

Air quality indicators from the Environmental Performance Index: potential use and limitations in South Africa

Rebecca M. Garland^{1,2}, Mogesh Naidoo¹, Bheki Sibiyi¹, and Riëtha Oosthuizen¹

¹Natural Resources and the Environment Unit, Council for Scientific and Industrial Research, Pretoria, South Africa

²Climatology Research Group, North West University, Potchefstroom, South Africa

Received: 12 January 2017 - Reviewed: 13 February 2017 - Accepted: 8 April 2017

Online early research article

Abstract

In responding to deteriorating air quality, many countries, including South Africa, have implemented national programmes that aim to manage and regulate ambient air quality, and the emissions of air pollutants. One aspect within these management strategies is effective communication to stakeholders, including the general public, with regard to the state and trend of ambient air quality in South Africa. Currently, information on ambient air quality is communicated through ambient mass concentration values, as well as number of exceedances of South African National Ambient Standards. However, these do not directly communicate the potential impact on human health and the ecosystem. To this end, the use of air quality indicators is seen as a potential way to achieve communication to stakeholders in a simplified, yet scientifically defensible manner. Air quality indicators and their source data from the Environmental Performance Index (EPI) were interrogated to understand their potential use in South Africa. An assessment of four air quality indicators, together with their source data, showed improvements in air quality over the time period studied, though the input data do have uncertainties. The source data for the PM indicators, which came from a global dataset, underestimated the annual PM_{2.5} concentrations in the Highveld Priority Area and Vaal Triangle Airshed Priority Area over the time period studied (2009-2014) by ~3.7 times. This highlights a key limitation of national-scale indicators and input data, that while the data used by the EPI are a well-thought out estimate of a country's air quality profile, they remain a generalised estimate. The assumptions and uncertainty inherent in such an ambitious global-wide attempt make the estimates inaccurate for countries without proper emissions tracking and accounting and few monitoring stations, such as South Africa. Thus, the inputs and resultant indicators should be used with caution until such a time that local and ground-truthed data and inputs can be utilised.

Keywords

Air quality indicators, air quality management

Introduction

Globally, air pollution is of concern and has deteriorated in many areas due to emissions from anthropogenic activities including industrial and vehicular activities, as well as from emissions from natural sources such as biomass burning. The current state of air quality has become a major threat to the health and wellbeing of people, as well as the environment in many areas around the world. In responding to deteriorating air quality, many countries, including South Africa, have implemented national programmes that aim to manage and regulate ambient air quality, and the emissions of air pollutants. In South Africa, the approach includes the declaration of priority areas for air quality management, development of national air quality standards, and an air quality monitoring programme. The aim of these programmes is to reduce air pollution-related illnesses and conditions.

One aspect within these management strategies is effective

communication to stakeholders, which includes the general public, with regard to the state and trend of ambient air quality in South Africa. Currently, information on ambient air quality is communicated through ambient mass concentration values, as well as number of exceedances of South African standards. However, these do not directly communicate the potential impact on human health and the ecosystem. To this end, the use of air quality indicators is seen as a potential way to achieve communication to stakeholders in a simplified, yet scientifically defensible manner.

Environmental Performance Indicator

The Environmental Performance Index (EPI) was developed by the Yale Centre for Environmental Law and Policy at Yale University and the Centre for International Earth Science Information Network at Columbia University (<http://epi.yale.edu>). The EPI aggregates over 20 indicators relating to national environmental data.

The EPI assesses two objectives, namely Environmental Health and Ecosystem Vitality. The Issue Category of “Air Quality” is within Environmental Health, though in previous EPI reports there have been Air Quality issues within the Ecosystem Vitality objective.

There have been a variety of indicators in the Air Quality issue category in the history of EPI, with recent years focussing on particulate matter (PM) and household air pollution. The EPI 2016 assessment includes an indicator on exposure to nitrogen dioxide (NO₂), and previous EPIs (e.g. 2008) have included indicators on ground-level ozone for health and for ecosystems considerations. For developing local indicators, this suite of present and historical EPI indicators should be assessed for their relevance to local air quality issues and policy priorities, and for the availability and reliability of local data.

For this study, the following indicators were selected to ground-truth air quality aspects for South Africa. This assessment is not comprehensive of all air quality indicators from the EPI, however focussed on selecting indicators that assessed different pollutants and data sources, as well as highlighting some of the pollutants and emission sources of concern in South Africa. These indicators provide information on the potential impact on human health and on ecosystems. In this analysis, only the following four indicators (that form part of the EPI) were considered. The objective that the indicator was included under in the EPI is listed in parenthesis below.

- HAP = Household Air Pollution (environmental health) - Percentage of population using solid fuel as the primary cooking fuel (%)
- SO₂CAP = Air pollution (ecosystem vitality) - Sulphur dioxide emissions per capita
- SO₂GDP = Air pollution (ecosystem vitality) - Sulphur dioxide emissions per Gross Domestic Product (GDP)
- PM_{2.5} = Air Pollution (environmental health) - Population weighted exposure to PM_{2.5} (µg/m³)

The SO₂CAP and SO₂GDP are not included in the current (2014 or 2016) estimations of EPI; they are nonetheless important indicators for South Africa and thus included here.

