity of SK on SO_2 levels is also signified by T being the most significant factor contributing to modelled SO_2 values at this site. The less distinct seasonal pattern at AF can be attributed to the proximity of AF to the industrial SO_2 sources, with the major point sources consistently emitting the same levels of SO_2 throughout the year. Therefore, the average monthly SO_2 concentrations measured with passive samplers at AF do not reflect the influence of local and regional meteorology on atmospheric SO_2 concentrations.

The slopes of the trend lines of SO_2 values calculated when only global factors were included in the model did not correspond with the trend lines of the measured SO_2 concentrations at all the sites, with the exception of LT that showed slightly better correlations, signifying the stronger influence of global factors at this site (Pane iii in Fig. 9a, b and c). However, the slopes of the linear regression trend lines for the measured SO_2 concentrations and the modelled SO_2

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

levels when all the factors are included in the model are exactly the same at AF, LT and SK when the same period is considered for both the modelled and measured values. A positive slope for the 19-year trend line for measured SO_2 concentrations is observed at AF (Fig. 9a(iii)), indicating an increase in SO_2 levels over the 19-year sampling period, i.e. 0.43 $\mu g.m^{-3}.y^{-1}$. An increase in SO_2 concentration, i.e. 0.09 $\mu g.m^{-3}.y^{-1}$ is also determined for the 16-year measurement period at SK (Fig. 9b(iii)), which is significantly smaller than the upwards trend at AF. In contrast to AF and SK, LT indicates a slight net negative slope with SO_2 decreasing on average by 0.03 $\mu g.m^{-3}.y^{-1}$ during the 21-year sampling period (Fig. 9c(iii)). The 19- and 21-year datasets at AF and LT also allowed for the calculation of decadal trends, which were determined to be 5.24 $\mu g.m^{-3}.dec^{-1}$ (average SO_2 concentrations from 1997 to 2006 were 7.20 $\mu g.m^{-3}$ and average SO_2 concentrations from 2007 to 2015 were 12.44 $\mu g.m^{-3}$) and 0.18 $\mu g.m^{-3}.dec^{-1}$ (average SO_2 concentrations from 1995 to 2004 were 1.64 $\mu g.m^{-3}$ and average SO_2 concentrations from 2005 to 2014 were 1.82 $\mu g.m^{-3}$), respectively, for the two decades. Trend lines are also presented for the periods characterised by increased (1995, 1997 to 2003) and decreased (2004 to 2008/2009) SO_2 concentrations at LT and AF. The average annual trend between 1997 and 2003 at AF was -0.53 $\mu g.m^{-3}.y^{-1}$, while the annual trend from 2004 to 2009 was 1.87 $\mu g.m^{-3}.y^{-1}$. At LT, the average annual SO_2 concentrations decreased by -0.26 $\mu g.m^{-3}.y^{-1}$ from 1995 to 2002, and increased by 0.37 $\mu g.m^{-3}.y^{-1}$ from 2003 to 2007.

3.2.2 Nitrogen dioxide (NO_2)

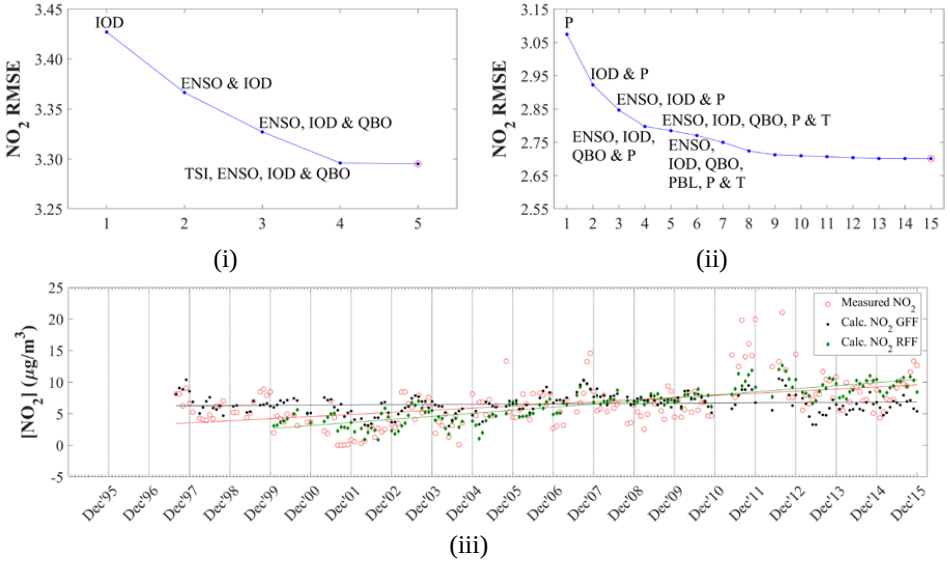
In Fig. 10, the measured NO_2 concentrations are related to the modelled NO_2 levels, while Table 2 presents the coefficients and RIW% of each of the independent variables included in the optimum MLR equation modelling NO_2 concentrations. Similar to SO_2, the relationships between measured and modelled NO_2 are also significantly improved when local, regional and global factors are included in the model at all three sites (Pane iii in Fig. 10a, b and c). However, inclusion of only global factors in the model yielded modelled NO_2 concentrations that mimicked the general measured NO_2 trend. The R^2 values, when only global factors are included, i.e. 0.171, 0.170 and 0.099 at AF, LT and SK, respectively, are enhanced to 0.498, 0468 and 0.362 at AF, LT and SK, respectively, when all factors are considered in the MLR model. The R^2 values, when all factors are included, especially AF and LT, can be considered relatively good correlations (Sheskin, 2003). In general, modelled NO_2 concentrations

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

500 corresponded well with the observed variances in measured NO_2 levels when all factors are
501 included in the model at all three sites, with the exception of very high NO_2 concentrations.

502



503 **Figure 10a:** (i and ii) RMSE differences between modelled and measured NO_2 concentrations
504 as a function of the number of independent variables included in the model, as
505 well as comparison between modelled and measured NO_2 levels (iii) for global
506 force factors only (GFF), and for global, regional and local factors (RFF)
507 determined for AF

508

509

510

511

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

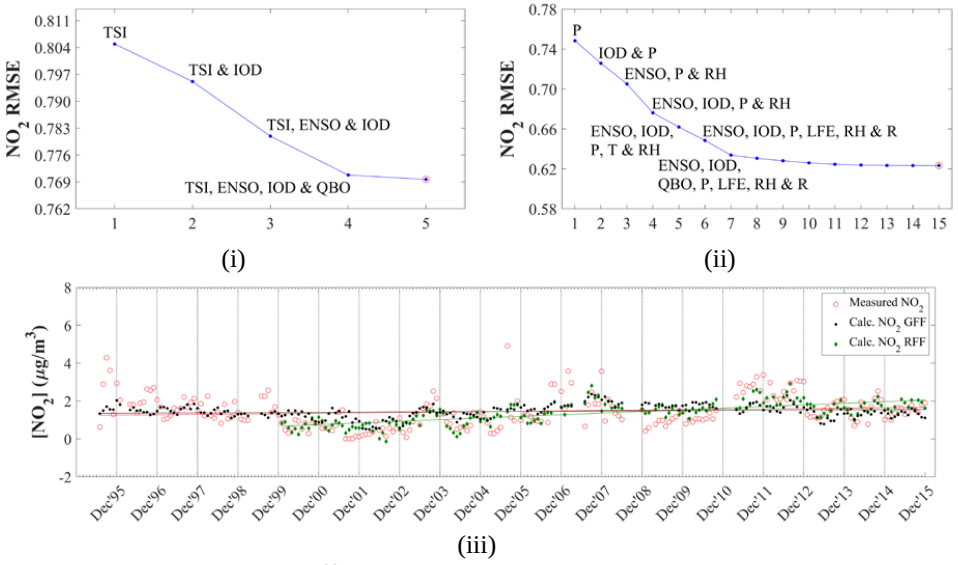


Figure 10b: (i and ii) RMSE differences between modelled and measured NO_2 concentrations as a function of the number of independent variables included in the model, as well as comparison between modelled and measured NO_2 levels (iii) for global force factors only (GFF), and for global, regional and local factors (RFF) determined for LT

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

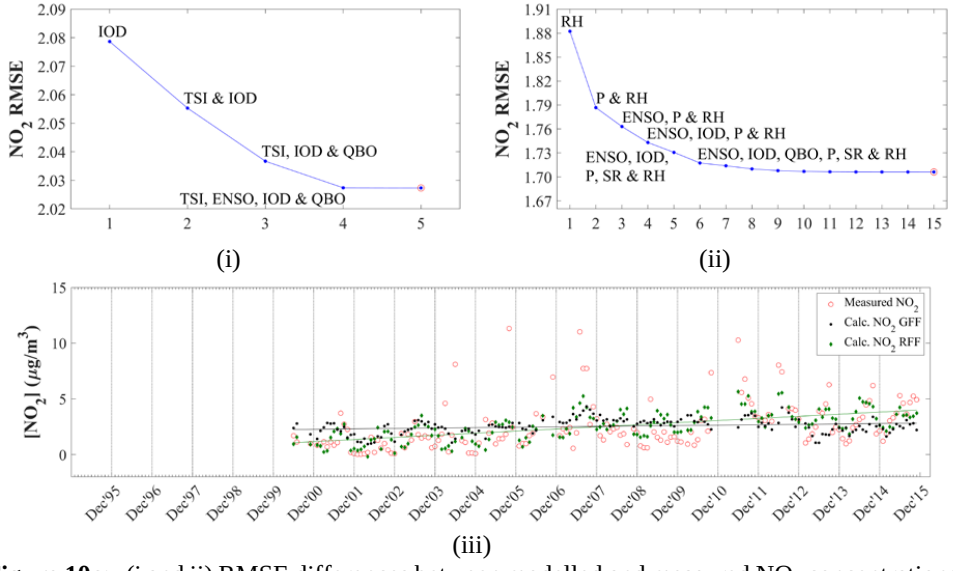


Figure 10c: (i and ii) RMSE differences between modelled and measured NO_2 concentrations as a function of the number of independent variables included in the model, as well as comparison between modelled and measured NO_2 levels (iii) for global force factors only (GFF), and for global, regional and local factors (RFF) determined for SK

