

Multiplayer Perceptron and Simple Regression Linear Approaches to Predict Photovoltaic Active Power Plant: Case Study

APALOO BARA Komla Kpomonè^{1,2}, APEKE Kodjo Séna³,
PALANGA Eyouleki Tcheyi Gnadi^{1,2},
BEDJA Koffi-Sa³

¹Département de Génie Électrique, Ecole Polytechnique de Lomé (EPL), Université de Lomé, Togo

²Laboratoire de Recherche en Sciences de l'Ingénieur (LARSI), Université de Lomé, Togo

³Département de Génie Électrique, Institut de Formation Technique Supérieur, Togo

Corresponding Author: APALOO BARA Komla Kpomonè

DOI : <https://doi.org/10.52403/ijrr.20231207>

ABSTRACT

This article deals with the forecast of electricity production from a 50 MW photovoltaic power plant in Blitta, a town in Togo. The objective is to use meteorological variables such as instantaneous irradiation (A), wind speed (B), ambient temperature (C) and module temperature (D) to predict the active power. Multilayer Perceptron architecture, Artificial Neural Networks and multiple linear regression are explored as methods in Python. A classification of variables is presented. Certain model performance evaluation criteria made it possible to observe the results of the models. 26989 data samples are used. The results give a strong correlation between the ambient temperature of the location and the temperature of the module, i.e. 87% and 40% between the wind speed and the instantaneous irradiation. Also, as results we have: MAE = 6.017; MSE = 67.392; RMSE = 8.209; RRMSE = 15.185% and $R^2 = 55.321$ by multilayer perceptron and 60 neurons under the hidden layer then MAE = 6.93; MSE = 77.37; RMSE = 8.80; RRMSE = 51.42%, $R^2 = 69.97\%$ obtained by linear regression. This shows that there is a strong correlation between the variables used but the high values of RRMSE will drive the need to use other algorithms.

Keywords: Artificial Neural Networks, Correlation, Photovoltaic Active Power Plant, Simple Regression Linear, Multi-layer Perceptron,

INTRODUCTION

The planet's demand for energy continues to grow; on the other hand, the current means of producing this energy causes many problems for nature. We realize in the literature that the most commonly used means of production are from fossil sources (oil, coal, gas, etc.), [1], [2]. The proof lies in the climatic anomalies observed almost everywhere in the world, the heatwave [6], the bush fires here and there [3], the abnormal floods [4], etc.

The solution to all these problems, for the moment, can be seen in changes in attitude regarding the production of electrical energy. For this, photovoltaic solar energy, [5], [6], the safest and most common source in tropical areas, provides relief.

Togo, a country in humid and coastal West Africa, is not an oil or gas producer, [15] but for its energy needs it imports these fossil fuels. National energy potential only takes into account renewable energy sources. Renewable energies are those whose resources are only renewed over time because they use inexhaustible sources. Togo only produces barely 40% of its energy consumption [23].

In 2018, the minimum coverage rate was 40% with a share in rural areas between 15 and 20%. In fact, the real objective of

Togolese policy is to move towards global electrification by 2030, [23]. This saw the birth in 2021 of the Blitta photovoltaic solar power plant planned for a capacity of 50 MVA. In the meantime, there was the project for individual household power supply based on photovoltaic solar power, [7]. The “CIZO” project (which means turning on “Guin” in the local language) covers the entire extent of the Togolese territory and aims, by 2022, to have access to electricity through the supply of individual solar kits at affordable costs at more than of 2 million citizens (or around 300,000 households). The social component of the project plans to equip around 800 health centers and 3,000 small farms with individual solar or irrigation kits, [8].

In this context, a problem arises: the power of a solar installation is dependent on the weather, more precisely on the level of sunshine (diffuse, direct and indirect radiation), temperatures and seasons. This means that electricity production is not linear but variable, [9], [10]. It varies with the seasons and during the day. Hence the need to explore these meteorological variables in order to predict the instantaneous power harnessable in photovoltaic solar fields, particularly that of Blitta.

The validity of a modern forecast requires the exploitation of artificial intelligence algorithms [11] and the results provided by model performance evaluation criteria [20]. Several algorithms exist for modeling a system. We can cite Adaptive Neuro-Fuzzy Inference Systems (ANFIS) [12], AutoRegressive Integrated Moving Average

(ARIMA) [13], Support Vector Machines (SVM) [14], Artificial Neural Networks (ANN), [15], [20], [21], Ant Colony, Recurrent Neural Networks [16], Fuzzy Inference Systems (FIS) [17], Support Vector Machines [18], Genetic Algorithms [20], etc. These algorithms can be incorporated into languages like Matlab, Python, Anaconda etc.

This work aims to use Artificial Neural Networks (ANN) with Multilayer Perceptron (MPL), incorporating it into Python. Then we will use certain performance evaluation criteria such as the Average Absolute Error (MAE), the Mean Square of Errors (MSE), the Square Root of the Mean Quadratic Error (RMSE), the Square Root of the Error Relative Mean Quadratic Expressed (RRMSE) and Correlation Coefficient (R^2), [20], [21]. All this in order to predict the active power available in the Blitta photovoltaic solar power plant in the central region of Togo.

The aim is to facilitate the CEET network manager with the production strategy in order to effectively plan the consumption of electrical energy.

MATERIALS & METHODS

The data used in this work were collected at the Blitta solar power plant which is a town in Togo, in the central region. Blitta is approximately 266 km from Lomé, capital of Togo. It covers an area of 723 km², with a density of 76 inhabitants per km², [8]. Figure 1 shows the map of Togo and the central region with the town of Blitta.

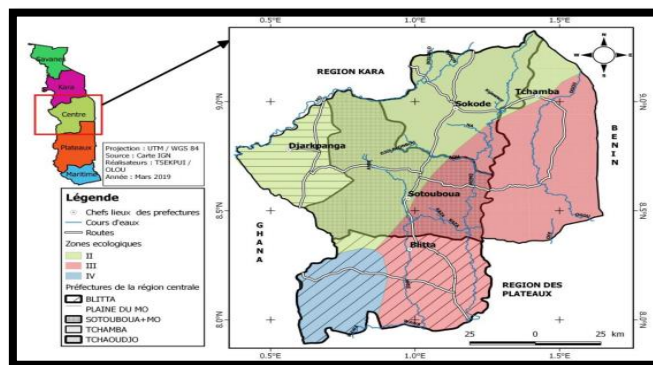


Figure 1: Central region of Togo containing the town of Blitta

Blitta is home to a photovoltaic solar power plant called Sheikh Mohamed Bin Zayed, built by AMEA Togo Solar, a subsidiary of AMEA Power, a company based in Dubai. It is made up of Jinko brand module 390 W (JKM-390M-72V) and 395 W (JKM-395M-72 V), 400 W (JKM-400M-72V), with 127,344 panels installed in total over an area of 252684 m², [22]. These modules have a power which varies between 390 and 400 Wp.

