Short Research Article

Correlation between Photovoltaic energy production and certain climate parameters: Case study in the Plateau Department in Southern Benin

ABSTRACT

The Plateau department, where the first 25 MWp grid-connected solar plant was installed, is also an industrial cement zone, with high-energy demand, located in the south of Benin. In this region, the equatorial climate oscillated between two dry and two rainy seasons, with a high relative humidity. This climate variability influences the electrical output of photovoltaic (PV) modules. The analysis of the impact of climatic parameters such as relative humidity, precipitation, wind speed, ambienttemperature, sunshine, on the photovoltaic production becomes necessary to optimize the energy generation in such a region. The objective of this study is to quantify the dependency relationship that exists between the variability of these parameters and the PV power generation using Pearson correlation method. The daily data collected for each parameter during the period from January 2011 to December 2020 were processed with Python 3.7.10 language. The results showed that relative humidity, with an average value of 80.14%, is the climatic has the highest negative impact (correlation coefficient of -0.42) on the performance of PV modules. Thus, the design and operation of a PV plant in this area should consider this parameter, especially with dust deposits, to improve the production yield.

Keywords: Equatorial, Pearson, Humidity, PV plant, Dust deposits

1. INTRODUCTION

The global energy demand is constantly growing. The rapid spread of photovoltaic (PV) technology allows electricity generation from solar radiation, even in rural areas where access to the electricity grid remains a challenge. The electrical output of PV modules is intermittent. Indeed, solar irradiance varies dynamically and randomly over time, from one country to the other. PV production is at its maximum when the solar irradiance is very high [1]. This irradiance is influenced by the climatic parameters which decrease its intensity [2]. This implies a variation in a PV cell's I-V (current-voltage) characteristic curve, thus decreasing its electrical output [3]. Climate variability is an important parameter to consider when planning an on/off-grid PV-based electrification project, especially in regions with significant among the various seasons. Benin has an equatorial climate with a single rainy season in the north and two rainy seasons in the south. The southern region is characterized by high humidity, two rainy seasons (April to mid-July and mid-September to October), interspersed with two dry seasons (November to March and mid-July to mid-September), and dry and hot wind from the Sahara [4]. The first ever 25 MWp grid-connected PV powerplant was built in the Plateau department, in the south of Benin [5].
To understand the impact of climate parameters such as rainfall, temperature, relative humidity, and wind speed on PV production, researchers often analyse, temporal data of phenomena or concepts using statistical methods. The collection of a series of data over a period of 3 months in the city of Sohar in Oman, on air temperature, relative humidity, and wind speed allowed determining a correlation with solar radiation by the Pearson method [2]. A study carried out in the province of Errachidia in Morocco, on

meteorological data over a period of 3 years, using the Pearson's coefficient, allowed to determine the parameter which best approaches the variation of solar energy [6]. The Pearson method is used to measure the linear correlation between two variables. The formula for calculating the correlation coefficient is integrated by default in most software and corresponds to light-tailed distributions [7]. Correlation studies have been performed in the literature but not to our knowledge in the equatorial climate in the plateau department with these natural phenomena, namely precipitation, humidity, ambienttemperature, wind speed, sunshine, and PV power output in the department of the plateau, in the South of Benin.

This study aims to determine the strength of the dependence relationship and the direction of the variability of these natural phenomena with the PV power output in the Plateau Department, using Pearson's coefficient. After this introduction, the plan for the remainder of this study includes a literature review, the material and methods used to determine the correlation coefficients, the discussions of the results obtained, and finally a conclusion.

2. STATE OF ART

Correlation is defined as an existing relationship between phenomena or objects or between mathematical or statistical variables that tend to vary, to be associated, or to occur together in a way that is not expected by chance alone [8].

There are several methods in the literature for qualifying this dependency relationship. The calculation of the Pearson correlation coefficient, developed in 1896 by Karl Pearson, allows for determining the linear relationship between two variables and the direction of progression of the values of statistical data. The Pearson coefficient is the most widely used and corresponds to hightailed distributions [7]. In 1904, Charles Sperman, introduced the calculation of the ρ (Rho) coefficient, often used when the correlation between variables is not linear. The ρ coefficient is a correlation of the ranks of the data values [7, 9]. Kendall's τ (Tau) coefficient, developed by Maurice Kendall in a paper in 1938, is used when the rank is repeated several times in a small data set. Kendall's τ is considered as an extension of Spearman's correlation [8].