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

527 **Table 2:** Regression coefficients (b) and relative important weight percentage (RIW%) of
528 each independent variable included in the MLR model to calculate NO_2
529 concentrations at AF, LT and SK

	AF			**LT**			**SK**	
	b	***RIW%***		***b***	***RIW%***		***b***	***RIW%***
i) Global forcing factors								
IOD	4.718	65.3	TSI	-0.625	52.4	IOD	1.954	49.4
TSI	-1.156	15.1	IOD	0.723	25.5	TSI	-0.698	27.6
QBO	-0.037	10.5	QBO	-9.326×10^{-3}	11.8	QBO	-0.018	15.4
ENSO	-0.798	8.6	ENSO	-0.186	8.9	ENSO	-0.301	7.1
SAM	0.047	0.5	SAM	0.025	1.4	SAM	-8.422×10^{-3}	0.5
ii) Global, regional and local factors								
P	1.444×10^{-3}	53.7	P	1.512×10^{-5}	29.9	P	1.366×10^{-5}	29.8
IOD	3.861	17.8	RH	-0.056	16.6	RH	-0.090	20.6
RH	-0.036	6.0	IOD	0.916	15.2	IOD	1.032	7.1
QBO	-0.028	3.5	TSI	-0.186	8.4	DFE	1.473×10^{-7}	6.9
PBL	5.119×10^{-3}	3.2	ENSO	-0.327	6.8	R	3.833×10^{-3}	6.1
TSI	0.040	2.8	QBO	-9.368×10^{-3}	6.5	LFE	3.800×10^{-6}	4.1
ENSO	-0.965	2.7	R	2.482×10^{-3}	3.8	SR	0.073	4.0
Ws	0.075	2.7	DFE	-6.055×10^{-7}	2.9	T	-0.072	3.8
T	-0.415	2.5	PBL	-1.225×10^{-3}	2.5	TSI	-0.160	3.7
R	0.014	1.5	T	0.069	1.9	QBO	-0.015	3.6
LFE	-1.229×10^{-4}	1.0	LFE	-2.134×10^{-4}	1.8	ENSO	-0.441	3.1
DFE	-5.044×10^{-6}	0.9	Ws	0.107	1.5	Ws	0.313	3.0
SR	0.028	0.6	SAM	0.021	0.8	Wd	4.912×10^{-4}	1.9
Wd	-1.419×10^{-3}	0.6	SR	0.010	0.8	PBL	1.567×10^{-4}	1.8
SAM	-0.141	0.5	Wd	-1.587×10^{-4}	0.6	SAM	-0.025	0.5

530

531 The annual trend calculated from the slope of the 19-year measured NO_2 dataset at AF indicates
532 an annual increase of 0.33 $\mu g.m^{-3}.y^{-1}$, while the 16-year measured NO_2 concentrations indicate
533 an upwards trend of 0.19 $\mu g.m^{-3}.y^{-1}$ at SK. The trend line of measured NO_2 concentrations at
534 LT also indicated a marginal increase, i.e. 0.02 $\mu g.m^{-3}.y^{-1}$ in NO_2 levels over the 21-year
535 sampling period. Decadal trends were determined to be 3.43 $\mu g.m^{-3}.dec^{-1}$ (average NO_2
536 concentrations from 1997 to 2006 were 4.86 $\mu g.m^{-3}$ and average NO_2 concentrations from 2007
537 to 2015 were 8.29 $\mu g.m^{-3}$) and 0.45 $\mu g.m^{-3}.dec^{-1}$ (average NO_2 concentrations from 1995 to
538 2004 were 1.23 $\mu g.m^{-3}$ and average NO_2 concentrations from 2005 to 2014 were 1.68 $\mu g.m^{-3}$),
539 respectively, for the two decades. Trend lines were also calculated for the periods coinciding
540 with increases and decreases in measured NO_2 concentrations at AF and LT. The average

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

annual trend between 1997 and 2003 at AF was -0.26 $\mu g.m^{-3}.y^{-1}$, while the annual trend from 2004 to 2009 was 0.37 $\mu g.m^{-3}.y^{-1}$. At LT, the average annual NO_2 concentrations decreased by -0.29 $\mu g.m^{-3}.y^{-1}$ from 1995 to 2002, and increased by 0.28 $\mu g.m^{-3}.y^{-1}$ from 2003 to 2007. Similar to SO_2, the slopes of the linear regression trend lines for the measured NO_2 concentrations and the modelled NO_2 levels when all the factors are included in the model are exactly the same at AF, LT and SK (Pane iii in Fig. 10a, b and c). However, with the exception of LT, the slopes of the trend lines of NO_2 levels calculated including only global factors in the model did not correspond with the trend lines of the measured NO_2 concentrations, indicating the significance of local and regional factors on measured NO_2 concentrations (Pane iii in Fig. 10a, b and c).

The RMSE differences between the modelled and measured NO_2 concentrations (Pane i Fig. 10a, b and c) indicated that the linear combination between most of the global force factors, i.e. IOD, TSI, QBO and ENSO, resulted in the largest decrease in RMSE when only global force factors were included. The RIW% listed in Table 2 for the optimum MLR equation, including only global factors, indicates that IOD (65.3% and 49.4%, respectively) was the most significant parameter at AF and SK, while TSI (52.4%) was the most important factor at LT. The inclusion of local, regional and global factors in the MLR model indicated that the interdependencies between P, IOD, QBO, ENSO and T at AF, P, RH, IOD, ENSO and T at LT, and P, RH, IOD and ENSO at SK, yielded the largest decrease in RMSE difference. The RIW% determined for each independent variable in the optimum MLR equation containing all parameters indicated the most important factors explaining variances in the dependent variable (i.e. NO_2 levels) were P (53.7%) and IOD (17.8%) at AF, P (29.9%), RH (16.6%) and IOD (15.5%) at LT, and P (29.8%) and RH (20.6%) at SK. It is evident from these interdependencies of the dependent variable and RIW% of parameters included in the MLR model that local and regional factors were more significant to NO_2 variability at AF, LT and SK, while global meteorological factors also contributed to variances in NO_2 levels.

Population growth made the most significant contribution to modelled NO_2 concentrations at all three sites, and not only at AF, as observed for SO_2. Therefore, the influence of increased population growth and associated anthropogenic activities is reflected in ambient NO_2 concentrations modelled for the entire north-eastern interior region. Therefore, the periods coinciding with decreased (up until 2002) and increased (2003 to 2007) NO_2 inter-annual variability can be attributed to similar variances in source contribution, as discussed above for SO_2, with regional circulation of air masses passing over major sources also influencing LT and SK (Fig. 2). However, the significant contribution of population growth to the modelled

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

574 NO_2 levels at two rural background sites (LT and SK) also points to increased household
575 combustion associated with enlarged populations within rural communities being a major
576 source of NO_2 in this part of South Africa. The influence of increased seasonal household
577 combustion is also indicated by higher NO_2 concentrations determined in June and July at SK
578 (Fig. 4), which also signifies the impacts of the growing rural communities in proximity of SK.

579 RH made the second most important contribution in explaining variances in modelled NO_2
580 concentrations at LT and SK, while it was the third most important factor at AF as indicated
581 by RIW%. Therefore, RH can be considered the factor representing the influence of changes
582 in local and regional meteorology at these sites. Although T was indicated as a factor included
583 in the linear combination of parameters yielding the largest decrease in RMSE at AF and SK,
584 its relative importance in explaining modelled variances is not indicated by its RIW% in Table
585 2. The strong negative correlation with RH is indicative of increased NO_2 corresponding with
586 months (or years) when dry meteorological conditions prevail, i.e. winter and early spring
587 months in the north-eastern interior of South Africa. As indicated in Fig. 4, higher NO_2
588 concentrations did correspond with dry months (August to November) associated with
589 increased biomass burning. However, the model does not reflect significant contributions of
590 the two parameters included in the model to represent biomass burning, i.e. LFE and DFE to
591 NO_2 variability with relatively higher RIW% observed for DFE (6.9%) and LFE (4.1%) only
592 at SK. Furthermore, higher annual average NO_2 concentrations observed in 2011 and 2012
593 (Fig. 7) at all the sites are also not explained by the MLR model.

594

595 **3.2.3 Ozone (O_3)**

596 Modelled and measured O_3 concentrations at AF, LT and SK are presented in Fig.11, while
597 Table 3 presents the coefficients and the RIW% of independent variables considered in the
598 optimum MLR equation. When only global factors are considered in the model, the linear
599 combinations between ENSO, TSI, IOD and SAM at AF, ENSO, TSI and SAM at LT, and
600 ENSO and IOD at SK resulted in the largest RMSE differences between measured and
601 modelled O_3 levels. However, according to RIW% values calculated, the most significant
602 global factor contributing to O_3 variability was ENSO at all three sites (84.1%, 41.8% and
603 96.7% at AF, LT and SK, respectively). The interdependencies between parameters when local,
604 regional and global factors were included in the models, as well as the RIW% contributions of
605 all factors included in the optimum MLR equation also indicated the significance of ENSO in

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

explaining variances in atmospheric O_3 concentrations at all three sites. Interdependencies between ENSO, IOD, PBL, LFE and R at AF, ENSO, PBL, T, RH and R at LT, and ENSO, PBL, T, RH and R at SK yielded the largest decrease in RMSE differences between measured and modelled O_3 levels, while RIW% indicated that the largest contributions made by factors explaining O_3 variability were ENSO (22.6%), R (14.6%) and Ws (10.1%) at AF, RH (23.1%), ENSO (16.8%) and T (10.5%) at LT, and T (24.6%), ENSO (19.5%), RH (11.3%) and DFE (10.1%) at SK when local, regional and global factors were included in the model.

