Characterization of satellite based proxies for estimating nucleation mode particles over South Africa

A.-M. Sundström[1], A. Nikandrova[1], K. Atlaskina[1], T. Nieminen[1], V. Vakkari[2,1], L. Laakso[2,3], J.P. Beukes[3], A. Arola[4], P.G. van Zyl[3], M. Josipovic[3], A.D. Venter[3], K. Jaars[3], J.J. Pienaar[3], S. Piketh[3], A. Wiedensohler[5], E.K. Chiloane[3,6], G. de Leeuw[2], and M. Kulmala[1]

[1]{Department of Physics, University of Helsinki, Helsinki, Finland}

[2]{ Finnish Meteorological Institute, Helsinki, Finland.}

[3]{Unit for Environmental Science and Management, North-West University, Potchefstroom, South Africa}

[4] {Finnish Meteorological Institute, Kuopio, Finland.}

[5] {Leibniz Institute for Tropospheric Research, Leipzig, Germany}

[6] {Eskom Holdigns SOC Ltd, Sustainability Division, South Africa}

Abstract

Proxies for estimating nucleation mode number concentrations and further simplification for their use with satellite data have been presented in Kulmala et al. (2011). In this paper we discuss the underlying assumptions for these simplifications and evaluate the resulting proxies over an area in South Africa based on comparison with a suite of ground-based measurements available from four different stations. The proxies are formulated in terms of sources (concentrations of precursor gases (NO_2 and SO_2), and UV-B radiation intensity near the surface), and a sink term related to removal of the precursor gases due to condensation on pre-existing aerosols. A-Train satellite data are used as input to compute proxies. Both the input data and the resulting proxies are compared with those obtained from ground-based measurements. In particular a detailed study is presented on the substitution of the local condensation sink (CS) with satellite aerosol optical depth (AOD) which is a column-integrated parameter. One of the main factors affecting the disagreement between CS and AOD is the presence of elevated aerosol layers. Overall, the correlation between proxies

calculated from the in situ data and observed nucleation mode particle number concentrations (N_{nuc}) remained low. At the time of the satellite overpass (13-14 LT) the highest correlation is observed for SO_2/CS (R^2=0.2). However, when the proxies are calculated using satellite data, only NO_2/AOD showed some correlation with N_{nuc} (R^2=0.2). This can be explained by the relatively high uncertainties related especially to the satellite SO_2 columns and by the positive correlation that is observed between the ground-based SO_2 and NO_2 concentrations. In fact, results show that the satellite NO_2 columns compare better with in situ SO_2 concentration than the satellite SO_2 column. Despite the high uncertainties related to the proxies calculated using satellite data, the proxies calculated from the in situ data did not predict significantly better $N_{nuc.}$ Hence, overall improvements in the formulation of the proxies are needed.

1 Introduction

Aerosol particles are key constituents in the Earth-Atmosphere system that can alter climate through their direct and indirect effects on the Earth's radiation budget. Aerosols affect the radiation budget directly by scattering and absorbing solar radiation, and indirectly by acting as cloud condensation nuclei or ice nuclei and modifying clouds' radiative properties and lifetimes. However, the quantification of the aerosol effects on climate is very complex and large uncertainties still exist due to the high spatial and temporal variability of aerosol mass and particle number concentrations (e.g. IPCC, 2013). Besides the climatic effects, aerosols affect human life by reducing the air quality and visibility as well as affecting human health especially in urban areas. Particulate air pollution has been associated with adverse cardiovascular and pulmonary diseases, and even with rises in the numbers of deaths among older people (e.g. Seaton et al., 1995 Uttel et al., 2000, Schnelle-Kreis, 2009).

Primary aerosol particles are emitted directly into the atmosphere; e.g. sea spray aerosol, desert dust, aerosol generated from biomass burning and fossil fuel combustion. Secondary particles are formed from precursor gases through gas-to-particle conversion. The formation of new particles is strongly connected to the presence of sulphuric acid and other vapours of very low volatility, as well as the magnitude of solar radiation (e.g. Kulmala et al., 2008, Kulmala et al., 2005). On the other hand pre-existing aerosol particles act as a sink for the vapours inhibiting new aerosol formation (e.g. Kulmala et al, 2008). These new nanometer-size aerosol particles grow through condensation and coagulation to sizes where they may act

as cloud condensation nuclei (particle diameter D_p> ~ 50 nm) or where they are large enough (D_p > ~ 100 nm) to scatter solar radiation and thus affect the Earth's radiation budget.

Several studies have shown that nucleation occurs frequently in the continental boundary layer and free troposphere from clean to polluted environments (Kulmala et al., 2004, Kulmala et al. 2008 and references therein). Laakso et al. (2008) and Vakkari et al. (2011) have studied new particle formation over moderately polluted savannah ecosystems in South Africa and found that nucleation takes place in the boundary layer almost every sunny day throughout the year with a frequency of as high as 69% of all analysed days (Vakkari et al. (2011)). Hirsikko et al. (2012) extended the studies in South Africa to a polluted measurement site and found an even higher frequency for the nucleation event days (86%), which is among the highest event frequencies reported in the literature so far. Hirsikko et al. (2013) also studied the causes for two or three consecutive daytime nucleation events, followed by subsequent particle growth during the same day. They concluded e.g. that the multiple events were associated with SO_2 rich air from industrial sources.

Satellite instruments have been providing global observations of the Earth's atmosphere for three decades (e.g. Lee et al., 2009, Kokhanovsky and de Leeuw, 2009, Burrows et al., 2011). Information about the spatial distribution of aerosols and trace gases can be obtained from multiple instruments with various temporal and spatial resolution and coverage. Passive remote sensing instruments such as NASA's Ozone Monitoring Instrument (OMI) onboard the AURA platform or the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Terra and Aqua platforms use solar radiation to detect either trace gases or aerosol and cloud properties. Trace gas remote sensing techniques using OMI are based on the trace gas absorption features in the UV-region (wavelength λ ~ 200-400 nm), whereas the remote sensing of aerosol particles is mainly based on measurements in the UV/visible and near infrared regions (λ ~ 500-2000 nm). Since the aerosol measurements utilize only the optically active size range of the solar spectrum, the detectable aerosol sizes are limited to particles with diameters greater than about 100 nm. Nucleation mode particles (smaller than about 25-30 nm in diameter), therefore, cannot be detected directly using satellite instruments. In 2011 Kulmala et al. introduced proxies, i.e. parameterizations for estimating the number concentrations of nucleation mode (N_{nuc}) simplified for the use with satellite data. These simplifications were made assuming that in situ parameters could be replaced with satellite-based observations. Their study was the first attempt to estimate the global nucleation mode aerosol concentrations using data derived from satellite measurements. The proxies were

defined in terms of sources and sinks. The nucleation source terms consist of precursor gas column densities (NO_2 or SO_2) and UV-radiation intensity near the surface (all from OMI as opposed to in situ data in the initial proxies) whereas the sink term, i.e. the condensation sink in the original proxy formulation related to the aerosol surface area concentration is assumed to be proportional to the aerosol optical depth (AOD, from MODIS). More recently Crippa et al. (2013) formulated a new proxy algorithm for ultrafine particle number concentrations based on satellite-derived parameters. They used multivariate linear regression approach to derive the proxy, where the source terms consisted of SO_2, UV (from OMI), and NH_3 (from Tropospheric Emission Spectrometer, TES). The sink term was formulated using MODIS (collection 5.0) AOD and the Ångström exponent, which expresses the spectral dependence of AOD on the wavelength. However, there are issues with the Ångström coefficient (e.g. Mielonen et al., 2011), and thus this parameter is no longer included in the most recent MODIS collection 6.0 land parameters (Levy et al., 2013).

In this work we evaluate the simplifications and underlying assumptions of the method introduced in Kulmala et al. (2011) to estimate the number concentration of nucleation mode particles from satellite-derived data. The study area is the north-eastern part of South Africa (25-28S, 25.5 -30.5E, Figure 1.). Even though the area is not very large, it comprises lots of contrasts from the emission point of view; the cities of Johannesburg and Pretoria, as well as highly industrialized areas especially east from the cities, versus a very clean background in the western part of the study area. The study period considered is Jan 2007- Dec 2010. There are also four different measurement stations located within the region of interest, where observations of various in situ parameters were available.

This work comprises of two parts:

1) A detailed investigation of replacing the condensation sink (CS, defined below in Eq. 8), a local parameter evaluated from in situ observations, with the AOD, a column-integrated aerosol property available from satellite.
2) The estimation of how well satellite data can be used to compute proxies for nucleation mode particle number concentrations. This comprises the analysis of both the satellite- and in situ-based proxy components and the proxies, as well as the comparison of the proxies with the measured concentration of nucleation mode particles. The influence of the uncertainties in the satellite-derived quantities on the proxy is also evaluated.

2 Data

In this study, a variety of data was used from satellite instruments and ground-based stations (see Table 1 for a summary). Satellite data used originate from NASA's Afternoon-Train (A-Train) constellation. The A-Train constellation consists of seven satellites that are on a same polar-orbiting track and follow each other closely enabling near-simultaneous observations of a variety of atmospheric parameters. The equatorial overpass for the A-Train satellites is around 1:30 p.m. local time. In this study we use OMI Level 2 products, i.e. the NO_2 tropospheric column (Bucsela et al., 2013), the SO_2 planetary boundary layer (PBL) product (Krotkov et al., 2006, Krotkov et al., 2008), and the 310 nm irradiance (UV-B) at surface at local noon (Tanskanen et al., 2006). It is noted that the OMI SO_2 PBL product describes the SO_2 concentration integrated over the whole atmospheric column, and PBL refers to the a priori profile assumed in the retrieval of this product. The OMI L2 products are provided with a nominal spatial resolution of 13 x 24 km^2. For the current study they were re-gridded onto a 3 km x 3 km geographical grid as in Fioletov et al. (2011). In this way the effective spatial resolution could be increased despite that the instrument resolution is coarser than the grid. For NO_2 and SO_2 only those observations were used where the (radiative) cloud fraction was below 20%.

According to Lamsal et al. (2014), and references therein, the uncertainty in the OMI NO_2 tropospheric column concentrations is about $0.75 \cdot 10^{15}$ molec./cm^2, whereas Krotkov et al. (2008) report that the SO_2 PBL product could be associated with noise as high as 1.5 DU. However averaging the SO_2 columns over longer a period and/or over a larger spatial area could reduce the noise to 0.3-0.6 DU. For OMI UV-B irradiance the relative uncertainty is on average 7%, but could be higher e.g. due to some episodic aerosol plumes (Tanskanen et al., 2006).

The AOD used in this study is the MODIS Aqua collection 6.0 AOD product at 3 km spatial resolution (Levy et al., 2013). The relative uncertainty for the MODIS AOD over land is reported as 0.05+15%. For selected cases also vertical aerosol extinction profiles from the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) (Winker et al., 2007) are used.

