

Langley Calibration of Sun Photometer at Kinabalu Park (1574 M a.s.l.) Using PDM Algorithm and Statistical Filter

Nur Hasinah Najiah Binti Maizan¹, Jackson Chang Hian Wui²,
Jumal Sulaiman³, Jedol Dayou^{1*}

¹ Energy, Vibration and Sound Research Group (e-VIBS), Faculty science and Natural Resources, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, MALAYSIA.

² Preparatory Center for Science and Technology, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, MALAYSIA

³ Mathematics with Computer Graphics Program, Faculty science and Natural Resources, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, MALAYSIA.

*Corresponding author: jed@ums.edu.my; Tel: +6088-320000; Fax: +6088-435324.

ABSTRACT: This paper reports the use of improved Langley plot for LED-type sun photometer calibration at four wavelengths 440, 500, 670 and 870 nm. The Langley plot is improved by series filtration using Perez-Dumortier (PDM) algorithm and statistical filter. Data was collected at a mid-altitude site, Kinabalu Park, Kundasang (1,574 m a.s.l.) using a portable ASEQ, LR-1 spectrometer. It is shown that with Langley plot alone, it is impossible to correctly identify or remove atmospheric variations within the calibration measurements. These variations are dominated by cloud cover, and short-cirrus cloud. However, result findings show that PDM algorithm and statistical filter are useful tool to improve the result by filtering data contaminated by cloud loading and remove possible drifts caused by instabilities of the instrument.

KEYWORDS: Langley calibration; Perez-Dumortier (PDM); Mid- altitude site; Statistical filter

Received 7 April 2016 Revised 26 May 2016 Accepted 31 May 2016 Inpress 3 June 2016 Online 20 December 2016
© Transactions on Science and Technology 2016

INTRODUCTION

Ground-based sun photometer is an electronic device which measures direct sunlight over a narrow range of wavelengths. It can be categorized into two main categories: filter-type sun photometer and LED-type sun photometer. The former uses interference filter to limit the amount of light reaching a photosensitive detector as little as 1 to 2 nm. The latter uses light emitting diode (LEDs) as detectors (Hamasha *et al.*, 2012; Xia *et al.*, 2014) which measures broader range of wavelengths than that with interference filter. They both play an important role in meteorological and atmospheric application. Despite the important potential of sun photometer in this area, accuracy of sun photometer is important. Sun photometer accuracy can be affected by several errors which include instrumental error, calibration error and atmospheric error (Shaw, 1976). However, instrument and atmospheric error will not be treated in this study. Coexisting with development of ground based sun photometer invention, a perfect, easy and economical calibration protocol is required in order to maintain sun photometer accuracy. Up until now, Langley method still becomes a favourable choice even though it suffers a few disadvantages. Langley method uses sunlight as the light source which is more stable than laboratory lamp but calibration measurement has to take place at high altitude for clear and stable atmosphere. Unfortunately, regular access to high altitude for Langley calibration is troublesome because it is not efficient in accessibility and economical prospect.

Due to inherent of limitation Langley calibration method, Chang *et al.* (2014) had developed the combination of Perez-Dumortier (PDM) algorithm and statistical filter as an alternative for sun photometers calibration. This improved Langley plot uses the combination of clear-sky selection model, Perez and Dumortier together with statistical filter to constrain the Langley extrapolation. Function of the clear- sky selection model is to ensure only cloudless and clear sky data is selected for the regression. Statistical filter is used to filter the residual regression data for improved instrument's response. As a whole, the main intention of this series filtration is to select appropriate data as imitation to the atmospheric condition at high altitude. In previous, this formulation had been successfully tested on calibrating spectrometer for aerosol optical depth (AOD) (Chang *et al.*, 2014; Dayou *et al.*, 2014), cloud optical depth (COD) (Maizan *et al.*, 2015) and investigation on diurnal evolution of solar radiation spectrum in UV, PAR, and NIR bands at near sea level (Chang *et al.*, 2013a; 2013b) retrieval at near sea level..However, its performance at altitude higher than near sea level remains unknown. Therefore, the objective of this study is to report the use of improved Langley plot for LED-type sun photometer calibration at altitude 1574 m above sea level.

BACKGROUND THEORY

In Perez model, sky type is categorized into different types from clear to overcast using discreet sky clearness index. The Perez's index can be computed using the relationship between the diffuse I_{ed} and global I_{eg} horizontal irradiance as shown (Perez *et. al.*, 1990)

$$\varepsilon = \frac{(I_{ed} + I_{dir}/I_{ed}) + 1.04 \phi_H}{1 + 1.04 \phi_H} \quad (1)$$

According to Eq (1), we need the component of diffuse I_{ed} , direct irradiance I_{dir} and global I_{eg} horizontal irradiance for estimating sky clearness index, and ϕ_H is the solar zenith angle in radian.