It is planned for a capacity of 50 MWp to generate approximately 90,255 MWh of energy per year, supplying energy to 158,333 Togolese homes per year, with 9% (approximately 8,123 MWh per year) of energy supplying the local distribution network. of Blitta, enough to meet demand in the region.

The power station is equipped with a device for automatic recording of active power, reactive power, solar irradiation, wind speed, ambient temperature of the environment and the temperature at the module level. It is made up of a photovoltaic field subdivided into 6 blocks, namely: 4 blocks of 30 MWp DC / 24 MW AC (at a rate of 6 MW each) and 2 blocks of 20 MWp DC / 16 MW AC (at a rate of 8 MW each). There is also a command and control room, a counting room, and an electric field including two 33 kV / 161 kV- 20 MWp transformers, connected to the line of the Electric Community of Benin.

Still at the level of the recording device, meteorological variables are also stored as shown in Figure 2, allowing for better planning.

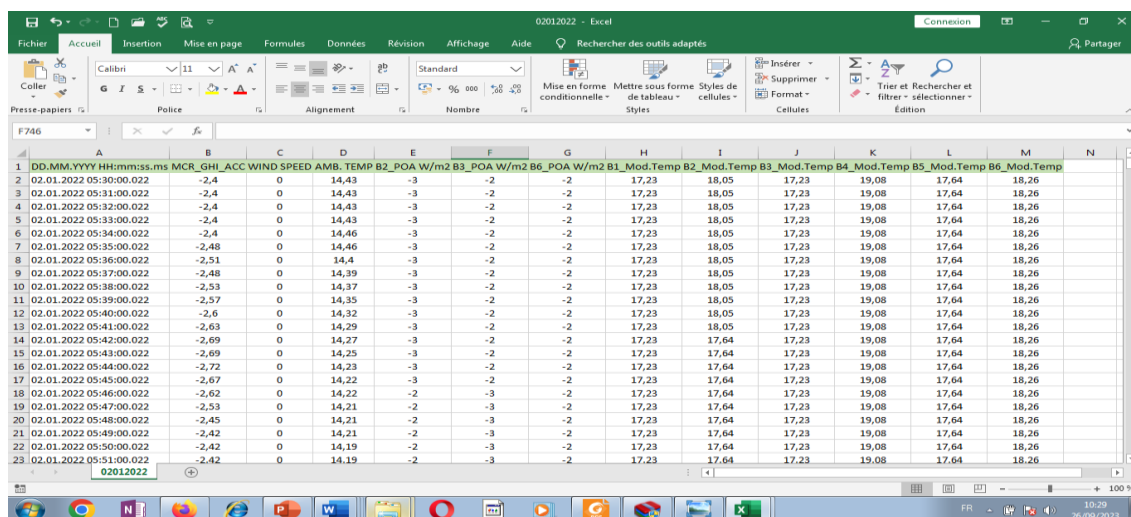


Figure 2: Weather Data Excel Processing Page

It is these data and variables that will be used in developing our planning.

Python is a free, interpreted, multi-paradigm and multi-platform programming language. We used it particularly because of its free license and its ease, especially since the codes are inspired by and close to C++. It promotes structured, functional and object-oriented imperative programming. It works on most computer platforms (smartphones, computer) and uses several operating systems (Windows, Unix, MacOS, Android, IOS, etc.). These codes can be easily translated into Java or .NET. Python is designed to optimize programmer

productivity by offering high-level tools and easy-to-use syntax.

When it comes to data prediction and analysis, several libraries must be incorporated into the Python language to facilitate certain functionalities. We can cite:

- Numpy which allows you to simply and efficiently create and manipulate matrices with ease;
- Matplotlib which facilitates the editing of graphs and diagrams;
- Scikit-learn which offers access to the codes of several ranges of machine learning algorithms, data

preprocessing tools, model evaluation criteria;

- Pandas which offers ease and flexibility in data analysis.

Artificial neural networks are often used for classification and pattern recognition. In this work, we use it for its forecasting ability. La sortie d'un réseau de neurones prend en compte la procédure d'apprentissage. Le processus d'apprentissage est basé sur la rétropropagation de l'erreur. Sa sortie est exprimée comme présentée à la relation (1), [20], [21] :

$$O_k = \sum_{j=1}^q W_{kj} b_j(x) - \theta_k \quad (1)$$

Where:

- $1 < k < m; m =$ The number of the nodes
- $O_k =$ The output of k_{th} node of the output layer

- $W_{kj} =$ The connection between the j_{th} neuron of hidden layer and k_{th} neuron of output layer
- $b_j(x) =$ The output of the j_{th} neurone of the hidden layer
- $\theta_k =$ The bias of the k_{th} neurone output layer

The architecture of the Multilayer Perceptron (MLP) model is illustrated in Figure 3. The result of this model is given by relation (2):

$$y = \beta_0 + \sum_{i=1}^n \beta_i h_i \quad (2)$$

Where

- $y =$ The predicted value with the neural network
- $n =$ The number of hidden layers
- $\beta_0 =$ The bias
- $\beta_i =$ The weighted coefficients
- $h_i =$ The result of the non-linear transformation of the i_{th} hidden unit

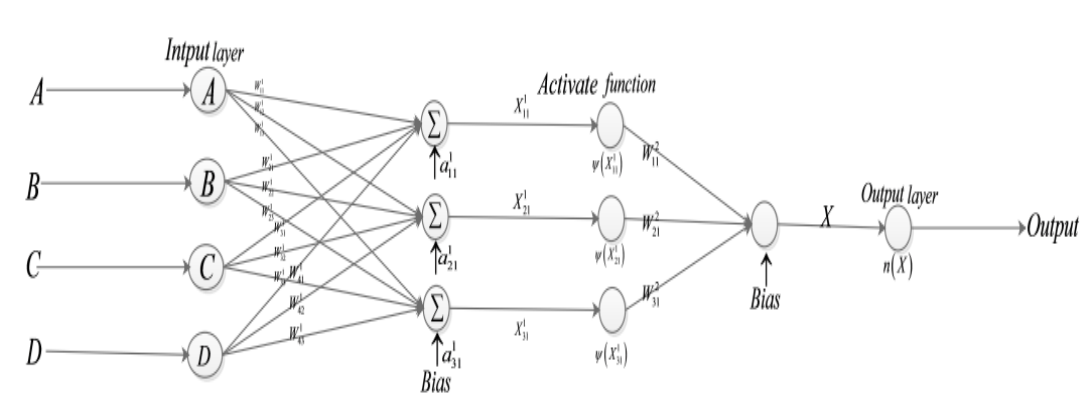


Figure 3: Neural architecture of the MLP model, [21]

The validity and effectiveness of ANN modeling requires the organization of data into 3 groups (training, validation and test). For this work, the distribution carried out is:

70% for training, 15% for validation and 15% for testing, all of which constitutes a total of 26989 data used. Figure 4 shows the data distribution for the study.