Other methods of calculating correlation were derived from Pearson's method. These are the Point Biserial Correlation, and the φ (Phi) correlation respectively used if one of the variables is quantitative and the other variable is nominal and dichotomous and if both variables are nominal and dichotomous. If the two variables are continuous and normally distributed, the Tetrachoric correlation coefficient is used when the two variables are dichotomous as in the case of φ correlation [10]. The values of the correlation coefficients are between –1 and +1. Table 1 summarizes the interpretation of the results of the correlation [9,10].

Table 1. Interpretation of a correlation coefficient

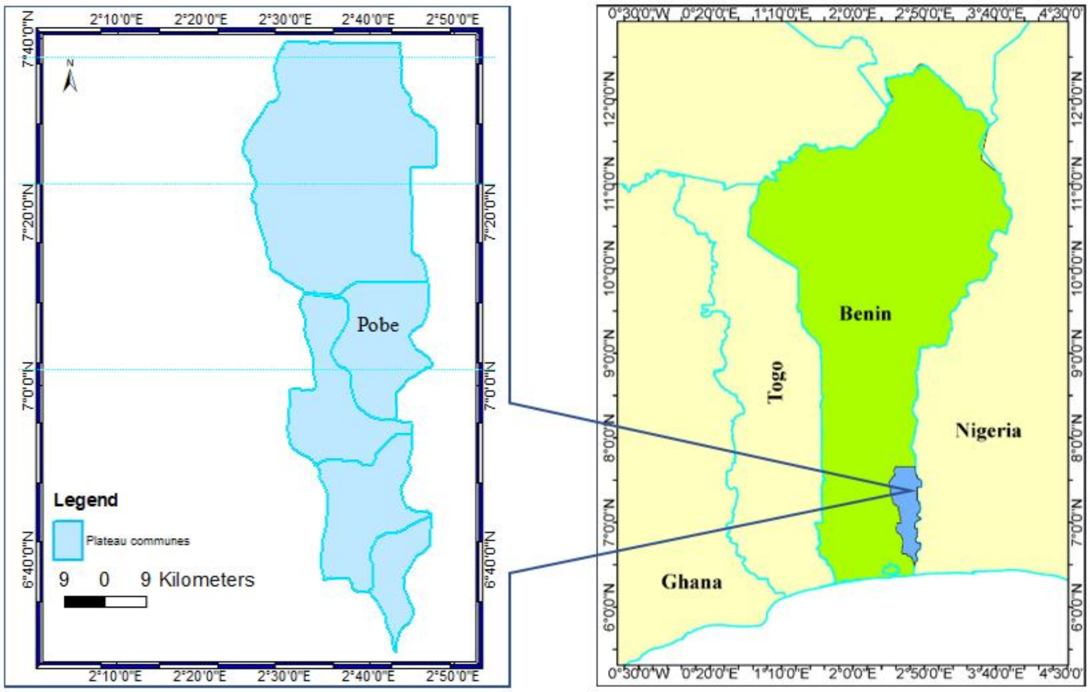
Absolute value of correlation coefficients	Interpretation
0	No Correlation
0.00 to 0.10	NegligibleCorrelation
0.10 to 0.39	Weak correlation
0.40 to 0.69	Moderatecorrelation
0.70 to 0.89	Strong correlation

3. MATERIALS AND METHODS

3.1. Description of study area

The Plateau Department is in the southeast of Benin and is bordered to the east by Nigeria. It is rich in minerals such as clay and limestone and attracts economic investors. Two clinker and cement factories have been set up. The energy demand is becoming more and more important in the area and the intermittent photovoltaic energy source, vulnerable to climate variability, is becoming the solution to this energy shortage. The plateau department has an area of 3264 Km², which is about 3% of the national territory. It is located between the meridians 2°23' and 2°47' East longitude and the parallels 6°32' and 7°47' North latitude. It is made up of 5 communes: Adja-Ouèrè, Ifangni, Kétou, Pobè, and Sakété. The meteorological data from the Météo-Bénin station used for this study is located in the commune of Pobè, capital of the Plateau department, at 2°40'East and 6°55'59'' North. Pobè is in the center of the Plateau department [11]. Figure 1 shows the location of the Plateau Department, with the commune of Pobe on the map of Benin.