This study interrogated the input data into the EPI and compared this input data to publically available local data. This comparison will help to gain a better understanding of the robustness of the input data, and in turn, the indicators. These findings assisted in understanding the potential uses and limitations of the indicators, as well as provided insight into the state of air quality in the country.

Methods

EPI input data

EPI output values for indicators are reported on a national scale. The website does give links to the underlying data sources, and those sources are described here (<http://epi.yale.edu/>).

EPI: Solid fuel use for cooking data

The HAP data are derived from Bonjour et al. (2013), which were based upon data from the WHO Household Energy Database (2012). The EPI indicator is defined as the percentage of population using solid fuel for cooking. The percentage of population that are exposed to household air pollution was assumed to be the same as the percentage of households using solid fuels; thus the percentage of households using solid fuels is assessed and compared. These data are based on national surveys, which do report percentage of households using solid fuels for cooking.

EPI: SO₂ emissions

The SO₂GDP and SO₂CAP were last reported in EPI 2012, and those are the input data reported here. The SO₂ aspect is represented by total anthropogenic emissions for a country. The input SO₂ emissions were based on the research detailed by Smith et al. (2011), in which global bottom-up inventories (primarily through mass balance for combustion and metal smelting) were created for each country and constrained by any available locally derived emissions measurements or estimates. The original inventory covers years 1850-2005 and is reported in 10 year increments in Smith et al. (2011). The source sectors considered were coal combustion, petroleum combustion, natural gas processing and combustion, petroleum processing, biomass combustion, shipping bunker fuels, metal smelting, pulp and paper processing, other industrial processes, and agricultural waste burning.

EPI: Population and GDP data

Indicator SO₂GDP requires that GDP be converted to international dollars using purchasing power parity rates; for the 2012 EPI, 2005 international dollars were used. These were sourced from the World Development Indicators (indicator NY.GDP.MKTP.PP.KD; World Bank, 2011) and covered the period 1980-2011. SO₂CAP requires country population data and this was also sourced from the World Development Indicators (indicator SP.POP.TOTL; World Bank, 2011) and covered the period 1960-2010.

EPI: Ambient PM_{2.5} simulated concentrations

The PM_{2.5} EPI indicator quantifies the population weighted exposure to PM_{2.5} for the country. This indicator uses PM_{2.5} ambient concentrations that were originally estimated by Van Donkelaar et al. (2015) and are available online (ACAG, 2016). The datasets used here were the “All composition” satellite-derived PM_{2.5} at a relative humidity of 35% for a three-year running median.

The methodology used to estimate surface PM_{2.5} concentrations was included in the method used in estimating the Global Burden of Disease that is attributable to PM (Burnett et al., 2014; Brauer et al., 2012). The methods are not the same, however, as Brauer et al. (2012) did use multiple data sources, including ground-based data.

The methodology followed to estimate PM_{2.5} ambient concentrations is detailed in Van Donkelaar et al. (2010), Van

Donkelaar et al. (2015) and Boys et al. (2014). Briefly, aerosol optical depth (AOD) from the combination of Moderate Resolution Imaging Spectroradiometer (MODIS), Multiangle Imaging Spectroradiometer (MISR) and Sea-viewing wide field-of-view sensor (SeaWiFS) satellite instruments were used together with global chemistry transport model simulations using the Goddard Earth Observing System model with Chemistry (GEOS-Chem). Ground-level concentrations of $PM_{2.5}$ were estimated by developing an AOD conversion factor (accounting for aerosol size, aerosol type, diurnal variation, relative humidity and the vertical structure of aerosol extinction) based on GEOS-Chem simulations. The results were daily values coinciding with satellite overpass time; these were aggregated into three year moving median values. The median values were used to reduce the noise in the data from the satellite retrievals (Van Donkelaar et al., 2015).

Local data

This section details the local data sources that were compared to the “international” data from the EPI. As discussed in the section below, the national SO_2 and solid fuel use data for the EPI and the “local” data have similar sources.

Local: Solid fuel use for cooking data

The local data were provided by the South African Department of Environmental Affairs (DEA) and included information on the distribution (in percentage) of households that use domestic fuels (paraffin, wood, and coal) for household activities such as cooking, heating and lighting. These data were compiled from the 2014 General Household Survey data from Statistics South Africa (Stats SA) (Statistics South Africa, 2015).

Figure 1 displays the percentage of households using paraffin, wood or coal for cooking in South Africa for 2002–2014. For comparison to the EPI, the percentage of households using solid fuels for cooking was defined as those using coal and wood. It should be noted that EPI includes the burning of crop residues, dung and charcoal, which were not included here due to lack of local data.

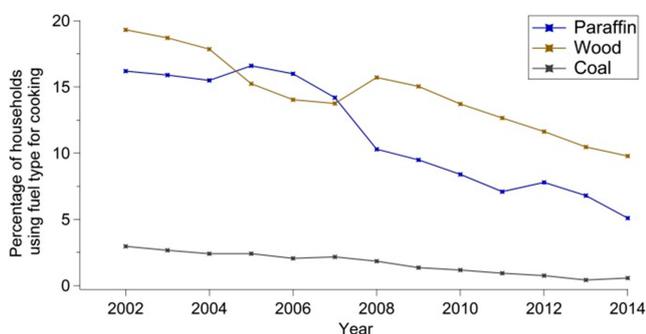


Figure 1: Stats SA data for percentage of households in South Africa using coal (grey line), wood (brown line) or paraffin (blue line) for cooking.