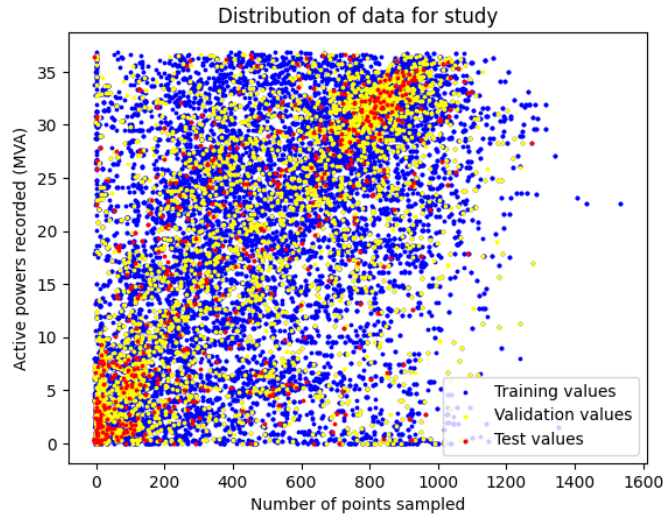


Figure 4: Distribution of study values

Linear regression is a supervised learning method used to model the relationship between a continuous dependent variable and one or more independent variables. It is based on the hypothesis that the dependent variable can be approximated by a linear combination of independent variables, with a certain margin of error. The linear regression model therefore tries to find the coefficients that minimize this margin of error, by adjusting the regression right to the training data.

The linear regression model can be formulated mathematically by expression (3), [24].

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad (3)$$

Where is the dependent variable:

- x_1, x_2, \dots, x_n are the independent variables;

- b_1, b_2, \dots, b_n are the coefficients of the model.

The linear regression algorithm seeks to estimate the values of the regression coefficients which minimize the error between the real values of the dependent variable and the values predicted by the model. This estimate is generally carried out using the least square method, which aims to minimize the sum of the squares of the differences between real values and predicted values. Once the regression coefficients are estimated, the linear regression model can be used to predict the values of the dependent variable using the values of the independent variables.

In order to facilitate the presentation of the results, we have carried out coding grouped in table 1.

Table 1: Coding table of input variables

Input variables	Input variables Code	Output Variable
Instantaneous irradiation	A	Active power
Wind speed	B	
Ambient temperature of the place	C	
module temperature	D	

From Table 1 we constituted the configurations for the design of the model. Table 2 sets out the models to consider.

Table 2: Configuration of combinations for the model

Combination numbers	Configurations associated with models
1	[AB]
2	[AC]
3	[AD]
4	[BC]

5	[BD]
6	[CD]
7	[ABC]
8	[ABD]
9	[ACD]
10	[BCD]
11	[ABCD]

The combinations of models are subject to performance evaluation criteria MAE (3), MSE (4), RMSE (5), RRMSE (6) and R² (7) which will allow us to choose the best model for prediction [20], [21].

$$MAE = \frac{1}{N} \sum_{j=1}^N |p_{p_i} - p_{m_i}| \quad (3)$$

$$MSE = \frac{1}{N} \sum_{j=1}^N (p_{p_i} - p_{m_i})^2 \quad (4)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^N (p_{p_i} - p_{m_i})^2} \quad (5)$$

$$RRMSE = \frac{\sqrt{\frac{1}{N} \sum_{j=1}^N (p_{p_i} - p_{m_i})^2}}{\sum_{j=1}^N p_{m_i}} \quad (6)$$

$$R^2 = \frac{\sum_{j=1}^N (p_{p_i} - p_{m_p}) * (p_{m_i} - p_{m_m})}{\sqrt{\left[\sum_{j=1}^N (p_{p_i} - p_{m_p})^2 \right] * \left[\sum_{j=1}^N (p_{m_i} - p_{m_m})^2 \right]}} \quad (7)$$

Where :

- p_{p_i} is the predicted power
- p_{m_i} is the measured power
- p_{m_p} is the predicted average power
- p_{m_m} is the average power measured
- N is the number of points sampled

RESULTS

Before proceeding with the modeling of our system, a characterization of the variables is carried out. Tables 3 to 8 summarize the variables studied. The statistical parameters taken into account in this characterization are the mean, the mode, the median, the minimum, the maximum, the standard deviation, skewness and kurtosis.

Table 3: Monthly characteristics of irradiation data in W/m²

	Mean	Mode	Median	Minimum	Maximum	Standard deviation	Skewness	Kurtosis
January	478.53	-3	580	-4	1005	315.31	-0.38	-1.42
February	496.43	-2	600.5	-4	984	325.86	-0.4	-1.41
March	473.08	-2	536	-4	1036	333.56	-0.18	-1.51
April	341.86	0	338.69	-4	931	269.27	0.24	-1.21
May	457.1	0	382	-1	1277	355.97	0.19	-1.46
June	422.61	0	346	-1	1195	338.76	0.4	-1.12
July	340.14	0	264	-2	1292	315.58	0.97	0.02
August	333.29	0	251.5	-1	1315	317.76	1.06	0.25
September	328.66	0	237	-1	1408	322.77	1.07	0.33
October	475.73	-1	402	-2	1534	383.17	0.25	-1.35
November	549.39	-2	708	-3	1055	353.5	-0.46	-1.41
Décember	522.11	-2	676	-4	959	331.31	-0.54	-1.38

Table 4: Monthly characteristics of speed data in km/h

	Mean	Mode	Median	Minimum	Maximum	Standard deviation	Skewness	Kurtosis
January	1.53	0	1.4	0	6.19	1.01	0.59	0.58
February	1.6	0	1.48	0	5.39	1.14	0.51	-0.02
March	1.62	0	1.48	0	6.19	1.08	0.64	0.67
April	1.09	0	1.07	0	4.35	0.83	0.69	0.64
May	1.69	0	1.56	0	6.27	1.06	0.93	1.98
June	1.7	0	1.56	0	6.59	1.11	0.96	1.63
July	1.73	0	1.64	0	6.19	1.25	0.56	0.07
August	1.64	0	1.56	0	6.03	1.32	0.53	-0.24
September	1.33	0	1.32	0	4.35	0.97	0.32	-0.45
October	1.49	0	1.32	0	6.91	1.08	0.77	1.46
November	1.59	0	1.48	0	5.39	1.06	0.59	0.2
Décember	1.6	0	1.48	0	4.83	1.01	0.49	-0.11

Table 5: Monthly characteristic of ambient temperature data in °C

	Mean	Mode	Median	Minimum	Maximum	Standard deviation	Skewness	Kurtosis
January	25.77	18.62	27.64	12.96	34.83	6.46	-0.43	-1.28
February	27.62	19.63	29.7	12.31	38.07	6.74	-0.48	-1.05
March	26.48	31.09	27.45	12.31	38.07	5.84	-0.39	-0.96
April	22.07	18.62	21.93	2.17	34.29	5.67	0.02	-0.71
May	26.26	19.82	26.91	18.7	34.09	4.26	-0.14	-1.28
June	24.55	19.36	24.99	17.02	32.59	4.04	-0.11	-1.29
July	23.75	26.98	24.29	17	29.79	3.02	-0.18	-1.19
August	23.32	19.93	23.81	17.12	30.84	3.13	-0.09	-1.06
September	23.32	19.76	23.53	17.88	30.81	3.12	0.14	-1.04
October	25.06	18.97	25.49	17	33.16	4.34	-0.12	-1.27
November	26.72	29.76	29.16	13.29	34.82	5.84	-0.76	-0.8
Décember	25.91	32.58	28.71	10.98	34.23	6.84	-0.66	-0.96

