



The Ångström–Prescott Regression Coefficients for Six Climatic Zones in South Africa

Brighton Mabasa ^{1,2,*}, Meena D. Lysko ^{2,3}, Henerica Tazvinga ¹, Sophie T. Mulaudzi ⁴, Nosipho Zwane ¹ and Sabata J. Moloi ²

- ¹ Research & Development Division, South African Weather Service, Pretoria 0001, South Africa; Henerica.Tazvinga@weathersa.co.za (H.T.); Nosipho.Zwane@weathersa.co.za (N.Z.)
- ² Department of Physics, University of South Africa, UNISA Preller Street, Muckleneuk, Pretoria 0001, South Africa; mlysko@gmail.com (M.D.L.); moloisj@unisa.ac.za (S.J.M.)
- ³ Move Beyond Consulting (Pty) Ltd., Pretoria 0001, South Africa
- ⁴ Department of Physics, School of Mathematical and Natural Sciences, University of Venda, Thohoyandou 0950, South Africa; Sophie.Mulaudzi@univen.ac.za
- * Correspondence: brighton.mabasa@weathersa.co.za; Tel.: +27-12-367-6088

Received: 31 July 2020; Accepted: 29 September 2020; Published: 16 October 2020



Abstract: The South African Weather Service (SAWS) manages an in situ solar irradiance radiometric network of 13 stations and a very dense sunshine recording network, located in all six macroclimate zones of South Africa. A sparsely distributed radiometric network over a landscape with dynamic climate and weather shifts is inadequate for solar energy studies and applications. Therefore, there is a need to develop mathematical models to estimate solar irradiation for a multitude of diverse climates. In this study, the annual regression coefficients, *a* and *b*, of the Ångström–Prescott (AP) model, which can be used to estimate global horizontal irradiance (GHI) from observed sunshine hours, were calibrated and validated with observed station data. The AP regression coefficients were calibrated and validated for each of the six macroclimate zones of South Africa using the observation data that span 2013 to 2019. The predictive effectiveness of the calibrated AP model coefficients was evaluated by comparing estimated and observed daily GHI. The maximum annual relative Mean Bias Error (*rMBE*) was 0.371%, relative Mean Absolute Error (*rMAE*) was 0.745%, relative Root Mean Square Error (rRMSE) was 0.910%, and the worst-case correlation coefficient (R^2) was 0.910. The statistical validation metrics results show that there is a strong correlation and linear relation between observed and estimated GHI values. The AP model coefficients calculated in this study can be used with quantitative confidence in estimating daily GHI data at locations in South Africa where daily observation sunshine duration data are available.

Keywords: South African Weather Services; radiometric network; climatic zone; Ångström–Prescott; global horizontal irradiance; sunshine duration

1. Introduction

Solar radiation data are important because they are required in many research fields such as meteorology, agriculture, hydrology, ecology, and environment [1–4]. Solar radiation data are also an important reference for many applications such as solar power plants, engineering designs, regional crop growth modelling, evapotranspiration estimation, and irrigation system development [1,3,5]. In relation to this, South African Weather Services (SAWS) re-established a global horizontal irradiance (*GHI*) radiometric network with 13 solar radiometric stations located in all 6 macroclimatic zones of South Africa [6]. The macroclimatic zones are regions with similar climatic conditions, and they were established to classify different areas based on their maximum energy demand and maximum energy



consumption [6]. The data collected from the SAWS network help in the validation of satellites as well as the development and verification of empirical models [7]. SAWS also manage a very dense sunshine duration recording network over South Africa to the extent that sunshine duration data have been continuously measured for several years [8]. SAWS *GHI* stations are sparse; according to [1-5,9], having dense radiometric networks is a worldwide challenge because of the high costs involved in the installation and maintenance of the solar radiation stations. To compensate for this, reliable measurements taken from a sparse network are needed to develop and validate empirical models that can be used to estimate and forecast the availability of solar energy at other locations [10]. The main objective of this study is to calibrate the Ångström–Prescott (AP) model regression coefficients *a* and *b* that could be used to estimate *GHI* in different climatic zones of South Africa, thus increasing the density of available solar radiation data in the country.

The AP model estimates daily GHI using daily extraterrestrial (top of the atmosphere) GHI radiation (GHI_{TOA}) , daily astronomical day length (N), daily measured sunshine duration (n), and Ångström model coefficients *a* and *b*. The model was first proposed by Angström [11] in 1924 before Prescott [12] modified it in 1940 by adding GHI_{TOA} to replace GHI on a clear sky day. The original AP coefficients were a = 0.25 and b = 0.75; these were calculated using data from Stockholm [13]. The regression coefficients a and b are site-dependent; therefore, there is a need to calibrate them using a linear relationship in Equation (1) at regions where they will be used to estimate *GHI* [1,4,13,14]. Martinez et al. [14] emphasised that AP regression coefficients with proven accuracy in one climatic region should not be assumed to be equally reliable in the other climatic region without additional evidence. Researchers such as those from the Chinese Academy of Sciences, the Indian National Academy of Agricultural Research Management, the Brazilian Federal University of Rio Grande do Norte, and Spanish Polytechnic University of Madrid [1–5] calibrated AP coefficients to their own climatic regions by using the linear relationship in Equation (1). The study from Spain by Almorox et al. [5] focused on only one station and the study from Brazil by De Medeiros et al. [4] focused on four stations, but all were located in only two climatic zones. According to the findings by Zhang et al. [3], the regression coefficients were different in different climate zones. The differences in *a* and *b* in different climatic zones could be due to variations in latitude, altitude, aerosols, and water vapor concentration, surface albedo, and mean solar altitude [14]. The study by Tsung et al. [15] focused on one location using n and GHI data collected from two different stations because of the unavailability of both GHI and n from one location. In this study, eight stations located in all six climatic zones of South Africa, and where both *n* and *GHI* data were collected from the same location, are considered so that the respective AP coefficients are representative of a climatic zone in the country and not the whole country.