This figure is obtained using ArcGIS 10.8 software with the shape data of Africa and Benin [12],[13].



(a) (b)

Figure 1. Study area of correlation of climatic parameters with PV output power (a) map of Plateau Department (b) Map of Benin

3.2. Study materials

The daily climate data used are from the period of january 2011 to december 2020, and are obtained from the national weather agency in Benin Meteo-Benin. The daily PV output power data obtained by simulation at a point in the area by a PV module over the same period, are obtained from the PVGIS-SARAH2 database using the PVGIS 5.2 software. The characteristics of the production simulation site are summarized in Table 2. Table 3, summarizes the studied parameters. All daily observations for each of the climate variables as well as the PV simulation output power result from 2011 to 2020, are processed with the Python 3.7.10 programming language. Details of the software used for data processing are recorded in Table 4.

Table 2: Parameters of the photovoltaic production data collection site

Site references	Values
PV technology	Crystalline silicon (c-Si)
Installed peak PV power	100 W
System loss	14%
Slope	10°
Azimuth	0°
Latitude	6.933
Longitude	2.667

Table 3: Study variables

Nr.	Data	Unit	Format	Data source
1	Global irradiance	kJ/cm^2	csv	Meteo-Benin
2	Temperature	° C	csv	
3	Wind speed	m/s	csv	
4	Relative Humidity	%HR	csv	
5	Precipitation	mm	csv	

6	PV output power	W	csv	PVgis 5.2

Table 4: Software list and data sources

Nr.	Softwares	Data format	Data sources
1	ArcGIS 10.8	Shapefile	https://www.diva-gis.org/gdata https://open.africa/dataset/africa-shapefiles
2	Anaconda navigator 1.10.0 with Jupiter Notebook 1.0.0 and Python 3.7.10 packages - Matplotlib 3.3.4 - Pandas 1.2.3 - NumPy 1.19.2	csv	Meteo-Benin
3	PVGIS 5.2	csv	PVGIS- SARAH2

3.3. Correlation analysis methodology

3.3.1. Processingflowchart

The methodology for analyzing the dependence of climate variability on PV production is presented in three essential steps. The first step in this correlation study is to determine the most representative site in the study area, with a weather station, in order to collect the corresponding data. The climatic data from 2011 to 2020, studied, were received from the Benin National Weather Agency, in charge of recording and preserving meteorological data at the national level. The estimated photovoltaic output of a 100W PV module over the same period was downloaded from the PVGIS-SAHA2 database, following the installation conditions specified in Table 2. The second step is to process the data using the Python 3.7.10 programming language in the Jupiter Notebook 1.0.0 environment under the Anaconda 1.10.0 browser. The Matplotlib library, is used to create high quality graphs to generate visualizations of data over 10 years of observations. With the Pandas package, the loading, manipulation and analysis of tables of observation data over its multiple time periods, were performed and the statistical calculations of monthly averages and Person correlation coefficients, using the Numpy package and the existing integrated function (1). The graphs are obtained using the results of the monthly averages of each parameter over the entire period. The superposition of the graphs on the same x-axis is made possible with the Subplot and Twinx functions. The last step is to display the graphs of the climate parameters with the PV output power, on the same x-axis as well as the Pearson correlation coefficients, allowing to visualize and appreciate the correlation relationship between the climate variability and the PV production. The data processing flow chart is shown in Figure2.