Table 6: Monthly characteristic of module temperature data in °C

	Mean	Mode	Median	Minimum	Maximum	Standard deviation	Skewness	Kurtosis
January	41.28	53.74	46.15	13.95	61.95	13.05	-0.66	-1
February	44.37	55.18	49.64	13.33	62.97	13.83	-0.71	-0.95
March	43.06	54.36	46.56	14.36	67.08	13.3	-0.45	-1.11
April	33.98	20.31	35.28	1	58.65	11.95	-0.14	-1
May	43.62	25.64	45.13	23.18	69.13	12.69	-0.11	-1.39
June	41.07	25.03	42.05	22.15	66.26	11.92	-0.03	-1.31
July	38.78	26.05	38.87	20.72	67.08	10.37	0.2	-1.09
August	38.01	25.44	37.74	22.77	64.82	10.04	0.29	-0.93
September	37.87	26.67	36.31	22.97	65.85	10.58	0.46	-0.92
October	42.06	25.03	44.31	20.51	69.13	12.5	-0.12	-1.41
November	44.25	53.13	50.05	16.41	73.23	13.22	-0.75	-0.83
Décember	42.47	54.15	48.62	13.13	63.18	13.71	-0.82	-0.78

Table 7: Monthly characteristics of active power data in MW

	Mean	Mode	Median	Minimum	Maximum	Standard deviation	Skewness	Kurtosis
January	18.51	0.3	22.71	0	36.35	12.64	-0.36	-1.5
February	18.94	0.36	23.79	0.08	34.74	12.53	-0.41	-1.44
March	18.2	0.37	21.17	0.05	35.85	13.26	-0.19	-1.62
April	18.71	0.34	20.98	0	36.85	13.43	-0.1	-1.58
May	17.48	0.26	18.09	0	36.54	12.23	0.01	-1.45
June	17.49	0.01	16.49	0	36.69	11.75	0.03	-1.36
July	13.87	0.25	13.82	0	36.21	10.58	0.26	-1.27
August	13.62	0.28	12.4	0	35.94	10.67	0.33	-1.14
September	13.21	0.27	10.99	0.08	35.84	10.5	0.38	-1.14
October	17.65	0	19.04	0	36.79	12.5	-0.04	-1.54
November	19.53	0	24.39	0	35.91	13.07	-0.28	-1.61
Décember	19.93	0.31	24.76	0.28	34.37	12.19	-0.46	-1.46

Table 8: Monthly characteristics of reactive power data in MVAR

	Mean	Mode	Median	Minimum	Maximum	Standard deviation	Skewness	Kurtosis
January	4.08	0.42	3.87	0.07	3.43	4.82	3.7	-1.27
February	3.11	0.38	3.08	0.02	7.68	2.37	0.19	-1.31
March	3.25	0.4	2.7	0.03	10.33	2.81	0.37	-1.29
April	3.61	0.41	3.22	0	10.3	3.1	0.34	-1.37
May	2.86	0.09	2.19	0	1.25	2.59	0.74	0.17
June	2.86	0.01	2.14	0	11.93	2.54	0.72	-0.37
July	1.75	0.11	1.14	0	6.42	1.76	0.93	-0.31
August	2.14	0.22	1.19	0	9.51	2.19	1.06	0.15
September	1.96	0.14	0.91	0.09	11.93	2.12	1.27	1.1
October	2.96	0.13	2.39	0	10.2	2.72	0.53	-1.05
November	3.2	0	3.46	0	8.81	2.56	0.16	-1.35
Décember	3.16	0.24	3.51	0.07	8.53	2.41	0.06	-1.44

Figure 5 presents the correlation map between the variables and Figure 6 illustrates the characterization results.

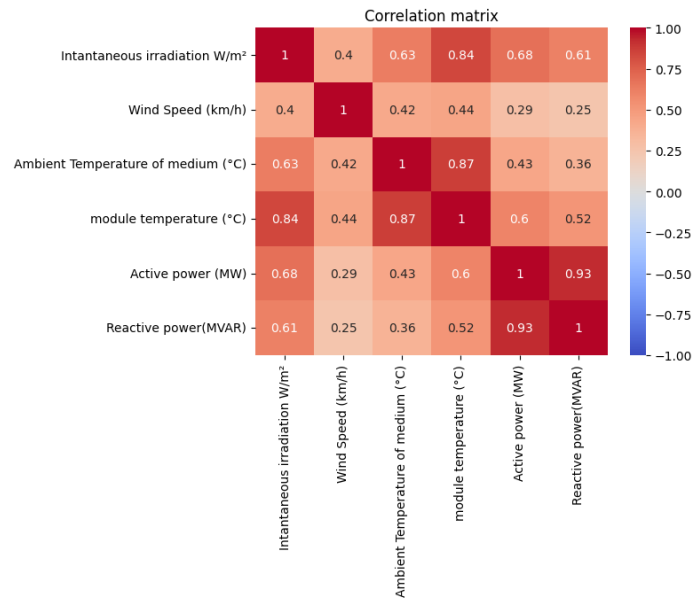


Figure 5: Correlation map between variables

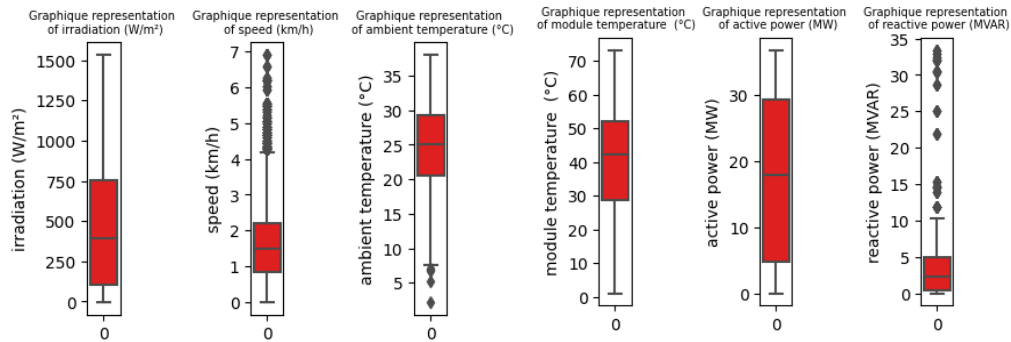


Figure 6: Graphical characterization of the variables studied

The results obtained by configuration are presented in tables from 9 to 19.