According to works by four different research groups [1–3,14,16], sunshine-based models provided better *GHI* estimates when compared to cloud and temperature-based models. This might be because the amount of *GHI* reaching the earth's surface is closely related to sunshine duration [3]. Cloud cover restricts the amount of *GHI* reaching the earth's surface and cloud-based models also perform better [3], but accurate cloud observation data to be used as an input in the model are scarce compared to sunshine duration data. The effect of temperature on *GHI* is lower than that of sunshine duration and cloud cover. This is because most of the long-wave solar radiation reaching the earth's surface is absorbed, emitted in the atmosphere, or reflected to space [3]. *GHI* is the total amount of short-wave solar radiation reaching the earth's surface and has little dependence on temperature. The availability of reliably measured sunshine duration data in all climatic zones and the performance of sunshine-based models motivated the focus on sunshine-based models in this study. The AP linear regression model was chosen because of its simplicity and also, as suggested by Tsung et al. [15], linear models were ranked high in global performance indicator (GPI) in comparison to other models in a review in 2015.

In South Africa, studies to calibrate AP coefficients were carried out by Eberhard [17] and Mulaudzi et al. [18]. The challenge, according to Mulaudzi et al. [18], was the unavailability of a long-term observation *GHI* dataset that covers all the climatic regions to calibrate and validate the AP coefficients. In this study, a large enough dataset with observations spanning 2013 to 2019 from

stations that cover all the climatological zones of South Africa was used to calibrate the AP coefficients, which were then used to estimate *GHI*. The estimated *GHI* was validated using the observed *GHI* daily averages, while the statistical metrics (10) to (16) from [19,20] were used to quantify the differences between observed and estimated *GHI*.

From the results of this study, annual AP coefficients a and b in all six macro-climatological regions could be used to estimate daily *GHI*, for the respective climate regions, using daily observation sunshine duration data. The knowledge of estimated daily *GHI* data can thereby be used to develop energy policies and solar energy programmes. They can also be used as benchmarks in climate analysis studies.

2. Materials and Methods

The observed 1-min *GHI* data used in this study were collected from 8 SAWS solar radiometric stations during the periods shown in Table 1, which also shows the geographical locations and the climatic zones in which the stations are located. The map in Figure 1 gives the macro-climate regions in South Africa. *GHI* data were collected using secondary standard, CMP11, Kipp and Zonen pyranometers.

Station	Latitude (°)	Longitude (°)	Altitude (m)	Period	Climatic Zone
Upington	-28.48	21.12	848	2014-02-01 to 2019-11-30	Arid Interior
De Aar	-30.67	23.99	1284	2014-05-01 to 2019-12-31	Cold Interior
Irene	-25.91	28.21	1524	2014-03-01 to 2019-12-31	Temperate Interior
Mthatha	-31.55	28.67	744	2014-07-01 to 2019-12-31	Subtropical Coastal
George	-34.01	22.38	192	2015-01-01 to 2019-12-31	Temperate Coastal
Durban	-29.61	31.11	91	2015-03-01 to 2019-12-31	Subtropical Coastal
Polokwane	-23.86	29.45	1233	2015-03-01 to 2019-12-31	Temperate Interior
Thohoyandou	-23.08	30.38	619	2015-03-01 to 2017-10-31	Hot Interior

Table 1. South African Weather Services radiometric station location, altitude, period covered, and climatic zones. (The climate zones correspond with the regions shaded in Figure 1).

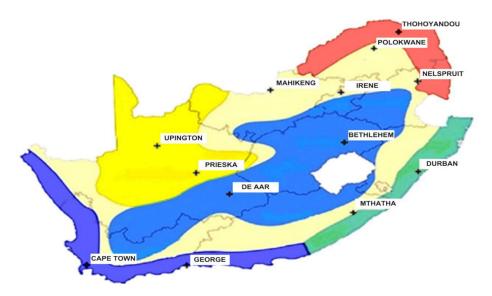


Figure 1. A map showing SAWS radiometric station location and climatic zones.

The methodology is provided in the flowchart in Figures 2 and 3. Daily *GHI* data were calculated from 1-min *GHI* data. First, the 1-min *GHI* data were quality controlled using a Baseline Solar Radiation Network (BSRN) quality control (QC) procedure outlined by Long and Dutton in [21]. *GHI* values that failed the QC test were regarded as outliers and were discarded; only the data that passed the test were used [6,7,20–22]. Minute values that passed the BSRN QC were averaged to 15 min and then, 4 slots of 15-min averages were averaged to obtain an hourly mean [6,7,22–25]. Hourly mean values were then averaged to obtain daily average values. Daily average values were further quality checked by subjecting them to HelioClim model QC, described by Geiger et al. in [26]; outliers, which were daily average points coded 1, were discarded before further analysis.