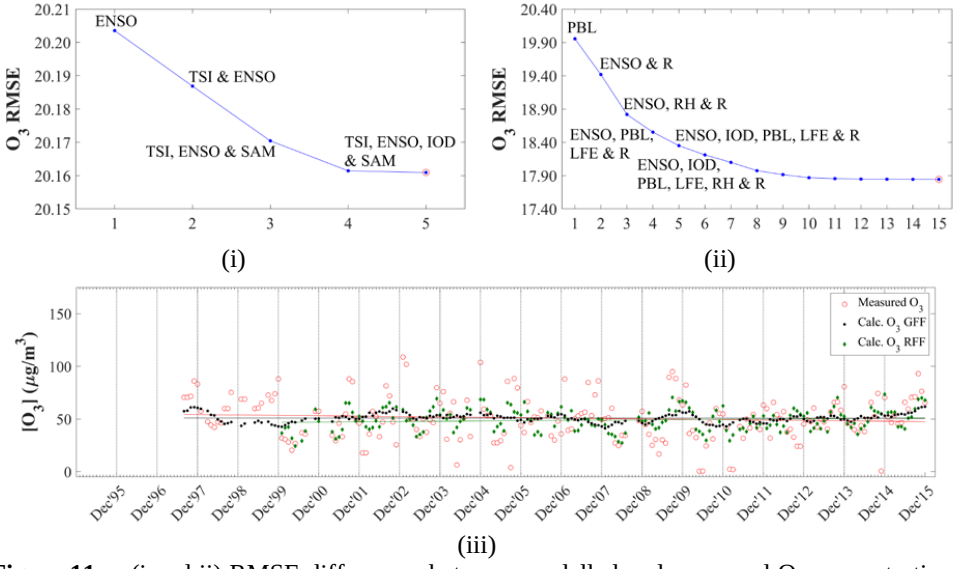


Figure 11a: (i and ii) RMSE differences between modelled and measured O_3 concentrations as a function of the number of independent variables included in the model, as well as comparison between modelled and measured O_3 levels (iii) for global force factors only (GFF), and for global, regional and local factors (RFF) determined for AF

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

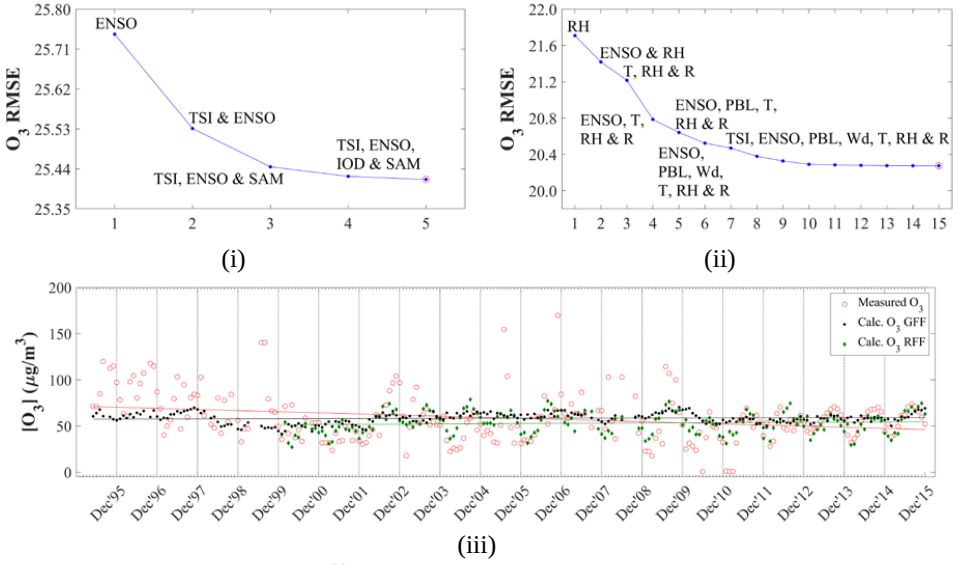
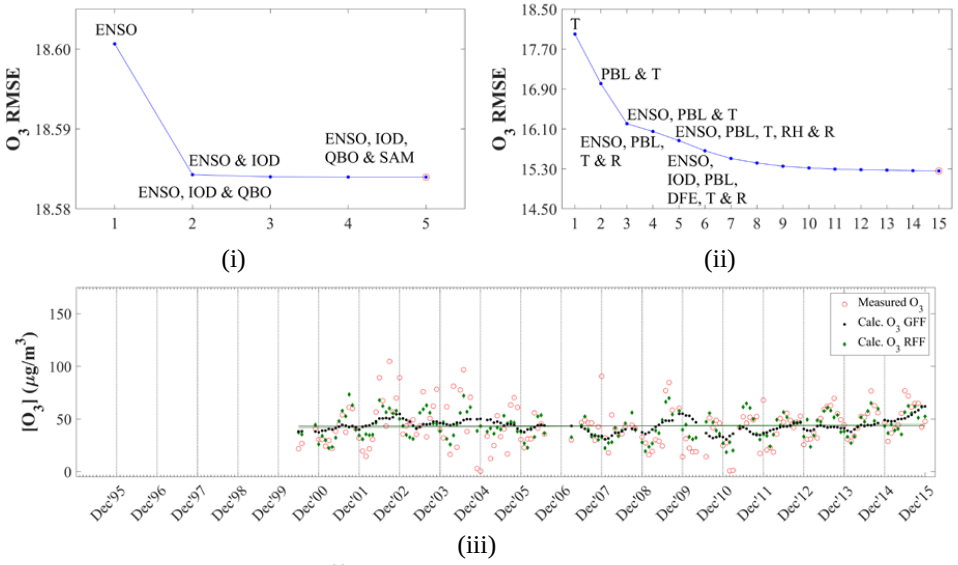


Figure 11b: (i and ii) RMSE differences between modelled and measured O_3 concentrations as a function of the number of independent variables included in the model, as well as comparison between modelled and measured O_3 levels (iii) for global force factors only (GFF), and for global, regional and local factors (RFF) determined for LT

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions



630 **Figure 11c:** (i and ii) RMSE differences between modelled and measured O_3 concentrations
631 as a function of the number of independent variables included in the model, as
632 well as comparison between modelled and measured O_3 levels (iii) for global
633 force factors only (GFF), and for global, regional and local factors (RFF)
634 determined for SK

635

636

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

637 **Table 3:** Regression coefficients (b) and relative important weight percentage (RIW%) of
638 each independent variable included in the MLR model to calculate O_3
639 concentrations at AF, LT and SK

	AF			**LT**			**SK**	
	b	***RIW%***		***b***	***RIW%***		***b***	***RIW%***
i) Global forcing factors								
ENSO	4.923	84.1	ENSO	4.732	41.8	ENSO	8.353	96.7
SAM	-0.539	7.9	TSI	-8.397	36.3	IOD	-3.151	1.5
IOD	-2.337	5.2	SAM	-1.313	18.0	TSI	-0.034	1.5
TSI	1.844	2.5	IOD	-4.231	2.6	SAM	-0.020	0.2
QBO	0.010	0.2	QBO	0.044	1.2	QBO	-6.823×10^{-3}	0.1
ii) Global, regional and local factors								
ENSO	7.478	22.6	RH	-0.966	23.1	T	-5.378	24.6
R	0.122	14.6	ENSO	5.135	16.8	ENSO	7.458	19.5
Ws	5.988	10.1	T	-3.542	10.5	RH	-0.276	11.3
SR	0.474	9.4	DFE	1.070×10^{-5}	9.7	DFE	3.886×10^{-5}	10.1
PBL	2.287×10^{-3}	7.7	PBL	0.043	7.2	PBL	0.070	8.6
T	0.306	7.5	R	0.166	6.5	SR	1.376	8.2
LFE	9.076×10^{-4}	6.8	Wd	-0.087	4.7	R	0.100	4.3
Wd	-0.029	5.1	SR	0.340	4.5	LFE	-5.803×10^{-4}	3.7
RH	-0.257	4.7	IOD	4.900	4.4	Wd	-0.036	3.3
DFE	1.185×10^{-5}	4.2	Ws	-0.601	4.2	Ws	-2.536	2.8
IOD	-12.736	3.7	TSI	-4.195	3.2	IOD	-11.527	1.4
P	6.657×10^{-4}	1.2	LFE	-5.076×10^{-3}	2.3	P	3.013×10^{-5}	1.0
SAM	-0.339	1.2	P	-1.834×10^{-4}	1.5	TSI	1.670	1.0
TSI	-2.989	0.6	SAM	0.101	0.9	QBO	0.038	0.1
QBO	0.018	0.4	QBO	0.031	0.1	SAM	-0.279	0.1

640

641 The significant contribution of ENSO on variances of the dependent variable (modelled O_3
642 concentrations) is evident at all three sites, with RIW% indicating ENSO to be the major factor
643 at AF, and the second most important factor at LT and SK when local, regional and
644 meteorological factors are included in the model. Therefore, inter-annual variability in O_3
645 concentrations can most likely be attributed to ENSO cycles. El Niño periods are associated
646 with drier and warmer conditions in the South African interior, which are conducive to O_3
647 formation, while cloudy and increased rainfall conditions related to La Niña hinder O_3
648 production (Balashov et al., 2014). Balashov et al. (2014) indicated that surface O_3
649 concentrations on the South African Highveld are sensitive to ENSO, with the El Niño period
650 amplifying O_3 formation. The influence of local and regional meteorological conditions is also

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

indicated by the substantial contributions of R and Ws at AF, as well as T and RH at LT and SK on modelled O_3 levels. At LT, RH made the most substantial contribution to the dependent variable, while T made the most significant contribution to modelled O_3 levels. The negative correlation to T and RH at LT and SK is indicative of higher O_3 concentrations corresponding with drier colder months, as indicated in Fig. 5. Laban et al. (2018) indicated the significance of RH to surface O_3 concentrations in the north-eastern part of South Africa through the statistical analysis of *in situ* O_3 measurements conducted in this region, with RH also negatively correlated to surface O_3 levels. The positive correlation to R and Ws at AF reflects higher O_3 concentrations measured during late spring and summer at AF, i.e. October to January, which is a period associated with increased rainfall and less stable meteorological conditions (Fig. 5). The influence of regional open biomass burning during late winter and spring (August to November) on surface O_3 concentrations in this part of South Africa is indicated by the relatively significant contribution of DFE on modelled O_3 concentrations at LT and SK. A recent paper reporting tropospheric O_3 levels measured at four sites in the north-eastern interior of South Africa indicated that O_3 is a regional problem, with O_3 concentration measured at these four sites being similar to levels thereof measured at AF, LT and SK (Laban et al., 2018). A time series of O_3 levels measured from 2010 to 2015 at one of the sites presented by Laban et al. (2018) also indicated higher O_3 concentration corresponding to drier years associated with the ENSO cycle.