Table 9: Modeling results using the [AB] configuration

Number of neurons under the hidden layer	MAE	MSE	RMSE	RRMSE (%)	R ² (%)
5	6.6299	73.7785	8.5894	50.7076	51.0877
10	6.6121	73.7916	8.5902	50.7121	51.0790
15	6.8467	75.6310	8.6966	51.3403	49.8596
20	6.4591	74.0800	8.6070	50.8111	50.8878
*25	6.4626	73.2349	8.5577	50.5205	51.4481
30	6.6244	74.0105	8.6029	50.7873	50.9339
40	6.9069	76.2481	8.7320	51.5493	49.4505
50	6.5774	73.5846	8.5781	50.6409	51.2163
60	6.3685	75.3111	8.6782	51.2316	50.0717
70	6.7243	74.1077	8.6086	50.8206	50.8695
80	6.8152	75.3272	8.6791	51.2371	50.0610
90	6.4256	73.3645	8.5653	50.5652	51.3622
100	6.4685	73.4577	8.5707	50.5973	51.3004
10	6.6713	74.3347	8.6218	50.8984	50.7189
130	6.4522	73.9816	8.6013	50.7774	50.9530
150	6.6534	74.8210	8.6499	51.0646	50.3965
150	6.6534	74.8210	8.6499	51.0646	50.3965

Table 10: Modeling results using the [AC] configuration

Number of neurons under the hidden layer	MAE	MSE	RMSE	RRMSE (%)	R ² (%)
5	6.5925	76.8490	8.7664	51.7520	49.0521
10	6.3116	74.2751	8.6183	50.8780	50.7584
15	6.2527	72.1966	8.4969	50.1611	52.1364
20	6.2407	75.1132	8.6668	51.1642	50.2028
25	6.3213	72.2609	8.5006	50.1834	52.0938

30	6.1871	73.4440	8.5699	50.5926	51.3094
40	6.4576	72.3295	8.5047	50.2072	52.0483
50	6.2185	72.2647	8.5009	50.1847	52.0913
60	6.5164	73.1803	8.5545	50.5016	51.4843
70	6.3892	71.7281	8.4692	49.9981	52.4470
80	6.3458	72.3159	8.5039	50.2025	52.0573
90	6.5379	72.2840	8.5020	50.1914	52.0785
100	6.3370	71.7595	8.4711	50.0090	52.4262
*110	6.2583	71.7135	8.4684	49.9930	52.4567
120	6.4455	73.1097	8.5504	50.4773	51.5311
130	6.8773	78.2240	8.8444	52.2130	48.1405
150	6.1617	72.3567	8.5063	50.2166	52.0303

Table 11: Modeling results using the [AD] configuration

Number of neurons under the hidden layer	MAE	MSE	RMSE	RRMSE (%)	R ² (%)
5	6.6355	72.7194	8.5276	50.3424	51.7898
10	6.1863	73.6565	8.5823	50.6657	51.1686
15	7.3011	82.2458	9.0689	53.5384	45.4742
20	6.6342	72.6711	8.5247	50.3256	51.8218
25	6.2943	71.6594	8.4652	49.9741	52.4926
30	6.1136	76.6627	8.7557	51.6893	49.1756
40	6.3005	70.9678	8.4242	49.7324	52.9510
50	6.3775	71.1385	8.4344	49.7921	52.8379
60	6.1833	71.7232	8.4690	49.9963	52.4503
70	6.3913	71.4631	8.4536	49.9056	52.6227
80	6.3472	70.9680	8.4242	49.7324	52.9510
*90	6.3574	70.8919	8.4197	49.7058	53.0014
100	6.2332	72.0514	8.4883	50.1106	52.2327
110	6.4026	70.9002	8.4202	49.7087	52.9959
120	6.2586	71.0379	8.4284	49.7569	52.9046
130	6.0823	73.3981	8.5673	50.5767	51.3399
150	6.4383	71.4037	8.4501	49.8849	52.6621

Table 12: Modeling results using the [BC] configuration

Number of neurons under the hidden layer	MAE	MSE	RMSE	RRMSE (%)	R ² (%)
5	9.4493	120.6230	10.9829	64.8371	20.0316
10	9.3424	119.1946	10.9176	64.4520	20.9786
15	9.1492	116.8504	10.8097	63.8151	22.5327
20	9.2315	117.4965	10.8396	63.9913	22.1043
25	9.1280	116.5123	10.7941	63.7227	22.7569
30	9.2376	116.1650	10.7780	63.6277	22.9871
40	9.1783	115.8205	10.7620	63.5333	23.2155
50	9.1684	115.8218	10.7621	63.5336	23.2146
*60	9.1457	114.3613	10.6940	63.1317	24.1829
70	9.1190	114.8577	10.7172	63.2686	23.8538
80	9.2664	116.1558	10.7776	63.6252	22.9931
90	9.1076	115.4663	10.7455	63.4360	23.4503
100	9.1322	115.2232	10.7342	63.3692	23.6114
110	9.1629	117.8815	10.8573	64.0961	21.8491
120	9.2120	114.7285	10.7111	63.2330	23.9394
130	9.1033	117.5263	10.8410	63.9994	22.0846
150	9.2752	116.4218	10.7899	63.6980	22.8168

Table 13: Modeling results using the [BD] configuration

Nombre de neurones sous la couche cachée	MAE	MSE	RMSE	RRMSE (%)	R ² (%)
5	8.8109	109.4969	10.4641	61.7745	27.4078
10	8.8462	109.4839	10.4635	61.7708	27.4164
15	8.5083	103.7757	10.1870	60.1390	31.2007
20	8.8055	109.3966	10.4593	61.7462	27.4743
25	8.5123	103.2363	10.1605	59.9825	31.5583
30	8.5069	103.4464	10.1709	60.0435	31.4190
40	8.4391	103.5465	10.1758	60.0725	31.3527
50	8.4788	102.9639	10.1471	59.9033	31.7389
60	8.4787	103.4176	10.1694	60.0352	31.4381
70	8.5020	103.0366	10.1507	59.9245	31.6907
80	8.5165	104.1752	10.2066	60.2546	30.9359
90	8.4619	103.1635	10.1569	59.9614	31.6066
100	8.3408	102.6324	10.1308	59.8068	31.9587
110	8.4892	103.6761	10.1821	60.1101	31.2668
120	8.4469	103.6634	10.1815	60.1064	31.2752

*130	8.4778	102.5807	10.1282	59.7918	31.9929
150	8.5193	103.0455	10.1511	59.9270	31.6848

Table 14: Modeling results using the [CD] configuration

Number of neurons under the hidden layer	MAE	MSE	RMSE	RRMSE (%)	R ² (%)
5	8.5576	103.7783	10.1872	60.1397	31.1990
10	7.9496	94.6193	9.7272	57.4246	37.2711
15	7.9869	93.7518	9.6825	57.1608	37.8462
20	7.8954	92.8095	9.6338	56.8728	38.4708
25	7.8761	92.7422	9.6303	56.8522	38.5155
30	7.8573	92.6214	9.6240	56.8151	38.5956
40	7.9480	92.5480	9.6202	56.7926	38.6443
50	7.8935	93.5630	9.6728	57.1032	37.9714
60	7.8818	92.5824	9.6220	56.8032	38.6214
70	7.9409	92.1554	9.5998	56.6720	38.9045
80	7.7816	92.5604	9.6208	56.7964	38.6360
*90	7.7868	91.8150	9.5820	56.5673	39.1302
100	7.9461	94.1196	9.7015	57.2728	37.6023
110	7.8696	91.9653	9.5899	56.6136	39.0305
120	7.9172	92.5958	9.6227	56.8073	38.6125
130	7.9558	92.9267	9.6399	56.9087	38.3931
150	7.9064	91.9292	9.5880	56.6024	39.0544