Local: SO_2 emissions

There are no locally derived data for a complete national SO_2 emission inventory. As detailed in section 2.1.2, the EPI used data from Smith et al. (2011). However, there is much room for

improvement regarding SO_2 emissions as there are high levels of uncertainty in Smith et al. (2011) estimates for South Africa. This is due to lack of local emissions reporting, and in uncertainty due to assumptions in bottom-up calculations such as fuel sulphur content and activity data (i.e. the actual amount of fuel used). Smith et al. (2011) specify uncertainties of up to 54% for South Africa (included in the “Other Countries” grouping for uncertainty analysis) for the sources included.

Klimont et al. (2013) built on the methodology and work by Smith et al. (2011) and estimated global SO_2 emissions through the Greenhouse Gas–Air Pollution Interactions and Synergies (GAINS; Amann et al. 2011) model for the period 2000–2011. This period is relevant to assessing a South African emissions profile due to rapid development. These published outputs are used in this comparison.

Klimont et al. (2013) also report a still significant amount of uncertainty may exist in this newer inventory; however, a quantitative estimate was not produced. The assumption around this large uncertainty was based on the inclusion of regional activity and fuel data from developing countries and within the international shipping sector. Ideally, a locally derived estimate, which includes local fuel specifications and activity data, must be provided to the general public such that researchers can include these data into their studies and indices.

Local: Population and GDP data

Economic and demographic data are readily available for most countries through either the World Bank or United Nations Populations Division. While it is possible to refine the World Bank Development Indicator population estimates using local census data, the difference is marginal for years 1996 (1.5% underestimate), 2001 (0.2% overestimate) and 2011 (0.4% underestimate) when compared to Stats SA census releases (Statistics South Africa, 2012). In order to have a full time series of population data, the EPI data source for population were used (World Bank, 2011).

In evaluating and comparing the SO_2 GDP indicator it is necessary to use the same units specified within the EPI methodology. The 2012 EPI methodology specifies GDP in 2005 international dollars. The only readily available data representing this are the World Development Indicators (same as used in EPI); these data could not be easily found from a local source directly. It is assumed here that these represent accurate local estimates of GDP and thus were used in calculating the local indicator (indicator SP.POP.TOTL; World Bank, 2011).

Ambient $PM_{2.5}$ monitored data

The South African Air Quality Information System (SAAQIS) data were used to ground truth the EPI $PM_{2.5}$ data to verify its level of accuracy. The SAAQIS data used were for the Vaal Triangle Airshed Priority Area (VTAPA) and DEA Highveld Priority Area (HPA) air quality monitoring networks. The VTAPA network has been running since 2007, and HPA network since 2008. The networks have a relatively continuous monitoring record, and

the measured data are quality controlled and managed by the South African Weather Service (SAWS) through the SAAQIS. The networks are in priority areas, and thus are in the more polluted areas of South Africa.

The PM_{2.5} data were provided by SAAQIS as one hour averages for 1 Jan 2008 – 1 October 2015. These data were quality checked (QC) by CSIR, and then averaged to monthly values, and processed to a corresponding three year moving median as reported in EPI data. The quality control included removing negative values and repeating values. Table 1 displays the number of data points (N) before the QC procedure was applied and after the QC procedure was applied. Further analyses were only performed on the valid hourly values (i.e. after the QC procedure was applied).

It is best practice when averaging monitored data to use a “data completeness threshold.” This threshold indicates the percentage of data that must be present in order to derive a representative average. For example, for a 70% data completeness rule, if fewer than 70% of 1-hour PM_{2.5} data were recorded in one day, then the daily average could not be calculated and would be left blank. In this analysis, thresholds from 75% to 50% were applied and tested to calculate a three year moving median. However the loss of data was large and analysis presented for this criterion would not have been possible at all sites. Thus, no threshold was applied when averaging the 1-hour values from SAAQIS for this analysis.

Table 1: Number of data points (N) per station before QC was applied and after QC was applied.

Site	Lat	Lon	N (before QC)	N (after QC)	
HPA	Ermelo	-26.4934°	29.9681°	55 949	55 949
	Hendrina	-26.1509°	29.7168°	39 611	39 592
	Middelburg	-25.7961°	29.4636°	53 438	53 419
	Secunda	-26.5486°	29.0801°	52 063	52 028
	Witbank	-25.8778°	29.1887°	48 789	48 789
VTAPA	Diepkloof	-26.2507°	27.9564°	42 939	42 737
	Kliprivier	-26.4203°	28.0849°	48 762	48 755
	Sebokeng	-26.5878°	27.8402°	41 739	41 725
	Sharpeville	-26.6898°	27.8678°	51 685	51 474
	Three Rivers	-26.6583°	27.9982°	45 852	45 803
	Zamdela	-26.8449°	27.8551°	57 731	57 580

Figure 2 shows the geographical location of the VTAPA and HPA monitoring stations (as points with names of stations indicated) on the 10 km x 10 km grid from the PM_{2.5} data used by the EPI. The PM_{2.5} 3-year running median from each site was compared

to the corresponding EPI value from the grid cell where the site is located.