In Dumortier model, sky type is classified into five types namely blue, intermediate blue, intermediate mean, intermediate overcast, and overcast. This classification using Nebulosity index (NI) as indicator and can be computed by (Zain *et. al.*, 2002)

$$NI = \frac{1 - I_d / I}{1 - CR} \quad (2)$$

where I_d is the diffuse irradiance, I is the global irradiance. CR is the cloud ratio given by

$$CR = \frac{I_{d,cl}}{[I_{d,cl} + \exp(-4mAr) \sin \alpha]}, \quad (3)$$

where, $I_{d,cl}$ represents the clear sky illuminance and α represents the solar altitude. $I_{d,cl}$ is calculate using equation

$$I_{d,cl} = 0.0065 + (0.255 - 0.138 \sin \alpha_o), \quad (4)$$

where, Ar is the Rayleigh scattering coefficient written as

$$Ar = \{5.4729 + m[3.0312 + m\{-0.6329 + m(0.091 - 0.00152m)\}]\}^{-1}, \quad (5)$$

with m as the optical air mass. The Perez's and Dumortier model are used together in this work to form a combined sky classification as shown in Table 1 (Chang *et. al.*, 2014). Combinations of indices for both models were sought to classify raw data into specific sky condition. Filtration of

contaminated raw data that possess cloudy and overcast sky condition will be explained later in the next section.

Table 1. Perez and Dumortier model classification of sky condition.

Value of Indices		
Clearness index, ε	Nebulosity index, NI	Sky Condition
$\varepsilon \geq 4.50$	$0.95 \leq NI \leq 1.00$	Ideal clear sky
$1.23 < \varepsilon < 4.50$	$0.70 \leq NI \leq 0.95$	Intermediate blue
	$0.20 \leq NI \leq 0.70$	*Cloudy overcast/ Intermediate mean
* $\varepsilon \leq 1.23$	$0.05 \leq NI \leq 0.20$	*Cloudy overcast/ Intermediate overcast
	$0.00 \leq NI \leq 0.05$	*Cloudy overcast/ Overcast

As the aim for this study is to obtaining clear data for Langley plot, all cloudy and overcast data need to be identified and removed. This removal process undergoes filtration using multiple permuted criteria for each indices combination between clearness index and nebulosity index until the best Langley plot is obtained. Then, statistical filtration is performed to remove outliers or data points exhibit $>2\sigma$ from the resulting regression line after PDM algorithm filtration. Details of the method can be referred in Chang *et al.* (2014).

METHODOLOGY

Raw data collection for this study was accomplished at Kinabalu Park from 26th to 30th August, 2015 using a portable radiometer (ASEQ LR-1 spectrometer). The measuring site is located at open area at Kinabalu Park (6.0N, 116 E, 1,574 m a.s.l.). There is no huge industry located near the experimental area. Therefore stable aerosol loading is expected due to less pollution emission. Measurements for four nominal wavelengths 440, 500, 670, and 870 nm was performed on visually clear morning starting at sunrise, between 0600 and 0900 local time at periodic interval of 3 minutes.

Each measurement series consists of global and diffuse irradiance component. The diffuse irradiance component is measured using sunshade (solid disk on pole) diffuser after each global irradiance is measurement. This sunshade has 0.09m diameter ($D=0.09m$), and is held 1.0m from the sensor. The direct irradiance component was determined by subtracting the value recorded by the spectrometer in this condition from the corresponding global irradiance scans.

RESULT AND DISCUSSION

During the measurement period, atmospheric pressure and ambient temperature presented diurnal variations in the order of 5 hPa and 5°C, respectively, while the columnar ozone content varied on average by about <5 DU on each measurement day. Although these are relatively small variations on daily basis, the importance of the air mass evolution contributes a significant optical depth on diurnal basis for both Rayleigh Optical Depth (ROD) and Ozone Optical Depth (OOD). Figure 1 shows the temporal evolution of ROD and OOD within the Langley measurement at each wavelength. Both ROD and OOD graph show similar pattern for all wavelength on how its effect

in varying airmass. Based on Figure 1, contributions for both optical depths are more significant at high air masses intervals and gradually decrease in linear form for lower air masses for all wavelength. However, contribution of ROD are more prone at short wavelength, 440 nm followed by 550, 670 and 870 nm as its show high values compare to longer wavelength. This is due to short wavelength are easier to scattered by air molecules. On the contrary, wavelength 670 nm are more prone to OOD compare to other wavelength. This is because, ozone are predominant gas absorber in Hartley & Huggins (277.8- 362.5 nm) and Chappuis band (407.5- 682.5) (Shaw, 1979; Nissen et al., 2007; Serdyuchenko et al., 2014). This indicates that the optical depth contributions due to Rayleigh and ozone are wavelength dependent where one can observe their own behaviour towards specific wavelength.