Table 15: Modeling results using the [ABC] configuration

Number of neurons under the hidden layer	MAE	MSE	RMSE	RRMSE (%)	R ² (%)
5	11.1227	150.9084	12.2845	72.5212	46.4782
10	8.9336	125.2172	11.1900	66.0603	16.9858
15	6.5154	73.3779	8.5661	50.5698	51.3533
20	6.7325	74.4625	8.6292	50.9421	50.6342
*25	6.3984	71.6650	8.4655	49.9760	52.4889
30	6.8411	74.7107	8.6435	51.0270	50.4697
40	6.0857	72.9228	8.5395	50.4127	51.6550
50	6.1948	72.7648	8.5302	50.3581	51.7597
60	6.4882	72.1764	8.4957	50.1541	52.1498
70	6.2276	71.7381	8.4698	50.0015	52.4404
80	6.5028	71.9372	8.4816	50.0709	52.3084
90	6.5025	73.8521	8.5937	50.7329	51.0389
100	6.3590	72.2401	8.4994	50.1762	52.1076
110	6.2962	71.8530	8.4766	50.0416	52.3642
120	6.4643	71.8955	8.4791	50.0564	52.3360
130	6.1160	73.6026	8.5792	50.6472	51.2043
150	6.2939	72.0617	8.4889	50.1142	52.2258

Table 16: Modeling results using the [ABD] configuration

Number of neurons under the hidden layer	MAE	MSE	RMSE	RRMSE (%)	R ² (%)
5	6.6719	88.4900	9.4069	55.5335	41.3346
10	6.2020	73.7381	8.5871	50.6937	51.1145
15	6.3668	71.4813	8.4547	49.9119	52.6107
20	6.4669	71.3053	8.4442	49.8505	52.7273
25	6.4113	71.6489	8.4646	49.9704	52.4995
30	6.1282	72.3861	8.5080	50.2268	52.0108
40	6.7238	73.0874	8.5491	50.4696	51.5459
50	6.3823	70.8965	8.4200	49.7074	52.9984
60	6.3393	71.0458	8.4289	49.7597	52.8994
70	6.5438	71.3700	8.4481	49.8731	52.6845
80	6.4710	71.0823	8.4310	49.7725	52.8751
90	6.1433	71.0254	8.4277	49.7525	52.9129
100	6.1881	71.7481	8.4704	50.0050	52.4338
*110	6.3461	70.7162	8.4093	49.6441	53.1179
120	6.8001	72.7038	8.5267	50.3369	51.8002
130	6.2480	71.3690	8.4480	49.8727	52.6851
150	6.4645	71.5573	8.4592	49.9385	52.5603

Table 17: Modeling results using the [ACD] configuration

Number of neurons under the hidden layer	MAE	MSE	RMSE	RRMSE (%)	R ² (%)
5	6.4834	72.0855	8.4903	50.1225	52.2101
10	6.6036	73.1172	8.5509	50.4799	51.5261
15	6.1221	68.3838	8.2694	48.8186	54.6642
20	6.3034	69.3843	8.3297	49.1744	54.0009
25	6.3445	68.4149	8.2713	48.8296	54.6436

30	6.1691	68.5259	8.2780	48.8693	54.5699
*40	5.9753	67.6962	8.2278	48.5725	55.1201
50	6.0099	67.7726	8.2324	48.5999	55.0693
60	6.2193	68.7736	8.2930	48.9575	54.4057
70	6.5132	69.3993	8.3306	49.1797	53.9910
80	6.4958	70.4836	8.3954	49.5624	53.2721
90	6.2117	68.6020	8.2826	48.8964	54.5195
100	6.3110	68.4680	8.2745	48.8486	54.6084
110	6.2152	67.9765	8.2448	48.6730	54.9342
120	6.1000	67.9557	8.2435	48.6655	54.9480
130	6.3746	69.2693	8.3228	49.1336	54.0771
150	6.3726	68.6758	8.2871	48.9227	54.4706

Table 18: Modeling results using the [BCD] configuration

Number of neurons under the hidden layer	MAE	MSE	RMSE	RRMSE (%)	R ² (%)
5	8.4964	102.4985	10.1242	59.7678	32.0474
10	7.9270	93.5418	9.6717	57.0967	37.9854
15	7.8606	92.4596	9.6156	56.7655	38.7029
20	7.8997	92.2646	9.6054	56.7056	38.8321
25	7.8709	91.7284	9.5775	56.5406	39.1876
30	7.7760	91.1940	9.5496	56.3756	39.5419
40	7.7082	90.5006	9.5132	56.1609	40.0016
50	7.7772	91.6566	9.5737	56.5185	39.2352
60	7.7495	91.0329	9.5411	56.3258	39.6487
*70	7.6266	90.1043	9.4923	56.0378	40.2643
80	7.8048	91.5593	9.5687	56.4884	39.2997
90	7.7880	90.7724	9.5275	56.2452	39.8214
100	7.8432	91.5649	9.5690	56.4902	39.2960
110	7.6742	90.9213	9.5353	56.2913	39.7227
120	7.8576	91.4218	9.5615	56.4460	39.3909
130	7.7101	90.2777	9.5015	56.0917	40.1493
150	7.6401	91.3671	9.5586	56.4291	39.4271

Table 19: Modeling results using the [ABCD] configuration

Number of neurons under the hidden layer	MAE	MSE	RMSE	RRMSE	R ²
5	6.4233	71.8063	8.4739	50.0253%	52.3952%
10	6.6037	72.2899	8.5023	50.1935%	52.0746%
15	6.9683	77.7964	8.8202	52.0701%	48.4240%
20	6.0095	71.3405	8.4463	49.8628%	52.7040%
25	6.3151	68.0552	8.2496	48.7011%	54.8820%
30	5.9837	68.5788	8.2812	48.8881%	54.5349%
40	6.1884	68.8909	8.3001	48.9992%	54.3280%
50	5.9351	67.5124	8.2166	48.5065%	55.2419%
*60	6.0178	67.3921	8.2093	48.4633%	55.3217%
70	6.0435	67.8271	8.2357	48.6195%	55.0332%
80	6.3464	69.3785	8.3294	49.1724%	54.0047%
90	5.8851	68.5348	8.2786	48.8724%	54.5641%
100	6.4360	71.1565	8.4354	49.7984%	52.8260%
110	6.2496	68.6157	8.2835	48.9013%	54.5105%
130	6.2793	68.7189	8.2897	48.9380%	54.4420%
150	6.1943	68.4850	8.2756	48.8547%	54.5971%

The best performance results of each model are grouped in table 20.