Hourly sunshine duration data were obtained by determining the burn made by the sun on a coated card in a Campbell–Stokes sunshine recorder [8]. Hourly data were then summed to obtain total daily sunshine duration (*n*). Daily top-of-atmosphere (*TOA*) irradiance (*GHI*_{*TOA*}) and theoretical sunshine duration (*N*) were calculated using Equations (1)–(9), from Iqbal [13], and the solar angles were calculated using the Solar Position Algorithm (SPA) on Python PVLIB [27,28] and Microsoft Excel. The coefficients *a* and *b* of the AP model were calculated by using the linear regression analysis between the irradiance fraction or clearness index, $\frac{GHI}{GHI_{TOA}}$ and daily sunshine fraction, $\frac{n}{N}$ for each day, based on a linear relationship shown by Equation (1) proposed by Ångström [11] and then, modified by Prescott [12].

$$\frac{GHI}{GHI_{TOA}} = a + b(n/N),\tag{1}$$

where GHI is the daily Global Horizontal Irradiance in W/m².

 GHI_{TOA} is an approximation of the top of the atmosphere GHI or extraterrestrial radiation on a horizontal surface, i.e., the amount of global horizontal radiation that a location on Earth's surface would receive if there was no atmosphere; it is given by Equation (2), as in Duffie and Beckman [29].

$$GHI_{TOA} = \left(\frac{24}{\pi}\right) I_{SC} E_o[(\pi/180) \cdot \omega_s \cdot (\sin \delta \ \sin \emptyset) + (\cos \delta \ \cos \emptyset \ \sin \omega_s)], \tag{2}$$

$$I_{SC} = \text{solar constant} = 1367 \text{ W/m}^2, \qquad (3)$$

(World Meteorological Organization recommendation, according to Gueymard in [30]),

$$E_o = \text{eccentricity factor} = 1 + 0.033 \cos\left[\left(\frac{2\pi D}{365}\right)\right],\tag{4}$$

where D is the Julian day,

$$\omega_s = \text{sunset hour angle} = \cos^{-1}(-\tan \emptyset \tan \delta),$$
 (5)

$$\emptyset =$$
degree of latitude, (6)

$$\delta = \text{solar declination} = -23.45 \sin\left[\frac{360(D+284)}{365}\right],\tag{7}$$

$$N = \frac{2}{15} \cos^{-1}(-\tan \varnothing \tan \delta) = \text{Astronomical sunshine duration,}$$
(8)

$$n =$$
daily recorded sunshine duration (in hours), (9)

where *a* and *b* represents Ångström–Prescott regression coefficients.

Annual AP coefficients were calculated for 8 stations. The observation periods for concurrent *GHI* and sunshine duration data for these stations are given in Table 1. Datasets up to the end of 2018 were used for determination of the AP coefficients, and the daily observation data for 2019 were used to validate the corresponding estimated daily *GHI* data. For Thohoyandou, the 2017 data were used to validate the coefficients.

The statistical metrics that were used to compare estimated daily *GHI* data with the observed daily *GHI* data were derived from the literature [19,20] and these are:

1. Mean Bias Error (*MBE*), which estimates the average error in the prediction. A positive *MBE* indicates that the prediction is overestimated and vice versa; the lower values of *MBE* indicate a strong correlation between the prediction and observation. A relative Mean Bias Error (*rMBE*), which measures the size of the error in percentage terms, was also calculated. The metrices are expressed as:

$$MBE = \frac{1}{n} \sum_{i=1}^{n} (Pi - Oi)$$
(10)

$$rMBE = 100 \times \frac{1}{n} \sum_{i=1}^{n} \frac{(Pi - Oi)}{\overline{O}i}$$
(11)

2. Mean Absolute Error (*MAE*), which measures the absolute value of the differences between the observed and the predicted values, gives a better idea of the prediction accuracy; relative Mean Absolute Error (*rMAE*), which measures the size of the error in percentage terms, was also calculated. The caution with *MBE* and *rMBE* is with the cancelling of positive and negative bias, which can lead to a false interpretation. The metrics are expressed as:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |Pi - Oi|$$
(12)

$$rMAE = 100 \times \frac{1}{n} \sum_{i=1}^{n} \frac{|Pi - Oi|}{\overline{O}i}$$
(13)

3. Root Mean Square Error (*RMSE*), which compares the predicted and observed datasets, measures the statistical variability of the prediction accuracy and is expressed as shown in Equation (14), while Equation (15) shows the relative Root Mean Square Error (*rRMSE*), which measures the size error in percentage terms. The *RMSE* and *rRMSE* are also indifferent to the direction of the error. They are considered in this study since these put extra weight on large errors. The metrices are expressed as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Pi - Po)^2}$$
(14)

$$rRMSE = \frac{100}{\overline{Oi}} \times \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Pi - Po)^2}$$
(15)

4. Coefficient of Determination (\mathbb{R}^2), which is a statistical measure of the strength of the relationship between the movement of predicted and observed. \mathbb{R}^2 also measures how well the regression line represents the data. The value of \mathbb{R}^2 is such that $0 \le \mathbb{R}^2 \le 1$. The closer \mathbb{R}^2 is to 1, the better the prediction. The metric is expressed as:

$$R^{2} = 1 - \frac{\sum_{i=1}^{n} (Pi - Oi)^{2}}{\sum_{i=1}^{n} (Pi - \overline{Oi})^{2}}$$
(16)

where Oi is the observation value, Pi is the estimated value, \overline{Oi} is the average of the observation values, i is the time point, and n is the total number of points used.

The results were converted from W/m^2 to MJ m⁻²d⁻¹ by dividing by 11.57415, a methodology used by Almorox et al. [5] to allow for easy comparison with other literature studies. Monthly averages of each metric were calculated and then aggregated to annual averages, and where observation data were not available, data were replaced by NaN. The annual AP coefficients *a* and *b* were calculated.

