INTERNATIONAL JOURNAL of RENEWABLE ENERGY RESEARCH
S. Huaquipaco et al., Vol.12, No.2, June 2022

Cross-validation of the operation of photovoltaic systems connected to the grid in extreme conditions of the highlands above 3800 meters above sea level.

Saul Huaquipaco [1‡], W. N. Macêdo [2], Henry Pizarro [3]

Reynaldo Condori [4], Jose Ramos [5], Oscar Vera [6]

Jose Cruz [7], Wilson Mamani [8], Norman Beltran [9]

[1]Postgraduate unit of EPG-UNA-PUNO, Faculty of Mecánica Eléctrica, Electrónica y Sistemas, Universidad Nacional del Altiplano de Puno. P.O. Box 291, Puno-Perú

[2]Grupo GEDAE, Universidade Federal de Pará (UFPA) P.O. BOX 6010, Belém-PA 66.025-972, Brazil

[3 4] School of Ingeniería en Energías Renovables, Faculty of Engineering, Universidad Nacional de Juliaca, Puno, Perú.

[5 9] School of Ingeniería Mecánica Eléctrica, Faculty of Mecánica Eléctrica, Electrónica y Sistemas, Universidad Nacional del Altiplano de Puno. P.O. Box 291, Puno-Perú

[7 8] School of Ingeniería Electrónica, Faculty of Mecánica Eléctrica, Electrónica y Sistemas, Universidad Nacional del Altiplano de Puno. P.O. Box 291, Puno-Perú

[6] School of Ingeniería de Sistemas, Faculty of Engineering, Universidad Nacional de Moquegua, Moquegua-Perú

(saul@pizdii.com, wmacedo@ufpa.br, h.pizarrov@gmail.com, rcondori@unaj.edu.pe, jramos@unap.edu.pe, overar@unam.edu.pe , josecruz@unap.edu.pe, wilmamanimac@est.unap.edu, nbeltran@unaj.edu.pe)

[‡] Corresponding Author; Saul Huaquipaco, Postgraduate unit of EPG-UNA-PUNO, Faculty of Mecánica Eléctrica, Electrónica y Sistemas, Universidad Nacional del Altiplano de Puno. P.O. Box 291, Puno-Perú, Tel: +51 977247764,

saul@pizdii.com

Received: 13.04.2022 Accepted: 06.05.2022

Abstract- Climatological factors influence the performance of grid-connected photovoltaic systems (PV). These factors vary according to the altitude above sea level. The present work aimed to compare the performance of a PV with a string inverter (String PV) versus a PV with DC-DC power optimizers (DC DC PV) at 3800 meters above sea level (m.a.s.l.) under natural conditions. For this purpose, two PV of 3 kW each were installed, and their performance was measured under the IEC 62053 standard. Then multiparametric regression models were applied for each of them. The results were Validate through linearity, normality of the error terms, correlation, autocorrelation and homoscedasticity. Subsequently, the cross-validation of both models has performed, whose results showed that the DC-DC PV has a better result in 6.09% over the String PV model, so we conclude that the DC-DC Pv converter performs better at 3800 m.a.s.l.

Keywords Cross-validation, Photovoltaic Systems, Extreme Conditions.

1. Introduction

Solar energy is an inexhaustible source of clean energy and Peru, due to its geographical location close to the equator and its topography due to the Andes, has radiation levels that vary from 5.5 to 6.5 kWh/m²; which favors the capture of this type of energy. Some of the factors that affect the quantity and quality of the energy captured in addition to climatic factors [1] [2] [3] are the type and location of the inverters and energy optimizers along with the way in which this energy is injected into the electrical grid to increase efficiency, reduce costs and generate as little pollution as possible. [4] One of the cross-cutting factors in energy harvesting and its integration into the power grid is the variability of insolation and weather conditions. [5] The amount of energy generated is directly proportional to the amount of insolation or sunlight at any given time, [6] which leads to under- or overgeneration phenomena that cause instability in the grid. According to [7] [8] [9] [10], ways to solve this problem are: using better forecasting tools to make forecasts that reduce these effects, increasing the amount of solar panels in large geographic areas to reduce the effects of cloud cover, switching to traditional electricity service to reduce these effects, and finally incentivizing customers to match their energy demand to the generation process. Therefore, in the present study, the implementation of two PV power generation models was carried out: the traditional one with an array of solar panels with a single DC to AC converter and the second one with the implementation of DC amplifiers, before injecting power into the DC. AC converter for grid integration. To determine which system is more efficient, we used cross-validation on two generated regression models. The present study was conducted in the city of Juliaca located at 3824 m.a.s.l. in the department of Puno Peru.