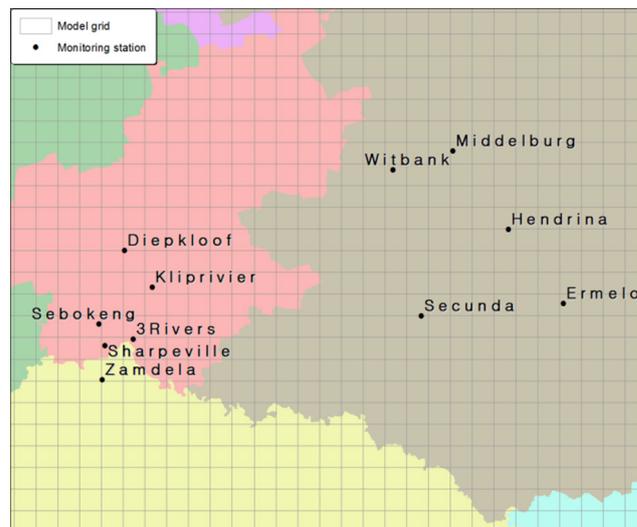


Figure 2: Map of SAAQIS monitoring stations as points with their names indicated on the 10km resolution model grid of the EPI PM_{2.5} input data. The colours indicate the provinces. The VTAPA monitoring stations are on the west and the HPA are on the east.

Results and Discussion

Overview of South Africa’s standing in the Air Quality Issue Category

The scoring of countries’ performance within the EPI is relative to the top-performing country, which receives a score of 100. The other countries’ indicator scores are normalised to this top-performer. Thus in the score, a larger number indicates better performance. The top-performing country also receives a rank of 1, thus in rank a lower score indicates better performance. Scoring and ranking occur at each level within the EPI (i.e. indicators, issue categories, objectives and total score).

As the EPI Issue Category Scores are normalised to the top-performing country, over time, a country can have a worsening score even if the air quality is improving, if other countries are improving at a faster rate.

In the 2016 EPI Air Quality issue category, South Africa’s score was reported as 88.84 (out of 100), which led to a ranking of 49 out of 180 countries.

South Africa was calculated to have improved its Air Quality issue category score over the past ten years from 74.47 (resulting in a rank of 61) in 2006. The rank of South Africa for PM_{2.5} was 60 in 2006, and fell to 69 in 2016. For the HAP indicators, the rank improved from 114 in 2006 to 94 in 2016.

Compared to sub-Saharan Africa’s performance as a region, South Africa’s 2016 Air Quality score was 18.98 points higher than the region’s average score. Thus according to EPI, in

general, South Africa for this issue category is performing well for the region, and has on average been improving.

HAP

Figure 3 displays the comparison of the EPI input data and local Stats SA data for the percentage of households using solid fuels for cooking in South Africa. The two datasets compare well, which was expected as the EPI input data source does rely on national surveys. Figure 3 also does highlight the large decrease in households who report using a solid fuel as their primary source for cooking, which can have positive implications for indoor and ambient air quality.

A limitation in these data is that the Stats SA data are limited to primary fuel only. In South Africa, low-income households rely on multiple fuels (e.g. Madubansi and Shackelton, 2007; Llyod et al., 2004; Thom, 2000; Davis 1998). A national survey found that 48% of South African households rely on multiple fuels for cooking (DoE, 2012). Thus, these data won't capture those who do use solid fuels but do not consider it their primary source, nor those who use multiple fuels.

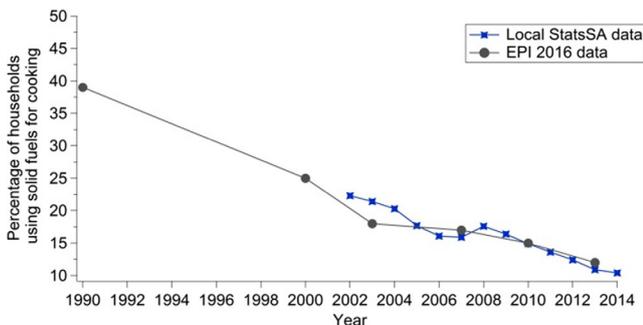


Figure 3: Percentage of households using solid fuel for cooking from the EPI 2016 input data (grey line) and local Stats SA data (blue line).

SO2CAP and SO2GDP

There are no local data to compare to the EPI 2012 SO₂ indicators. However, Klimont et al. (2013) has updated the emission estimates from the methodology used in the EPI 2012 indicator; those two datasets are explored here. The economic and population data used to calculate these indicators were the same for EPI 2012 and “Klimont” reported results (Figure 4).

As seen in Figure 4, anthropogenic SO₂ emissions from the datasets agree well, though the Smith et al. (2011) data had a larger dip in emissions in 2001-2002. The anthropogenic SO₂ emissions have been estimated to have increased from 1860 Gg in 1980 to 2795 Gg in 2011. However, during this time, SO2CAP and SO2GDP have both had an overall decreasing trend, though recently the SO2CAP has appeared to hover ~55 kg/person.

The SO2GDP indicator in particular highlights that South Africa’s SO₂ economic intensity has decreased, i.e. the GDP growth has been decoupled from the growth emissions in SO₂. This is a positive trend; however as SO₂ emissions are still increasing, there is a continuing negative impact on air quality. However, without a comprehensive national emissions inventory, it is not possible to validate this trend using bottom-up local data.