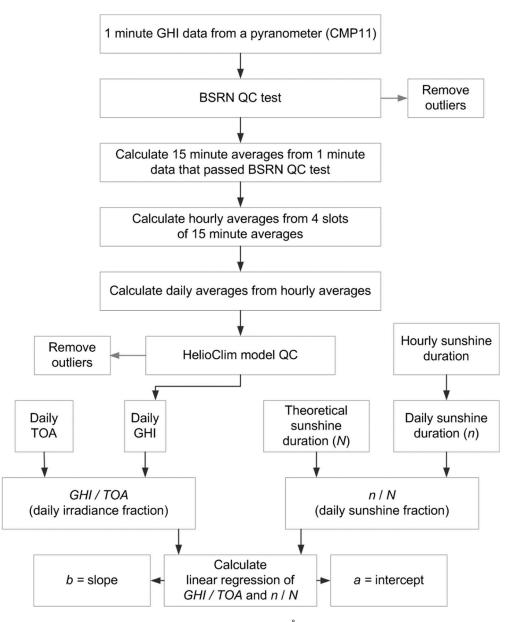


Figure 2. The flowchart of the approach used to calculate Ångström–Prescott regression coefficients *a* and *b*.

```
import numpy as np
import pandas as pd
import matplotlib
import matplotlib.pyplot as plt
%matplotlib inline
from matplotlib import style
from statistics import mean
import seaborn as sns
from sklearn.linear model import LinearRegression
from sklearn import metrics
from scipy import stats
from scipy.stats import linregress
import matplotlib.pylab as pylab
matplotlib.rcParams['figure.figsize'] = (12,8)
sns.set_palette(sns.color_palette("hls", 20))
sns.set style("whitegrid")
obs file = 'Station Name.csv'
columns = ['DATE', 'X', 'Y']
obs = pd.read csv(obs file, header = None, skiprows = 1, names = columns)
obs['DATE'] = pd.to datetime(obs['DATE'], format='%Y-%m-%d')
obs = obs.replace(np.inf, np.nan).replace(-np.inf, np.nan).dropna()
x = obs['X']
y = obs['Y']
def best fit slope(x,y):
    b = (((mean(x)*mean(y)) - mean(x*y)) /
         ((mean(x)**2) - mean(x**2)))
    return b
b = best fit slope(x,y)
a = mean(y) - b*mean(x)
def xyline(x, y):
    xypts = min(x.min(), y.min()), max(x.max(), y.max())
    return xypts, xypts
np.polyfit(obs['X'], obs['Y'], 1)
regression_line=np.polyfit(obs['X'], obs['Y'], 1)
plt.plot(*xyline(x, y), 'k--')
plt.plot(x,regression line[1]+regression line[0]*x)
plt.title('Station name', fontsize=25)
plt.xlabel('n/N',fontsize=25)
plt.ylabel('GHI/TOA', fontsize=25)
plt.plot(x, (a + b*x), label = 'y = {:.3f} + {:.3f}*x'.format(a, b))
plt.legend(loc='lower left',fontsize=25)
plt.scatter(x, y, c=x, cmap='jet')
plt.colorbar()
plt.savefig("Station name")
```

Figure 3. A Python code or script that was written and used to calculate and plot regression coefficients *a* and *b* from $X = \frac{n}{N}$ and $Y = \frac{GHI}{GHI_{TOA}}$.

3. Results and Discussion

3.1. Annual AP Results

In this study, the annual AP regression coefficients *a* and *b* were calculated using Equation (1) and the following variables daily *n*, daily mean *GHI*, daily mean *GHI*_{TOA}, and daily *N* were used as inputs. The calculated *a* and *b* were then used together with daily *n* and *N* to estimate daily *GHI*, which was then compared to corresponding observed daily *GHI*. Statistical metrics in Equations (10)–(16) were used to quantify the errors between the two datasets; the results are shown in Table 2 and Figures 4–7.

In Figures 4 and 5, the annual AP coefficients and the data points that were used to derive them are displayed. The values of the AP coefficients ranged from 0.188 to 0.243 for *a*, while those for *b* ranged from 0.515 to 0.6. Values of a = 0.25 and b = 0.5 were recommended by Allen et al. [31] to be used when there is no local observation *GHI* data to calibrate the coefficients. The minimum value of *a* in this study was less than 0.25 and the maximum value was greater than 0.25; the minimum and maximum values of *b* were greater than 0.5. The difference in default AP coefficients and calibrated AP coefficients proved that calibrating the coefficients locally is a necessity.

