

Analysis of Aggregated Solar PV Output in South Africa

Jyotsna Singh, Bhishma Tyagi, Lucky Ntsangwane, and Joel Botai

Abstract—Present work deals with the analysis of aggregated solar photovoltaic plants output (ASPV_op in kWh) (South Africa) for the year 2015. The box plot of monthly ASPV_op has shown lower values in winter and higher values in spring and summer. The histograms have also been plotted to understand the variation of ASPV_op in different seasons. The histograms of ASPV_op have confirmed that it is not normally distributed. Monthly variation of ASPV_op was similar to monthly surface incident shortwave flux (SISW in Wm^{-2}). Morlet wavelet has been used for the investigation ASPV_op. The wavelet power spectrum shows significant (at 5 % level) power concentration between 2-16 days. Wavelet analyses of ASPV_op suggested that days of mid-summer to early autumn have significant power concentrated from 8 to 16 days band. In other seasons power is concentrated almost less than 8 days. Almost all days of spring have shown significant power concentration (< 8days).

Index Terms—Surface incident shortwave flux, photovoltaic output, wavelet transform, Morlet wavelet, South Africa.

I. INTRODUCTION

Eskom fulfils approximately 90% of South Africa's electricity demand and rest is fulfilled by the independent power producers (IPPs) and municipalities [1]. Even after this energy crisis is evident in this country. Eskom is facing challenges to meet the energy demand of this country [2]. People are facing several problems because of the load shedding. South African government have shown concern about the issue of the energy crisis and introduced many plans and policies to deal with this problem (Integrated Energy Plan, Renewable Energy Independent Power Producer Procurement Programme, Integrated Resource Plan and Independent System and Market Operator). South Africa is among the top three countries in the world with 2500 hours of the sunshine per year and direct solar radiation ranged between 4.5 and 6.5 kWhm⁻² per day [3].

Use of solar energy has now become the epitome of sustainable development. Solar energy is not only the clean energy source but is also abundant. Now it is a big challenge for us to use our scientific and technical skills to harness this free form of energy. There are two major categories of solar technology – solar photovoltaic (SPV) modules and concentrating solar power [4].

Photovoltaic (PV) cells are made up of semiconducting materials that convert sunlight into electricity (photovoltaic

effect). The output of PV plants depends on various weather factors like incoming solar radiation, temperature, humidity and wind speed, and environmental factors. There is a possibility to develop a model for forecasting solar photovoltaic output (SPV_op) using the output of PV plants and weather variables as mentioned above. The analysis of SPV_op data is very helpful for understanding SPV_op. The solar energy potential in South Africa is very high [5]. There are nine provinces and in many provinces, the Solar PV power plants have been installed (mainly in Northern Cape). The detailed study of the SPV_op help us to assess the efficiency of SPV plant. Through the analysis of SPV_op data we could also get to know in which season output is more and in which season it is less. This information is useful for SPV plant owners to take decision about the SPV plant.

To understand ASPV_op data we have used descriptive statistics. In the present work, we have also used wavelet transform. This is widely used in signal processing. The concept of 'wavelet' was first introduced in 1982 by Jean Morlet a French geophysicist [6]. To extract the more information from the data, the signal processing methods have been widely used in environmental studies [7]. The paper is divided into following sections: Introduction (Section I), Study region and datasets (Section II), Methodology (Section III), Results and Discussion (Section IV) and Conclusion (Section V).

II. STUDY REGION AND DATASETS

The present study has been carried for ASPV_op of Eskom operated and maintained SPV plants in South Africa. The geographic location of South Africa (22°-34°S Latitude and 16°-32°E Longitude) is very suitable for SPV plants installation. We have obtained hourly ASPV_op data for South Africa from Eskom for the year 2015 (kWh). To understand the variation of incoming solar radiation in 2015, we have also used Modern Era Retrospective Analysis for Research and Applications (MERRA 2D) monthly SISW. The spatial and temporal resolution of this data is 2/3° x 1/2° and monthly, respectively. More detail about this data can be found in the given website [8].

III. METHODOLOGY

We have studied the diurnal variation of ASPV_op on annual and seasonal scales. We have divided our data in four seasons-summer (December, January and February), autumn (March, April and May), winter (June, July and August), and spring (September, October and November). The variation of ASPV_op in each month has been analysed using box plots. We were also interested to see the probability distribution of ASPV_op. Histograms and cumulative distribution function

Manuscript received June 30, 2016; revised December 13, 2016. This work was supported by South African Weather Service.