As indicated in Fig. 8, inter-annual O_3 concentrations at LT decreased from 1995 to 2001, which corresponded to the period when SO_2 and NO_2 concentrations decreased, as discussed in section 3.1. This period of inter-annual decrease in O_3 levels is not reflected in the statistical model. Since LT is a rural background site with low NO_x emissions, it can be considered to be located in a NO_x-limited O_3 production regime where O_3 concentrations correspond with NO_x concentrations, i.e. increase/decrease with increasing/decreasing NO_x. Therefore, the decrease in O_3 concentrations from 1995 to 2001 can be attributed to decreasing NO_2 concentrations during this period, and the factors influencing NO_2 concentrations at LT, i.e. mainly population growth, as discussed above (section 3.2.2).

The comparisons between modelled and measured O_3 concentrations (Pane iii in Fig. 11a, b and c) also indicated, as observed for SO_2 and NO_2, that the correlations are significantly improved when local, regional and global factors are included in the model. The R^2 values, when only global factors are included, i.e. 0.042, 0.048 and 0.094 at AF, LT and SK, respectively, are improved to 0.259, 0.241 and 0.389 at AF, LT and SK, respectively. These

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

correlations can be considered relatively weak, with the exception of a moderate correlation at SK (Sheskin, 2003). These generally weaker correlations can be attributed to the complexity associated with tropospheric O_3 chemistry. Tropospheric O_3 is a secondary atmospheric pollutant with several factors contributing to its variability. In addition, Laban et al. (2018) indicated the significance of the precursor species CO to surface O_3 concentrations in the north-eastern interior of South Africa, which were not measured at any of the sites and included in the model. Swartz et al. (2019) also compared passively derived O_3 concentrations with active O_3 measurements and illustrated limitations associated with the use of passive samplers to determine O_3 concentrations. However, the general trend of measured O_3 concentrations is mimicked by the modelled O_3 values when local, regional and global factors are included in the model, while the overall trend is weakly followed when only global factors are included. Higher and lower O_3 concentrations are underestimated by the MLR model.

The trend lines for the O_3 concentrations measured during the entire sampling periods indicate slight negative slopes at AF and LT (Fig. 11a(iii) and 11b(iii), respectively), and a small positive slope at SK (Fig. 11c(iii)). Annual average decreases in O_3 levels of 0.37 $\mu g.m^{-3}.y^{-1}$ and 1.20 $\mu g.m^{-3}.y^{-1}$ were calculated at AF and LT, respectively, while an average annual increase of 0.21 $\mu g.m^{-3}.y^{-1}$ was calculated at SK. However, in general, it seems that O_3 concentrations remained relatively constant at all three sites for the entire 19-, 21- and 16-year sampling periods at AF, LT and SK, respectively. Decadal trends of -3.46 (average O_3 concentrations from 1997 to 2006 were 52.56 $\mu g.m^{-3}$ and average O_3 concentrations from 2007 to 2015 were 49.10 $\mu g.m^{-3}$) and -9.15 $\mu g.m^{-3}.dec^{-1}$ (average O_3 concentrations from 1995 to 2004 were 63.16 $\mu g.m^{-3}$ and average O_3 concentrations from 2005 to 2014 were 53.01 $\mu g.m^{-3}$) were calculated for AF and LT, respectively, for two decades. Similar to SO_2 and NO_2, the slopes of the linear regression trend lines for the measured and modelled O_3 concentrations when local, regional and global factors are included are exactly the same at AF, LT and SK (Pane iii in Fig. 11a, b and c), which indicates that measured and modelled O_3 trends compares well in spite of low R^2 values. In addition, relatively good correlations are observed between the slopes of the trend lines of measured O_3 concentrations and modelled O_3 values calculated when only global factors are included at all the sites, signifying the influence of global factors, especially ENSO, as indicated above, on O_3 variability (Pane iii in Fig. 11a, b and c).

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

3.3 Contextualisation

In order to contextualise the long-term SO_2, NO_2 and O_3 concentrations measured with passive samplers at AF, LT and SK located in the north-eastern interior of South Africa, the statistical spread of the concentrations of these species determined during the entire sampling period at each site are compared to average concentrations of these species determined with passive samplers during other studies in South Africa and Africa, as well as regional sites in other parts of the world. SO_2, NO_2 and O_3 concentrations determined in this study are related to levels reported elsewhere in Fig. 12, 13 and 14, respectively.

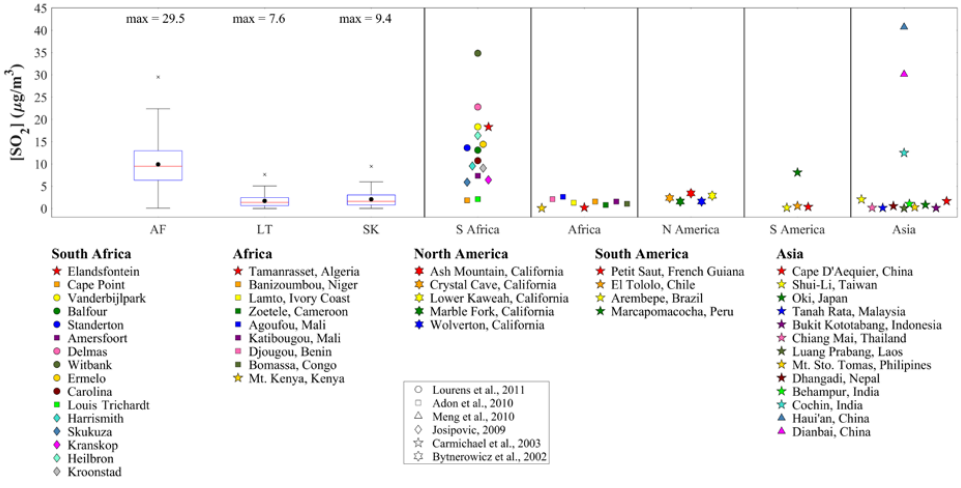


Figure 12: Statistical spread of SO_2 concentrations determined during the entire measuring period at each site compared to mean levels determined with passive samplers elsewhere. The red line of each box represents the median, the top and bottom edges of the box the 25th and 75th percentiles, respectively, the whiskers ± 2.7σ (99.3% coverage if the data has a normal distribution) and the black dots the average concentrations

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

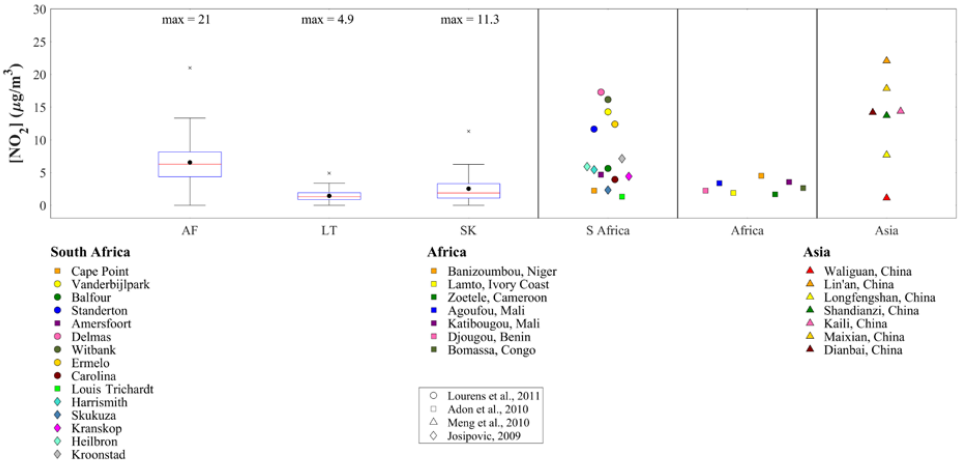


Figure 13: Statistical spread of NO_2 concentrations determined during the entire measuring period at each site compared to mean levels determined with passive samplers elsewhere. The red line of each box represents the median, the top and bottom edges of the box the 25th and 75th percentiles, respectively, the whiskers ± 2.7σ (99.3% coverage if the data has a normal distribution) and the black dots the average concentrations

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

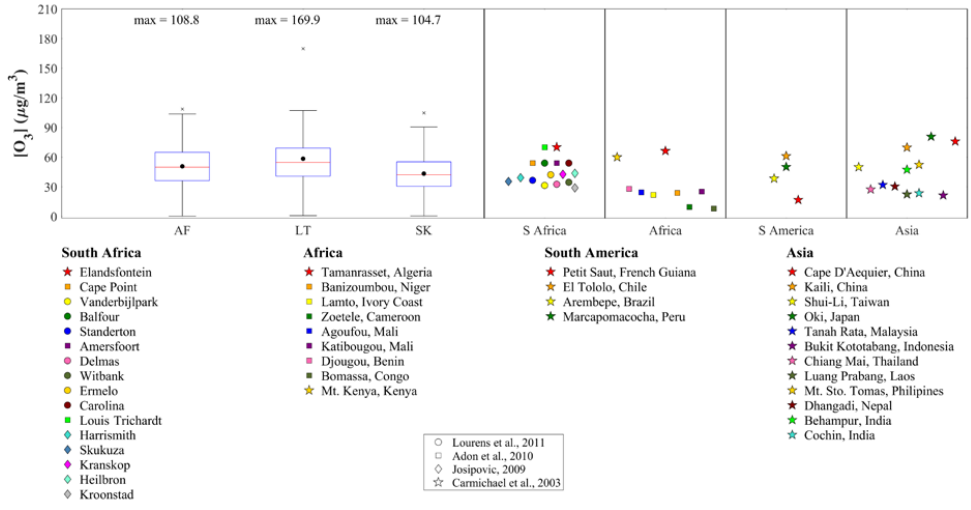


Figure 14: Statistical spread of O_3 concentrations determined during the entire measuring period at each site compared to mean levels determined with passive samplers elsewhere. The red line of each box represents the median, the top and bottom edges of the box the 25th and 75th percentiles, respectively, the whiskers ± 2.7σ (99.3% coverage if the data has a normal distribution) and the black dots the average concentrations