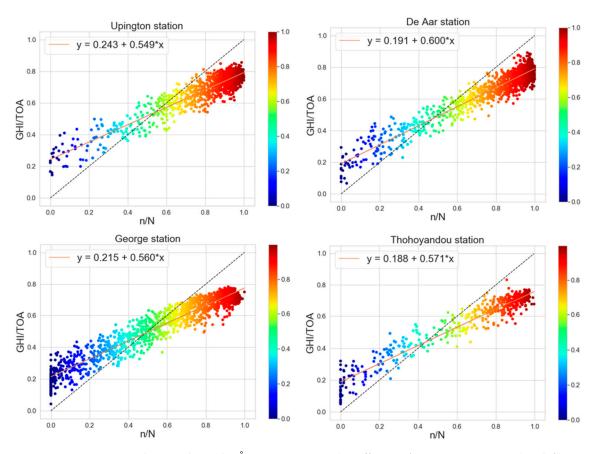
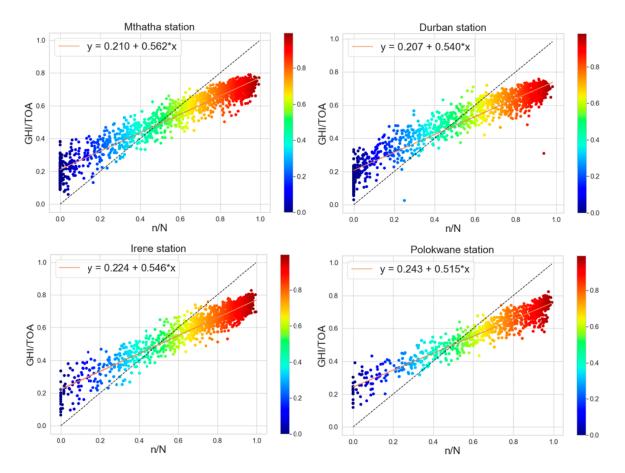


Figure 4. Regression lines and AP (the Ångström–Prescott) coefficients for Upington station (top left), De Aar (top right), George (bottom left), and Thohoyandou (bottom right).

Studies done by Zhang et al., De Medeiros et al., Almorox et al., and Tsung et al. in [3–5,14] also found different results to Allen et al. [31] when they did a local calibration. The AP coefficients from this study are in line with the coefficients from similar studies done elsewhere in the world.

When comparing the factors for stations that are located in the same climatological zone such as Irene and Polokwane located in the Temperature Interior climatic zone, and Mthatha and Durban in the Tropical Coastal climatic zone (Table 1), the difference was less than 0.05 for both *a* and *b*, which is a very small difference. This means that the AP coefficients *a* and *b* calibrated for a climatic zone could be



used as a representative for an entire climatic zone to estimate *GHI* when observed sunshine duration data for the location are available.

Figure 5. Regression lines and AP coefficients for Mthatha station (top left), Durban station (top right), Irene station (bottom left), and Polokwane (bottom right).

3.2. Validation Results

Estimated *GHI* values were compared to the measured *GHI* values; errors were quantified by validation metrics in Equations (10)–(16) and the results were tabulated in Table 2. It can be seen that in Table 2, the *rMBE* ranged from –1.20 to 0.371%, *rMAE* from 0.311 to 0.745%, *rRMSE* from 0.393 to 0.910%, and R² from 0.910 to 0.948. De Aar, Irene, and Thohoyandou had a positive *MBE*, meaning that the model overestimated *GHI*, while Upington, Durban, Mthatha, George, and Polokwane had a negative *MBE*, meaning that the model underestimated *GHI* values at these locations. The values of *MBE* and *rMBE* for all the stations were less than 1, indicating that there was a strong correlation between the predicted and observed *GHI* values. The worst case R² value was 0.910, suggesting that there is a very strong linear relation between observed and predicted values.

The maximum *RMSE* of 1.741 MJ m⁻²d⁻¹ was less than 1.94 and 1.9 MJ m⁻²d⁻¹ that De Medeiros et al. [4] and Tsung et al. [14], respectively, determined. The maximum *MAE* of 1.425 MJ m⁻²d⁻¹ was less than 1.8 MJ m⁻²d⁻¹ that Tsung et al. [14] determined. The maximum *MBE* of 0.733 MJ m⁻²d⁻¹ was less than 1.040 and 0.85 MJ m⁻²d⁻¹ that De Medeiros et al. [4] and Tsung et al. [14], respectively, determined, and the worst case R² of 0.910 was greater than 0.875, 0.74, and 0.58 that Zhang et al. [3], Adamala et al. [2], and De Medeiros et al. [4], respectively, determined. As shown in Table 2, the results across the range of climate zones (including the ones from the literature), all differ significantly.

Station	а	b	RMBE	rMBE (%)	MAE	rMAE (%)	RMSE	rRMSE (%)	R ²
Upington	0.243	0.549	-0.360	-0.120	0.841	0.311	1.061	0.393	0.930
De Aar	0.191	0.600	0.733	0.371	1.136	0.506	1.375	0.598	0.930
Irene	0.224	0.546	0.689	0.353	1.328	0.608	1.618	0.729	0.912
Mthatha	0.210	0.562	-0.104	-0.013	1.168	0.582	1.474	0.735	0.951
George	0.215	0.560	-0.270	-0.036	1.261	0.636	1.520	0.769	0.948
Durban	0.207	0.540	-0.322	-0.106	1.425	0.745	1.741	0.910	0.915
Polokwane	0.243	0.515	-0.286	-0.085	1.272	0.488	1.572	0.606	0.910
Thohoyandou	0.188	0.571	0.286	0.168	1.071	0.550	1.433	0.746	0.937
Almorox et al.	0.287	0.452	-0.002	-	-	-	1.260	-	-
De Medeiros et al.	0.39	0.29	1.040	6.29	-	-	1.94	-	0.58
Tsung et al.	0.5	0.11	0.85	3.4	1.8	-	1.9	-	-
Zhang et al.	0.214	0.552	-	-	2.249	-	0.214	-	0.875
Adamala et al.	0.28	0.52	-	-	-	-	7.04	-	0.74

Table 2. Calibration coefficients *a* and *b* and validation metrics results in (MJ $m^{-2}d^{-1}$).