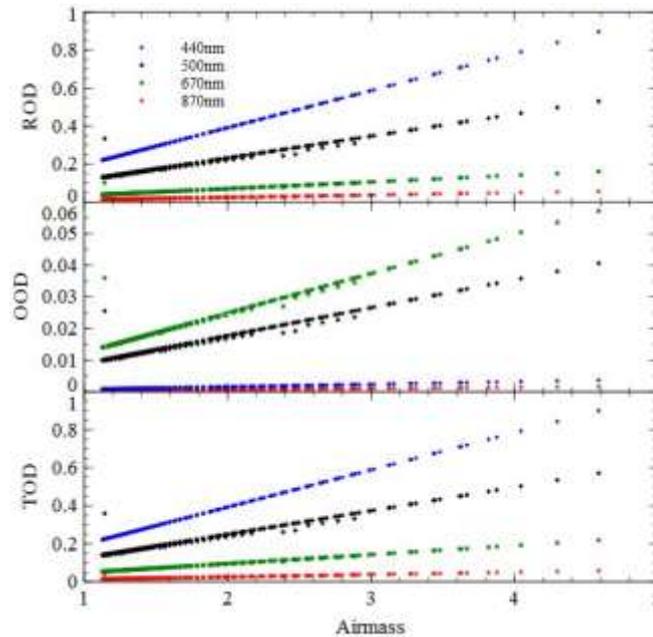


Figure 1. Temporal evolution of Rayleigh optical depth (upper), ozone optical depth (middle) and total optical depth (lower) within the Langley measurement at each wavelength.

To obtain reliable values of extraterrestrial constant, it is appropriate to consider only calibration periods presenting very stable atmospheric conditions that have limited variations in atmospheric pressure and columnar gas contents. To achieve such condition, we applied a refined Langley technique for the calibration of sun photometer to account for the variations due to atmospheric constituents such as air molecules and ozone. The refined method consists of the use of the signal $P(\lambda)$ that would be measured if aerosol was the only attenuator. It can be obtained from the measured $P(\lambda)$ by removing the contributions from molecular scattering and absorption where in this case the dominant attenuators were Rayleigh and ozone contributions. The refined Langley plot technique has the advantage that temporal changes of ROD and OOD contributing to AOD will not affect the determination of $P_0(\lambda)$, provided they can be estimated simultaneously to the measurements.

Within the Langley measurement, a total of $n=219$ raw data have been collected. However, not all data can be used for Langley plot because some of the data are contaminated by cloud loading and some might be inhibited by instrumental drifts. As previously mentioned, these errors are triggered by the atmospheric variations that occurred naturally during the Langley measurement. With Langley plot alone, it is impossible to correctly identify or remove these atmospheric variations. To improve the goodness of the best-fitted line, we applied the PDM algorithm and

statistical filter on all available data to yield a perfect Langley plot suitable for extraterrestrial constant retrieval. The highest correlation coefficient obtained is 0.7914 when the Langley plot is constrained by PDM algorithm at $NI \geq 0.99$ and $\varepsilon \geq 1.31$. After that, statistical filter is used to remove outliers that have standard deviation greater than 2σ from the resulting Langley regression after PDM filtration. These outliers are undetectable due to the fact that they are fictitiously predicted as clear-sky data by the PDM algorithm but apparently fall outside the acceptable range $<2\sigma$ from the resulting regression line. One reason that can explain this discrepancy is the instabilities of the instrument itself. Hence, statistical filter is necessary as precaution to account for the instrumental drifts. For illustration purpose, Figure 2(a) shows the Langley plot with unfiltered data $n=219$, followed by Figure 2(b) after filtration using PDM algorithm ($n=147$), and Figure 2(c) after statistical filter ($n=131$) at wavelength 440 nm.