As expected, the average and median SO_2 concentrations determined at the industrially impacted AF (9.91 μg.m^{-3} and 9.48 μg.m^{-3}, respectively) site were higher compared to average and median SO_2 levels determined at the rural background sites LT (1.70 μg.m^{-3} and 1.35 μg.m^{-3}, respectively) and SK (2.07 μg.m^{-3} and 1.60 μg.m^{-3}, respectively) for the entire sampling period at each site. Geospatial maps of SO_2 column amount in the planetary boundary layer and NO_2 tropospheric column density averaged over the period 2005 to 2015 over southern Africa (Fig. A4 and A5 respectively) indicate higher average SO_2 and NO_2 concentrations being observed over the region where AF is located. Much lower average SO_2 and NO_2 concentrations are observed over the northernmost parts of the country, where LT is located, as well as the western region where SK is situated. Therefore, the influence of coal-fired power stations on SO_2 (and NO_2) levels measured at AF is evident. The average SO_2 levels at AF were similar to average SO_2 concentrations determined at other sites located in the Mpumalanga Highveld, for which the measurement period was from August 2007 to July 2008 (Lourens et al., 2011). However, the average SO_2 level at AF was significantly lower than the mean SO_2 levels at Elandsfontein, Delmas and Witbank. Elandsfontein and Delmas are situated

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

763 within closer proximity to major industrial activities in the Mpumalanga Highveld, while
764 Witbank is a relatively large urban area with numerous large industrial point sources (Lourens
765 et al., 2011). In addition, the average SO_2 concentrations at Vanderbijlpark – an urban area
766 located within the highly industrialised Vaal Triangle region – were also higher compared to
767 levels thereof at AF. Average SO_2 concentrations determined at regional sites in South America
768 and India, i.e. Marcapomacocha and Cochin, respectively, were also similar to mean SO_2 levels
769 determined at AF (Carmichael et al., 2003). The measurement period of the Carmichael et al.
770 (2003) study was 12 months, starting in September 1999 (Carmichael et al., 2003). SO_2
771 concentrations reported for two rural sites in China, i.e. Dianbai and Haui'an were similar to
772 SO_2 levels determined at Witbank (Meng et al., 2010). Meng et al. (2010) presented results
773 obtained during a two-year study that commenced in January 2007. The mean SO_2
774 concentrations determined at LT and SK were similar to average SO_2 concentrations
775 determined at regional background sites in west- and central African sites (Carmichael et al.,
776 2003; Adon et al., 2010), as well as mean SO_2 levels determined at most of the regional sites
777 in North America – measured between May and November 1999, South America and Asia
778 (Bytnerowicz et al., 2002; Carmichael et al., 2003). Adon et al. (2010) presented ambient SO_2,
779 NO_2 and O_3 concentrations measured from 1998 to 2007 at Katibougou in Mali, Banizoumbou
780 in Niger, Lamto in Ivory Coast and Zoetele in Cameroon. The measurement periods for
781 Agoufou in Mali and Djougou in Benin was from 2005 to 2007, while for Bomassa in Congo
782 measurements were reported between 1998 and 2006 (Adon et al., 2010).

783 Similar to SO_2, the mean and median NO_2 levels determined for the respective sampling
784 periods at each site were higher at AF (6.56 $\mu g.m^{-3}$ and 6.29 $\mu g.m^{-3}$, respectively) compared to
785 mean and median levels thereof at LT (1.45 $\mu g.m^{-3}$ and 1.32 $\mu g.m^{-3}$, respectively) and SK (2.54
786 $\mu g.m^{-3}$ and 1.89 $\mu g.m^{-3}$, respectively). Relatively higher NO_2 concentrations were determined
787 at SK compared to LT, which can be attributed to the influence of growing rural communities
788 on the border of the Kruger National Park (Maritz et al., 2019). The mean NO_2 concentrations
789 at AF were lower compared to most of the average NO_2 levels determined at other sites located
790 in the Mpumalanga Highveld within closer proximity to industrial sources, while being similar
791 to mean NO_2 concentrations measured at Balfour and Carolina. In addition, average NO_2 levels
792 at AF were also lower than average NO_2 concentrations determined in the Vaal Triangle
793 (Lourens et al., 2011). Average NO_2 concentrations determined at rural and regional sites in
794 China were higher than mean NO_2 levels at AF, with the exception of Longfengshan that had
795 similar NO_2 concentrations to AF (Meng et al., 2010), which reflects the scale of atmospheric

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

pollution in China. The average NO_2 concentrations at LT and SK were also similar to mean NO_2 levels determined at regional sites in west- and central African sites (Carmichael et al., 2003; Adon et al., 2010), as well as a remote site (Waliguan) in China (Meng et al., 2010).

The statistical distribution of O_3 concentrations determined at AF, LT and SK indicates similar surface O_3 levels at all three sites with marginally higher O_3 concentrations determined at LT (58.44 $\mu g.m^{-3}$ and 54.67 $\mu g.m^{-3}$, respectively) compared to AF (50.77 $\mu g.m^{-3}$ and 49.84 $\mu g.m^{-3}$, respectively) and SK (43.36 $\mu g.m^{-3}$ and 42.20 $\mu g.m^{-3}$, respectively). Higher O_3 levels are expected at the rural background LT site due to decreased O_3 titration compared to polluted regions, while LT is also impacted by aged air masses passing over the Mpumalanga Highveld source region as previously indicated. However, the regional O_3 problem in the South Africa interior is reflected by high O_3 concentrations also measured at the industrially influenced AT site, as well as similar O_3 levels determined at other sites in the Mpumalanga Highveld (Lourens et al., 2011). Laban et al. (2018) attributed high regional O_3 concentrations in the north-eastern interior of South Africa to the influence of household combustion and widespread open biomass burning impacting this region. In addition, the influence of rural communities is also reflected by the slightly lower average O_3 levels at SK. O_3 concentrations measured at west- and central Africa sites were lower than South African O_3 levels (Adon et al., 2010), with the exception of Mt Kenya and a site in northern Africa that had similar O_3 concentrations (Carmichael et al., 2003). Similar O_3 concentrations were determined at the South American regional sites, except for Petit Saut that had lower O_3 concentrations (Carmichael et al., 2003). Average O_3 levels determined at some of the regional Asian sites were in the same range as O_3 concentrations over the interior of South Africa, while certain sites in Asia had lower mean O_3 levels (Carmichael et al., 2003).

4. Summary and conclusions

In this study, long-term trends of atmospheric SO_2, NO_2 and O_3 concentrations measured with passive samplers at three sites located in the north-eastern interior of South Africa are presented. This paper illustrates the value of low-cost atmospheric sampling techniques in order to obtain long-term data, especially for regions restricted by logistical accessibility and limited capacity. A 19-year (1997 to 2015), 21-year (1995 to 2015) and 16-year (2000 to 2015) dataset for AF, LT and SK could be evaluated. Long-term temporal trends indicated seasonal and inter-annual variability at all three sites, which could be ascribed to changes in meteorological

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

828 conditions and/or variances in source contribution. Inter-annual variability indicated periods
829 up until 2003/2004 and 2002 during which SO_2 and NO_2 concentrations, respectively,
830 decreased, followed by periods during which SO_2 and NO_2 levels increased up until 2009 and
831 2007, respectively. These long-term trends were assessed with an MLR model in order to
832 establish the influence of sources, as well as local, regional and global meteorology on
833 atmospheric SO_2, NO_2 and O_3 concentrations.

834 Interdependencies between local, regional and global parameters included in the statistical
835 model indicated the influence of global meteorology on SO_2 variability at all three sites,
836 especially at the rural background site LT. However, population growth was the most
837 substantial factor in the statistical model at the industrially impacted AF site, while the
838 significance of local and regional meteorology was also evident with T being the most
839 significant factor at SK. The important contribution of population growth on modelled SO_2
840 levels at AF was indicative of the impact of increased anthropogenic activities and energy
841 demand in the north-eastern interior of South Africa. Higher SO_2 concentrations associated
842 with lower temperatures reflected the influence of pollution build-up during winter, while the
843 influence of air masses passing over the source region is also evident at SK and LT. Although
844 global parameters contributed to variances in NO_2 concentrations, local and regional factors
845 made more substantial contributions to modelled NO_2 levels. The most significant factor
846 explaining NO_2 variability at all three sites was population growth, while RH was the most
847 important local and regional meteorological factor. Therefore, similar to SO_2, the influence of
848 population growth and associated increases in anthropogenic activities in the north-eastern
849 interior is also reflected in NO_2 levels, while the impacts of increased household combustion
850 associated with growing rural communities are also evident, especially at SK. The negative
851 correlation to RH indicates higher NO_2 levels associated with drier months, i.e. winter, which
852 contribute to seasonal variances. ENSO was shown to make a significant contribution to
853 modelled O_3 levels at all three sites, while the important influence of local and regional
854 meteorological factors was also evident, especially through significant negative correlations
855 with T and RH at SK and LT. Inter-annual O_3 variability in this part of South Africa can
856 therefore most likely be attributed to ENSO cycles, while seasonal patterns are attributed
857 changes in local and regional meteorology.

858 The decreases in SO_2 and NO_2 concentrations from 1995 were attributed to the implementation
859 of mitigation policies by industries post the establishment of the new democracy in South
860 Africa. However, these improvements were offset from 2002 due to rapid economic growth

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020


Atmospheric Chemistry and Physics Discussions

associated with increased industrial activities, as well as the increase in population growth accompanied by higher energy demand. The 19-year trend lines for SO_2 and NO_2 at AF indicated an increase in SO_2 and NO_2 concentrations over the 19-year sampling period. In addition, an upwards trend in NO_2 levels was also evident at SK, signifying the influence of the growing rural communities on the border of the Kruger National Park. Marginal trends were observed for SO_2 at SK, as well as SO_2 and NO_2 at LT. Trend analysis of O_3 at all three sites indicated that O_3 concentrations remained relatively constant at all three sites for the entire 19-, 21- and 16-year sampling periods at AF, LT and SK, respectively.

As expected, SO_2 and NO_2 concentrations were higher at AF compared to levels thereof at the rural background sites LT and SK. SO_2 levels at AF were similar to levels of these species determined with passive samplers at other sites within the Mpumalanga Highveld with the exception of sites closer to the major industrial sources. NO_2 levels at AF were generally lower than NO_2 concentrations determined at sites within the source region, as well as than regional sites in China. SO_2 and NO_2 concentrations determined at LT and SK were similar to levels thereof determined with passive samplers at regional and rural sites in Africa and other parts of the world. The regional problem of O_3 in the interior of South Africa was also evident, with similar O_3 levels determined at all three sites.