Table 20: Summary of best performances by configuration

Configuration	Nombre de neurones sous la couche cachée	MAE	MSE	RMSE	RRMSE (%)	R ² (%)
[AB]	25	6.4626	73.2349	8.5577	50.5205	51.4481
[AC]	110	6.2583	71.7135	8.4684	49.9930	52.4567
[AD]	90	6.3574	70.8919	8.4197	49.7058	53.0014
[BC]	130	8.4778	102.5807	10.1282	59.7918	31.9929
[BD]	90	7.7868	91.815	9.582	56.5673	39.1302
[CD]	25	6.3984	71.665	8.4655	49.976	52.4889
[ABC]	110	6.3461	70.7162	8.4093	49.6441	53.1179
[ABD]	40	5.9753	67.6962	8.2278	48.5725	55.1201
[ACD]	70	7.6266	90.1043	9.4923	56.0378	40.2643
[ABCD]	*60	6.0178	67.3921	8.2093	48.46%	55.3217

For the best result obtained by the ABCD configuration, Figure 7 presents a superposition of the measured values compared to the predicted values.

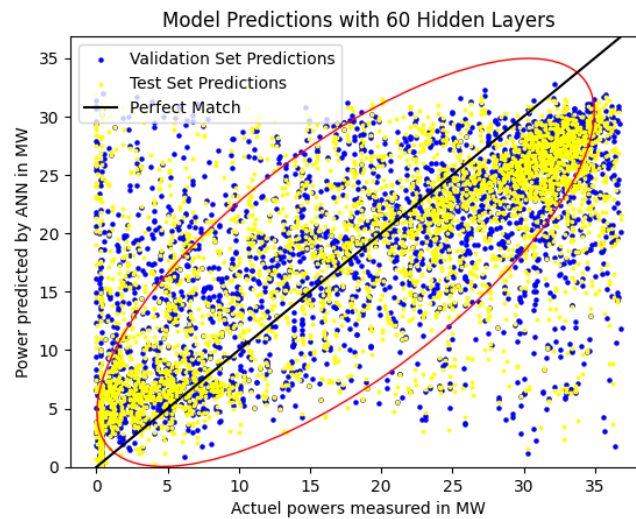


Figure 7: Visualization of predicted values versus measured values

The results obtained by simple linear regressions did not take into account the different configurations, given its algorithm. All the variables considered for the study are implemented at the same time. The results obtained by performance evaluation criteria considered are grouped in Table 21. Figure 8 exhibits graphic visualization.

Table 21: Summary of the results obtained by simple linear regression

Performance assessment criteria	MAE	MSE	RMSE	RRMSE (%)	R ² (%)
Results	6.93	77.37	8.80	51.42	69.97

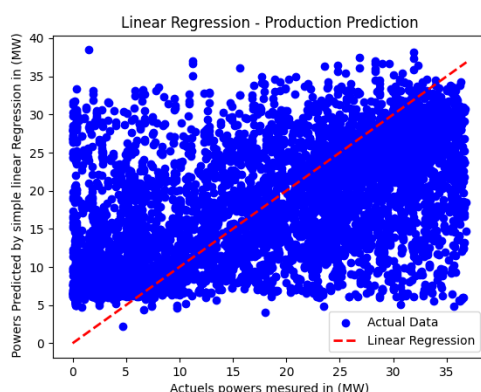


Figure 8: Graphical view of active powers predicted by Simple Linear Regression compared to actual measured values

DISCUSSION

To meet the energy needs of populations and the economy, the government's ambition is to

reduce dependence on electrical energy from 50% in 2015 to 35% by 2022, to bring the rate of access to electricity to the level national from 36% in 2016 to 60% in 2022, to reduce the rate of losses on the network from 16.8% to 10% by 2022 and to improve the carbonization efficiency from 15% to 25% in 2022, [23].

The data processed in this work are those collected between 5 a.m. and 5 p.m. Given that, during other periods of the day, solar irradiation is almost zero. There is virtually no active power production. Looking at Figure 7, the disparity in the data is very visible, which explains a fairly low correlation ($R^2 = 55.3217\%$). Weather conditions can change suddenly from one moment to the next, creating unpleasant effects on irradiation, speed, ambient temperature and module temperature. This shows the random nature of meteorological variables, which are difficult to identify in forecast analyses.

Indeed, the result of the best performances by the multilayer architecture of artificial neural networks: MAE = 6.0178; MSE = 67.3921; RMSE = 8.2093; RRMSE = 48.4633%; $R^2 = 55.3217\%$; obtained through the ABCD configuration justifies that all the parameters considered have a direct effect on the production of electrical energy from solar

photovoltaic energy in the Blitta solar photovoltaic power plant. These results are confirmed by the prediction with simple linear regression which gives: MAE = 6.93; MSE = 77.37; RMSE = 8.80; RRMSE = 51.42%, $R^2 = 69.97\%$ which constitutes the best result on the two algorithms used, just if we consider the correlation coefficient.

on the other hand, if we take into account the square root of the relative mean square error, we will conclude that the results are very bad because these values are greater than 30 %. Confirmation of this observation remains linked to the correlation table in Figure 5.

However, it should be emphasized that the disparity in the power produced can easily be filled by the storage batteries installed in the power plant. Additionally, the one-year period of variable collection should also be reviewed to see the improvement or status of the performance evaluation criteria. It should be added that this study will have to be repeated a few years later to reassess the effectiveness of the model.

CONCLUSION

The work accumulated in this document focused on modeling the production of photovoltaic solar energy from meteorological data such as instantaneous irradiation (A), wind speed (B), ambient temperature of the location (C) and the temperature of the modules (D). The Multilayer Perceptron architecture of Artificial Neural Networks operated in a Python environment is the first method used. The second method is simple linear regression always in the same environment. We took into account a classification of the aforementioned variables as well as the active and reactive power recorded in the plant. A correlation study between the variables is also presented. The data samples used in our study come from the photovoltaic power plant of Blitta, a town located in the central region of Togo.

Codifications were carried out on the variables in order to facilitate the presentation of the models to be studied. Thanks to some performance evaluation

criteria for models such as MAE, MSE, RMSE, RRMSE, and R^2 we were able to observe the results of each configuration of the models and then of each algorithm.

Following this, the results obtained with the performance evaluation criteria give for artificial neural networks: MAE = 6.017; MSE = 67.392; RMSE = 8.209; RRMSE = 15.185% and $R^2 = 55.321\%$. This is the best result obtained with 60 neurons under the hidden layer by the ABCD configuration. This means that all variables have direct effects on the production of active power in the Blitta solar photovoltaic plant. This statement is confirmed by the results obtained by simple linear regression giving: MAE = 6.93; MSE = 77.37; RMSE = 8.80; RRMSE = 51.42%, $R^2 = 69.97\%$. Contrary to these results, we find a rather unfavorable model with the ABC configuration, with the multilayer perceptron architecture of artificial neural networks using 10 neurons under the hidden layer and the performances give: MAE = 8.933; MSE = 125.217; RMSE = 11.190; RRMSE = 66.060% and an $R^2 = 16.985\%$.