The in situ data used in this study are collected at four different stations in South Africa: Elandsfontein (ELA), Marikana (MAR), Botsalano (BOT), and Welgegund (WEL). All of these stations are located in the north eastern part of the country shown in Fig. 1. Depending

on the station, the measured parameters included e.g. particle size distribution, extinction coefficient and trace gas concentrations. More detailed description of the in situ measurements at the Marikana station can be found e.g. in Venter et al. (2012), at the Welgegund station in Beukes et al. (2014), at the Elandsfontein station in Laakso et al. (2012), and at the Botsalano station in Vakkari et al. (2013). Also data from the Aerosol Robotic Network (AERONET, http://aeronet.gsfc.nasa.gov, Holben et al., 1998) at the Elandsfontein station is used. AERONET is a global ground-based sunphotometer network, providing observations of aerosol optical, microphysical, and radiative properties that are available in a public domain. The aerosol optical properties in the total atmospheric column are derived from the direct and diffuse solar radiation measured by the Cimel sunphotometers.

3 Proxies

Kulmala et al. (2011) derived the N_{nuc} proxies for regional scale nucleation and nucleation from primary emissions. The proxies were determined as the ratio of a source and a sink term. Regional scale nucleation is associated with photochemistry, and typically occurs over a spatial scale of hundreds of kilometres, whereas nucleation from primary emissions occur in the vicinity of local sources such as industrial or urban areas (Kulmala et al., 2011; and references therein). On a regional scale it was assumed that sulphuric acid acts as the driver of the regional nucleation process. Sulphuric acid is formed by oxidation of sulphur dioxide (SO_2) with the hydroxyl radical (OH), which, on the other hand, is mainly formed via photolysis of ozone and UV-radiation. The main sink for sulphur acid is collisions with pre-existing aerosols. Petäjä et al. (2009) derived the proxy for the ambient sulphuric acid as $UV \cdot [SO_2]/CS$, which was considered as the source term in the regional scale nucleation proxy. Taking into account that in addition to sulphuric acid, the pre-existing aerosols are also the sink for the newly formed particles (N_{nuc}), the regional scale nucleation proxy is determined as (Kulmala et al., 2011):

$$P_{Nnuc,regional} = \frac{UV \cdot [SO_2]}{CS^2} \quad (1)$$

where CS denotes the condensation sink of pre-existing aerosols.

Nucleation from primary emissions can be extremely rapid process. The source term of the corresponding proxy is related to the concentration of nitrogen dioxide (NO_2) or sulphur

dioxide while the sink term is determined by the condensation sink. For nucleation from primary emissions two proxies are defined as (Kulmala et al., (2011)):

$$P_{N_{nuc,prim.}} = \frac{[NO_2]}{CS}, \quad (2)$$

$$P_{N_{nuc,prim.}} = \frac{[SO_2]}{CS}, \quad (3)$$

In each of the proxies the source terms are estimated from the satellite measurements by replacing the SO_2 and NO_2 concentrations at the surface with the column densities from the satellite. The amount of global UV radiation is also available from satellite measurements e.g. as a local noon irradiance at 310 nm wavelength (UV-B-radiation) at the surface. For the sink parameter (CS), Kulmala et al. (2011) proposed to use the AOD which describes the total aerosol extinction in the atmospheric column. The relation between the CS and the AOD will be discussed in the following section. By replacing CS with AOD the simplified proxy for using satellite data for primary nucleation becomes:

$$P^{Sat.}_{N_{nuc}} = \frac{[NO_2]_{column}}{AOD} \quad (4)$$

$$P^{Sat.}_{N_{nuc.}} = \frac{[SO_2]_{column}}{AOD} \quad (5)$$

For regional nucleation the proxy expressed in terms of satellite data becomes

$$P^{Sat.}_{N_{nuc.}} = \frac{UV[SO_2]_{column}}{AOD^2} \quad (6)$$

In addition we also considered

$$P^{Sat.}_{N_{nuc.}} = \frac{UV}{AOD^2} \quad (7)$$

as a potential proxy for the number concentration of nucleation mode particles. This proxy corresponds to the case shown in Kulmala et al. (2011), where the sulphur dioxide concentration was assumed to be constant. In this work the proxy defined in Eq. 7 is considered mainly to study how large effect the satellite-based SO_2 has on the performance of the regional scale nucleation proxy.

3.1 Condensation sink and aerosol extinction

As indicated in the previous section, Kulmala et al. (2011) proposed AOD as a substitute for CS. Both parameters are also roughly proportional to the aerosol surface area distribution. According to e.g. Lehtinen et al. (2003) the condensation sink is defined as

$$CS = 2\pi\rho_{diff}\int_0^\infty D_p\beta_M(D_p)n(D_p)dD_p \quad (8)$$

, where D_p is the particle radius, $n(D_p)$ is the particle number size distribution function, ρ_{diff} is the diffusion coefficient of the condensing vapour, and $\beta_M(D_p)$ is the transitional correction factor for mass flux (Fuchs and Sutugin, 1971).

Aerosol optical depth describes quantitatively the column-integrated extinction of solar light caused by atmospheric aerosols and it is one of the standard aerosol parameters that is retrieved from the satellite radiance observations. At a height z and for a wavelength λ the aerosol extinction is defined as

$$\sigma_{ext,z,\lambda} = \frac{1}{4}\pi\int_0^\infty Q_{ext}(\lambda, D_p, m){D_p}^2 n(D_p)dD_p \quad , \quad (9)$$

where Q_{ext} is the extinction efficiency describing aerosols ability to scatter and absorb solar light. At a fixed wavelength the extinction efficiency is a complex function of aerosol size and complex refractive index m (which in turn depends on the aerosol particle composition). Also the particles shape affects somewhat on Q_{ext}, but this is not considered in this study. If the particles are assumed to be spherical, Q_{ext} can be calculated using a computer code based on the Lorenz-Mie theory (Mishchenko et al., 2002). AOD is obtained by integrating σ_{ext} over the total atmospheric column.

The differences between CS and σ_{ext} (at a certain height) as a function of particle size are illustrated in Fig. 2. Both parameters are derived using the same aerosol size distribution (Fig. 2, left panel). The σ_{ext} is calculated using a refractive index of m=1.48+0.003i and wavelengths of 0.55 and 0.45 μm. As Fig. 2 shows, particles with D_p about 0.05-0.1 μm have the largest contribution to CS, whereas for σ_{ext} the largest contribution is coming from particles with D_p about 0.2-0.8 μm. The notable difference between the two quantities is that particles D_p < 0.1 μm can have a contribution to CS which is several orders of magnitude larger than that to σ_{ext}. On the other hand, σ_{ext} is significantly more sensitive to particles with

D_p > 1.0 μm than CS. It is clear that e.g. a large change in number concentration of the smaller particle sizes would change the value of total CS when integrated over the size distribution, but would have a minor effect on the value of σ_{ext}, and vice versa, if e.g. the number concentration of large particles increased there would be little effect on CS. It is noted that in addition to the theoretical differences the possibility of elevated aerosol layers affect the column integrated values of σ_{ext}, i.e. the AOD, which must be considered when comparing the satellite based AOD with in situ CS.

The response of σ_{ext} to changes in the particle size distribution depends to a certain extent on the particle composition and the measurement wavelength. If the particle absorption is high (i.e. the imaginary part of m ~0.1i), the contribution of particles D_p < 0.1 μm to σ_{ext} would be somewhat higher than in Fig. 2. Shorter wavelengths increase the sensitivity to smaller particles, but as Fig. 2 illustrates, a 0.1 μm decrease in wavelength does not improve the sensitivity significantly. Much shorter wavelengths would be needed to increase the sensitivity of σ_{ext} to particles D_p < 0.1 μm, but such measurements could not be carried out in a real atmosphere.

4 Results

The proxies as defined in Sect. 3 are formulated in terms of parameters which are either obtained from ground-based in situ measurements (Eqs. 1-3) or from satellite data (Eqs. 4-7). In this section the performance of these proxies is critically evaluated and in particular each of the satellite-based parameters is critically examined.

4.1 Comparison of condensation sink and aerosol optical depth

Replacing CS with AOD is perhaps the most crucial assumption when determining the proxies using satellite data, as indicated in Kulmala et al. (2011). Apart from the sensitivity of these parameters for different particle sizes discussed in Section 3.1, other differences play a role such as the vertical variation of the aerosol concentrations, the particle size range considered and the dependence of aerosol particle size on relative humidity. CS is determined from measured dry particle size distributions with a correction for ambient humidity. CS at Botsalano and Marikana has been estimated from submicron size distribution while at Elandsfontein size distributions up to 10 μm were used. In contrast, the AOD is an integrated quantity with contributions from all optically active aerosols throughout the whole

atmospheric column. To assess the effect of these different factors on the relation between the AOD and CS, the following comparisons are made:

1) In situ CS with nephelometer aerosol scattering coefficient
2) In situ nephelometer aerosol scattering coefficient with AOD from AERONET
3) In situ CS with AOD from both AERONET- and satellite measurements.

Coincident measurements of size distributions to derive the CS and aerosol scattering coefficients from a nephelometer are only available from the Elandsfontein measurement station. The comparison between CS and scattering coefficient serves to eliminate effects of the vertical variation of the aerosol concentrations on the comparison. The nephelometer measures the dry particle scattering at 0.525 μm wavelength and the results are presented at Standard Temperature and Pressure (STP) atmosphere. The maximum particle size is limited to $D_p \sim 10$ μm. It is noted that the nephelometer considers only aerosol scattering, and not the total extinction which would also require information on absorption. However, the contribution of absorption to the total aerosol extinction is generally much smaller than scattering. Laakso et al. (2012) reported that at Elandsfontein the absorption was increased during the coldest months (May-Oct.) due to biomass burning, domestic burning of coal for heating and cooking, contributing about 15-20% to the total aerosol extinction whereas during the warmer months (Nov.-Apr.) absorption contributed ~10% of the total aerosol extinction. To take the seasonal variation of absorption into account, the CS and the scattering coefficients were compared separately for the periods May-Oct. and Nov.-Apr. The results in Fig. 3 show that for both periods scattering coefficients and CS were well-correlated with R^2=0.67 for Nov.-Apr., and R^2=0.71 for May-Oct. The R^2 values were somewhat higher than those from measurements at a clean continental boreal forest measurement site in Hyytiälä, Southern Finland (R^2=0.62, Virkkula et al., 2011).

The next step is to compare the nephelometer scattering coefficient to the AOD to evaluate effects of the possible occurrence of elevated aerosol layers and/or boundary layer mixing. Also the presence of large dust particles might have some effect on the comparison due to the limited particle size in the nephelometer inlet. In this comparison we first compare with AERONET measurements of AOD at Elandsfontein, which are more accurate than those retrieved from satellite data. As Fig. 4 shows, the correlations between the AERONET AOD and the in situ scattering coefficient (warm season R^2=0.46, cold season R^2=0.24) are lower than those between the CS and the scattering coefficient. This indicates that the elevated

aerosol layers and boundary layer mixing might affect more than the theoretical differences when estimating the sink of pre-existing aerosols by using the AOD.