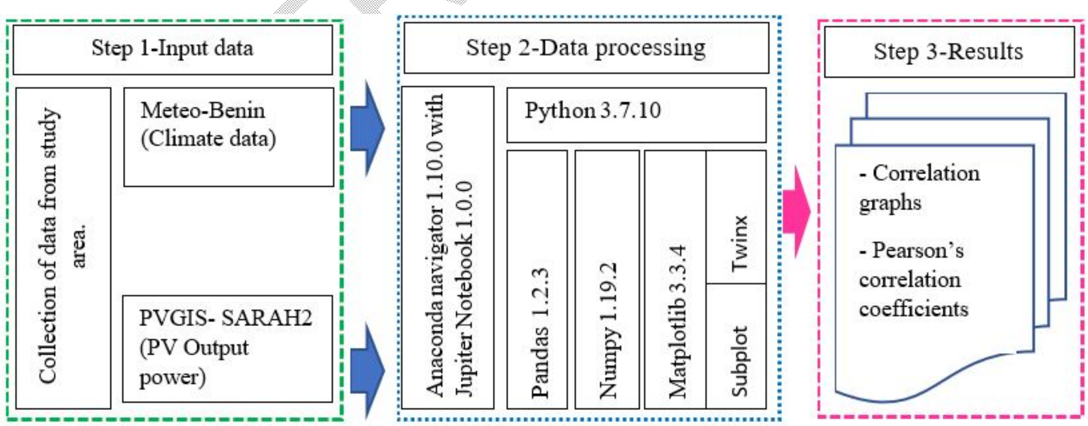


Figure 2. Data processing flow diagram

3.3.2. Pearson correlation

To measure the dependence between the variability of the parameters, the calculation of the correlation coefficient, makes it possible to appreciate the linear relation between these two parameters. The Pearson correlation coefficient is the most commonly used. It is often implemented by default in several statistical calculation software [7]. The absolute value of the coefficient quantifies the magnitude of the linear relationship between the two variables and the sign (+ or -), the direction of variation, by the following formula [6].

$$r_{xy} = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}\sqrt{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}} \qquad (1)$$

With, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, (2)

And, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, (3)

r_{xy}, is the Pearson's coefficient, with $-1 \leq r_{xy} \leq +1$.

This Pearson formula is integrated into the Python programming language, by the expression of the Corr function.

4. RESULTS AND DISCUSSION

The curves obtained (Figure 3, Figure 4) show a similarity of periodicity between PV power and the variation of irradiance and ambienttemperature over the 10-year period (2011 to 2020) of the daily data processed.The meteorological data used for this study are obtained from field measurements. The lowest PV production, according to the curves, is obtained around the middle of the year of each year, during the month of June, corresponding to the rainiest month of the year in southern Benin. The highest PV production occurs during the months of December to January, corresponding to the middle of the long dry season, from November to March. Each year, over the study period, an increase in the solar radiation leads to an increase in PV power and vice versa. Also, an increase in the ambient temperature leads to an increase in the output power. However, at a fixed level of radiation, an increase in temperature leads to a decrease in PV output power [14]. The variation of the monthly average values of sunshine over the study period is between 9.5 kJ/cm2 and 25.7 kJ/cm2 with an average of 18 kJ/cm2 and that of the ambienttemperature, between 23.10°C and 31.75°C with an average of 28.09°C (Table 5). Figure 3 and Figure 4 show the existence of a dependency relationship in the same direction of variation between photovoltaic production, irradiance and ambienttemperature.