As a background, it is indicated that [11] for the implementation in a 31-level asymmetrical switching diode based multilevel DC link inverter uses a disturbance and observation-based voltage regulator with a capacitor compensating circuit (DC) managing to deliver 97.21% of the theoretical maximum power to the system. [12] It also proposes a robust current (DC) controller and a DC link voltage controller based on the μ and H ∞ synthesis method to ensure system robustness against weak grid uncertainties, as well as to compensate for system delays arising from computation, pulse width modulation, and zero-order hold. It also seeks to minimize the bus voltage fluctuations caused by variations in the power generation of the photovoltaic system, indicating that the results of the simulation and experiments in controlled environment are auspicious. On the other hand [13] proposes a simulation topology to achieve a better control of solar panels in which the voltage-current characteristic can be reproduced in the desired atmospheric conditions, such as solar radiation, temperature and partial shade; as well as it can be applied to high power panels with an open circuit voltage of 1000 V and a short circuit current of 16.5 A, finally indicating that the results obtained through MatLab Simulink are verified. Similarly [14] [15] using MatLab Simulink to test the results of a proposed transformer less solar PV inverter system integrating a solar panel, battery, DC link, DC load and the AC grid with a 300W resistive load using DC-DC to control the power flow between the battery and the DC rail by controlling the flow using fuzzy logic. Also, [16] [17] proposes a photovoltaic system to charge batteries through solar panels using a three-stage constant current/constant voltage (CC/CV) decentralized charging strategy and a DC-DC step-down converter with maximum power point tracking (MPPT) to charge these battery cells, achieving an 88.37% utilization of the generated energy, this result is also simulated. In another type of article [18] proposes two types of optimal charge controllers, adjusting their parameters by genetic optimization using neural networks and a Sugeno-type fuzzy logic controller (SFLC), applied to constant current voltages (CCCV) for charging photovoltaic battery systems evaluating it under conditions of rapid climatic changes, obtaining improvements over the classical on/off control system. On the other hand, for the extraction of the fundamental component of the static load current for a distribution static compensator [19], proposes a control technique based on a complex variable filter and a generalized second-order integrator composed of a two-stage system. consisting of a full-bridge boost converter cascaded with an inverter, again the experiment is performed under controlled conditions. [20] proposes a solar charge controller to power a DC motor and a rechargeable battery of an electric vehicle, for which he proposes a controller operating in two modes: constant current (DC) charging mode and battery tracking mode. maximum power point (MPPT) indicating that in the first mode what is sought are short charging times, while in the second is to extract the maximum power. [21] proposes a fractional open-loop MPPT tracking algorithm for improved photovoltaic applications. [22] proposes a phase-shifted interleaved modulation scheme for a step-down step-up converter to avoid the disadvantages of dc-dc conversion, using interleaved modulation to reduce the inductor current; it is tested on a 150W prototype indicating that it has feasible results. [23] proposes the collection, modeling and prediction of a multivariate SFV, using a multiparametric regression model, presenting five regression models with machine learning: three using shrinkage regularization and two using eXtreme Gradient Boosting (XGBoost). [24] [25] On the other hand [AC] and [AD] highlight the effects of the application of intelligent networks on interconnected systems; to achieve the monitoring, control and management of energy from generation to distribution.