For the comparison of CS with the AOD retrieved from MODIS, daily AOD values were used which are spatial averages of the observations within 3 km radius from each measurement station. As Fig. 5 shows, the CS- vs. satellite AOD data are scattered all over the graph and although there is a tendency of increasing CS with increasing AOD there is no apparent correlation ($0.03 \leq R^2 \leq 0.06$). As an alternative, a bivariate method (York et al., 2004) was applied to account for the uncertainties associated to both CSs and MODIS AODs in the fitting. For CS the uncertainty was assumed to be 10% (Petäjä et al., 2013) and for MODIS AOD an uncertainty of 0.05+15% was used (Levy et al., 2013). This means that for low AOD the relative uncertainty is rather high, e.g. for AOD=0.1 the relative uncertainty would be 65%. As Fig. 5 shows the bivariate method gave very different results than LSQ.

At Marikana and Elandsfontein the largest observed AODs are not related to lagest CS, which could be due to the presence of elevated aerosol layers. In a recent study by Giannakaki et al., (2015) data from a ground-based lidar at Elandsfontein are analyzed and the results show that the mean contribution of elevated aerosol layers to the AOD is 46%. To estimate the effect of elevated aerosol layers on the CS-AOD comparison at Marikana, CALIPSO observations of aerosol vertical extinction profiles are used. All CALIPSO daytime overpasses between 8.2.2008 and 17.5.2010 within 50 km from the Marikana station were considered. Due to the small CALIPSO swath width only 48 days of data are available. At Marikana the median MODIS AOD is 0.15 for the whole measurement period, and as Fig. 5 shows, the CS values are less scattered when AODs are smaller than the median. Therefore the vertical aerosol extinction profiles from CALIPSO are studied separately for the cases where MODIS AOD $\leq$ 0.15 and AOD > 0.15. As Fig. 6 shows, for higher AODs the median extinction profile indicates an elevated aerosol layer, which supports the result that high AODs also at Marikana are likely to be associated with an elevated aerosol layer.

4.2 Proxies defined from the in situ data and comparison with N_{nuc}

The proxies are first computed using in situ measurements from Marikana and Elandsfontein following Eqs. 1-3 to evaluate how well each of them could predict the nucleation mode number concentration within our study area. It is noted that due to different instrumentation N_{nuc} from Marikana consists of particles with D_p < 30 nm, but at Elandsfontein N_{nuc} consists

of particles with D_p 10-30 nm. In addition, CS at Marikana is defined from submicron particles whereas at Elandfontein CS is defined from particles with D_p<10 μm.

Figure 7 shows the diurnal variation of each of the in situ proxy components and the number concentration of nucleation mode particles. At Marikana the N_{nuc} median peaks about 10 a.m., and at Elandsfontein about an hour later. At the time of the satellite overpass the median of N_{nuc} is lower than before noon at both locations, and about the same order of magnitude. The diurnal variation of NO_x-NO and SO_2 concentrations show somewhat different characteristics at Marikana than at Elandsfontein, The morning and evening peaks of NO_x-NO at Marikana are most likely associated with household combustion and traffic whereas the single SO_2 peak in the morning is most likely related to the industrial emissions and the break-up of the inversion layers that form quite regularly in the South African Highveld (Venter et al., 2012). At Elandsfontein, where the major emission source is heavy industry, an increase in the NO_x-NO and the SO_2 concentration medians are seen about 10 a.m. The median of SO_2 concentration decreases in the late afternoon while the median of NO_x-NO concentration does not vary much. At the time of the satellite overpass the NO_x-NO and SO_2 medians are much higher at Elandsfontein than at Marikana. Results show also that at the time of the satellite overpass NO_x-NO and SO_2 are positively correlated; at Elandsfontein R^2=0.58, and at Marikana R^2=0.32 are obtained. At Elandsfontein CS does not show any clear diurnal variation and it is systematically lower than at Marikana. Also at Marikana the diurnal variation of the CS is rather weak during the daytime but a peak in the median is seen in the evening.

Figure 8 shows the diurnal variation of the in situ proxies at Marikana and Elandsfontein. The comparison of the diurnal variation of the proxies and N_{nuc} indicates that the proxy-N_{nuc} relation depends on the time of the day. At the time of the satellite overpass (13-14 LT) the highest correlation with N_{nuc} at Marikana is obtained with the SO_2/CS-proxy (R^2=0.22, Fig.9), but at Elandsfontein the correlation remains below 0.1. At Marikana the correlation of N_{nuc} with $SO_2 \cdot UV./CS^2$ - proxy (Eq. 1) is less good at the time of the satellite overpass but at 9-10 a.m. R^2 =0.25. On the other hand the (NO_x-NO)/CS and UV/CS^2 proxies do not perform well in predicting N_{nuc}. Also, it is noted that at the time of the satellite overpass all the proxy values show much higher median values at Elandsfontein than at Marikana while the median for N_{nuc} is about the same at both locations. At Elandsfontein somewhat better correlations with N_{nuc} are observed if only the source terms of the proxies are considered. For example, the values of R^2 between N_{nuc} and SO·UV are 0.35 at 10-11 LT, and 0.14 at 13-14 LT,

respectively, but when the sink-term CS^2 is included in the proxy there is no correlation. At Marikana CS doesn't have as high influence on the proxy performance as at Elandsfontein.

The difference with the results reported for Southern Finland (Kulmala et al. (2011)) is that in our study SO_2 has a strong effect on the performance of the proxy: without SO_2 the UV/CS^2 – term does not correlate with N_{nuc}. Given that the satellite data are associated with much higher uncertainties than the in situ measurements, these in situ-based results can be considered as upper limit for the overall performance of the proxies computed using satellite data (Eqs. 4-7).

.

4.3 Proxies using satellite data

4.3.1 Spatial pattern of the satellite-based proxies

Each of the satellite based parameter is analyzed from Jan. 2007 to Dec. 2010. Figure 11 shows the four year medians of SO_2 and NO_2 column densities obtained from the OMI instrument, as well as the AOD at 550 nm from MODIS Aqua observations. Daily satellite data is used to define the satellite-based proxies over the study area (Eq. 4-7). Figure 12 shows the four year median spatial patterns for the four satellite-based proxies. The spatial patterns of these four proxies are quite different and in particular there is large difference between the spatial variation of the regional proxies and that of the proxies for nucleation from primary emissions. As expected, the latter strongly reflect the spatial distributions of the precursor gases with high concentration over the Highveld industrial area, where the values of NO_2 and SO_2 columns are high and the sink (AOD) is low. For the NO_2/AOD proxy also elevated values are observed over the Johannesburg-Pretoria area while for the other proxies a local minimum occurs over these cities.

All the four satellite proxies show larger values at Elandsfontein than at Marikana, which is consistent with the results obtained for the in situ proxies. Based on the in situ results the SO_2-related proxies are expected to predict N_{nuc} at the time of the satellite overpass better than the other proxies. Comparison of the spatial patterns of each proxy calculated using satellite data in the vicinity of the in situ measurement stations shows that there are not very much difference between the spatial pattern of SO_2- and NO_2-related proxies.

The propagation of relative uncertainty associated with the proxies using satellite data can be estimated by comparing the uncertainties related to each satellite parameter (Sect. 3) and the observed median values shown in Fig. 11. For example, over background areas where both

AOD and SO_2 are low, the SO_2 ·UV-B /AOD^2 -proxy can have an uncertainty of over 90%. On the other hand, over source areas where both NO_2 and AOD are slightly elevated the NO_2/AOD proxy would have an uncertainty of about 50%. Generally over South Africa the uncertainty in satellite-based proxies is high, especially over areas where both low values of NO_2, SO_2 and AOD are frequently observed.

4.3.2 Comparison of satellite and in situ proxy components

Before evaluating the performance of the proxies using satellite data, first the quality of the parameters used in these proxies should be examined. The CS/AOD comparison was discussed in Sect. 4.1. Here we compare satellite data for NO_2, SO_2 and UV-B with in situ data at each of the measurement stations. The satellite data for each station is collected within a 12 km (NO_2, SO_2, UV-B) or a 3 km (AOD) radius from the station and the results are compared with hourly means of the in situ data extracted between 13-14 LT, i.e. ± 30 min within the approximate satellite overpass.

The satellite NO_2 column densities and the in situ NO_x-NO concentrations are reasonably well correlated as are the satellite UV-B irradiances and the global radiation measured at each station. The highest correlation for NO_2 were obtained at Marikana (R^2=0.55), and lowest at Elandsfontein (R^2=0.26). For UV-B and global radiation the correlations were $0.61 \leq R^2 \leq 0.77$. In Kulmala et al. (2011) a constant value was assumed for the satellite-based SO_2 when defining the global proxy maps, because the SO_2 product they used (middle tropospheric SO_2) did not show a reasonable spatial pattern. In this study the middle - troposphere SO_2 data was replaced by the OMI boundary layer product (Sec. 3), which improved the characterization of the SO_2 spatial variation (Fig. 10). However, the relative uncertainty in the satellite-based SO_2 remains still high, unless the data is averaged over a long time period/ large spatial area. At all three stations lower correlation between the satellite and in situ based SO_2 measurements were obtained than for the other source parameters, at Marikana there is practically no correlation. Similar results were obtained when the satellite- and in situ-based proxies were compared (Table 2, figures in the supplementary material). Overall large differences exist between the satellite proxies and in situ proxies.

Since at Marikana and Elandsfontein the in situ data showed correlation between the (NO_x-NO) and the SO_2 concentrations, the satellite NO_2 column density is also compared with the

in situ SO_2. Results show that in fact the OMI NO_2 compares better with the in situ SO_2 than the actual OMI SO_2 product. At Elandsfontein R^2=0.25, and at Marikana R^2=0.31 are obtained between the satellite NO_2 column and in situ SO_2 concentration.

4.3.3 Comparison of satellite-based proxies with N_{nuc}

To further evaluate the performance of the satellite-based proxies, they are compared with the in situ N_{nuc}. Only data from Elandsfontein and Marikana are included in the comparison since the number of coincident N_{nuc} and satellite proxy observations was too low at the other stations. As expected, neither of the two satellite-based SO_2 – proxies are able to predict N_{nuc}. Interestingly, the only case where weak correlation is obtained between a proxy using satellite data and N_{nuc} is for the NO_2/AOD (Fig 12). This result is very different than what is expected based on the comparison of the in situ proxies and N_{nuc}. In fact, the connection between NO_2/AOD and N_{nuc} is most probably related to the correlation between the satellite NO_2 column density and the in situ SO_2 concentration. If the source term in the SO_2·UV-B/AOD^2 proxy was replaced by NO_2·UV-B, the correlation with N_{nuc} at Elandsfontein would be R^2=0.23, and at Marikana R^2=0.06. This implies that over areas where SO_2 and NO_2 are affected by some common factors, e.g. emission sources, the satellite NO_2 could be a better estimate for the source term than SO_2.