Taking these results into account makes it possible to confirm that simple linear regression and artificial neural networks, using multilayer perception, are suitable for modeling the production of active power in the Blitta photovoltaic solar power plant. Meteorological variables measured at the plant level, recommended by a renowned company, are well suited to the implementation of this model. However, simple linear regression gives a higher correlation than neural networks. On the other hand, the square root of the relative mean square error of 51.42% for neural networks and 66.06% for simple linear regression shows that the results are bad.

For this, it would be necessary to explore other architectures of neural networks or other algorithms to check if the results obtained with the performance evaluation criteria cannot be improved. We also recommend taking into account, for future studies, a database that extends over more than one year.

Declaration by Authors

Acknowledgement: None

Source of Funding: None

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

REFERENCES

1. Smith J. A., Johnson M. B. and Williams C. D., "Fossil Fuel-Based Production of Chemicals and Polymers", *Environmental Science & Technology*, 2011, Vol. 45, n°15, pp. 6709-6717.
2. Brown Alice, Martinez Carlos, "Impacts des émissions de gaz à effet de serre provenant des énergies fossiles sur l'intensification des vagues de chaleur" *Revue du changement climatique*, 2022, Vol. 20, Numéro 4, pp. 250-265.
3. Allard, M., Michaud, Y., Ruz, M.H. et Héquette, A. "Ice foot, freeze-thaw of sediments, and platform erosion in a subarctic microtidal environment, Manitousuk Strait, northern Quebec, Canada", *Canadian Journal of Earth Sciences*, Février 2011, vol. 35, 1998, p. 965–979.
4. Wahiba Menad, Johnny Douvinet, Gérard Beltrando et Gilles Arnaud-Fassetta, Evaluer l'influence de l'urbanisation face à un aléa météorologique remarquable : les inondations des 9-10 novembre 2001 à Babel-Oued (Alger, Algérie), *Open Edition journals, revue trimestrielle Géomorphologie : Relief, Processus, Environnement*, 2012, Vol. 18, p. 337-350
5. Chun Sing Lai et Malcolm D. McCulloch, « Levelized cost of electricity for solar photovoltaic and electrical energy storage », *Applied Energy*, vol. 190, 15 mars 2017, p. 191-203.
6. Green, M. A.; Emery, K.; Hishikawa, Y.; Warta, W, "Solar cell efficiency tables (version 45)", *Progress in Photovoltaics: Research and Applications* , article , 2021, Vol 29, p. 3-15
7. Sabas SONOU AGOSSOU, Moussaoui TASSOU, études de faisabilité technico-économique et environnementale stratégiques pour l'électrification rurale décentralisée par minicentrales solaires photovoltaïques de 317 localités au Togo, 2019
8. Ministère des mines et des énergies, agence togolaise d'électrification rurale et des énergies renouvelables (at2er), rapport final, Projet d'Electrification Rurale CIZO, Togo, juillet 2019
9. ZERDOUDI A., CHENNI R, "étude de l'influence des différents paramètres sur un module photovoltaïque", *Sciences & Technologie A – N°41*, Juin 2015 , p. 49-54.
10. BENSALÉM S. "Effets de la température sur les paramètres caractéristiques des cellules solaires", *Mémoire de Master*, Université Fehat Abbas Sétif, 2009
11. Bellahsen A., "L'intelligence artificielle au service de l'optimisation de l'énergie électrique dans un réseau intelligent", *Mémoire de Maîtrise*, Ecole Polytechnique de Montréal, 2020
12. Jang J.-S. R, "ANFIS: Adaptive-network-based fuzzy inference system", *IEEE Transactions on Systems, Man, and Cybernetics*, 1993, Vol 23, n°3, p. 665-685
13. Box, G. E. P., Jenkins, G. M., & Reinsel, G. C., *Analyse des séries temporelles : La prévision et la modélisation des processus stochastiques*. Dunod, 2008.
14. Rifai, Salah; Dhibi, Farès; Boukadoum, Mounir, "Apprentissage par machines à vecteurs de support (SVM) : théorie et applications", *Technique et science informatiques (Volume 27, Numéro 6, pp. 725-752)*, 2008.
15. Azzouz Elhassan, Gagnaire Pierre, "Introduction aux réseaux de neurones artificiels", *Technique et science informatiques (Volume 16, Numéro 3, pp. 385-405)*, 1997.
16. El Abed, Haikal; Mokbel, Chafic; Gilloux, Michel, "Les réseaux de neurones récurrents : une application à la reconnaissance de l'écriture manuscrite", *Traitement du Signal (Volume 14, Numéro 5, pp. 387-398)*, 1997.
17. Ross, T. J., *Fuzzy Logic with Engineering Applications*. Wiley, 2010.
18. Cortes, C., & Vapnik, V., Support-vector networks. *Machine Learning*, 20(3), pp. 273-297, 1995.
19. E. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, 1989.
20. APALOO BARA Komla K., "Modélisation de la résistivité électrique des sols dans les zones tropicales par les méthodes d'analyse des données imparfaites : étude et réalisation d'un outil d'aide à la caractérisation et à la prédiction des résistivités électriques au Togo", *Thèse de Doctorat Unique en*

- Sciences de l'Ingénieur, Option Génie Electrique, Université de Lomé, 17 septembre 2020.
21. Komla Kpomonè Apaloo-Bara and al., "Estimation of Soils Electrical Resistivity using Artificial Neural Network Approach", American Journal of Applied Sciences, 2019, (16) 2, pp. 43-58.
 22. Joel Musikingala, Marina Plaza, AMEA POWER, "Etude d'impact environnemental et social du projet d'installation d'une centrale solaire de 50 MW à Blitta-Losso, dans le canton de Blitta-Village (préfecture de Blitta)" Office Space 3301 Dubai, United Arab Emirates, P.O. Box : 37669 ; +971 4 222 2499, 2021, p. 16.
 23. République Togolaise, "Plan National de Développement (PND) 2018-2022", 03 août 2018.
 24. Mark Tranmer, Jen Murphy, Mark Elliot and Maria pampaka, "Multiple Regression linear", 2nd Edition, Cathie Marsh Institute Working Paper 2020-01, January 2020.
 25. Magalie Fromont Renoir, "Modèles de régression linéaire", Master Statistique Appliquée, Mention Statistique pour l'Entreprise, Université Rennes, consulté le 05 septembre 2023.

How to cite this article: APALOO BARA Komla Kpomonè, APEKE Kodjo Séna, PALANGA Eyouleki Tcheyi Gnadi, BEDJA Koffi-Sa. Multiplayer perceptron and simple regression linear approaches to predict photovoltaic active power plant: case study. *International Journal of Research and Review*. 2023; 10(12): 42-56. DOI: <https://doi.org/10.52403/ijrr.20231207>