Regarding validation as a Machine Learning technique, it is indicated that for the validation of the crime prediction model in India [26] uses K-fold cross-validation applied on six different types of learning algorithms: KNN and decision trees, Naïve Bayes, CART linear regression classification, regression) and SVM indicating that accuracy can be improved using cross-validation. Similarly [27] also uses a retention accuracy estimator and K-fold cross-validation accuracy estimator to determine the validity of 5 models: support vector machine, naïve bayes, decision tree algorithm, random forest and k-nearest neighbors for classification of 154 Hindi web poetries obtaining the best results with SVM. On the other hand [28] also uses 5-fold cross-validation to determine the validity of 5 Naïve Bayes support vector machine models, decision tree algorithm, random forest and k-nearest neighbors to improve the prediction accuracy and generalization ability for the design of three types of elastic moduli to predict the electronic work function of pure metals. Also [29] uses cross-validation to determine the validity of three support vector machine (SVM) models, k-nearest neighbors and Naïve-Bayes classifier for the development of predictive models to categorize rainfall amount intensity from ambient noise achieving accuracies around 99%. To optimize the kernel Ridge regression model to lower the correlation index, [30] uses cross-validation to reduce the optimal model justifying non-asymptotically for the proposed model. Similarly [31] uses cross-validation techniques in regression analysis to estimate the accuracy matrix to select the appropriate bandwidth for signal processing. In the field of malware detection on Android OS [32] uses four regression techniques: linear, multilayer neural network, additive regression and minimal sequential optimization, using cross-validation to determine that the best technique is linear regression with a Pearson correlation coefficient of 0.8655. In the field of agriculture to predict crop yields [33] uses four regression models: decision tree, linear regression, lasso regression and ridge regression using cross-validation for the validation of the errors: mean absolute, mean square; indicating that the best method is the decision tree method.

In the application of cross validation techniques to the field of solar panels [34] applies simple validation to test two conventional arrays of solar panels: parallel series and total cross tied in four different conditions of partial shading, in a MATLAB/Simulink simulation environment, to determine the most efficient configuration to face the problem of partial shading, indicating that total cross tied offers better results. In the same way to forecast the photovoltaic energy collected by photovoltaic panels in high temporal resolution using low temporal resolution meteorological variables [235] proposes three similarity-based models: basic, categorical and hierarchical, using cross validation to eliminate variables to obtain a more accurate model. Likewise [36] for the prediction of the energy generated by photovoltaic panels as a function of the solar radiation generated the previous day, using random forests with bagging and cross-validation to eliminate the most irrelevant variables in a real Australian scenario. On the other hand [37] for a PV system located in an area with snow accumulation predicts the snow accumulation on the solar panels for the energy harvesting forecast based on three-year historical data and applying 5-k fold cross-validation for the hyperparameters adjustment achieving an accuracy of 96%.

The contributions of this article are:

Implementation, data acquisition, data storage of two photovoltaic systems: String PV and DC-DC PV, connected to the grid in altitude conditions above 3,800 m.a.s.l.

Performing multiparametric regression models for each system: String PV and DC-DC PV, validating the results by linearity, normality of error terms, correlation, autocorrelation and homoscedasticity.

Cross-form validation of the above models to determine which model fits the height conditions.

2. Methodology

The Ordinary Least Squares (OLS) regression algorithms contain a balance of bias and variance providing the model with a correct and reliable prediction, for this purpose a pre-processing of the data was carried out with the purpose of equalizing the variables of both systems, the independent variables and the dependent variable, verifying that they had the same quantity and the same magnitude. After preprocessing and loading the data, the model of the algorithm used (OLS) was established with the parameters configured for a better performance, then the model was trained for both systems; demonstrating that the model has a correct performance and validating the results obtained with the tests of linearity, normality of error terms, correlation, autocorrelation, and homoscedasticity. Then, the CC-CC PV system was tested with the String PV model, and the String PV system with the CC-CC PV model, determining the performance through cross-validation and thus predicting one system with the other based on these two proposed models. Finally, the performance of such cross-validation was measured; see Fig. 1.