The overall validation results from this study are comparable and even better than what was found in similar studies like [2–5,15], which concluded that the AP coefficients could be used to estimate *GHI* with confidence based on those validation results. The data used in the study were collected using secondary standard pyranometers (CMP11), which, according to Urraca et al. [23], generate high quality records of GHI. Urraca et al. [23] found large and unstable validation errors from stations that uses second class pyranometers and silicon-based photodiodes compared to the ones using CMP11. This might be because CMP11 gives an integrated measurement of the total GHI available under all conditions due to its capability to measure the total solar spectrum from 0.3 to 3 micrometres wavelength. CMP11 pyranometers' use very high-quality quartz double domes, which improve the stability of a calibration factor over time; it also improves directional response and reduces thermal offsets. The CMP11 can give correct integrated values over a day, with the use of smaller sampling intervals. At SAWS, a sampling interval of 5 s is used and this enables sudden changes such as passing small clouds, birds sitting on top of the pyranometers, and some other factors that result in shading of pyranometers to be identified and factored in the data. It is noted that smaller sampling intervals are not possible with CMP11 predecessors. The CMP11 has a specified expected daily uncertainty of less than 2% [32,33].

GHI data were subjected to robust quality control methodologies BSRN QC [21] and HelioClim model QC [26] before any analysis and outliers which might affect the results were discarded. The use of Python codes in data analysis enabled big data to be handled much more efficiently and execution of a code in data analysis resulted in correct and consistent outputs; these are some of the reasons why the results in this study are better. This means that the AP coefficients results from this study could also be used with confidence to estimate *GHI* in different climatological zones of South Africa.

In Figure 6, 2019 monthly *GHI* observed data were compared to corresponding estimated 2019 monthly *GHI* data. Thohoyandou is the only station where validation was done using 2017 monthly datasets and the observation data were only available from February to October (January, November, and December 2017 datasets were not available). The need to fill in missing data further motivates this study, i.e., development and validation of models, and the results of this study can be used to fill any missing monthly mean *GHI* values for South African locations.

Similarly, in Figure 7, the 2019 monthly *GHI* observed data were compared to corresponding estimated 2019 monthly *GHI* data. In Durban, *GHI* observation data for September were not available. In Polokwane, the *GHI* observation data for March, April, May, and June were not available. The results from the study can be used to fill those missing monthly mean *GHI* values.

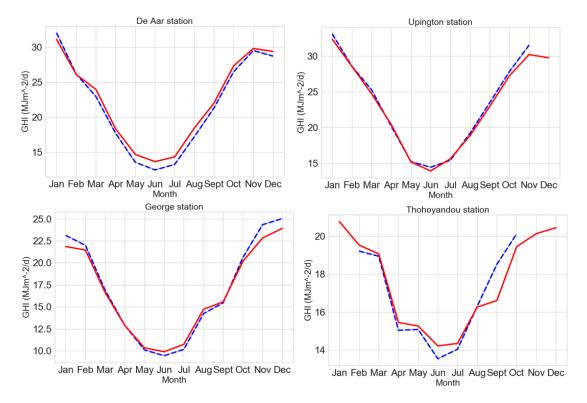


Figure 6. Comparison between measured and predicted monthly *GHI* values in De Aar (top left), Upington (top right), George (bottom left), and Thohoyandou (bottom right). Blue dotted line represents observed, *GHI* red solid line represents estimated *GHI*.

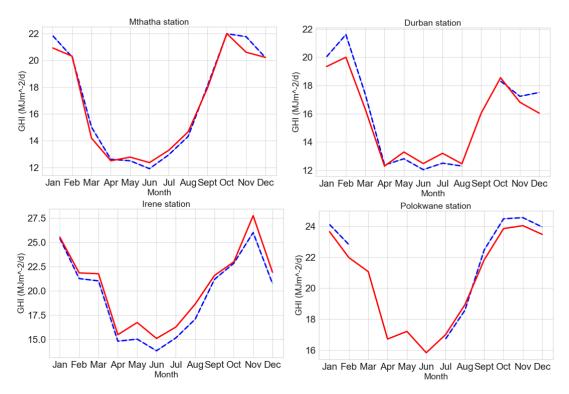


Figure 7. Comparison between measured and predicted *GHI* values in Mthatha (top left), Durban (top right), Irene (bottom left), and Polokwane (bottom right). Blue dotted line represents observed, red solid line represents estimated *GHI*.

4. Conclusions

The annual Ångström–Prescott coefficients *a* and *b* were calculated using the linear relationship between ratio of daily global radiation on a horizontal surface to the daily projected extraterrestrial radiation on that surface and the ratio of daily sunshine duration to the theoretical sunshine duration. They were used to estimate global horizontal irradiance and there was a very close agreement with the corresponding observation global horizontal irradiance. The agreement was quantified by statistical metrics in Equations (10)–(16), i.e., relative Mean Bias Error, relative Mean Absolute Error, relative Root Mean Square Error, and correlation coefficient (\mathbb{R}^2). The results were in good agreement with what other studies found.

The methodology used in the study can be applied elsewhere, where there is a station that records global horizontal irradiance and sunshine duration. Practitioners need to cross check against their climate zones and not use *a* and *b* from one site to represent the entire country, as it varies per climatic zone due to variations in latitude, cloud cover, aerosols, surface albedo, and day lengths. The unavailability of confident observation of daily sunshine data in some areas might be a drawback for other practitioners.