Acknowledgements

The authors would like to thank the International Global Atmospheric Chemistry programme for endorsing the DEBITS programme, as well as Sasol and Eskom for financial support of the South African INDAAF project. Assistance with sample deployment and collection by Ms Carin van der Merwe is also acknowledged, as well as the site operators, who include: Memory Deacon at AF; Chris James at LT; and Navashni Govender, Walter Kubheka, Eva Gardiner and Joel Tleane at SK.

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

887 **References**

888 Adon, M., Galy-Lacaux, C., Delon, C., Yoboue, V., Solmon, F. & Kaptue Tchuente, A. T.
889 2013. Dry deposition of nitrogen compounds (NO_2, HNO_3, NH_3), sulfur dioxide and ozone in
890 west and central African ecosystems using the inferential method. *Atmospheric Chemistry and*
891 *Physics*, 13**,** 11351-11374, doi: 10.5194/acp-13-11351-2013.

892 Adon, M., Galy-Lacaux, C., Yoboué, V., Delon, C., Lacaux, J. P., Castera, P., Gardrat, E.,
893 Pienaar, J. J., Al Ourabi, H., Laouali, D., Diop, B., Sigha-Nkamdjou, L., Akpo, A., Tathy, J.
894 P., Lavenu, F. & Mougin, E. 2010. Long term measurements of sulfur dioxide, nitrogen
895 dioxide, ammonia, nitric acid and ozone in Africa using passive samplers. *Atmospheric*
896 *Chemistry and Physics*, 10**,** 7467-7487, doi: 10.5194/acp-10-7467-2010.

897 Balashov, N. V., Thompson, A. M., Piketh, S. J. & Langerman, K. E. 2014. Surface ozone
898 variability and trends over the South African Highveld from 1990 to 2007. *Journal of*
899 *Geophysical Research: Atmospheres*, 119**,** 20, doi: 10.1002/2013JD020555.

900 Bencherif, H., Diab, R. D., Portafaix, T., Morel, B., Keckhut, P. & Moorgawa, A. 2006.
901 Temperature climatology and trend estimates in the UTLS region as observed over a southern
902 subtropical site, Durban, South Africa. *Atmospheric Chemistry and Physics*, 6**,** 5121-5128, doi:
903 10.5194/acp-6-5121-2006.

904 Booyens, W., Beukes, J. P., Van Zyl, P. G., Ruiz-Jimenez, J., Kopperi, M., Riekkola, M.-L.,
905 Josipovic, M., Vakkari, V. & Laakso, L. 2019. Assessment of polar organic aerosols at a
906 regional background site in southern Africa. *Submitted to Journal of Atmospheric Chemistry*.

907 Bytnerowicz, A., Tausz, M., Alonso, R., Jones, D., Johnson, R. & Grulke, N. 2002. Summer-
908 time distribution of air pollutants in Sequoia National Park, California. *Environmental*
909 *Pollution*, 118**,** 187-203, doi: 10.1016/S0269-7491(01)00312-8.

910 Carmichael, G. R., Ferm, M., Thongboonchoo, N., Woo, J.-H., Chan, L. Y., Murano, K., Viet,
911 P. H., Mossberg, C., Bala, R., Boonjawat, J., Upatum, P., Mohan, M., Adhikary, S. P., Shrestha,
912 A. B., Pienaar, J. J., Brunke, E. B., Chen, T., Jie, T., Guoan, D., Peng, L. C., Dhiharto, S.,
913 Harjanto, H., Jose, A. M., Kimani, W., Kirouane, A., Lacaux, J.-P., Richard, S., Barturen, O.,
914 Cerda, J. C., Athayde, A., Tavares, T., Cotrina, J. S. & Bilici, E. 2003. Measurements of sulfur
915 dioxide, ozone and ammonia concentrations in Asia, Africa, and South America using passive
916 samplers. *Atmospheric Environment*, 37**,** 1293-1308, doi: 10.1016/S1352-2310(02)01009-9.

917 Connell, D. W. 2005. *Basic concepts of environmental chemistry*, CRC Press.

918 Conradie, E. H., Van Zyl, P. G., Pienaar, J. J., Beukes, J. P., Galy-Lacaux, C., Venter, A. D. &
919 Mkhatshwa, G. V. 2016. The chemical composition and fluxes of atmospheric wet deposition
920 at four sites in South Africa. *Atmospheric Environment*, 146**,** 113-131, doi:
921 10.1016/j.atmosenv.2016.07.033.

922 Dhammapala, R. S. 1996. *Use of diffusive samplers for the sampling of atmospheric pollutants.*
923 M.Sc, Potchefstroom University for CHE.

924 Draxler, R. R. & Hess, G. D. 2014. Description of the HYSPLIT_4 modelling system. 7 ed.
925 Silver Spring, Maryland: Air Resources Laboratory.

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Ferm, M. 1991. A Sensitive Diffusional Sampler. *IVL Report L91*. Göteborg, Sweden: Swedish Environmental Research Institute.

Fowler, D., Pilegaard, K., Sutton, M. A., Ambus, P., Raivonen, M., Duyzer, J., Simpson, D., Fagerli, H., Fuzzi, S., Schjoerring, J. K., Granier, C., Neftel, A., Isaksen, I. S. A., Laj, P., Maione, M., Monks, P. S., Burkhardt, J., Daemmgen, U., Neirynck, J., Personne, E., Wichink-Kruit, R., Butterbach-Bahl, K., Flechard, C., Tuovinen, J. P., Coyle, M., Gerosa, G., Loubet, B., Altimir, N., Gruenhage, L., Ammann, C., Cieslik, S., Paoletti, E., Mikkelsen, T. N., Ro-Poulsen, H., Cellier, P., Cape, J. N., Horváth, L., Loreto, F., Niinemets, Ü., Palmer, P. I., Rinne, J., Misztal, P., Nemitz, E., Nilsson, D., Pryor, S., Gallagher, M. W., Vesala, T., Skiba, U., Brüggemann, N., Zechmeister-Boltenstern, S., Williams, J., O'dowd, C., Facchini, M. C., De Leeuw, G., Flossman, A., Chaumerliac, N. & Erisman, J. W. 2009. Atmospheric composition change: Ecosystems–Atmosphere interactions. *Atmospheric Environment*, 43**,** 5193-5267, doi: 10.1016/j.atmosenv.2009.07.068.

Garstang, M., Tyson, P. D., Swap, R., Edwards, M., Kållberg, P. & Lindesay, J. A. 1996. Horizontal and vertical transport of air over southern Africa. *Journal of Geophysical Research Atmospheres*, 101**,** 16, doi: 10.1029/95JD00844.

He, J. & Bala, R. 2008. Draft Report on passive sampler inter-comparison under Malé declaration. *Malé Declaration on Control and Prevention of Air Pollution and its Likely Transboundary Effect for South Asia.* Singapore: National University of Singapore.

Hewitson, B. C. & Crane, R. G. 2006. Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa. *International Journal of Climatology*, 26**,** 1315-1337, doi: 10.1002/joc.1314.

ICDA (International Chromium Development Association) 2012. High carbon charge grade ferrochromium Statistics. *Statistical Bulletin 2012*. Paris, France: International Chromium Development Association.

ICDA 2013. Statistical Bulletin 2013 (based on 2012 data). International Chromium Development Association.

ISO Survey 2015. Available: http://www/iso/org/iso/iso-survey [Accessed 23 January 2017].

Inglesi-Lotz, R. & Blignaut, J. 2011. Estimating the price elasticity for demand for electricity by sector in South Africa. *South African Journal of Economic and Management Sciences, Vol 14, Iss 4, Pp 449-465 (2011)***,** 449, doi: 10.4102/sajems.v14i4.134.

Jaars, K., Beukes, J. P., Van Zyl, P. G., Venter, A. D., Josipovic, M., Pienaar, J. J., Vakkari, V., Aaltonen, H., Laakso, H., Kulmala, M., Tiitta, P., Guenther, A., Hellén, H., Laakso, L. & Hakola, H. 2014. Ambient aromatic hydrocarbon measurements at Welgegund, South Africa. *Atmospheric Chemistry and Physics Discussions*, 14**,** 7075-7089, doi: 10.5194/acp-14-7075-2014.

Josipovic, M. 2009. *Acidic deposition emanating from the South African Highveld: A critical levels and critical loads assessment.* PhD, University of Johannesburg.

Josipovic, M., Annegarn, H. J., Kneen, M. A., Pienaar, J. J. & Piketh, S. J. 2011. Atmospheric dry and wet deposition of sulphur and nitrogen species and assessment of critical loads of acidic

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

966 deposition exceedance in South Africa. *South African Journal Science*, 107**,** 10, doi:
967 10.4102/sajs.v107i3/4.478.

968 Kaufman, Y. J., Ichoku, C., Giglio, L., Korontzi, S., Chu, D. A., Hao, W. M., Li, R. R. &
969 Justice, C. O. 2003. Fire and smoke observed from the Earth Observing System MODIS
970 instrument - products, validation, and operational use. *International Journal of Remote*
971 *Sensing*, 24**,** 1765-1781, doi: 10.1080/01431160210144741.

972 Kleynhans, E., Beukes, J. P., Van Zyl, P. G., Bunt, J., Nkosi, N. & Venter, M. 2017. The Effect
973 of Carbonaceous Reductant Selection on Chromite Pre-reduction. *Metallurgical & Materials*
974 *Transactions B*, 48**,** 827-840, doi: 10.1007/s11663-016-0878-4.

975 KMNI. 2016a. *monthly DMI HadISST1* [Online]. Available:
976 http://climexp.knmi.nl/getindices.cgi?WMO=UKMOData/hadisst1_dmi&STATION=DMI_
977 HadISST1&TYPE=i&id=someone@somewhere [Accessed 22 December 2016].

978 KMNI. 2016b. *monthly measured total solar irradiance* [Online]. Available:
979 http://climexp.knmi.nl/getindices.cgi?WMO=PMODData/tsi&STATION=measured_total_so
980 lar_irradiance&TYPE=i&id=someone@somewhere [Accessed 22 December 2016].