5 Conclusions

This work explores the use of proxies using satellite data to obtain information on the concentration of nucleation mode aerosol particles (N_{nuc}). These proxies have been formulated using relations derived from data on ground-based nucleation and precursor gases, which were simplified for the use of satellite data in Kulmala et al. (2011). The simplifications and associated assumptions are critically examined. In this study data were used over part of South Africa where ground-based observations are available from four experimental sites, for comparison with both the satellite-based parameters used in the proxy formulations and for comparison of the proxies with ground-based measurements of the nucleation mode aerosol particle number concentrations. For the computation of the proxies, data from the A-train satellites are used. The NO_2, SO_2 and UV-B radiation are obtained from the OMI instrument and AOD from the MODIS instrument. The NO_2 and UV-B data are the same than what was used in Kulmala et al. (2011), but the AOD was upgraded to the newest

Collection 6, three- km product. Also the SO_2 product was changed to the planetary boundary layer product (OMI SO_2 PBL) that represents the total column values with a priori assumption that the emissions are mainly in the boundary layer. The satellite observations are also extensively compared with in situ data.

Based on the proxies derived from the in situ data it is expected that the SO_2-related proxies would be the best predictors of N_{nuc} within the study area at the time of the satellite overpass (13-14 LT). It is also noted that even though the in situ NO_2/CS proxy did not do well in predicting N_{nuc}, a positive correlation between the SO_2 and NO_2 concentrations is found at the measurement stations (at 13-14 LT). The R^2 between in situ SO_2/CS and N_{nuc} is 0.22 and this value could be considered as some kind of "upper limit" for the satellite proxies, for which uncertainties are much higher than for the in situ proxies. Using ground-based data, Kulmala et al. (2011) reported that SO_2 had only moderate influence on the performance of the $SO_2 \cdot UV/CS^2$ proxy in Southern Finland. The overall correlation between this proxy and N_{nuc} over South Africa was even lower (R2= 0.13) than over Southern Finland (R^2=0.29), yet our results clearly indicate a strong influence of SO_2 on the performance of the proxy. If the SO_2 was excluded from the proxy, no correlation with in situ proxies and N_{nuc} was found.

Kulmala et al., (2011) emphasized that the most crucial assumption in deriving the satellite based proxies was the replacement of the CS with AOD. This assumption is further evaluated in the current study using several tests. A fundamental reason for differences between CS and AOD is the intrinsic dependence on different aerosol size ranges, with CS more sensitive to very small particles (smaller than about 200 nm) and AOD more sensitive to particles larger than that. Yet, good correlation is obtained between measured scattering coefficients for dry aerosol and CS evaluated from collocated particle size distribution measurements. When the in situ scattering coefficients or CS are compared with collocated AOD measurements the correlation decreases. This may be due to several effects. In particular the presence of elevated aerosol layers and/or large dust particle increases the AOD but does not affect the CS. However, overall the AOD is rather low (<0.1) over the major part of the study area, which means that these values are also associated with substantial relative uncertainty, which needs to be accounted for when deriving the satellite-based proxies.

Even though the OMI SO_2 PBL data product showed a distinct improvement in describing the spatial patterns of SO_2 as compared to the dataset used in Kulmala et al. (2011), the satellite-based SO_2 did not describe well the day-to-day variations at the measurement stations. In addition, the observed SO_2 column values were often close to the noise level associated with

a single column retrieval reported by Krotkov et al. (2008). The only relation between a satellite-based proxy and N_{nuc} was obtained for NO_2/AOD (at Elandsfontein R^2=0.24, and at Marikana R^2=0.09). The result is different than what was expected based on the in situ proxies. The most probable explanation is the positive correlation between the ground-based NO_2 and SO_2 concentrations within the study area. It is found that in fact the satellite NO_2 column correlates better with in situ SO_2 concentration than the satellite SO_2 column, where no correlation was found.

Overall this study shows that the uncertainties related to the satellite products remain a major issue in this satellite-based proxy approach, especially over areas like South Africa, where the AOD and the SO_2, and NO_2 concentrations are generally relatively low. Throughout the whole study the relative uncertainties related to the satellite-based proxies were well above 50%. In For the NO_2/AOD proxy the largest relative uncertainties were often related to AOD. Otherwise SO_2 was clearly the m**o**st uncertain component in the proxies calculated using satellite data. Despite these uncertainties related to the satellite data, the in situ data did not do significantly better in predicting N_{nuc} within our study area. This indicates that overall improvements in the formulation of the proxies are needed.

Acknowledgements

This work is supported by Academy of Finland (1251427, 1139656, Finnish Centre of Excellence in Atmospheric Science 272041), European Research Council (ATMNUCLE), The European Integrated project on Aerosol Cloud Climate and Air Quality Interactions (EUCAARI), and the ESA projects Aerosol-cci (ESA-ESRIN project AO/1-6207/09/I-LG and ALANIS-Aerosols (Contract no. 4200023053/10/I-LG, STSE-ALANIS Atmosphere-Land Interactions Study Theme 3: Aerosols). Eskom and Sasol supplied logistical support for measurements at Elandsfontein, while the town council of Rustenburg supplied support to the measurement at Marikana. The OMI NO_2, SO_2 and UV data were obtained from the NASA Mirador service maintained by Goddard Earth Sciences Data and Information Services Center (GES DISC). The OMI surface UV data were obtained from the NASA Aura Validation Data Center (AVDC). The MODIS Aqua data were provided by NASA LAADS Web, and the CALIPSO data was obtained from NASA Atmospheric Science Data Center (ASDC).

References

Beukes, P., Vakkari, V., van Zyl, P. G., Venter, A., Josipovic, M., Jaars, K., Tiitta, P., Kulmala, M., Worsnop, D., Pienaar, J., Virkkula, A., and Laakso L.: Source region plume characterisation of the interior of South Africa, as measured at Welgegund, Clean Air Journal, Vol. 23, No. 1, 2013.

Bucsela, E. J., Krotkov N., Celarier, E., Lamsal, L., Swartz, W., Bhartia, P., Boersma, K., Veefkind, J.P., Gleason, J., and Pickering, K.: A new stratospheric and tropospheric NO_2 retrieval algorithm for nadir-viewing satellite instruments: applications to OMI. Atmos. Meas. Tech, 6, 2607-2626 , doi: 10.5194/amt-6- 2607-2013, 2013.

Burrows, J.P., U. Platt and P.Borrell (Editors) The Remote Sensing of Tropospheric Composition from Space, 536 pp., Springer-Verlag Berlin Heidelberg 2011. ISBN: 978-3-642-14790-6 doi: 10.1007/978-3-642-14791-3., pp. 359-313.

Crippa, P., Spracklen, D., and Pryor, S.C.: Satellite-derived estimates of ultrafine particle concentrations over Eastern North America. J. Geophys. Res., 118, 9968-9981, 2013.

Dal Maso, M., Kulmala, M., Riipinen, I., Wagner, R., Hussein, T., Aalto, P. P. & Lehtinen, K. E. J.: Formation and growth of fresh atmospheric aerosols: eight years of aerosol size distribution data from SMEAR II, Hyytiälä, Finland. Boreal Env. Res. 10, 323–336, 2005.

Fioletov, V.E., McLinden, C.A., Krotkov, N., Moran, M.D., and Yang, K.; Estimation of SO_2 emissions using OMI retrievals. Geophys. Res. Lett., vol. 38, L21811, doi:10.1029/2011GL049402, 2011.

Fuchs, N. A., and Sutugin, A. G.: in: Topics in current aerosol research (Part 2), (Eds.) Hidy, G. M., and Brock, J. R., Pergamon, New York, 1971.

Giannakaki, E., Pfüller, A., Korhonen, K., Mielonen, T., Laakso, L., Vakkari, V., Baars, H., Engelmann, R., Beukes, J. P., Van Zyl, P. G., Josipovic, M., Tiitta, P., Chiloane, K., Piketh, S., Lihavainen, H., Lehtinen, K. E. J., and Komppula, M.: One year of Raman lidar observations of free tropospheric aerosol layers over South Africa. Atmos. Chem. Phys. Discuss., 15, 1343-1384, doi:10.5194/acpd-15-1343-2015, 2015.

Hirsikko, A., Vakkari, V., Tiitta, P., Manninen, H.E., Gagné, S., Laakso, H., Kulmala, M., Mirme, A., Mirme, S., Mabaso, D., Beukes, J.P., and Laakso, L.: Characterisation of sub-micron particle number concentrations and formation events in the western Bushveld Igneous Complex, South Africa. Atmos. Chem. Phys, 12, 3951-3967, 2012.

Hirsikko, A., Vakkari, V., Tiitta, P., Hatakka, J., Kerminen, V.-M., Sundström, A.-M., Beukes , J.P., Manninen, H.E., Kulmala M., and Laakso, L.: Multiple daytime nucleation events in semi-clean savannah and industrial environments in South Africa: analysis based on observations. Atmos. Chem. Phys, 13, 5523-5532, 2013.

Holben, B., Eck, T.F., Slutsker, I., Tanre, D., Buis, J.P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J. , Nakajima, T. , Lavenu, F. , Jankowiak, I. and Smirnov, A.: AERONET - A Federated Instrument Network and Data Archive for Aerosol Characterization. Rem. Sens. Environ., vol. 66, 1-16, 1998.

IPCC, Intergovernmental Panel on Climate Change, Fifth Assessment Report: Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Kokhanovsky, A.A., and de Leeuw Gerrit (editors): Satellite Aerosol Remote Sensing over Land, Springer, Berlin, Germany, 2009.

Krotkov, N., Carn, S., Krueger A., Bhartia, P., and Yang, K.: Band residual difference algorithm for retrieval of SO_2 from the Aura Ozone Monitoring Instrument (OMI). IEET T. Geosci. Remote, 44, 12591266, 2006.

Krotkov, N., McClure, B., Dickerson, R., Carn, S., Li, C., Bhartia, P., Yang, K., Krueger, A., Li, Z., Levelt, P., Chen, H., Wang, P., and Lu, D.: Validation of SO_2 retrievals from the Ozone Monitoring Instrument over NE China. Journal of Geophys. Res., Vol. 113, D16S40, doi:10.1029/2007JD008818, 2008.

Kulmala, M., Vehkamäki, H., Petäjä, T., Dal Maso, M., Lauri, A., Kerminen, V.-M., Birmili, W., and McMurry, P.H.: Formation and growth rates of ultrafine atmospheric particles: a review of observations. J. Aerosol Sci., 35, 143-176, 2004.

Kulmala, M., Petäjä, T., Mönkkönen, P., Koponen, I.K., Dal Maso, M., Aalto, P.P., Lehtinen, K.E.J., and Kerminen, V.-M.: On the growth of nucleation mode particles: source rates of condensable vapor in polluted and clean environments. Atmos. Chem. Phys. 5, 409-416, 2005.

Kulmala, M., and Kerminen, V.-M.: On the formation and growth of atmospheric nanoparticles. Atm. Res., 90, 132-150, 2008.

Kulmala, M., Arola, A., Riuttanen, L., Sogacheva, L., de Leeuw, G., Kerminen, V.-M. and Lehtinen, K.E.J.: The first estimates of global nucleation mode aerosol concentrations based on satellite measurements. Atmos. Chem. Phys. 11, 10791-10801, 2011.