Jyotsna Singh, Lucky Ntsangwane, Joel Botai are with the South African Weather Service, 442 Rigel Avenue South, Erasmusrand, Pretoria, South Africa-0181 (e-mail: jsinghenv@gmail.com, Lucky.Ntsangwane@weathersa.co.za, Joel.Botai@weathersa.co.za).

Bhishma Tyagi is with National Institute of Technology, Rourkela, Odisha, India (e-mail: bhishmatyagi@gmail.com).

(CDF) have also been analysed on the annual and seasonal scale. We have used wavelet transform to analyse ASPV_op information in both time and frequency domain. Wavelet is a powerful tool and has many applications (Signal Processing, Data Compression, Fingerprinting verification, DNA analysis, Speech recognition and others). With the use of wavelet technique, we can be able to detect localized (spatial and temporal) variations in the time series data [9]. The main strength of wavelet analysis as mentioned above is that it can decompose time series in both time and frequency domain whereas Fourier transform fails because of the uncertainty principle.

Wavelet transform method also does not depend on the condition of stationarity of time series. Under this analysis, the signals/time series $x(t)$ are decomposed over dilated and translated functions known as “mother wavelet” $\psi(t)$, given as

$$W_x(a, \tau) = 1/\sqrt{a} \int_{-\infty}^{\infty} x(t)\psi^*((t-\tau)/a)dt \quad (1)$$

where τ is time position, a is the scale of the wavelet, ‘*’ is the complex conjugate form.

There are different mother wavelet like – Morlet, Paul and DOG (Derivative of a Gaussian). In wavelet analysis the shape of wavelet never changes, only the size scale changes with the size of window. Present time series data is continuous so we have used Morlet wavelet function (Eq. 2)

$$\psi(t) = \pi^{-1/4} \exp(-i2\omega_0 t) \exp(-t^2/2) \quad (2)$$

ω_0 is the non-dimensional frequency, in present study we have taken $\omega_0=6$ to satisfy the admissibility condition. The scaled wavelet can be given as (Eq. 3)

$$\psi[(n' - n)\delta t / s] = (\delta t / s)^{1/2} \psi_0[(n' - n)\delta t / s] \quad (3)$$

In the present work we have used the daily data for the year 2015. The daily values have been obtained from the mean of hourly data. The scale of the wavelet is given as the fractional power of two [9].

$\delta t = 1$ day, $\delta j = 0.25$, $s_0 = 2\delta t$, $\omega_0 = 6$ for Morlet wavelet.

Where s the dilation parameter for changing scale is, n is the translation parameter to slide in time [10]. To test the significance for wavelet spectra we have used red noise ($\alpha = 0.72$) as a background spectra. Other technical detail about wavelet is given in [9].

We have also calculated the coefficient of determination (R^2) between monthly ASPV_op (kWh) and monthly SISW (Wm^{-2}). For this, we have calculated the monthly ASPV_op (Wm^{-2}) from the daily average of the hours from 6hrs to 17hrs.

IV. RESULTS AND DISCUSSIONS

A. Histogram and CDF

The histograms of ASPV_op have been shown in Fig. 1. From the histogram of ASPV_op, it is clear that the power output from the PV plant on annual and four seasons is not

normally distributed. It indicates ASPV_op is skewed and it could be due to the weather variability and extremes. It has raised the need to calculate empirical CDF. The empirical CDF and normal CDF for ASPV_OP are shown in Fig. 1. The comparison of empirical CDF and normal CDF for ASPV_op indicated that normal distribution underestimated the probability of smaller and higher ASPV_op and overestimated the ASPV_op of intermittent values. In Fig. 2, we have shown the diurnal variation of ASPV_op on annual and in different seasons (summer, autumn, winter and spring).