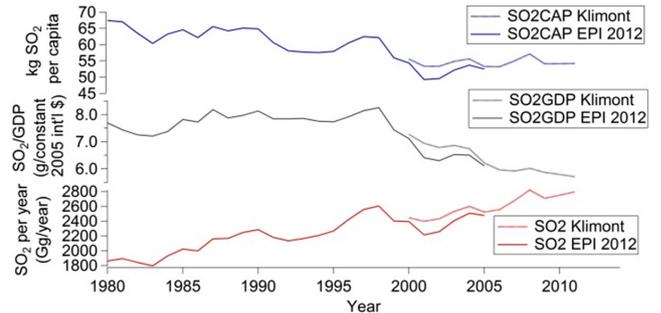


Figure 4: SO₂ emissions per year (red lines), SO₂ emissions per GDP (black lines) and SO₂ emissions per capita (blue lines)

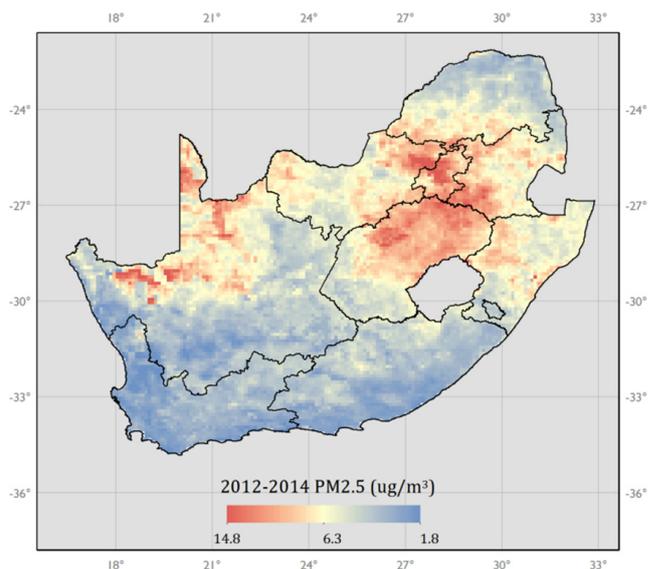


Figure 5: 2013 three year running median of surface PM_{2.5} mass concentrations that were used as input into the EPI.

PM25

Figure 5 displays the 2013 three-year running median of ground-level PM_{2.5} ambient mass concentrations that was used as input in the PM25 indicator. The running median will be reported by the midpoint year (i.e. 2012-2014 in Figure 5). In this study, the medians for 2010-2013 were compared for all sites. While the magnitude of the medians does change in these averaging periods, the general spatial distribution is similar to Figure 5 across years. This spatial distribution is what would be expected, with higher PM loadings in Gauteng, HPA, and VTAPA, where anthropogenic emissions of air pollution are high. In addition, peaks are seen in the Northern Cape, which in the input dataset are attributed to dust.

Figure 6 displays the average three-year running mean PM_{2.5} concentration of the HPA and the VTAPA stations’ for both observations (blue) and EPI database (red). Table 2 displays the PM_{2.5} mass concentrations for the EPI input dataset and from the monitored data from SAAQIS for each station within the HPA and VTAPA (labelled as “Monitored” data). The EPI uses a three-year running median of annual averages. The monitored data from SAAQIS did not have consistent data completeness at all sites across years. This inconsistent completeness may bias the median, as well as inter-annual comparisons. This could particularly have impact in areas that have a strong seasonal

Table 2: Ground-level PM_{2.5} concentrations (µg/m³) for EPI input data and monitored data from SAAQIS per site of the HPA and the VTAPA

Network	Site	Year (midpoint three-year running median)	EPI Input annual PM _{2.5} concentrations (µg/m ³) (Three-year running median)	Monitored annual PM _{2.5} concentrations (µg/m ³) (Three-year running median)	Monitored annual PM _{2.5} concentrations (µg/m ³) (Annual average)	Number of monthly values used in annual average	
HPA	Ermelo	2010	7.6	30.0	30.0	12	
		2011	7.1	29.2	29.2	12	
		2012	7.1	28.5	28.5	11	
		2013	7.4	24.6	24.6	12	
	Hendrina	2010	7.0	20.5	23.3	12	
		2011	7.0	18.8	18.8	11	
		2012	6.8	18.4	18.4	11	
		2013	6.6	18.4	12.5	6	
	Middelburg	2010	7.9	22.5	22.2	12	
		2011	7.5	23.8	26.0	12	
		2012	7.4	23.8	23.8	11	
		2013	7.4	19.2	19.2	12	
	Secunda	2010	10.2	41.0	40.2	12	
		2011	9.3	40.2	46.9	12	
		2012	8.9	34.7	34.7	11	
		2013	9.3	29.5	29.5	12	
	Witbank	2010	8.4	30.2	29.5	12	
		2011	7.8	29.5	30.2	12	
		2012	7.6	24.5	24.4	11	
		2013	7.3	24.3	24.5	12	
	VTAPA	Diepkloof	2010	10.4	46.8	64.6	7
			2011	10.3	29.7	29.7	9
			2012	9.7	27.0	27.0	11
			2013	9.7	23.0	23.0	12
Kliprivier		2010	8.6	52.4	52.4	8	
		2011	8.0	44.4	44.4	5	
		2012	7.4	37.8	37.8	11	
		2013	7.5	35.2	30.0	12	
Sebokeng		2010	8.2	51.5	51.5	12	
		2011	8.2	51.5	65.0	6	
		2012	7.6	33.9	33.9	11	
		2013	7.4	30.1	29.3	12	
Sharpeville		2010	9.8	43.8	48.0	9	
		2011	9.8	43.0	43.0	5	
		2012	9.0	40.7	40.7	11	
		2013	8.9	38.7	34.8	12	
Three Rivers		2010	9.6	27.5	32.2	12	
		2011	8.8	26.3	21.8	3	
		2012	8.7	24.3	26.3	11	
		2013	8.7	25.0	24.3	12	
Zamdela		2010	10.4	31.9	36.4	12	
		2011	9.8	31.9	31.9	10	
		2012	9.2	29.8	29.8	11	
		2013	9.2	29.8	29.0	11	