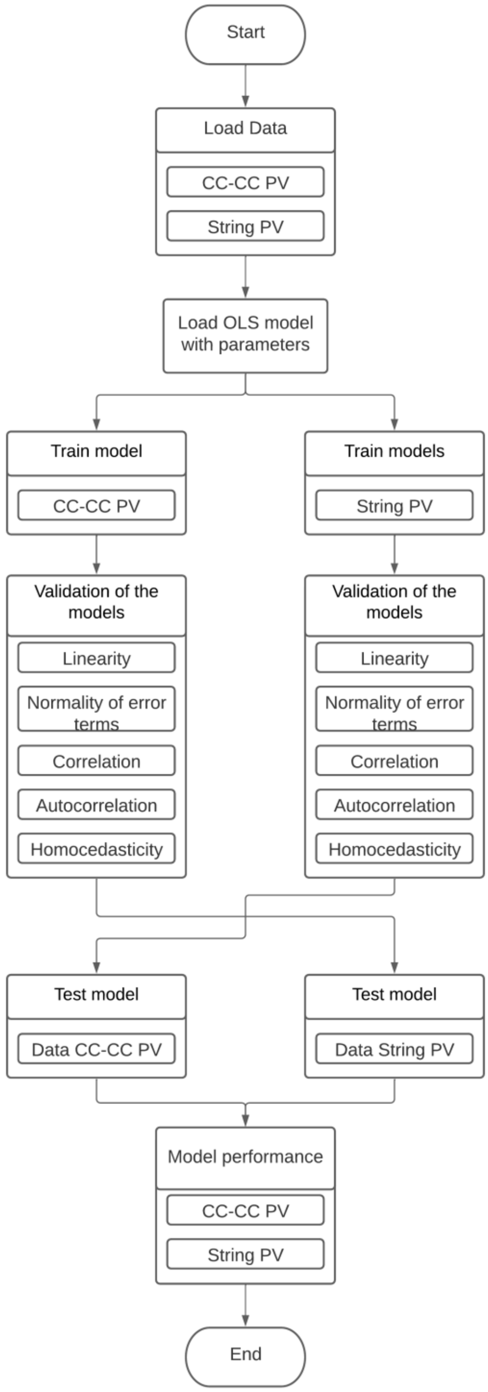

Fig. 1. Methodology Flowchart.

2.1. Power System Model

The measurements were obtained according to the European standard IEC 61724-20170, during the month of September 2021 with average temperatures of 10.5 C° and solar radiation of 6.4KWh/m^2 in the city of Juliaca, province of San Roman, department of Puno in Peru; which is located at an altitude of 3827 m.a.s.l. whose coordinates are: 15 ° 29 ′ 27 ″ S 70 ° 07 ′ 37 ″ W, this project was developed with funding from CONCYTEC-FONDECYT after winning a contest. The PV systems compared in this paper are two PV arrays, CC-CC PV and String PV, see Fig. 2.

Fig. 2. Photovoltaic solar system installed.

A. DC-DC PV

The PV with DC-DC optimizers has 10 370W monocrystalline photovoltaic modules of the ERA SOLAR brand model ESPSC370 with 10 Edge P370 solar DC-DC converters that support up to an input power of 370W and a single-phase inverter with HD technology -Wave Solar Edge SE3000H with an output power of 3000W

The PV has the following configuration shown in the following Fig. 3

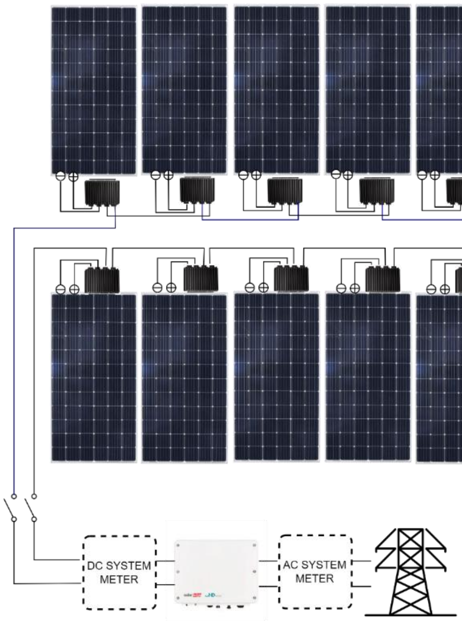

Fig. 3. Diagram of DC-DC PV.

B String PV

The PV with string-type single-phase inverter, has 12 polycrystalline photovoltaic modules of 270 W brand TALESUN model TP660P- and a String-type inverter of 3 kW. brand SUNNY BOY 3.0 The system diagram is shown in the following Fig. 4.