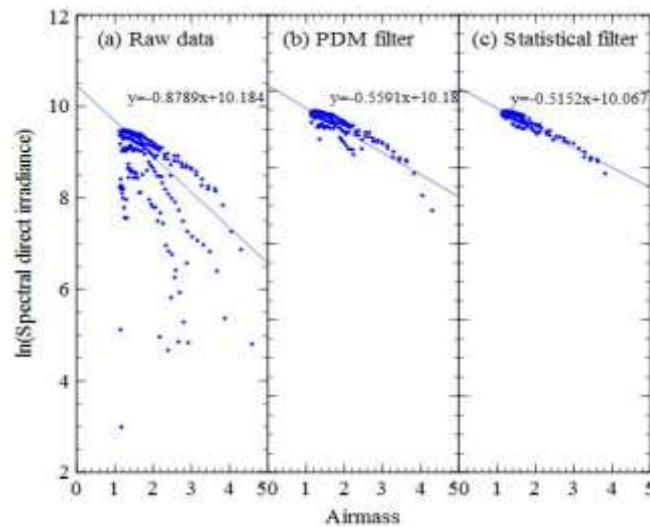


Figure 2. Langley plot at 440 nm (a) unfiltered data $n=219$, (b) after PDM clear-sky filtration, $n=147$, (c) after statistical filtration, $n=131$.

Figure 3 shows the final product of Langley plot after series filtration using PDM algorithm and statistical filter for wavelength 440, 500, 670 and 870 nm. Differences in term of correlation coefficients R^2 , before and after filtration, as well as their linear regression line for each wavelength are depicted in Table 2.

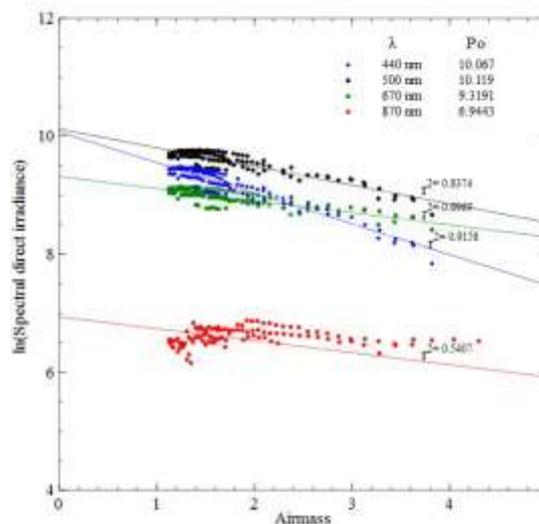


Figure 3. Final product of Langley plot after series filtration using PDM algorithm and statistical filter for each wavelength 440, 500, 670, and 870 nm.

Table 2. Improved Langley plot after series filtration using PDM algorithm and statistical filter.

Wavelength	Filtration	Criteria	N	Regression line	R ²
440 nm	-	-	219	$y = -0.8789x + 10.184$	0.2564
	PDM	$NI \geq 0.99$ and $\varepsilon \geq 1.31$	147	$y = -0.5591x + 10.105$	0.7914
	Statistical filter	$\sigma < \pm 0.2$	131	$y = -0.5152x + 10.067$	0.9158
500 nm	-	-	219	$y = -0.6646x + 10.305$	0.2529
	PDM	$NI \geq 0.99$ and $\varepsilon \geq 1.31$	147	$y = -0.365x + 10.166$	0.6657
	Statistical filter	$\sigma < \pm 0.2$	120	$y = -0.3134x + 10.119$	0.8374
670 nm	-	-	219	$y = -0.5687x + 9.4721$	0.1576
	PDM	$NI \geq 0.99$ and $\varepsilon \geq 1.31$	147	$y = -0.2437x + 9.3362$	0.4808
	Statistical filter	$\sigma < \pm 0.2$	124	$y = -0.2048x + 9.3191$	0.6969
870 nm	-	-	219	$y = -0.5003x + 7.02$	0.1423
	PDM	$NI \geq 0.99$ and $\varepsilon \geq 1.31$	147	$y = -0.21355x + 6.9555$	0.4358
	Statistical filter	$\sigma < \pm 0.19$	111	$y = -0.2049x + 6.9443$	0.5407

n: number of available data; R²: coefficient of correlation

On classical Langley plot, the slope of the regression line represents the daily aerosol optical depth. However, in this work, the slope of the regression line on the improved Langley plot is estimated from a pool of data that includes all days' measurements. Therefore, we can use this value as an index to represent the atmospheric variation after series filtration where lower slope indicates more stable atmospheric condition and higher slope basically indicates the opposite. At this point, we point out that when a more stringent filtration is imposed, a more stable atmospheric variation is expected. This pattern is consistent for all wavelengths where the slope value decreases from 0.88 before filtration to 0.52 after statistical filtration at 440 nm, 0.66 to 0.31 (500 nm), 0.57 to 0.20 (670 nm) and 0.50 to 0.20 (870 nm). Another parameter that has the similar pattern is the betterment on correlation R² after series filtration from PDM algorithm to statistical filter. For the final product, the greatest correlation is obtained for wavelength 440 nm at 0.92 followed by 500 nm (0.84), 670 nm (0.70) and 870 nm (0.54). As a whole, the series filtration has yielded a better dataset for Langley plot particularly on the cloud loading contamination and instrumental drifts. Future investigations will be focused on the insusceptible of the improved Langley plot at different altitude levels and eventually establish a robust calibration method which inhibits the necessity of travelling to high altitude for classical Langley plot.