Laakso, L., Laakso, H., Aalto, P.P, Keronen, P., Nieminen, T., Pohja, T., Siivola, E., Kulmala, M., Kgabi, N., Molefe, M., Mabaso, D., Phalatse, D., Piennaar, K., and Kerminen, V.-M.: Basic characteristics of atmospheric particles, trace gases and meteorology in a relatively clean Southern African Savannah environment, Atmos. Chem. Phys., 8, 4823-4839, 2008.

Laakso, L., Vakkari, V., Virkkula, A., Laakso, H., Backman, J., Kulmala, M., Beukes, J.P., van Zyl, P.G., Tiitta, P., Josipovic, M., Pienaar, J.J., Chiloane, K., Gilardoni, S., Vignati, E., Wiedensohler, A., Tuch, T., Birmili, W., Piketh, S., Collett, K., Fourie, G.D., Komppula, M., Lihavainen, H., de Leeuw, G., and Kerminen, V.-M.: South African EUCAARI measurements: seasonal variation of trace gases and aerosol optical properties. Atmos. Chem. Phys., 12, 1847-1864, 2012.

Lamsal, L. N., Krotkov, N. A., Celarier, E. A., Swartz, W.H., Pickering, K. E., Bucsela, E. J. , Gleason, J. F., Martin, R.V., Philip, S., Irie, H., Cede, A., Herman, J., Weinheimer, A., Szykman, J.J., Knepp, T. N.: Evaluation of OMI operational standard NO_2 column retrievals using in situ and surface-based NO_2 observations. Atmos. Chem. Phys.,14, 11587–11609, doi:10.5194/acp-14-11587-2014, 2014.

Lee, K. H., Li, Z., Kim, Y. J., & Kokhanovsky, A.: Atmospheric aerosol monitoring from satellite observations: a history of three decades. In Atmospheric and biological environmental monitoring, Springer Netherlands, 2009.

Lehtinen, K. E. J., Korhonen, H., Dal Maso, M., and Kulmala, M.: On the concept of condensation sink diameter. Boreal Env. Res., 8, 405-411, 2003.

Levy, R.C., Mattoo, S., Munchak, L.A., Remer, L.A., Sayer, A.M., Patadia, F., and Hsu N.C. : The Collection 6 MODIS aerosol products over land and ocean. Atmos. Meas. Tech., 6, 2989-3034, 2013.

Lourens, A. S. M., Beukes, J. P., Van Zyl, P. G., Fourie, G. D., Burger, J. W. , Pienaar, J. J., Read, C. E., and Jordaan, J. H.: Spatial and Temporal assessment of gaseous pollutants in the Highveld of South Africa, S. Afr. J. Sci., 107, No. 269, 8 pp., doi:10.4102/sajs.v107i1/2.269, 2011.

Lourens, A. S. M., Butler, T. M., Beukes, J. P., Van Zyl, P. G., Beirle, S., Wagner, T., Heue, K-P., Pienaar, J. J., Fourie, G. D., Lawrence, M.G.: Re-evaluating the NO_2 hotspot over the South African Highveld, S. Afr. J. Sci., 108(9/10), Art. #1146, 6 pages. DOI: 10.4102/sajs. v108i11/12.1146, 2012.

Maritz, P., Beukes, J. P., Van Zyl, P. G., Conradie, E.H., Liousse, C., Galy-Lacaux, C., Castéra, P., Ramandh, A., Mkhatshwa, G., Venter, A.D. and Pienaar, J.J.: Spatial and temporal assessment of organic and black carbon at IDAF sites in South Africa, submitted to Environ. Mon. Assess., 2014.

Mielonen, T., Levy, R.C., Aaltonen, V., Komppula, M., de Leeuw, G., Huttunen, J., Lihavainen, H., Kolmonen, P., Lehtinen, K.E.J., and Arola, A.: Evaluating the assumptions of surface reflectance and aerosol type selection within the MODIS aerosol retrieval over land: the problem of dust type selection. Atmos. Meas. Tech., 4, 201–214, 2011.

Mishchenko, M. I., L. D. Travis, and A. A. Lacis: Scattering, Absorption, and Emission of Light by Small Particles. Cambridge University Press, Cambridge, 2002.

Petäjä, T., Maudin III, R.L., Kosciuch, E., McGrath, J., Nieminen, T., Paasonen, P., Boy, M., Adamov, A., Kotiaho, T., and Kulmala, M.: Sulfuric acid and OH concentrations in a boreal forest site. Atmos. Chem. Phys., 9, 7435-7448, 2009.

Petäjä, T., Vakkari, V., Pohja, T., Nieminen, T., Laakso, H., Aalto, P.P., Keronen, P., Siivola, E., Kerminen, V.-M., Kulmala, M., and Laakso, L.: Transportable Aerosol Characterization Trailer with Trace Gas Chemistry: Design, Instruments and Verification. Aerosol Air Qual. Res., 13: 421–435, 2013.

Schnelle-Kreis, J., Küpper, U., Sklorz, M., Cyrys, J., Briedé, J. J., Peters, A., and Zimmerman, R.: Daily measurements of organic compounds in ambient particulate matter in Augsburg, Germany: new aspects on aerosol sources and aerosol related health effects. Biomarkers, 14 (SI), 39-44, 2009.

Seaton, A., MacNee, W., Donaldson, K., Godden, D.: Particulate air pollution and acute health effects. The Lancet, 345, 176-178, 1995.

Shi, Y., Zhang, J., Reid, J.S., Hyer, E.J., and Hsu, N.C.: Critical evaluation of the MODIS Deep Blue aerosol optical depth product for data assimilation over North Africa. Atmos. Meas. Tech., 6, 949-969, 2013.

Tanskanen, A., Krotkov, N.A., J.R. Herman, and Arola, A.: Surface Ultraviolet Irradiance From OMI. IEEE Trans. Geosci. Remote Sens., Vol. 44, No. 5, 1267-1271, 2006.

Utell, M.J., and Frampton, M.W.: Acute Health Effects of Ambient Air Pollution: The Ultrafine Particle Hypothesis. J Aerosol Med Pulm Drug Deliv., Vol. 13, 4, 355-359, 2000.

Vakkari, V., Laakso, H., Kulmala, M., Laaksonen, A., Mabaso, D., Molefe M., Kgabi, N., and Laakso, L.: New particle formation events in semi-clean South African savannah. Atmos. Chem. Phys., 11, 3333-3346, 2011.

Vakkari, V., Beukes, J.P., Laakso, H., Mabaso, D., Pienaar, J.J., Kulmala, M., and Laakso, L.: Long-term observations of aerosol size distributions in semi-clean and polluted savannah in South Africa. Atmos. Chem. Phys., 13, 1751–1770, 2013.

Veefkind, J.P., Boersma, K.F., Wang, J., Kurosu, T.P., Krotkov, N., Chance, K., and Levelt, P.F.: Global satellite analysis of the relation between aerosols and short-lived trace gases. Atmos. Chem. Phys., 11, 1255-1267, 2011.

Venter, A.D., Vakkari, V., Beukes, J.P., van Zyl, P.G., Laakso, H., Mabaso, D., Tiitta, P., Josipovic, M., Kulmala, M., Pienaar, J.J., and Laakso. L.,: An air quality assessment in the industrialised western Bushveld Igneous Complex, South Afric S. Afr. J. Sci., 108(9/10), No. 1059, doi:10.4102/sajs.v108i9/10.1059, 2012.

Virkkula, A., Backman, J., Aalto, P.P., Hulkkonen, M., Riuttanen, L., Nieminen, T., dal Maso, M., Sogacheva, L., de Leeuw, G., and Kulmala, M.,: Seasonal cycle, size dependencies, and source analyses of aerosol optical properties at the SMEAR II measurement station in Hyytiälä, Finland. Atmos. Chem. Phys., 11, 4445-4468, 2011.

Winker, D. M., Hunt, W. M., and McGill, M.J.: Initial performance assessment of CALIOP. Geophys. Res. Lett., Vol. 34, L19803, doi:10.1029/2007GL030135, 2007.

York, D., Evensen, N., Martinez, M., and Delgado, J.: Unified equations for the slope, intercept, and standard errors of the best straight line. Am. J. Phys., 72(3), 367–375, 2004.

1

2 Table 1. A summary of the measurements used in this study. Here are listed only
3 measurements between the study period 1.1.2007-31.12.2010.

Instrument	Measurement area/ Location	Measurement period	Measured parameters
Ozone Monitoring instrument OMI (satellite)	25.0-28.0S, 25.5-30.5E (whole study area)	Jan. 2007-Dec. 2010, obs. appr. once/day, only cloud-free obs.	NO_2 and SO_2 column densities, UV-B irradiance
Moderate Imaging Spectroradiometer MODIS (Aqua, satellite)	25.0-28.0S, 25.5-30.5E (whole study area)	Jan. 2007-Dec. 2010, obs. appr. once/day, only cloud-free obs.	Column integrated aerosol optical depth AOD at 550 nm wavelength
Cloud-Aerosol Lidar with Orthogonal Polarization CALIOP (satellite based lidar)	Selected locations within the study area	Selected days between Jan. 2007-Dec. 2010	Vertical profile of aerosol extinction at 532 nm waelength
Aerosol Robotic Network AERONET Sunphotometer (in situ)	Elandsfontein (26.25S, 29.42E)	Mar. – Dec. 2010, only cloudfree obs. during daylight.	Column integrated aerosol optical depth AOD at 500 nm wavelength.
Nephelometer (in situ)	Elandsfontein	Mar. – Dec. 2010	Aerosol scattering coefficient
Differential Mobility Particle Sizer DMPS (in situ)	Marikana (25.70S,27.48E) Botsalano (25.54S, 25.75E) Welgegund (26.57S, 26.94E)	Marikana: Feb 2008-May 2010 Botsalano: Jan. 2007-Feb. 2008 Welgegund: May-Dec. 2010	Particle size distribution, condensation sink, event classification
Scanning Mobility Particle Sizer SMPS (in situ)	Elandsfontein	Mar. – Dec. 2010	Particle size distribution, condensation sink
	All in situ stations	dates/station as above	NO_x, and NO, SO_2, global radiation, T, RH

Table 2. Correlations between in situ- and satellite-based proxies. The number of coincident observations is denoted with "N". Scatter plots for each of the case are provided as a supplementary material.

Station	$(NO_x\text{-}NO)/CS$ vs. NO_2/AOD	SO_2/CS vs. SO_2/AOD	$SO_2 \cdot UV\text{-}B/cs^2$ vs. $SO_2 \cdot UV\text{-}B/AOD^2$	$Glob./CS^2$ vs $UV\text{-}B/AOD^2$
Elandsfontein	R^2= 0.11, N=46	R^2= 0.20, N=41	R^2=0.13, N=39	R^2= 0.30, N=52
Marikana	R^2= 0.38, N=93	R^2= 0.005, N=76	R^2=0.13, N=76	R^2=0.22, N=117
Botsalano	R^2=0.004, N=16	R^2= 0.12, N=14	R^2=0.30, N=14	R^2=0.11, N=18

Figures

Figure 1. The study area and locations of the in situ measurement stations; BOT = Botsalano, MAR = Marikana, WEL = Welgegund, and ELA = Elandsfontein.