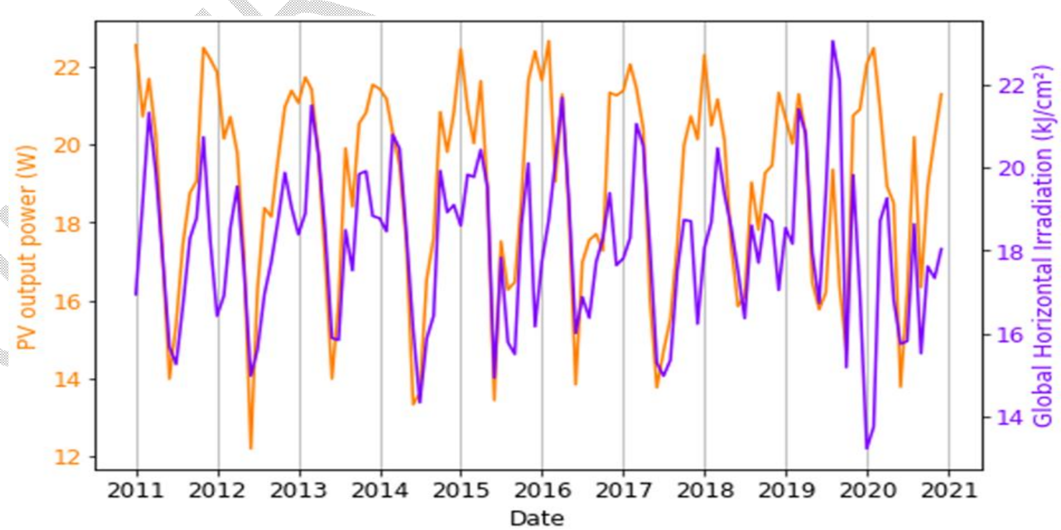


Figure 3. Correlation between PV output power and solar irradiance

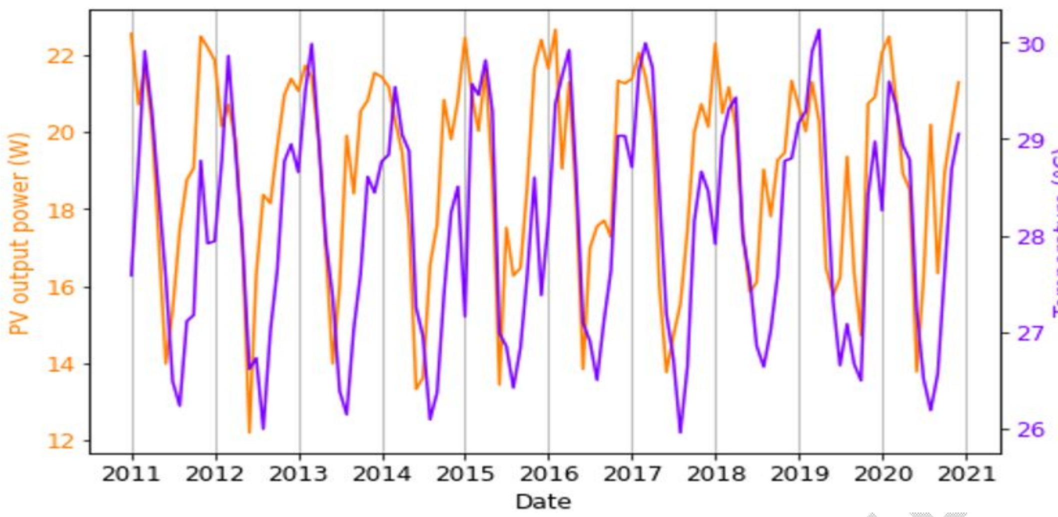


Figure 2. Correlation between PV output power and ambienttemperature

Figure 5 shows that wind speed variability does not have a significant influence on the PV power produced. The wind speed over the period varies from 0 m/s to 10.5 m/s with an average of 4 m/s (Table 5). On the other hand, the curves (Figure 6 and Figure 7) clearly show that the PV production decreases when the humidity and the amount of rainfall increase and conversely, the production increases when the humidity and the amount of rainfall decrease. The average relative humidity varies from 12% to 93.5% with a high monthly average value of 80.14% (Table 5). The amount of rainfall varies on average from 0 to 8.66 mm. The water droplets in the air change the direction of the light by the phenomenon of reflection, refraction or diffraction and thus decreases the intensity of the light captured by the PV cells and consequently the decrease of the PV production power. The statistical results of the study parameters are shown in Table 5.

Table 5: Summary of the values collected and averaged over the analysis period.