CONCLUSION

In this work, an improved Langley plot was used to calibrate a LED-type sun photometer at Kinabalu Park (1574 m a.s.l.). The improved Langley plot is a filtration series of Perez-Dumortier algorithm and statistical filter until the best fitted regression line is obtained. The regression and correlation analysis of the improved Langley plot confirms the betterment in both the slope and coefficient of determination after the series filtration. The results indicate that combination of PDM algorithm and statistical filter has the potential to appropriately select useful data for Langley plot even for a given pool of data. This selection is completely objective and automated where qualitative observation at distinct air mass is no longer necessary. While this method can serve as an objective algorithm for consistent AM0 extrapolated value, it is likely useful to accurately select data close to ideal Langley condition without travelling to high altitude.

ACKNOWLEDGEMENTS

This research was supported by the Malaysian Ministry of Education under the research grant number RAG0071-SG-2015, and is greatly acknowledged.

REFERENCES

- [1] Chang, J. H. W, Dayou, J. & Sentian, J. (2014). Development of Near-Sea-Level Calibration Algorithm for Aerosol Optical Depth Measurement Using Ground-Based Spectrometer. *Aerosol and Air Quality Research*, **14**(1), 380-395
- [2] Chang, J. H. W, Dayou, J. & Sentian, J. (2013a). Diurnal evolution of solar radiation in UV, PAR and NIR bands in high air masses. *Nature Environment and Pollution Technology*, **12**(1), 1 – 6.
- [3] Chang, J. H. W, Dayou, J. & Sentian, J. (2013b). Investigation of Short Time Scale Variation of Solar Radiation Spectrum in UV, PAR, and NIR Bands due to Atmospheric Aerosol and Water Vapour. *World Academy of Science, Engineering and Technology*, **7**(1), 143 - 148.
- [4] Dayou, J., Chang, J. H. & Sentian, J. (2014). *Ground-Based Aerosol Optical Depth Measurement Using Sunphotometers*. Springer. ISBN 978-981-287-101-5.
- [5] Hamasha, K. M., Abu Mostafa, H. M. & Alexander, L. T. (2012). Aerosol Optical Thickness at Tabuk City, SA. *International Journal of Applied Science and Technology*, **2**(10), 69-91
- [6] Maizan, N. H. N., Dayou, J. & Yusoff, R. (2015). Measurement of Cloud Optical Depth Using Sunphotometer Calibrated By PDM Algorithm. *Advances in Environmental Biology*, **9**(10), 19-25
- [7] Nissen, K. M., Matthens, K., Langematz, U. & Mayer, B. (2007). Towards a Better Representation of the Solar Cycle in General Circulation Models. *Atmospheric Chemistry and Physics*. **7**, 5391- 5400
- [8] Perez, R., Ineichen, P. & Seals, R. (1990). Modelling daylight availability and irradiance components from direct and global irradiance. *Journal of Solar Energy*, **44**, 271-289
- [9] Serdyuchenko, A., Weber, M., Chehade, W. & Burrows, J. P. (2014). High Spectral Resolution Ozone Absorption Cross- sections- Part 2: Temperature Dependence. *Atmospheric Measurement Techniques*, **7**, 625- 636
- [10] Shaw, G. E. (1976). Error Analysis of Multi-wavelength Sunphotometry. *Pure and Applied Geophysics*, **114**, 1-4
- [11] Shaw, G. E. (1979). Atmospheric Ozone: Determination by Chappuis- Band Absorbption. *Journal of Applied Meteorology*. **18**, 1335- 1339
- [12] Xia, M., Li, J., Li, Z., Gao, D., Pang, W., Li, D. & Zheng, X. (2014). Research on calibration method in lab of direct solar channels of SunPhotometers. *Chinese Optic Letters*, **12**, 121-201.
- [13] Zain-Ahmad, A., Sopian, K., Zainol Abidin, Z. & Othman, M.Y.H. (2002). The Availability of Daylight from Tropical Skies: A Case Study of Malaysia. *Renewable Energy*, **25**, 21-30.