981 Korhonen, K., Giannakaki, E., Mielonen, T., Pfüller, A., Laakso, L., Vakkari, V., Baars, H.,
982 Engelmann, R., Beukes, J. P., Van Zyl, P. G., Ramandh, A., Ntsangwane, L., Josipovic, M.,
983 Tiitta, P., Fourie, G., Ngwana, I., Chiloane, K. & Komppula, M. 2014. Atmospheric boundary
984 layer top height in South Africa: measurements with lidar and radiosonde compared to three
985 atmospheric models. *Atmospheric Chemistry and Physics*, 14**,** 4263-4278, doi: 10.5194/acp-
986 14-4263-2014.

987 Kraha, A., Turner, H., Nimon, K., Reichwein Zientek, L. & Henson, R. K. 2012. Tools to
988 support interpreting multiple regression in the face of multicollinearity. *Frontiers In*
989 *Psychology*, 3**,** 1-16, doi: 10.3389/fpsyg.2012.00044.

990 Laakso, L., Vakkari, V., Virkkula, A., Laakso, H., Backman, J., Kulmala, M., Beukes, J. P.,
991 Van Zyl, P. G., Tütta, P., Josipovic, M., Pienaar, J. J., Chiloane, K., Gilardoni, S., Vignati, E.,
992 Wiedensohler, A., Tuch, T., Birmili, W., Piketh, S., Collett, K. & Fourie, G. D. 2012. South
993 African EUCAARI measurements: seasonal variation of trace gases and aerosol optical
994 properties. *Atmospheric Chemistry & Physics*, 12**,** 1847-1864, doi: 10.5194/acp-12-1847-2012.

995 Laban, T. L., Van Zyl, P. G., Beukes, J. P., Vakkari, V., Jaars, K., Borduas-Dedekind, N.,
996 Josipovic, M., Thompson, A. M., Kulmala, M. & Laakso, L. 2018. Seasonal influences on
997 surface ozone variability in continental South Africa and implications for air quality.
998 *Atmospheric Chemistry and Physics* 15491-15514, doi: 10.5194/acp-18-15491-2018.

999 Lacaux, J. P., Tathy, J. P. & Sigha, L. 2003. Acid wet deposition in the tropics: Two case
1000 studies using DEBITS measurements. *IGACtivities Newsletter of the International Global*
1001 *Atmospheric Chemistry Project.*

1002 Lorenzo-Seva, U., Ferrando, P. J. & Chico, E. 2010. Two SPSS programs for interpreting
1003 multiple regression results. *Behavior Research Methods*, 42**,** 29-35, doi:
1004 10.3758/BRM.42.1.29.

1005 Lourens, A. S., Beukes, J. P., Van Zyl, P. G., Fourie, G. D., Burger, J. W., Pienaar, J. J., Read,
1006 C. E. & Jordaan, J. H. 2011. Spatial and temporal assessment of gaseous pollutants in the

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Highveld of South Africa. *South African Journal of Science*, 107**,** 1-8, doi: 10.4102/sajs.v107i1/2.269.

Maritz, P., Beukes, J. P., Van Zyl, P. G., Liousse, C., Gardrat, E., Ramandh, A. & Mkhatshwa, G. V. 2019. Spatial, temporal and source assessments of organic and elemental carbon at DEBITS sites in South Africa. *Submitted to Environmental Monitoring and Assessment*.

Marshall, G. 2018. *An observation-based Southern Hemisphere Annular Mode Index* [Online]. United Kingdom. Available: http://www.nerc-bas.ac.uk/icd/gjma/sam.html [Accessed 28 August 2018].

Martins, J. J., Dhammapala, R. S., Lachmann, G., Galy-Lacaux, C. & Pienaar, J. J. 2007. Long-term measurements of sulphur dioxide, nitrogen dioxide, ammonia, nitric acid and ozone in southern Africa using passive samplers. *South African Journal of Science*, 103**,** 336-342.

Meng, Z.-Y., Xu, X.-B., Wang, T., Zhang, X.-Y., Yu, X.-L., Wang, S.-F., Lin, W.-L., Chen, Y.-Z., Jiang, Y.-A. & An, X.-Q. 2010. Ambient sulfur dioxide, nitrogen dioxide, and ammonia at ten background and rural sites in China during 2007–2008. *Atmospheric Environment*, 44**,** 2625-2631, doi: 10.1016/j.atmosenv.2010.04.008.

Meth, O. 2018. *New satellite data reveals the world's largest air pollution hotspot is Mpumalanga – South Africa* [Online]. Available: https://www.greenpeace.org/africa/en/issues/inspirethemovement/4202/new-satellite-data-reveals-the-worlds-largest-air-pollution-hotspot-is-mpumalanga-south-africa/ [Accessed 17 January 2019].

Monks, P. S., Granier, C., Fuzzi, S., Stohl, A., Williams, M. L., Akimoto, H., Amann, M., Baklanov, A., Baltensperger, U., Bey, I., Blake, N., Blake, R. S., Carslaw, K., Cooper, O. R., Dentener, F., Fowler, D., Fragkou, E., Frost, G. J., Generoso, S., Ginoux, P., Grewe, V., Guenther, A., Hansson, H. C., Henne, S., Hjorth, J., Hofzumahaus, A., Huntrieser, H., Isaksen, I. S. A., Jenkin, M. E., Kaiser, J., Kanakidou, M., Klimont, Z., Kulmala, M., Laj, P., Lawrence, M. G., Lee, J. D., Liousse, C., Maione, M., Mcfiggans, G., Metzger, A., Mieville, A., Moussiopoulos, N., Orlando, J. J., O'dowd, C. D., Palmer, P. I., Parrish, D. D., Petzold, A., Platt, U., Pöschl, U., Prévôt, A. S. H., Reeves, C. E., Reimann, S., Rudich, Y., Sellegri, K., Steinbrecher, R., Simpson, D., Ten Brink, H., Theloke, J., Van Der Werf, G. R., Vautard, R., Vestreng, V., Vlachokostas, C. & Von Glasow, R. 2009. Atmospheric composition change – global and regional air quality. *Atmospheric Environment*, 43**,** 5268-5350, doi: 10.1016/j.atmosenv.2009.08.021.

Mphepya, J. N., Galy-Lacaux, C., Lacaux, J. P., Held, G. & Pienaar, J. J. 2006. Precipitation Chemistry and Wet Deposition in Kruger National Park, South Africa. *Journal of Atmospheric Chemistry*, 53**,** 169-183, doi: 10.1007/s10874-005-9005-7.

Mphepya, J. N., Pienaar, J. J., Galy-Lacaux, C., Held, G. & Turner, C. R. 2004. Precipitation Chemistry in Semi-Arid Areas of Southern Africa: A Case Study of a Rural and an Industrial Site. *Journal of Atmospheric Chemistry*, 47**,** 24, doi: 10.1023/B:JOCH.0000012240.09119.c4.

Nathans, L. L., Oswald, F. L. & Nimon, K. 2012. Interpreting Multiple Linear Regression: A Guidebook of Variable Importance. *Practical Assessment, Research & Evaluation*, 17**,** 1-19.

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

1047 NOAA. 2015a. *Climate Indices: Monthly Atmospheric and Ocean Time Series* [Online].
1048 Available: https://www.esrl.noaa.gov/psd/data/climateindices/list/ [Accessed 22 December
1049 2016].

1050 NOAA. 2015b. *Monthly Atmospheric and SST Indices* [Online]. Available:
1051 http://www.cpc.ncep.noaa.gov/data/indices/ [Accessed 22 December 2016].

1052 Rorich, R. P. & Galpin, J. S. 1998. Air quality in the Mpumalanga Highveld region, South
1053 Africa. *South African Journal of Science*, 94**,** 109.

1054 Sheskin, D. J. 2003. *Handbook of Parametric and Nonparametric Statistical Procedures*, Boca
1055 Raton, Chapman and Hall/CRC Press.

1056 Swartz, J.-S., Van Zyl Pieter, G., Beukes Johan, P., Labuschagne, C., Brunke, E.-G., Portafaix,
1057 T., Galy-Lacaux, C. & Pienaar Jacobus, J. 2019. Twenty-one years of passive sampling
1058 monitoring of SO_2, NO_2 and O_3 at the Cape Point GAW station, South Africa. *Submitted to*
1059 *Atmospheric Environment*.

1060 Tiitta, P., Vakkari, V., Croteau, P., Beukes, J. P., Zyl, P. G. V., Josipovic, M., Venter, A. D.,
1061 Jaars, K., Pienaar, J. J., Ng, N. L., Canagaratna, M. R., Jayne, J. T., Kerminen, V. M., Kokkola,
1062 H., Kulmala, M., Laaksonen, A., Worsnop, D. R. & Laakso, L. 2014. Chemical composition,
1063 main sources and temporal variability of PM_1 aerosols in southern African grassland.
1064 *Atmospheric Chemistry and Physics*, 14**,** 1909-1927, doi: 10.5194/acp-14-1909-2014.

1065 Toihir, A. M., Portafaix, T., Sivakumar, V., Bencherif, H., Pazmiño, A. & Bègue, N. 2018.
1066 Variability and trend in ozone over the southern tropics and subtropics. *Annales Geophysicae*,
1067 36**,** 381-404, doi: 10.5194/angeo-36-381-2018.

1068 Tyson, P. D., Garstang, M. & Swap, R. 1996. Large-Scale Recirculation of Air over Southern
1069 Africa. *Journal of Applied Meteorology*, 35**,** 18, doi: 10.1175/1520-
1070 0450(1996)035<2218:LSROAO>2.0.CO;2.

1071 Vet, R., Artz, R. S., Carou, S., Shaw, M., Ro, C.-U., Aas, W., Baker, A., Bowersox, V. C.,
1072 Dentener, F., Galy-Lacaux, C., Hou, A., Pienaar, J. J., Gillett, R., Forti, M. C., Gromov, S.,
1073 Hara, H., Khodzher, T., Mahowald, N. M., Nickovic, S., Rao, P. S. P. & Reid, N. W. 2014. A
1074 global assessment of precipitation chemistry and deposition of sulfur, nitrogen, sea salt, base
1075 cations, organic acids, acidity and pH, and phosphorus. *Atmospheric Environment*, 93**,** 3-100,
1076 doi: 10.1016/j.atmosenv.2013.10.060.