cycle, and thus a missing month(s) would strongly impact the annual average and thus the three-year running median. Thus, the annual average per year from the monitored data and the number of monthly values used to calculate each average are also shown in Table 2. While there are differences between the median and the mean values (e.g. 2010 for Diepkloof), it is clear from Table 2 that both values at all sites at all years are much

higher than the annual values in the EPI dataset.

Both Table 2 and Figure 6 highlight that the EPI input data underestimates the ground-level PM_{2.5} at all of these sites. On average, the monitored PM is ~3.7 times that used by the EPI, with a range of 2.4 to 6.3. Due to varying data completeness, comparisons between sites and between years were not made.

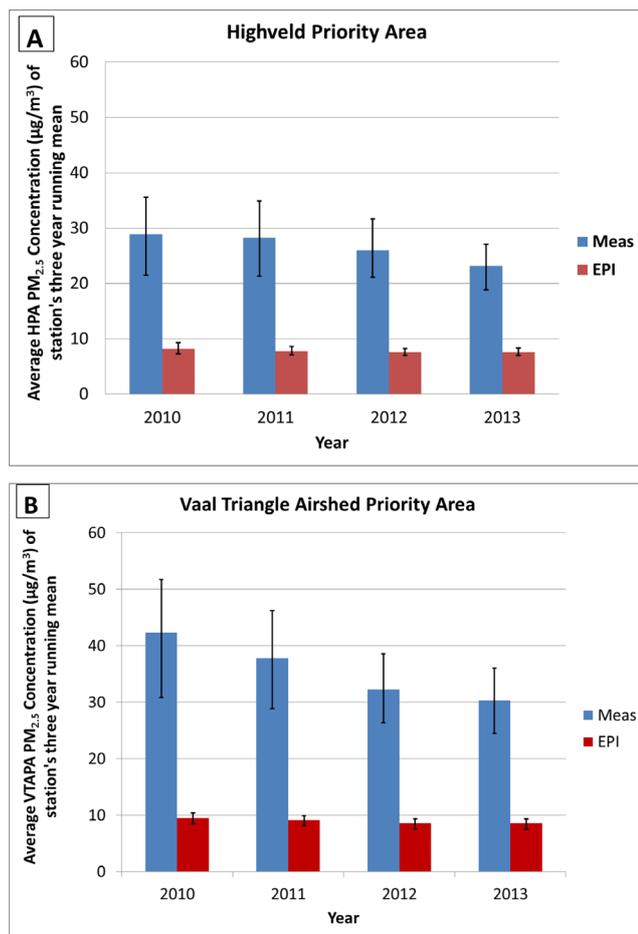


Figure 6: Average $\text{PM}_{2.5}$ concentration of (A) the Highveld Priority Area and (B) the Vaal Triangle Airshed Priority Area stations' three-year running means from observations (blue) and EPI database (red). The bars extend from 25th to 75th percentile values of the stations' averages.

At all sites, the EPI and the monitored data do show a decrease over the time period studied; however, due to the short length of the data set and the poor data recovery across years, the significance of this trend was not tested.

If it is assumed that this ~3 to 4 times underestimation of the EPI is valid for all of South Africa, then the resultant national $\text{PM}_{2.5}$ values become similar to Laos, which has a rank of 172 for $\text{PM}_{2.5}$ (out of 180). This assumption is a simplification; however, it provides a point of comparison of how such underestimation in $\text{PM}_{2.5}$ could impact South Africa's score and assessment of the quality of the air.

The method to derive the EPI input $\text{PM}_{2.5}$ data was also incorporated into the method used in the Global Burden of Disease (Burnett et al., 2014; Brauer et al., 2012). The methods are not the same, however, as Brauer et al. (2012) did use multiple data sources, including ground-based data. The underestimation of ground-level $\text{PM}_{2.5}$ does still occur in the Brauer et al. (2012) dataset (Supplementary Material Table S1). Thus, it is likely that the health impacts attributable to PM in such studies are underestimated; and further research is required to refine those estimates.

Conclusion

The EPI uses indicators at a national level, as it is a comparison of 180 countries across the globe. For air quality management in South Africa, such national indicators could also be useful to understand and communicate the national state of air quality. In order to identify hotspots, however, it would be necessary to spatially resolve indicators. In addition, as can be seen here, trend analyses of indicators over time are particularly useful to understand progress. Thus, for domestic purposes, indicators that can be spatially resolved and calculated over multiple years are ideal.