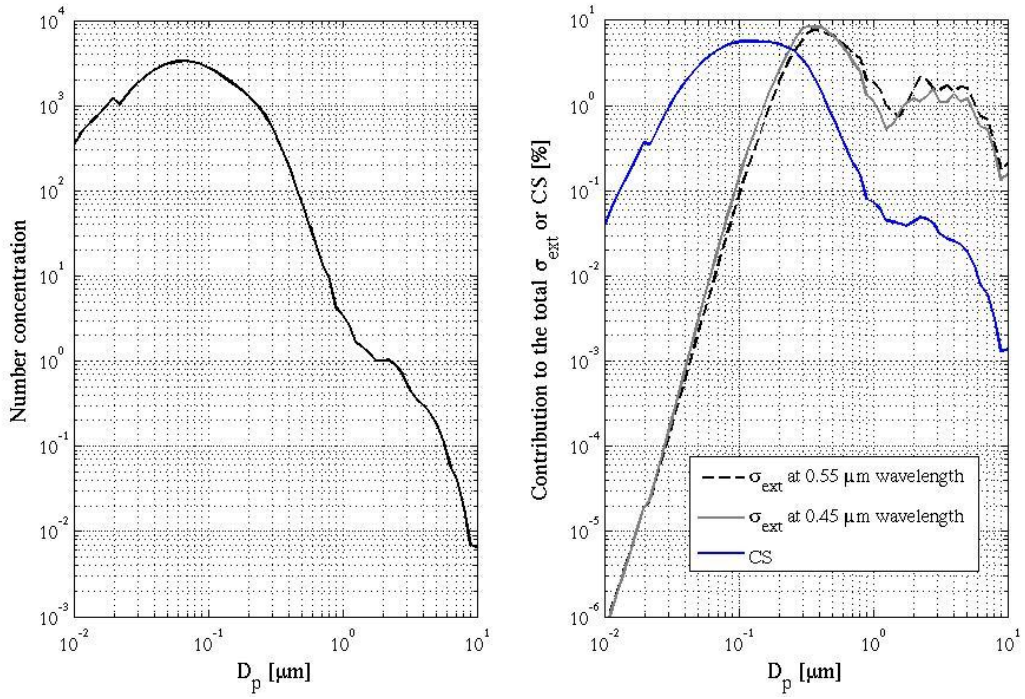

Figure 2. The sensitivity of CS and aerosol extinction coefficient to different particle sizes. In the left panel is shown the aerosol size distribution that is used to calculate CS and σ_{ext} is calculated for two wavelengths (0.55 and 0.45 μm) assuming spherical particles with a refractive index of m=1.48+0.001i. In the right panel is shown the contribution of each particle size to the total CS and σ_{ext}. The σ_{ext} is calculated for two wavelengths (0.55 and 0.45 μm) assuming spherical particles with a refractive index of m=1.48+0.001i.

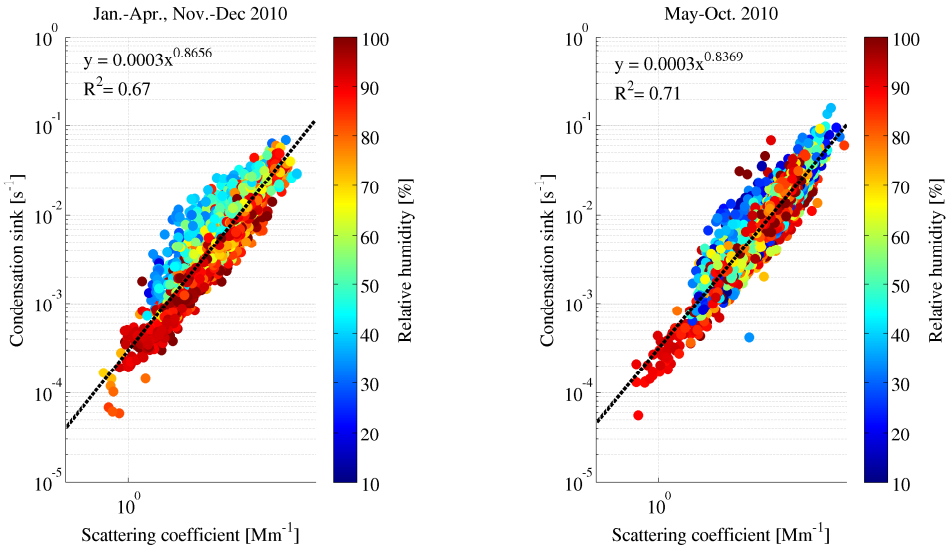

Figure 3. Comparison between condensation sinks derived from particle size distributions, as described in the text, and nephelometer scattering coefficients measured at Elandsfontein station in 2010 for the warm (Jan-Apr., Nov.-Dec), and the cold (May-Oct.) seasons. CS has been corrected to the ambient relative humidity but the scattering coefficient was measured from dry particles. The data are colour-coded according to ambient relative humidity (RH) and the strong influence of RH on the relation between CS and scattering coefficient is evident.

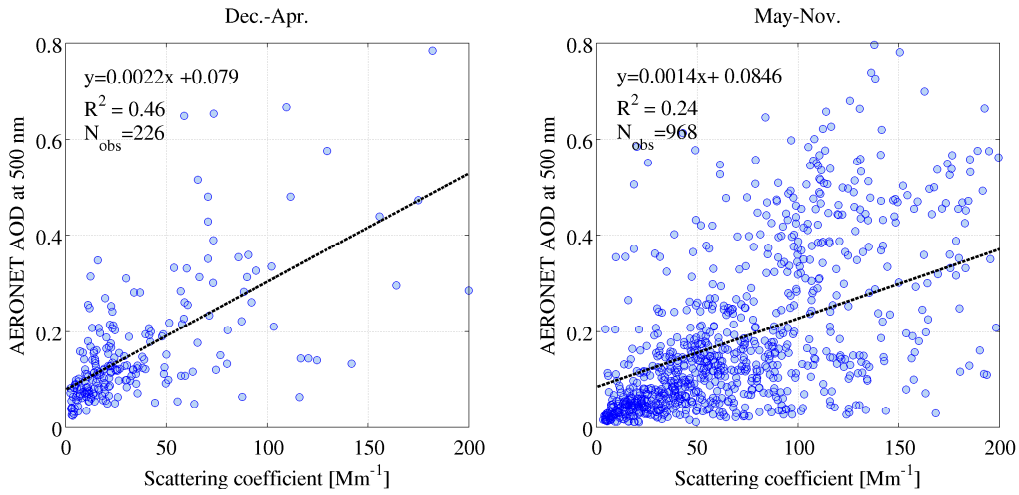

Figure 4. Comparison between AOD at 500nm available from AERONET (see text) and in situ scattering coefficients measured at the Elandsfontein station. The AOD is the column integrated value of aerosol extinction (scattering + absorption) obtained from sunphotometer measurements. The in situ scattering coefficient is measured with a nephelometer.

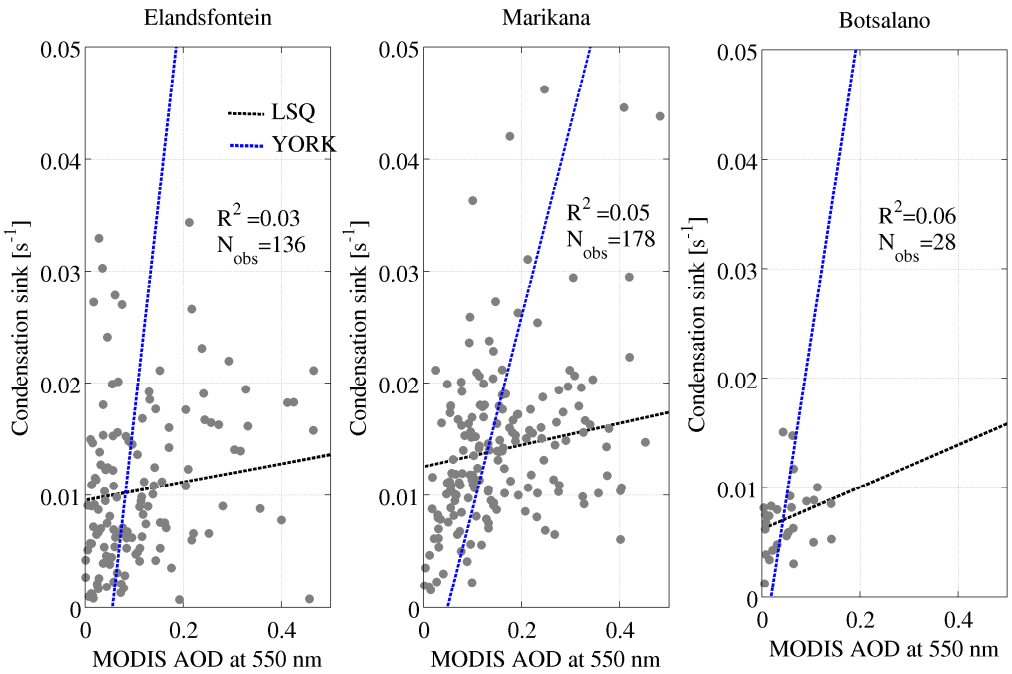

Figure 5. Comparison between MODIS AOD and in situ CS. The MODIS AOD values are spatial averages calculated from the observations within 3 km distance from the measurement station, whereas the CS values are one hour averages (13:00-14:00 LT). The black lines represent the slope from least squares linear fitting (LSQ). The blue lines represent the fitting method where the uncertainties related to CS and AOD values have been taken into account (YORK, York et al. 2004). The uncertainty for CS was set to 10 %, and for AOD to 0.05+15% (Levy et al., 2013).