Further research will focus on extended monitoring of the stability of the coefficients over time in each climate zone. Further research will also focus on calibrating and validating rainfall, cloud, temperature, and humidity-based models in areas where sunshine data are not recorded to make sure that daily global solar radiation data can be estimated in those areas in South Africa. The Python script used in calculating linear regression coefficients and validation of observation and estimated *GHI* data is available on request from the correspondence author.

The Ångström–Prescott coefficients calibrated for each station can be used as a representative for the climatic zone where that station is located. The Ångström–Prescott coefficients calculated in this study could enable the estimation of daily global horizontal irradiance data at any location in South Africa, where daily observation sunshine duration data are available and the climate is correctly classified. The knowledge of estimated daily global horizontal irradiance data can thereby be used to support energy policies and solar energy programmes. They can also be used as benchmarking in climate analysis studies.

Author Contributions: Conceptualization, B.M., M.D.L. and S.T.M.; methodology, B.M. and M.D.L.; software, B.M.; validation, M.D.L. and S.T.M.; formal analysis, B.M.; investigation, B.M.; data curation, B.M.; writing—original draft preparation, B.M.; supervision, H.T., M.D.L. and S.J.M.; writing—review and editing, H.T., M.D.L., S.J.M., N.Z. and S.T.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the South African Weather Service.

Acknowledgments: The author thanks the South African Weather Services (SAWS) for providing the Global Horizontal Irradiance (*GHI*) data and sunshine duration data used in the study. The author thanks the South African Weather Services (SAWS) for granting a bursary to study Master of Science in Physics at University of South Africa

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

Nomenclature table with abbreviations, full description and units of symbols used in the text.

Abbreviation	Full Description	Units
SAWS	South African Weather Service	
AP	Ångström–Prescott	
а	Ångström–Prescott regression coefficient	
Ь	Ångström–Prescott regression coefficient	
GHI	Global Horizontal Irradiance	W/m ²
GHI _{TOA}	Daily extraterrestrial or Top of the atmosphere global horizontal irradiance	W/m ²
TOA	Top of the Atmosphere	W/m ²
Ν	Daily astronomical day length	Hours
п	Daily measured sunshine duration	Hours
BSRN	Baseline Solar Radiation Network	
QC	Quality control	
SPA	Solar Position Algorithm	
I _{SC}	Solar constant	W/m ²
Eo	Eccentricity factor	Degrees
ω_s	Sunset hour angle	Degrees
Ø	Degrees of latitude	Degrees
δ	Solar declination	Degrees
MBE	Mean Bias Error	$MJ m^{-2}d^{-1}$
rMBE	relative Mean Bias Error	Percentage (%)
MAE	Mean Absolute Error	$MJ m^{-2} d^{-1}$
rMAE	relative Mean Absolute Error	Percentage (%)
RMSE	Root Mean Square Error	$MJ m^{-2} d^{-1}$
rRMSE	relative Root Mean Square Error	Percentage (%)
R ²	Correlation coefficient	
NaN	Not a Number	
D	Day of the year	
CMP11	Secondary standard Kipp & Zonen pyranometers	

References

- Tang, W.; Yang, K.; He, J.; Qin, J. Quality control and estimation of global solar radiation in China. *Sol. Energy* 2010, *84*, 466–475. [CrossRef]
- 2. Adamala, S.; Reddy, Y.V.K. Evaluation of Different Solar Radiation Estimation Methods for Indian Locations. *Water Resour. Environ. Eng. II* **2018**, 47–56. [CrossRef]
- 3. Zhang, Q.; Cui, N.; Feng, Y.; Jia, Y.; Li, Z.; Gong, D. Comparative Analysis of Global Solar Radiation Models in Different Regions of China. *Adv. Meteorol.* **2018**, *2018*, 1–21. [CrossRef]
- 4. De Medeiros, F.J.; Silva, C.M.; Bezerra, B.G. Calibration of Ångström-Prescott Equation to Estimate Daily Solar Radiation on Rio Grande do Norte State, Brazil. *Rev. Bras. Meteorol.* **2017**, *32*, 409–416. [CrossRef]
- 5. Almorox, J.; Benito, M.; Hontoria, C. Estimation of monthly Angström-Prescott equation coefficients from measured daily data in Toledo, Spain. *Renew. Energy* **2005**, *30*, 931–936. [CrossRef]
- 6. Mabasa, M.B.; Botai, J.; Ntsangwane, M.L. Update on the Re-Establishment of the South African Weather Services (SAWS) Radiometric Network in All Six Climatological Regions and the Quality of the Data. In Proceedings of the South African Solar Energy Conference (SASEC), Blue Waters Hotel, KwaZulu-Natal, South Africa, 25–27 June 2018. Available online: https://www.sasec.org.za/full_papers/68.pdf (accessed on 20 January 2020).
- Mabasa, B.; Tazvinga, H.; Zwane, N.; Botai, J.; Ntsangwane, L. Assessment of Global Horizontal Irradiance in South Africa. In Proceedings of the South African Solar Energy Conference (SASEC), Mpekweni Beach Resort, Eastern Cape Province, South Africa, 25–27 November 2019. Available online: https: //www.sasec.org.za/papers2019/47.pdf (accessed on 25 January 2020).