Potential uses and limitations of indicators in South Africa

For the indicators assessed here, the potential use and limitations are indicator-dependent.

HAP – This index has the strongest local data sources, and in addition, spatially and temporally resolved local data are available. As domestic burning impacts indoor and ambient air pollution, an indicator of this type can be useful in South Africa to track high-level progress of solid fuel use as a proxy for air pollution exposure. This analysis could be tailored to fuels and uses in South Africa (e.g. cooking and heating assessed separately). Local Stats SA data on this are available and could be used; however, in order to understand trends the same questions must be used across surveys or else Stats SA must “backcast” usage when the question changes. However, it must be noted that Stats SA data are limited to primary fuel only, and thus do not capture the fact that a large proportion of households, in particular low-income households, rely on multiple fuels. This use of multiple fuels should be included in a local indicator.

SO₂GDP and SO₂CAP – These provide a helpful perspective on the intensity of SO_2 emissions. However, there are no locally derived data for comprehensive national SO_2 emissions; thus there is a strong need for local, bottom-up estimates to understand how robust findings using international data are. The trend in SO_2 GDP looks promising; these numbers should be ground-truthed with known emission sources.

PM_{2.5} – There are no locally derived and validated products of $\text{PM}_{2.5}$ concentrations for South Africa with comprehensive spatial coverage that could be used to estimate this indicator. Thus, in order to quantify the national $\text{PM}_{2.5}$ indicator, global products for air quality would be needed together with local gridded data of population (such as in the EPI). However, from this study, it is clear that these underestimate $\text{PM}_{2.5}$ concentrations in the two priority areas.

It is not clear why there is this underestimation, though there may be many potential reasons for error. This comparison is comparing one sampling point to a grid cell, which assumes that the sampling point is representative of the full grid cell. The sampling stations have been sited to avoid strong local sources; however there would be spatial differences in the PM

concentrations across the grid cell. As emissions are not well-quantified on a national level for South Africa, there would be uncertainties in the emissions information used in GEOS-Chem simulations. In addition, the AOD-ground-level relationship is not well quantified for South Africa, and that may lead to uncertainties in deriving ground-based concentrations from satellite information (Hersey et al., 2015). Ford and Heald (2016) estimated an uncertainty of ~20% in deriving PM_{2.5} burden of mortality from satellite retrieved data due to uncertainties in the AOD-ground-level relationship alone. In addition, there are a lack of freely available and continuous ambient PM_{2.5} measurements in South Africa, that can be used for ground-truthing. Even this comparison is constrained to a few sites in heavily polluted areas in South Africa.

Since PM is a pollutant of concern in South Africa, indicators based on PM exposure are key to understanding and tracking air quality. Thus, there is a critical research need to develop input data for a national assessment, as well as at disaggregated spatial scales to identify hotspots and trends in such areas. This would need more continuous measurements of PM_{2.5} and modelling.

Data needs and recommendations

Basing indicators on locally measured and derived data is important. However, collecting and compiling local data for national indicators is not trivial. Bottom-up emission estimates need data and input from a variety of sources at a national scale. In addition, as can be seen here, an important analysis of such data and indicators are the assessment of their trends; this requires regular data collection and analysis for emissions estimates, and continuous monitoring for ambient concentration analysis. This can be resource intensive. However, without such data and analyses, it will not be possible to fully understand the state and trends of air quality and its impacts in South Africa. A starting point may be to focus on a small number of locally important indicators where local data are missing, and work to collect the necessary information for a first bottom-up estimate. Such estimates can then be compared to international estimates and data, which can help to identify missing sources (McLinden et al., 2016) and to decrease the uncertainties in both the local and international estimates.

It is recommended to focus on developing local information for a small number of indicators that are considered key for South Africa (e.g. SO₂ and PM_{2.5}). These indicators would be useful to South Africa and air quality management as they do present additional information than just ambient concentrations and exceedances. However, the strength of the indicator, and its trends, are in the underlying data.

Acknowledgements

This study was funded by the Department of Environmental Affairs (DEA) under the Rapid Response Research Component of the DEA-CSIR MOU. The authors would like to acknowledge Peter Lukey and Judy Beaumont from DEA for their assistance in

formulating the research question and approach. The authors would like to thank SAAQIS and SAWS for the air quality data.