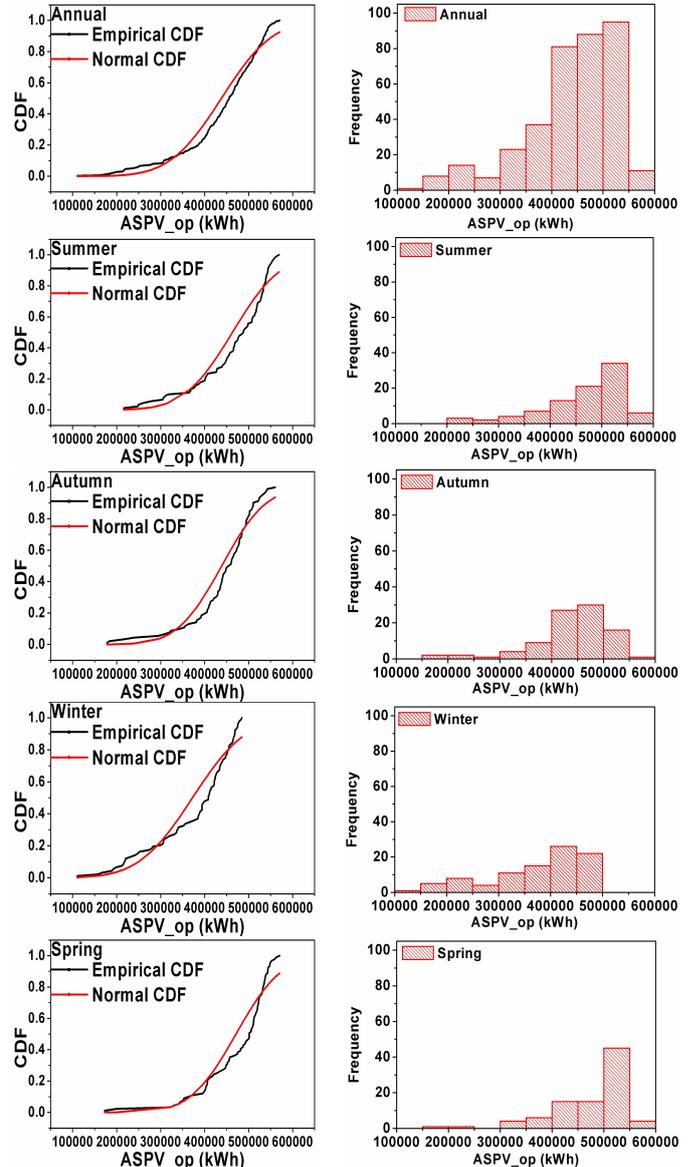


Fig. 1. Normal and empirical cumulative distribution function (CDF) and histograms of the aggregated solar PV output (ASPV_op) on annual and in the different seasons of South Africa.

Values of ASPV_op in summer and spring seasons have almost matched. We can see the higher values of ASPV_op from 10 hrs to 15 hrs. ASPV_op has shown maximum values in summer and spring followed by autumn and lowest values in winter. The more values of ASPV_op in spring season indicates that there might be the clearer sky condition during this month. It is causing the availability of more solar radiation for electricity generation from the solar panels. Mean and standard deviation (SD) of ASPV_op are presented in Table I. The mean value of ASPV_op is more in spring

474583.5 kWh) than summer (466123.4 kWh). Even SD is less in spring compared to summer.

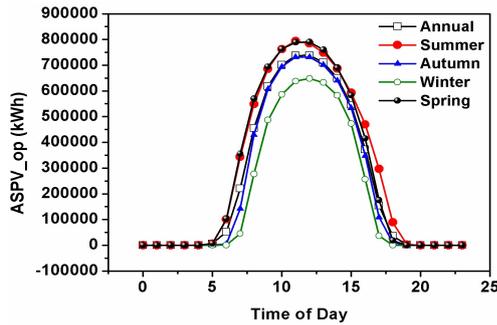


Fig. 2. Diurnal Variation of the aggregated solar PV output (ASPV_op) on annual and in the different seasons of South Africa.

TABLE I: MEAN AND STANDARD DEVIATION (SD) OF AGGREGATED SOLAR PV OUTPUT (ASPV_op) ON ANNUAL AND IN THE DIFFERENT SEASONS OF SOUTH AFRICA

	N	Mean (kWh)	SD (kWh)
Annual	365	439190.4	90340.9
Summer	90	466123.4	82827.86
Autumn	92	441718	75076.06
Winter	92	375306.9	92363.41
Spring	91	474583.5	75692.06

*N=Number of Data Points

The box plots of ASPV_op have been shown in Fig. 3. The box plot gives the proper overview of the data distribution. The box plot for each month for ASPV_op shows the values and variability of output of PV plant in each month. The red, blue, green and black boxes represent summer (December, January, and February), autumn (March, April, and May), winter (June, July, and August) and spring (September, October, and November), respectively. Boxes denote 1st and 3rd quartiles. Line inside the box represents the median value of ASPV_op. Whiskers represent highest and lowest values within 1.5 times the interquartile range. Data beyond this range (outliers) represented by stars. The length of whiskers represents the variability of ASPV_op. These box plots show remarkable variation among the ASPV_op values. The length of the box is a measure of the spread of the data. From the box plots of ASPV_op it is clear that in almost all the months ASPV_op data are skewed (Fig. 3). Monthly ASPV_op have shown similar seasonal changes like SISW (southern hemisphere), minimum in winter and maximum in summer (Fig. 4). The length of the box in the month of August is bigger than all the other months. It shows more variation in PV output in this month. It might be due to variation in weather parameters.