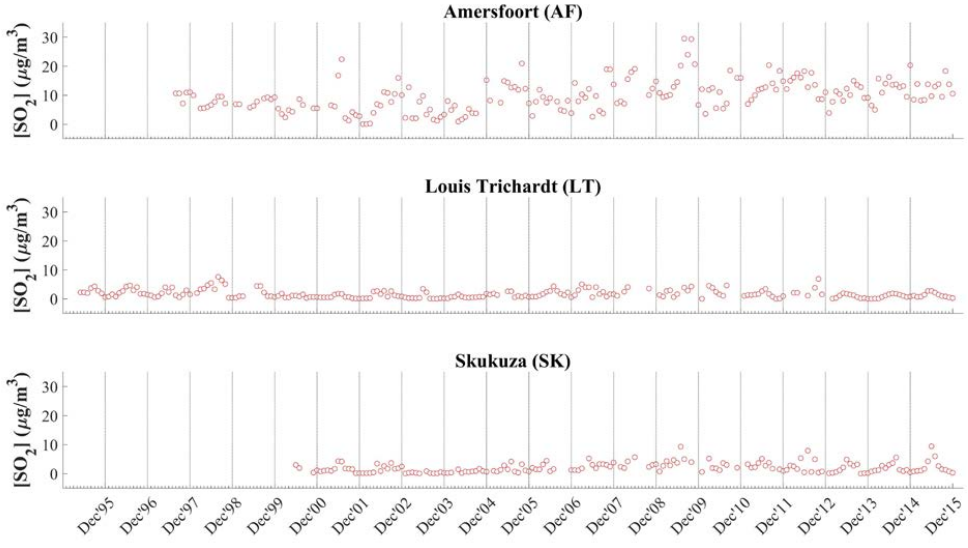
1077 Westcott, G., Tacke, M., Schoeman, N. & Morgan, N. 2007. Impala Platinum Smelter,
1078 Rustenburg: An integrated smelter off-gas treatment solution. *The Journal of the Southern*
1079 *African Institute of Mining and Metallurgy*, 107**,** 7.

1080

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

1081 # Appendix



1082

1083 **Figure A1:** Time series of monthly average SO_2 concentrations measured at Amersfoort
1084 (AF), Louis Trichardt (LT) and Skukuza (SK) using passive samplers over the
1085 relevant measurement periods

1086

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

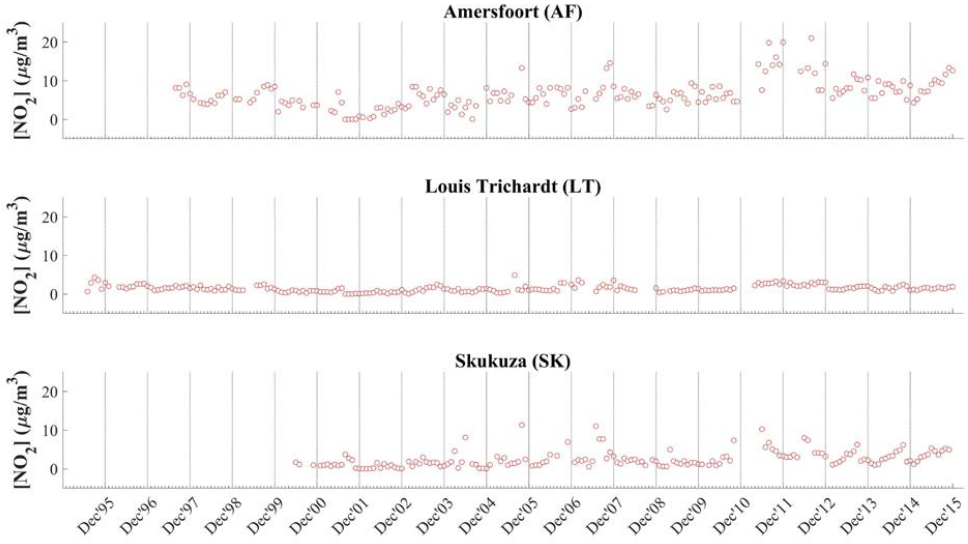


Figure A2: Time series of monthly average NO_2 concentrations measured at Amersfoort (AF), Louis Trichardt (LT) and Skukuza (SK) using passive samplers over the relevant measurement periods

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

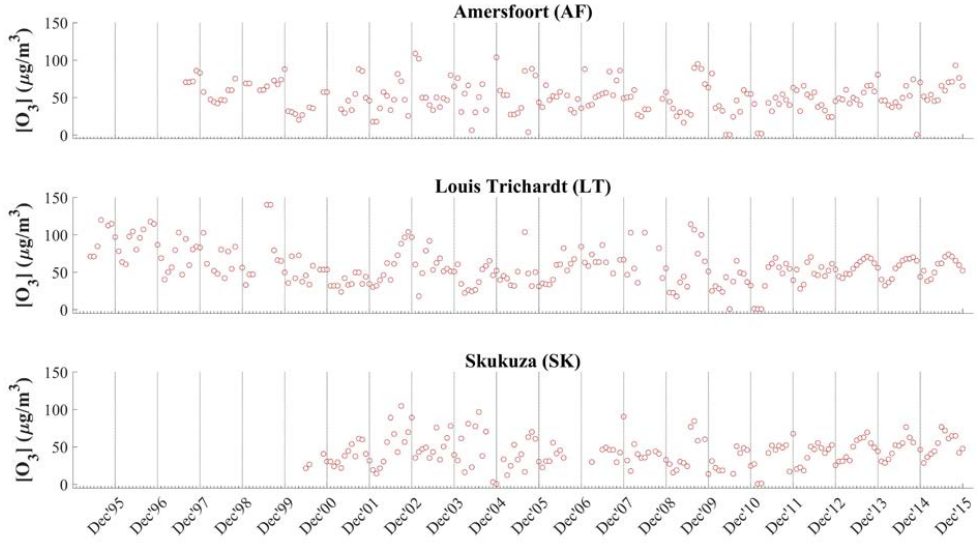


Figure A3: Time series of monthly average O_3 concentrations measured at Amersfoort (AF), Louis Trichardt (LT) and Skukuza (SK) using passive samplers over the relevant measurement periods

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

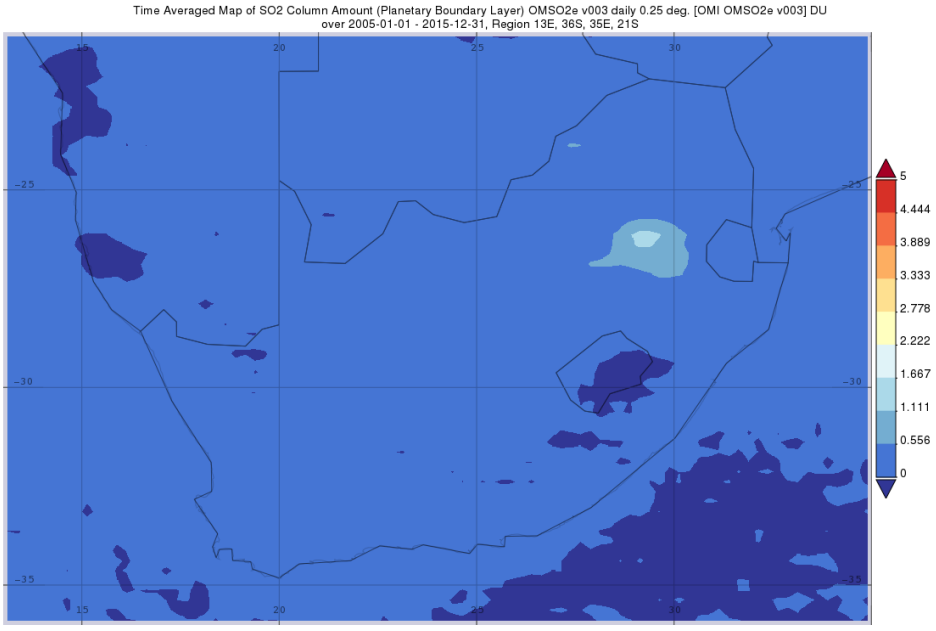


Figure A4: Geospatial map of southern Africa depicting the SO_2 column amount averaged over the period 2005 to 2015 obtained using the data from the NASA Giovanni satellite (https://giovanni.gsfc.nasa.gov/giovanni/)

https://doi.org/10.5194/acp-2020-166
Preprint. Discussion started: 6 April 2020

Atmospheric Chemistry and Physics Discussions

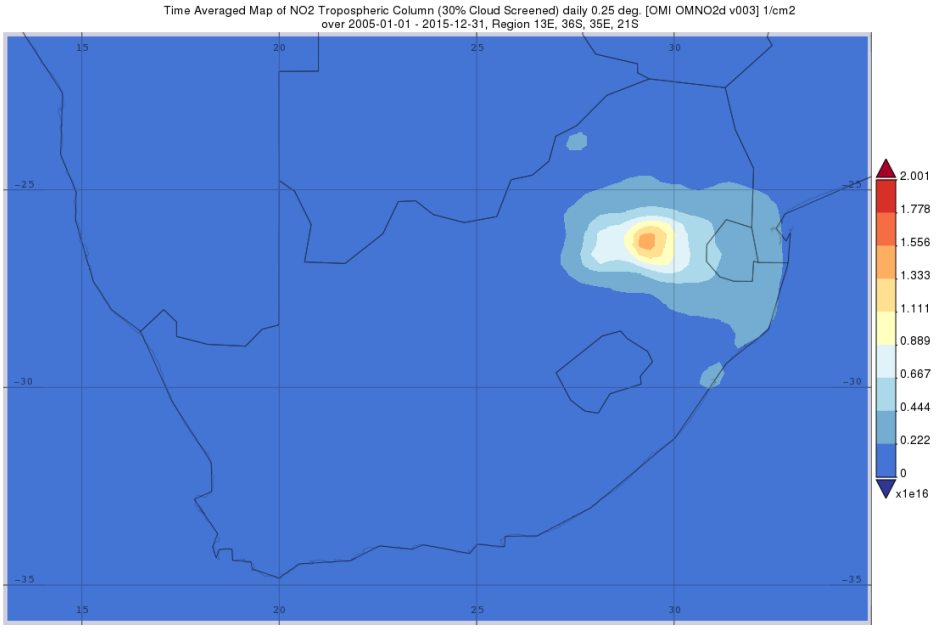


Figure A5: Geospatial map of southern Africa depicting the NO_2 tropospheric column density averaged over the period 2005 to 2015 obtained using the data from the NASA Giovanni satellite (https://giovanni.gsfc.nasa.gov/giovanni/)