Study period values (10 years)	AmbientTemperature (°C)	Global irradiance (kJ/cm^2)	Wind speed (m/s)	Relative Humidity (%)	Precipitation (mm)
Minimum daily value	16.70	9.5	0	6.00	0
Miximum daily value	36.6	25.9	10.5	99.00	121.3
Minimum monthly average	23.10	13.21	2.71	12	0
Miximum monthly average	31.75	25.7	6.84	93.50	8.66
Average over 10 years	28.09	18	4	80.14	2.87

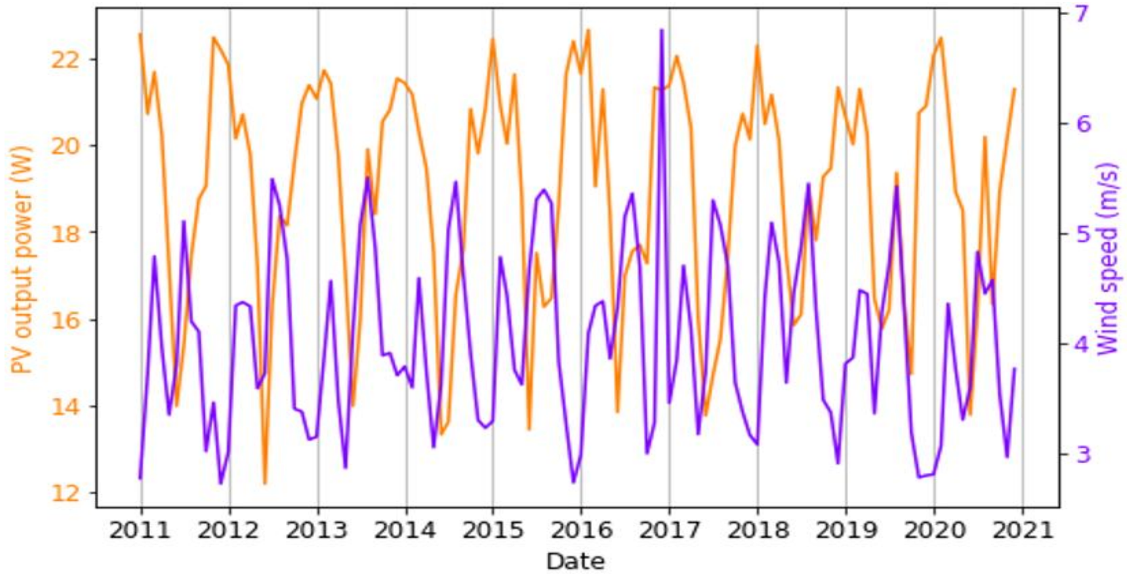


Figure 5. Correlation between PV output power and Wind speed

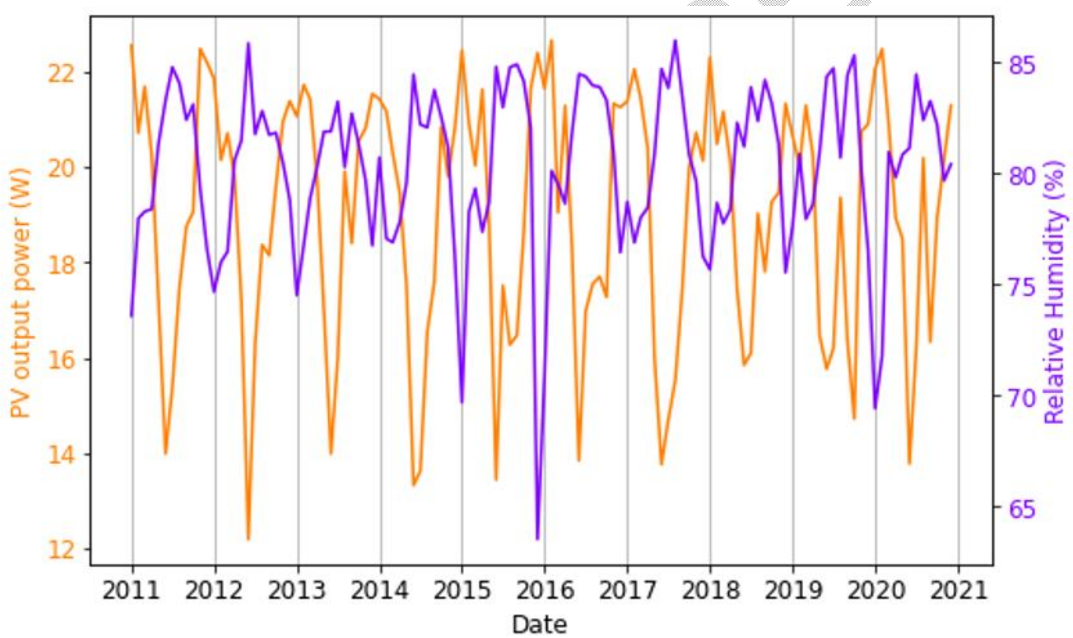


Figure 3. Correlation between PV output power and Relative humidity

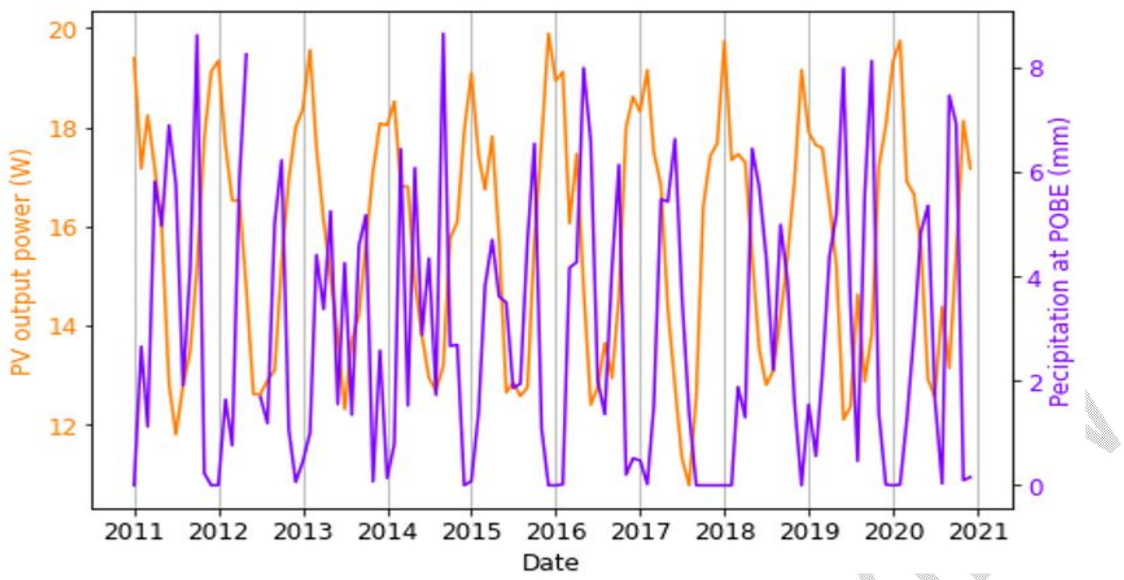


Figure 4: Correlation between PV output power and precipitation

The result of the Pearson correlation coefficient calculations (Table 6), gives positive values for solar radiation, ambienttemperature and negative values, for wind speed, precipitation and relative humidity. The positive values of the correlation coefficient r obtained confirm that there is a positive relationship between the solar radiation, the ambient temperature and the PV power output. The negative correlation coefficient for humidity, precipitation or wind speed increases denotes a negative relationship between these parameters and the PV power output. The magnitude of the correlation relationship r varies from one weather parameter to another and denotes of the strength of the relationship between the respective parameter and the PV power output. Solar radiation has a strong correlation (+0.76) with PV power output while ambienttemperature has a moderate correlation (+0.52). Relative humidity moderately decreases (-0.42) the PV output and precipitation has a relatively small negative influence (-0.22) on the energy produced. The effect of wind speed on PV production is negligible (-0.02). These negative correlation values show that relative humidity affects the performance of PV modules, more than precipitation, and wind speed in the study area according to the processed data.