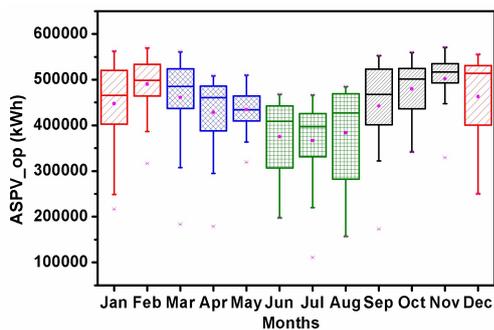


Fig. 3. Box plot of the aggregated solar PV output (ASPV_op) in different months.

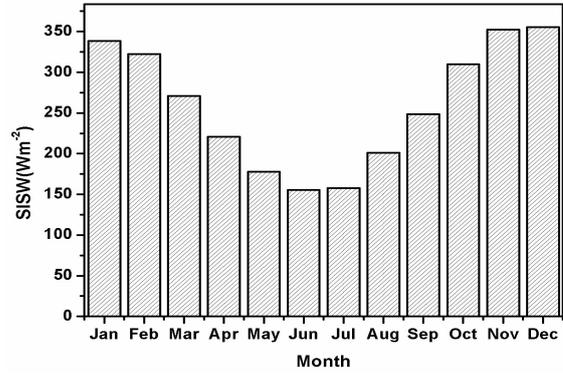


Fig. 4. Monthly mean of the surface incident shortwave flux (SISW) for the year 2015 over South Africa.

B. Wavelet Power Spectrum

The energy in a time series is preserved in its transform coefficients. It leads to the decomposition of the sample variance for a time series into components that can be associated with different scales. With multiresolution technique, we are able to identify local features and correlation properties in the time series [11]. Fig. 5a shows the time series of ASPV_op. The wavelet power spectrum using Morlet wavelet for a Gaussian red noise, normalized by standard deviation has been shown in Fig. 5b. The zero padding has been used in plotting wavelet power spectrum, and hence the area under the curve (marked by a solid black line) will have effects of this zero padding. This curve is known as Cone of Influence (COI). It is the region, below which we have to neglect our results because of error induced due to the edge effect [9]. The color bar indicates the coefficients of wavelet transform at particular frequency and time location, respectively.

From the Fig. 5b, it is clear that more concentration of power is between 2 to 16 day. In the upper portion of graphs, blue shades represents high frequencies. In the region of high concentration of power (2-16 day) the yellow color contours have shown the mixed behavior. Some of them are horizontally wide and some of them are vertically wide. ASPV_op in the late autumn to late winter (120-180 days) have not shown any contours. From November to mid of December we can see the significant ASPV_op (at 5 % level). During this time-frequency resolution is also higher. They are represented with the yellow contours on the upper portions of the graphs. However, the February and early March, ASPV_op is in the upper region, corresponding to a period between 8 to 16 days. In other months contours are located less than the 8- day period. The colour towards orange indicates frequencies with high energies. The scale of 2 to 16 days has not shown any significant contour from May to June. Spring and summer days have shown the higher energy of the high-frequency signal components in scale from 2-16 days. During winter, high-frequency components were missing. It might be due to special weather conditions that persist during this season. Fig. 5c shows the unbiased estimation of the true power spectrum of ASPV_op time series (global wavelet spectrum). Global wavelet spectra describe ASPV_op variability. It is useful for summarising region's temporal variability. Global wavelet spectrum is showing that dominant scales are between 2 and 8, as they are crossing the dashed line of 5 % level (Fig. 5c). The majority of contours fall under

this band. The variance of significant power zone (2 to 16 days) has been shown in Fig. 5d. This figure is made by the average of Fig. 5b over all scales between 2 and 16 days. It

gives a measure of average day variance versus time. Here, we can see in most of the cases ASPV_op have shown the significant peaks at 5% level.