References

- Amann, M., Bertok, I., Borcken-Kleefeld, J., Cofala, J., Heyes, C., Höglund-Isaksson, L., Klimont, Z., Nguyen, B., Posch, M., Rafaj, P., Sandler, R., Schöpp, W., Wagner, F. and Winiwarter, W. 2011, 'Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications', *Environmental Modelling & Software*, 26:1489-1501
- Atmospheric Composition Analysis Group (ACAG), 2016, Dalhousie University, Surface PM_{2.5}, http://fizz.phys.dal.ca/~atmos/martin/?page_id=140. Accessed February 2016.
- Bonjour, S., Adair-Rohani, H., Wolf, J., Bruce, N.G., Mehta, S., Prüss-Ustün, A., Lahiff, M., Rehfuess, E.A., Mishra, V. and Smith, K.R. 2013, 'Solid fuel use for household cooking: Country and regional estimates for 1980-2010', *Environmental Health Perspectives*, 121:784-790.
- Burnett, R.T., Pope III, C.A., Ezzati, M., Olives, C., Lim, S.S. et al. 2014, 'An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure' *Environmental Health Perspectives*, 112:391-403.
- Boys B.L., Martin, R.V., van Donkelaar, A., MacDonnell, R.J., Hsu, N.C., Cooper, M.J., Yantosca, R.M., Lu, Z., Streets, D.G., Zhang, Q. and Wang, S.W. 2014. 'Fifteen year global time series of satellite-derived fine particulate matter'. *Environmental Science and Technology*, 48:11109-11118.
- Brauer M., Amann, M., Burnett, R.T., Cohen, A., Dentener, F., Ezzati, M., Henderson, S.B., Krzyzanowski, M., Martin, R.V., Van Dingenen, R., van Donkelaar, A., and Thurston, G.D. 2012, 'Exposure assessment for estimation of the global burden of disease attributable to outdoor air pollution', *Environmental Science and Technology*, 46: 652-660.
- Center for International Earth Science Information Network - CIESIN - Columbia University. EPI 2014: Environmental Health Objective- Air Quality Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). <http://sedac.ciesin.columbia.edu/downloads/maps/epi/epi-environmental-performance-index-2014/epi2014-eh-air-quality.png>. Accessed February 2016.
- Emerson, J.W., Hsu, A., Levy, M.A., de Sherbinin, A., Mara, V., Esty, D.C. and Jaiteh, M. 2012. The 2012 Environmental Performance Index and Pilot Trend Environmental Performance Index. New Haven: Yale Center for Environmental Law and Policy.
- Hsu, A. et al. 2016, '2016 Environmental Performance Index'. New Haven, CT: Yale University. Available: www.epi.yale

Klimont, Z., Smith, S.J. and Cofala, J. 2013. The last decade of global anthropogenic sulfur dioxide: 2000–2011 emissions. *Environmental Research Letters*, 8-014003. 6pp

McLinden C.A., Fioletov, V., Shephard, M.W., Krotkov, N., Li, C., Martin, R.V., Moran, M.D., and Joiner, J. 2016, ‘Space-based detection of missing sulfur dioxide source of global air pollution’, *Nature Geoscience*, DOI: 10.1038/NGeo2724

Smith, S.J., van Aardenne, J., Klimont, Z., Andres, R.J., Volke, A. and Delgado Arias, S. 2011. ‘Anthropogenic sulfur dioxide emissions: 1850–2005’, *Atmospheric Chemistry and Physics*, 11:1101–1116.

Statistics South Africa. 2012. ‘Census 2011 – Census in brief’, *Public report number 03-01-41*. http://www.statssa.gov.za/census/census_2011/census_products/Census_2011_Census_in_brief.pdf (accessed 09/02/2016)

Statistics South Africa. 2015. ‘General household survey 2014’ report P0318.

Van Donkelaar, A., Martin, R.V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., and Villeneuve, P.J. 2010, ‘Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application’, *Environmental Health Perspectives*, 118:847–855.

Van Donkelaar, A., R. V. Martin, M. Brauer and B. L. Boys. 2015, Use of Satellite Observations for Long-Term Exposure Assessment of

Global Concentrations of Fine Particulate Matter’, *Environmental Health Perspectives*, doi: 10.1289/ehp.1408646

World Bank. 2011. ‘World Development Indicators’, International Bank for Reconstruction and Development/ The World Bank: Washington DC, USA. Data source: <http://users.cla.umn.edu/~erm/data/sr486/data/rawdata/wdi/csv/>

World Health Organization (WHO) 2012, ‘Household Energy Database http://www.who.int/indoorair/health_impacts/he_database/en/

Supplementary Material

Global gridded pollution and population estimates used in the Global Burden of Disease 2013 are available as Supporting Information from Brauer et al. (2012). These data were downloaded (Supporting Information 005) and the PM_{2.5} concentrations from the VTAPA and HPA sites were extracted for 2005, 2010, 2012 and 2013.

Table S1 below highlights the annual average PM_{2.5} concentrations from this database (significant figures as reported in database). The 2005 concentrations were used in Burnett et al. (2014).

Table S1: : Ground-level annual averaged PM_{2.5} concentrations (µg/m³) from Brauer et al. (2012). Significant figures are as reported in database.

		Annual average PM _{2.5} concentrations (µg/m ³)			
	Site	2005	2010	2012	2013
HPA	Ermelo	12.91724	12.81185	12.39297092	12.18869538
	Hendrina	11.9542	12.44675	12.4082598	12.38905936
	Middelburg	12.80226	13.46134	13.17713612	13.03729226
	Secunda	14.44574	14.57748	14.65277698	14.6905712
	Witbank	14.13451	15.9503	15.78367744	15.70102002
VTAPA	Diepkloof	24.72513	26.35033	26.69545272	26.86970548
	Kliprivier	15.9039	16.79124	16.38826719	16.19042202
	Sebokeng	16.0523	17.43382	17.58054743	17.65437359
	Sharpeville	19.50312	21.43347	22.18445547	22.56975871
	Three Rivers	19.45811	21.39796	21.41572369	21.42461106
	Zamdela	14.60568	15.44814	14.81980258	14.51528373