Table 2: The Results of the Pearson correlation coefficients between the PV output power and the different variables

Criteria	Global Horizontal Irradiation	AmbientTemperature	Wind speed	Precipitation	Relative Humidity
Coefficient de Pearson	+ 0.76	+ 0.52	– 0.02	– 0.22	– 0.42

5. CONCLUSION

Photovoltaic production in southern Benin is influenced by climatic phenomena. The equatorial climate in the Plateau department, with two rainy seasons and humidity, decreases the performance of PV modules. The data processing analysis of the observations from Janvier 2011 to December 2020 using Python 3.7.10 programming language, shows that the power of the PV production increases when the sunshine intensity and the ambienttemperature increase. At a constant solar access value, the PV power decreases when the temperature increases. A reduction in sunshine andambient temperature causes a decrease in PV power output. Over the study period, the minimum and maximum values of sunshine and ambienttemperature are 9.5 kJ/cm^2 and 25.9kJ/cm^2; 16.70°C and 36.6°C respectively. The positive Pearson correlation values of solar irradiance (+0.76) and ambienttemperature (+0.52) show the magnitude and progression in the same direction as the PV power output. On the other hand, an increase in humidity and rainfall amount leads to a decrease in energy output and vice versa. The average values of relative humidity, rainfall amount and wind speed are 80.14%, 2.87 mm and 4 m/s respectively. The negative result of the Pearson correlation coefficient of relative humidity (-0.42), shows the extent of its negative influence on the photovoltaic performance especially in an industrial area of cement production with dust deposits.

REFERENCES

1. H. E. Colak, T. Memisoglu, and Y. Gercek, “Optimal site selection for solar photovoltaic (PV) power plants using GIS and AHP: A case study of Malatya Province, Turkey,” *Renew Energy*, vol. 149, pp. 565–576, Apr. 2020.
2. H. A. Kazem, M. T. Chaichan, and M. T. Chaichan, “Effect of humidity on photovoltaic performance based on experimental study,” *International Journal of Applied Engineering Research (IJAER)*, 2015, vol. 10, no 23, p. 43572-43577.
3. J. ho Choi, J. H. Hyun, W. Lee, B. G. Bhang, Y. K. Min, and H. K. Ahn, “Power performance of high-density photovoltaic module using energy balance model under high humidity environment,” *Solar Energy*, vol. 219, pp. 50–57, May 2021.
4. Présidence de la république du bénin, “Géographie du Bénin.” (2023) [online]: https://presidence.bj/home/le-benin/geographie/ (accessed Jan. 18, 2023).
5. Government of Benin, Actualite, (2022) [Online] Available: (https://www.gouv.bj/actualite/1858/inauguration-centrale-solaire-photovoltaique-25-illoulofin-benin-poursuit-marche-vers-autonomie-energetique/.).
6. I. Jebli, F. Z. Belouadha, M. I. Kabbaj, and A. Tilioua, “Prediction of solar energy guided by pearson correlation using machine learning,” *Energy*, vol. 224, Jun. 2021.
7. Joost C.F. de Winter, Samuel D. Gosling, and Jeff Potter, “Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data,” *Psychological methods*, 2016, vol. 21, no 3, p. 273.
8. AkogluHaldun. “User's guide to correlation coefficients,” *Turkish journal of emergency medicine*, 2018, vol. 18, no 3, p. 91-93.
9. Patrick Schober, Boer Christa, and Lothar A. Schwarte, “Correlation coefficients: appropriate use and interpretation,” *Anesthesia & Analgesia*, 2018, vol. 126, no 5, p. 1763-1768.
10. Esezi Isaac Obilor, and Eric ChikweruAmadi, “Test for significance of Pearson’s correlation coefficient,”. *International Journal of Innovative Mathematics, Statistics & Energy Policies*, 2018, vol. 6, no 1, p. 11-23.
11. Ministry of Decentralization and Local Governance, *Plateau*, (2023) [Online] Available: https://decentralisation.gouv.bj/prefecture/11/plateau/
12. OpenAFRICA, Africa Shapefiles, 2023 [Online] Available: (https://africaopendata.org/dataset/africa-shapefiles).
13. DIVA-GIS, *Download data by country*, 2022. [Online] Available: (https://www.diva-gis.org/gdata/).
14. C. Coskun, U. Toygar, O. Sarpdag, and Z. Oktay, “Sensitivity analysis of implicit correlations for photovoltaic module temperature: A review,” *Journal of Cleaner Production*, vol. 164. Elsevier Ltd, pp. 1474–1485, Oct. 15, 2017. doi: 10.1016/j.jclepro.2017.07.080.