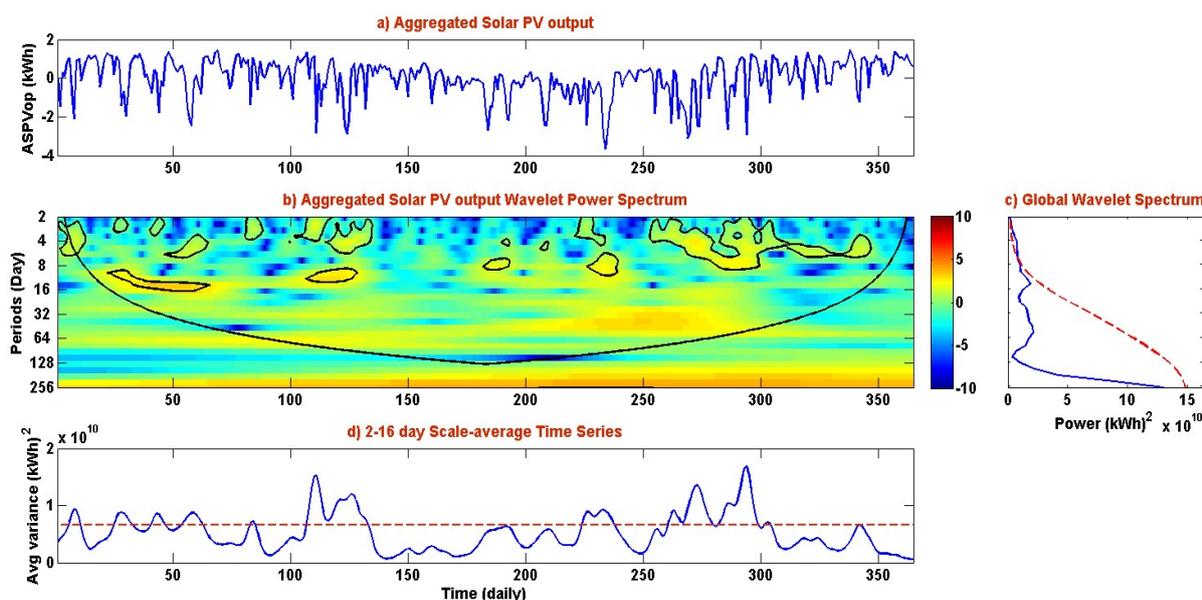


Fig. 5. (a) Time series of aggregated solar photovoltaic data (ASPV_op) normalized by standard deviation. (b) The wavelet power spectrum. The power has been scaled by the global wavelet spectrum (at right). The region below the curve is the cone of influence, where zero padding has reduced the variance. Black contour is the 5% significance level, using a red-noise (autoregressive lag1) background spectrum. (c) The global wavelet power spectrum (continuous line). The dashed line is the significance for the global wavelet spectrum, assuming the same significance level and background spectrum as in (b), and (d) The scale-average wavelet power spectrum over the 2 to 16 day for ASPV_op (continuous line). The dashed line is the 95% confidence level assuming red noise ($\tau = 0.72$).

C. Correlation of ASPV_op and SISW

Fig. 6 shows the scatter plot of ASPV_op and SISW. The correlation was very high $R^2 = 0.94$. This result is critical as we can use this information for forecasting the SPV_op.

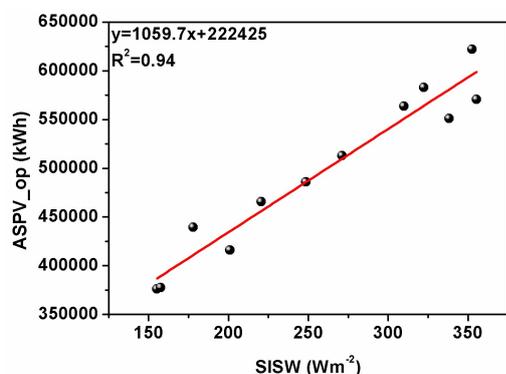


Fig. 6. Scatter Plot between the monthly aggregated solar PV output (ASPV_op) and the surface incident shortwave flux (SISW) for the year 2015 over South Africa.