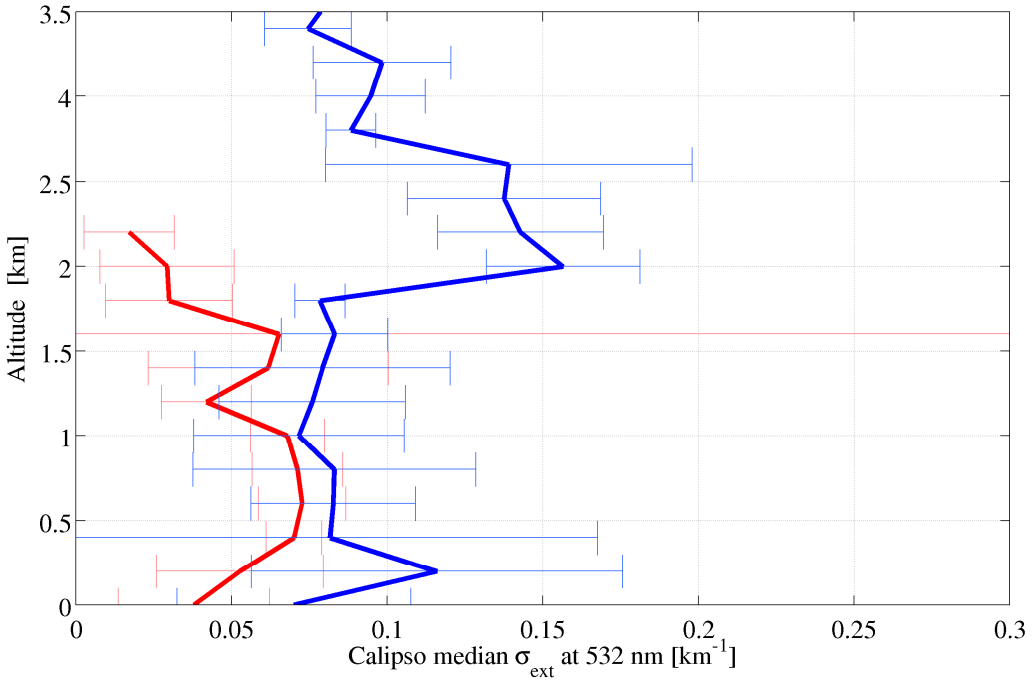

Figure 6. Median CALIPSO extinction profiles for days when MODIS AOD >0.15 (blue) and AOD≤0.15 (red) . The CALIPSO profiles are collected within 50 km radius from the Marikana station. The horizontal bars represent the interquartile ranges. The median extinction profile for MODIS AOD ≤0.15 cases extends only up to 2.2 km because the quality of the data above 2.2 km was too low.

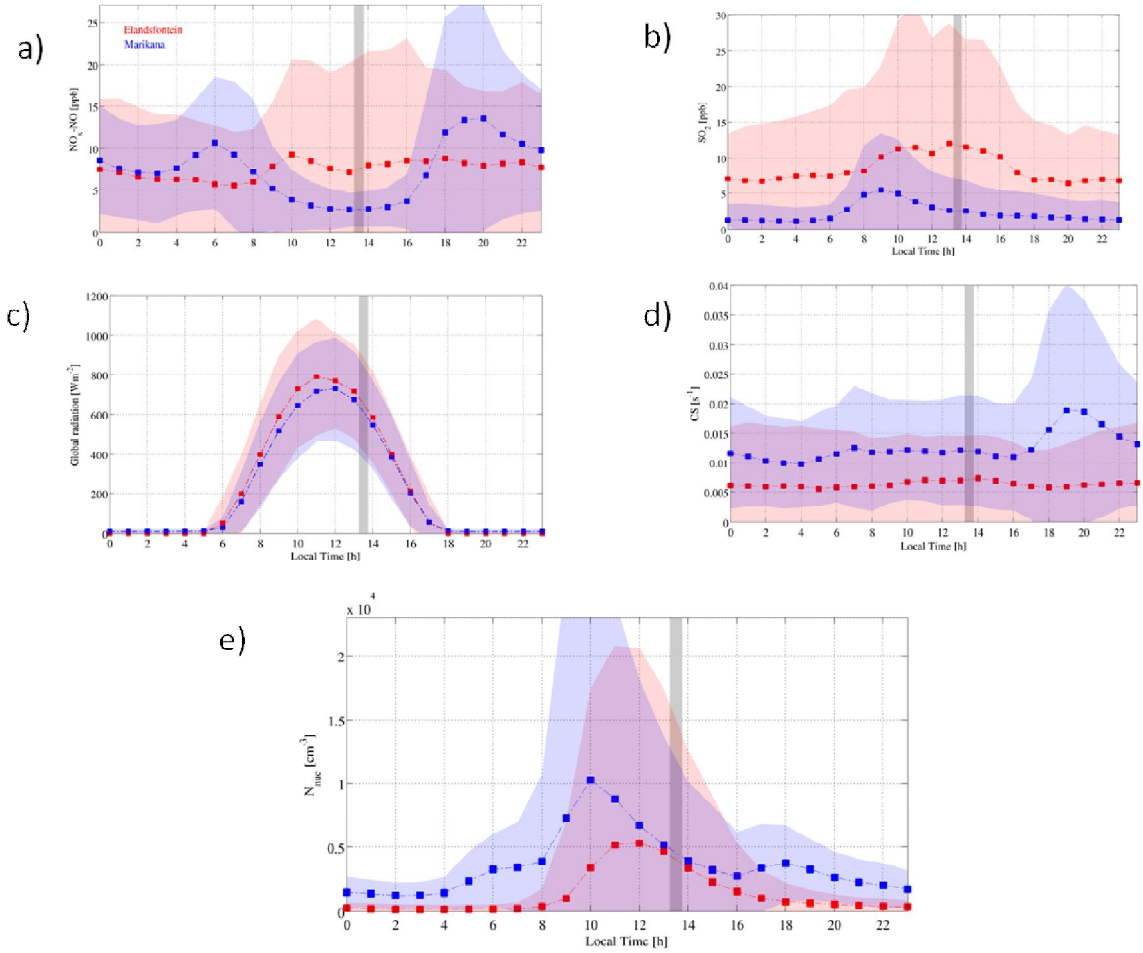

Figure 7. Diurnal variation of a) NO_x-NO, b) SO_2, c) global radiation, d) CS, and e) N_{nuc} at Elandsfontein (red) and Marikana (blue) stations. The grey columns represent the time window for the satellite overpass. The blue and red shading denote the 75th and 25th percentiles. It is noted that CS at Elandsfontein is defined with particles D_p<10µm, and at Marikana with particles D_p<1µm. N_{nuc} at Marikana represents particles D_p < 30 nm while at Elandsfontein N_{nuc} represents particles D_p 10-30nm.

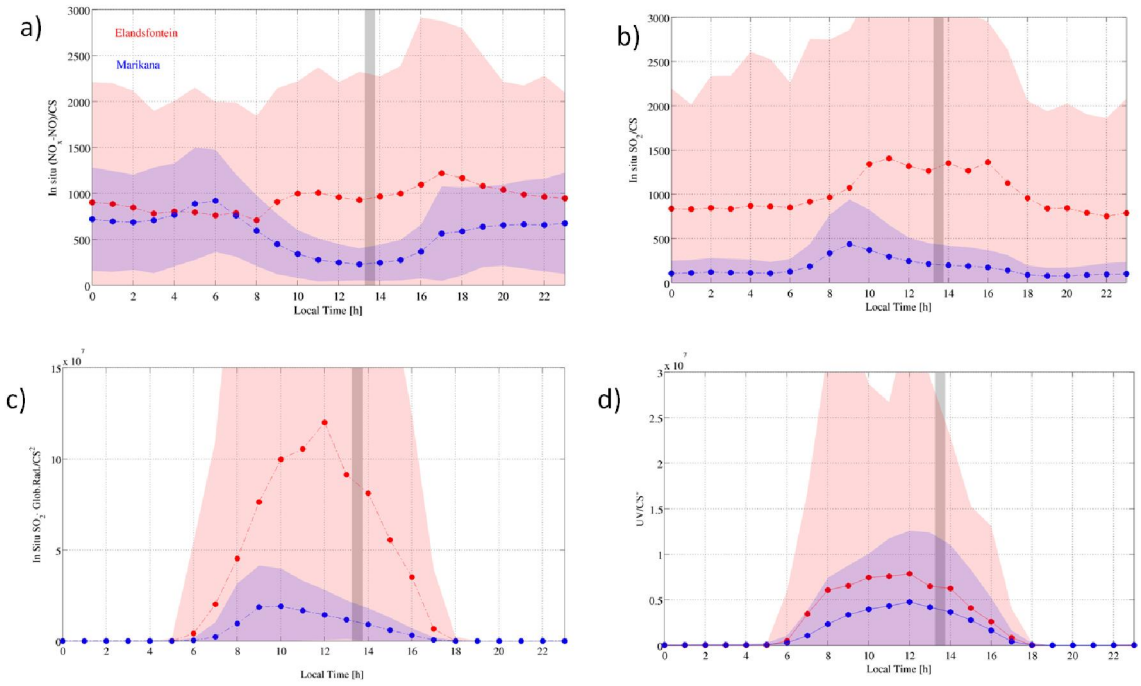

Fig 8. Diurnal variation of the proxies calculated using in situ data at Elandsfontein (red) and at Marikana (blue) stations. The red and blue shaded areas denote the 75th and 25th percentile ranges. The grey column represents the time of the satellite overpass.

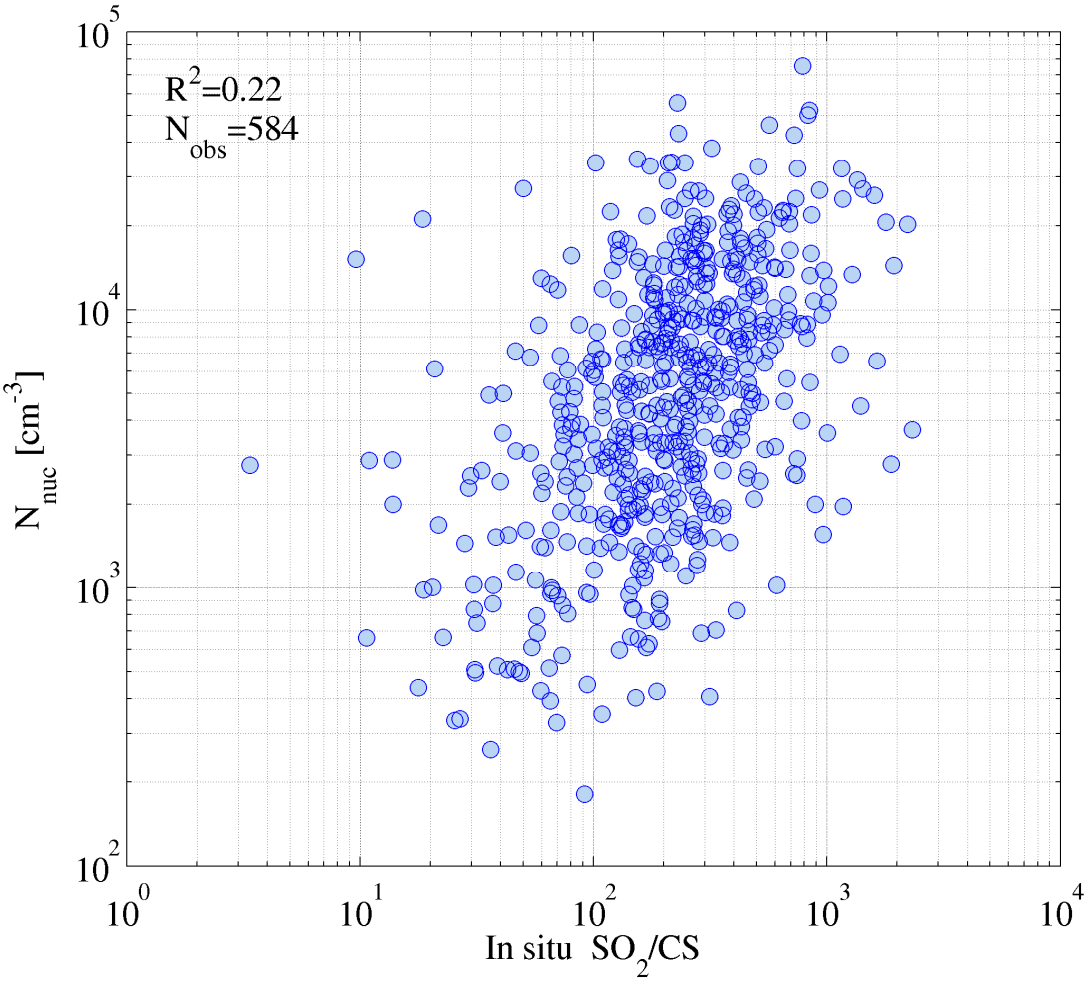

Fig. 9. Correlation between nucleation mode number concentration and SO_2/CS proxy calculated using in situ data at Marikana measurement station at the time of the satellite overpass (13-14 LT).