- 8. Singh, J.; Kruger, A. Is the summer season losing potential for solar energy applications in South Africa? *J. Energy South. Afr.* **2017**, *28*, 52. [CrossRef]
- 9. Wilbert, S.; Stoffel, T.; Myers, D.; Wilcox, S.; Habte, A.; Vignola, F.; Wood, J.; Pomares, L.M. Measuring Solar Radiation and Relevant Atmospheric Parameters. In *Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications*; National Renewable Energy Laboratory: Golden, CO, USA, 2017. Available online: https://hal-mines-paristech.archives-ouvertes.fr/hal-01184753 (accessed on 12 March 2020).
- Garg, H.P.; Garg, S.N. Measurement of solar radiation—II. Calibration and standardization. *Renew. Energy* 1993, 3, 335–348. [CrossRef]
- 11. Angstrom, A. Solar and terrestrial radiation: Report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation. *Quart. J. Roy. Meteorol. Soc.* **1924**, *50*, 121–126. [CrossRef]
- 12. Prescott, J.A. Evaporation from a water surface in relation to solar radiation. *Trans. Roy. Soc. S. Aust.* **1940**, 46, 114–118.
- 13. Iqbal, M. An Introduction to Solar Radiation; Academic Press: Toronto, ON, Canada, 1983.
- 14. Martínez-Lozano, J.A.; Tena, F.; Onrubia, J.E.; De La Rubia, J. The historical evolution of the Angström formula and its modifications: A review and bibliography. *Agric. For. Meteorol.* **1984**, *33*, 109–128. [CrossRef]
- 15. Tsung, K.Y.; Tan, R.; Ii, G.Y. Estimating the global solar radiation in Putrajaya using the Angstrom-Prescott model. *IOP Conf. Series: Earth Environ. Sci.* **2019**, *268*, 012056. [CrossRef]
- 16. Besharat, F.; Dehghan, A.A.; Faghih, A.R. Empirical models for estimating global solar radiation: A review and case study. *Renew. Sustain. Energy Rev.* **2013**, *21*, 798–821. [CrossRef]
- 17. Eberhard, A.A. A Solar Radiation Data Handbook for Southern Africa; Elan Press: Cape Town, South Africa, 1990.
- 18. Malaudzi, S.T.; Sankaran, V.; Lysko, M.D. Solar radiation analysis and regression coefficients for the Vhembe Region, Limpopo Province, South Africa. *J. Energy South. Afr.* **2017**, *24*, 2–7. [CrossRef]
- 19. Maupin, K.A.; Swiler, L.P.; Porter, N.W. Validation Metrics for Deterministic and Probabilistic Data. *J. Verificat. Valid. Uncertain. Quantif.* **2018**, *3*, 031002. [CrossRef]
- Paulescu, M.; Stefu, N.; Calinoiu, D.; Tulcan-Paulescu, E.; Pop, N.; Boata, R.; Mares, O. Ångström–Prescott equation: Physical basis, empirical models and sensitivity analysis. *Renew. Sustain. Energy Rev.* 2016, 62, 495–506. [CrossRef]
- 21. Long, C.N.; Dutton, E.G. BSRN Global Network Recommended QC Tests, V2. 2010. Available online: https://epic.awi.de/30083/1/BSRN_recommended_QC_tests_V2.pdf (accessed on 11 December 2019).
- 22. Thomas, C.; Wey, E.; Blanc, P.; Wald, L. Validation of three satellite-derived databases of surface solar radiation using measurements performed at 42 stations in Brazil. *Adv. Sci. Res.* **2016**, *13*, 81–86. [CrossRef]
- Urraca, R.; Gracia-Amillo, A.M.; Huld, T.; Martinez-De-Pison, F.J.; Trentmann, J.; Lindfors, A.V.; Riihelä, A.; Sanz-Garcia, A. Quality control of global solar radiation data with satellite-based products. *Sol. Energy* 2017, 158, 49–62. [CrossRef]
- 24. Riihelä, A.; Kallio, V.; Devraj, S.; Sharma, A.; Lindfors, A.V. Validation of the SARAH-E Satellite-Based Surface Solar Radiation Estimates over India. *Remote Sens.* **2018**, *10*, 392. [CrossRef]
- 25. Roesch, A.; Wild, M.; Ohmura, A.; Dutton, E.G.; Long, C.N.; Zhang, T. Assessment of BSRN radiation records for the computation of monthly means. *Atmos. Meas. Tech.* **2011**, *4*, 339–354. [CrossRef]
- 26. Geiger, M.; Diabaté, L.; Ménard, L.; Wald, L. A web service for controlling the quality of measurements of global solar irradiation. *Sol. Energy* **2002**, *73*, 475–480. [CrossRef]
- 27. Holmgren, W.F.; Hansen, C.; Mikofski, M. Pvlib python: A python package for modeling solar energy systems. *J. Open Source Softw.* **2018**, *3*. [CrossRef]
- 28. Reda, I.; Andreas, A. Solar position algorithm for solar radiation applications. *Sol. Energy* **2004**, *76*, 577–589. [CrossRef]
- 29. Duffie, J.A.; Beckman, W.A. *Solar Engineering of Thermal Processes*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1980; pp. 6–13, 22.
- 30. Gueymard, C.A. The sun's total and spectral irradiance for solar energy applications and solar radiation models. *Sol. Energy* **2004**, *76*, 423–453. [CrossRef]
- 31. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. *Crop Evapotranspiration-Guidelines for Computing Crop Requirements-Irrigation and Drainage Paper 56*; FAO: Rome, Italy, 1998; Volume 300, p. 05109. Available online: https://appgeodb.nancy.inra.fr/biljou/pdf/Allen_FAO1998.pdf (accessed on 15 February 2020).

33. CMP11 Pyranometer. Available online: https://www.kippzonen.com/Product/13/CMP11-Pyranometer# .X1UGh-fivIV (accessed on 12 May 2020).

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).