V. CONCLUSION

In present work, ASPV_op has been analysed using statistical and wavelet spectrum method for the year 2015 in South Africa. The mean ASPV_op for 2015, was higher in spring followed by summer. It can be concluded that in South Africa during spring months ASPV_op was higher. The probability distribution of ASPV_op is not normal which indicated ASPV_op is skewed because of the weather variability. The empirical CDF is better for explaining the variation of ASPV_op. The high correlation between ASPV_op and SISW indicates that output of PV plants is

highly dependent on the incoming solar radiation. The wavelet power spectra showed a high power concentration between the 2–16-day band. The average variance in this region presents the important peaks, most of them were significant at 5% level. There is a scope of applying wavelet analysis tool for longer ASPV_op time series to get more important information. This could be useful for increasing the efficiency of SPV plants and forecasting of SPV_op.

ACKNOWLEDGMENT

We are grateful to Eskom for giving us the aggregated solar photovoltaic output data for South Africa. We want to acknowledge South African Weather Service, Pretoria for support during the entire work. Analyses used in this study were produced with the Giovanni online data system, developed and maintained by the NASA GES DISC.

REFERENCES

- [1] Eskom, *Eskom Integrated Report*, Eskom holdings Limited, 2016.
- [2] P. J. C. Nel, M. J. Booysen, and B. van der Merwe, "Energy perceptions in South Africa: An analysis of behaviour and understanding of electric water heaters," *Energy for Sustainable Development*, vol. 32, pp. 62-70, 2016.
- [3] State of Renewable Energy in South Africa, Department of Energy, 2015.
- [4] M. M. Aman *et al.*, "A review of Safety, Health and Environmental (SHE) issues of solar energy system," *Renewable and Sustainable Energy Reviews*, vol. 41, pp. 1190-1204, 2015.
- [5] R. Munzhedzi and A. B. Sebitosi, "Redrawing the solar map of South Africa for photovoltaic applications," *Renewable Energy*, vol. 34, no. 1, pp. 165-169, 2009.
- [6] M. Sifuzzaman, M. R. Islam, and M. Z. Ali, "Application of wavelet transform and its advantages compared to Fourier transform," *Journal of Physical Sciences*, vol. 13, pp. 121-134, 2009.

- [7] A. Araghi *et al.*, "Using wavelet transforms to estimate surface temperature trends and dominant periodicities in Iran based on gridded reanalysis data," *Atmospheric Research*, vol. 155, pp. 52-72, 2015.
- [8] Giovanni. [Online]. Available: <http://giovanni.gsfc.nasa.gov/giovanni/>
- [9] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bulletin of the American Meteorological Society*, vol. 79, no. 1, pp. 61-78, 1998.
- [10] S. M. A. Ideião and C. A. G. Santos, "Analysis of precipitation time series using the wavelet transform," *Sociedade & Natureza*, vol. 1, no. 1, pp. 736-745, 2005.
- [11] D. B. Percival, M. Wang, and J. E. Overland, "An introduction to wavelet analysis with applications to vegetation time series," *Community Ecology*, vol. 5, no. 1, pp. 19-30, 2004.



Jyotsna Singh has done her PhD degree from Birla Institute of Technology, Mesra Ranchi, India in Environmental Sciences and Meteorology. She has also done Post-graduate Diploma Course in remote sensing and GIS from Indian Institute of Remote Sensing, Dehradun, India. She was working as an Assistant Professor in environmental sciences in Manipal University Jaipur, India, 2015. Currently, she is working as a Senior Scientist in South African

Weather Service, Pretoria, South Africa. She is working in the field of renewable energy. Her research interest includes solar radiation and its modelling, remote sensing and GIS.



Bhishma Tyagi has obtained his PhD degree in atmospheric science from Indian Institute of Technology, Kharagpur, India. He is working as an Assistant Professor at the Department of Earth and Atmospheric Science, National Institute of Technology, Rourkela, Odisha, India. His area of research includes dynamics of the atmospheric boundary layer, air pollution study and atmospheric modeling.



Lucky Ntsangwane is a Research Manager at the South African Weather Service where he is involved in research projects related to solar energy research and applications, climate change and variability, and their socio-economic impacts, particularly in South Africa. He started working on climate related issues when he joined the South African Weather Service as a Research Scientist working on seasonal climate prediction and applications. At present, he is a

member of the South African Society for Atmospheric Sciences (SASAS), the National Association for Clean Air (NACA) and HIOC. He is responsible for several projects as a project manager and principal investigator.



Joel Botai did his MSc degree in engineering from Chalmers University, Sweden, 2005. He completed his PhD degree in 2011 from the University of Pretoria. He was teaching remote sensing at the University of Pretoria. He joined South African Weather Service, Pretoria, South Africa in October 2015 as a Chief Scientist. Currently, he is working in the field of hydrometeorology, agro meteorology and renewable energy.