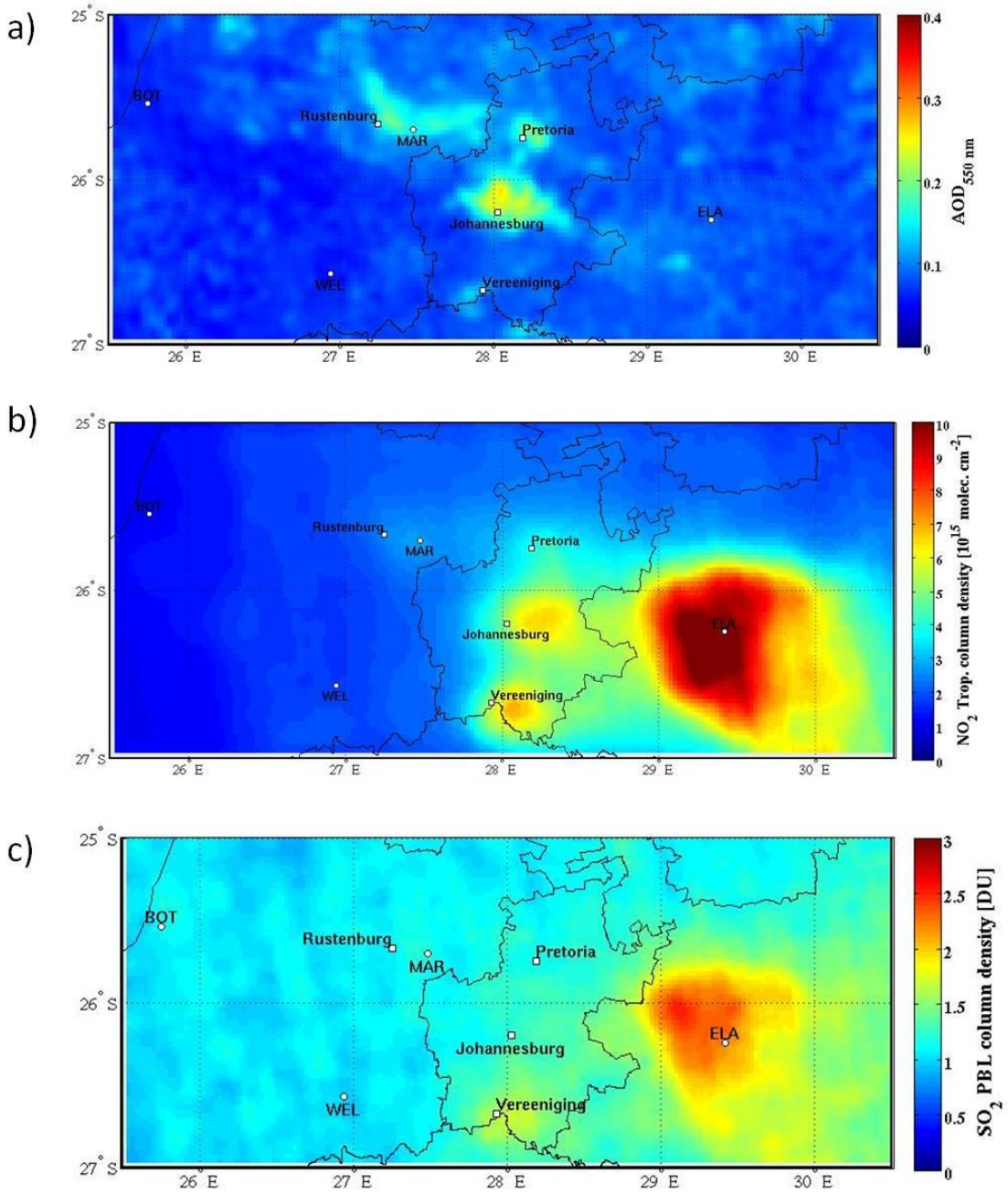

Figure 10. MODIS AOD (a) , OMI NO_2 (b) and SO_2 (c) column density medians for a four year period from Jan. 2007 to Dec. 2010. The locations of the in situ measurement stations (ELA= Elandsfontein, MAR=Marikana, BOT=Botsalano, and WEL=Welgegund) are marked with white dots.

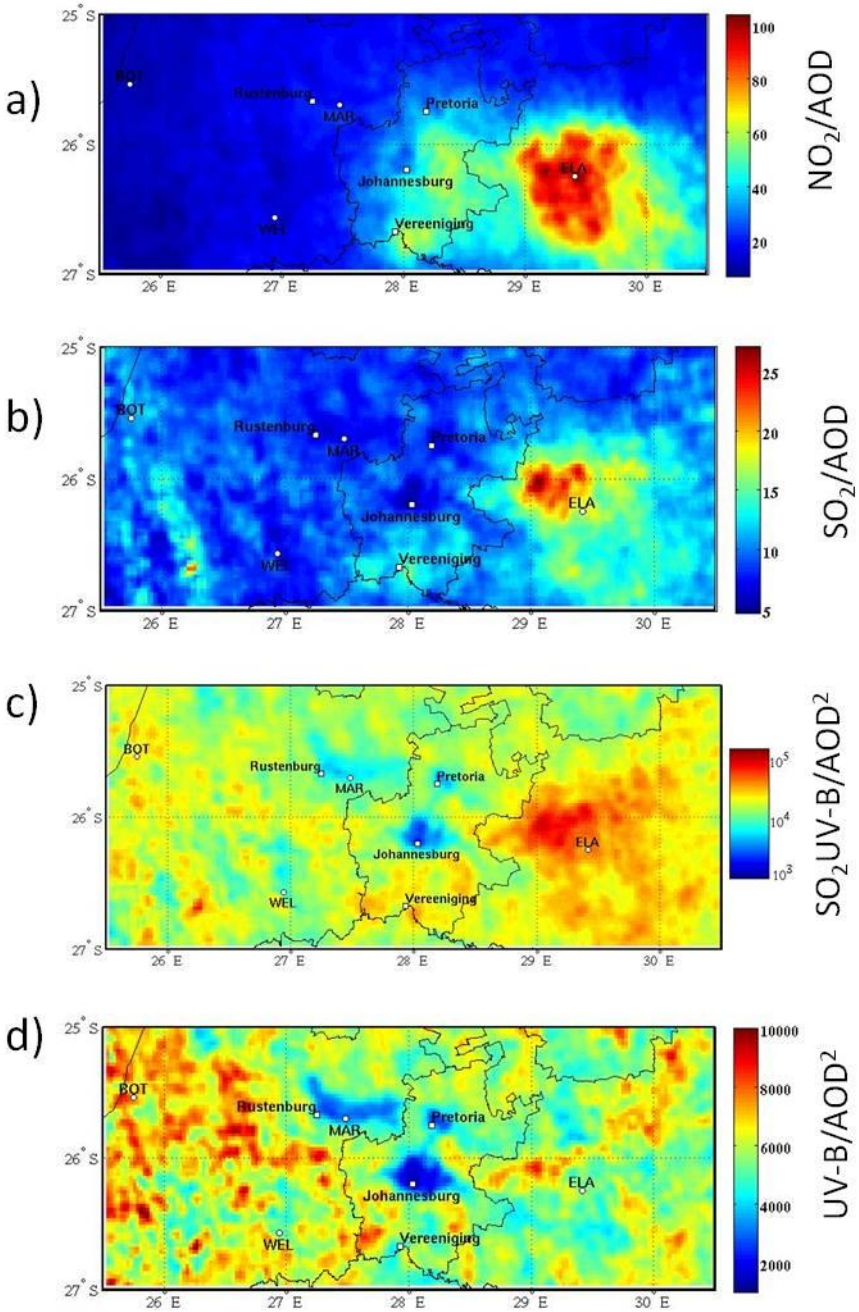

Fig. 11. Spatial pattern of proxy medians for 2007-2010 calculated using satellite data. The proxies are a) NO_2/AOD, b) SO_2/AOD, c) SO2·UV-B/AOD², and d) UV-B/AOD².

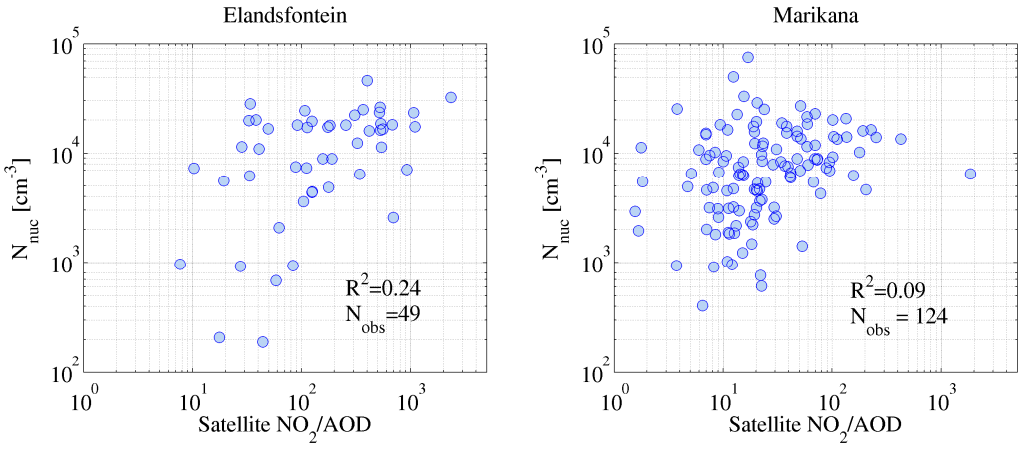

Figure 12. The comparison between the number concentration of nucleation mode particles and NO_2/AOD calculated from the satellite data at Marikana and at Elandsfontein stations. The number concentrations are one hour averages (13-14 LT) representative of the satellite overpass time. It is noted that at Elandsfontein Nnuc represents particles with D_p 10-30 nm, and at Marikana particles with D_p<30 nm. N_{obs} denotes the number of coincident observations.