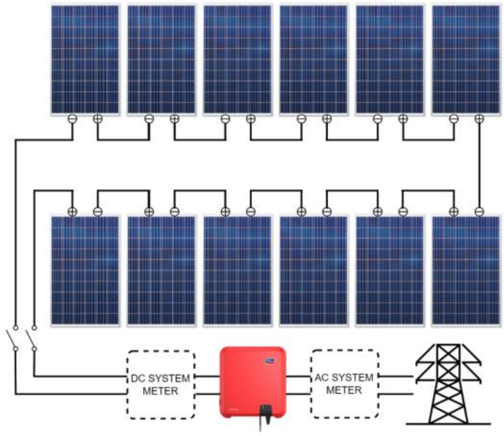

Fig. 4. Diagram of String PV.

2.2. Data Collection.

The instrumentation system for data acquisition uses current and voltage transducers Zelio Analog brand Schneider and power meter HIKING TOMZN with standard IEC 62053-21 that allow data to be recorded through a PLC micro-LOGO version 8.3 through Modbus RS485 communication protocol with Accuracy Class 1 and LAVBIEW software as shown in Fig. 5 below.

The instrument calibration system was carried out with FLUKE meters with valid calibration certificates for this year and applying the IEC 61724-1 standard with a Class A monitoring level with AC and or DC uncertainty including Instrumentation < 1% of Reading and Data log every 60 seconds.

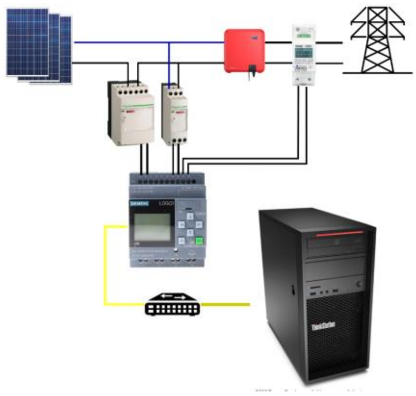

Fig. 5. Data collection instrumentation diagram

The following Fig. 6 shows the instrumentation system through which the data is collected.

Fig. 6. Data collection instrumentation installed.

The following Fig. 7 shows the user interface through the LabView program through which the data of the aforementioned photovoltaic systems can be viewed.

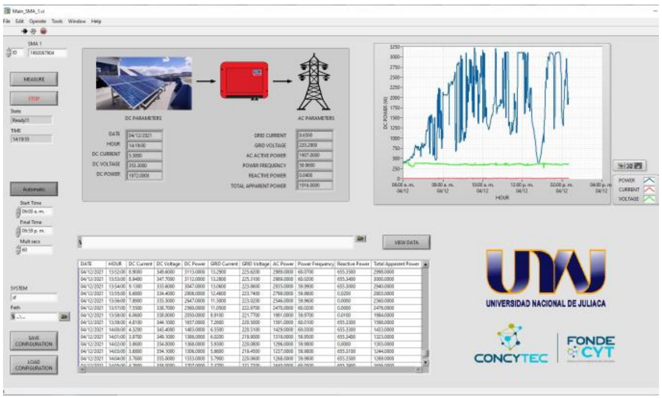

Fig. 7. Monitoring system user interface.

2.3. Models.

OLS Regression

Starting from a statistical model based on the relationship of two variables, we have the following equation (1).

$$y_i = \alpha + \beta x_i + \varepsilon_i \qquad (1)$$

Where:

ε_i: It is the error term.

α,β: They are the hidden parameters of the regression.

The objective is to find the values of α and β to minimize the error term. So that our positive errors are not compensated by the negative ones we have the following equation (2).:

$$S(\alpha,\beta) = \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2 \qquad (2)$$

We now have the ordinary least squares (OLS) error. S(α,β) reaches a minimum point at (α,β), so we derive equation (3).:

$$\left.\frac{\partial S}{\partial \alpha}\right|_{\alpha,\beta} = \left.\frac{\partial S}{\partial \beta}\right|_{\alpha,\beta} = 0 \qquad (3)$$

Evaluating α and β we have the equations (4)(5):

$$-2\sum_{i=1}^{n} (y_i - \alpha - \beta x_i) = 0 \qquad (4)$$

$$-2\sum_{i=1}^{n} (y_i - \alpha - \beta x_i) x_i = 0 \qquad (5)$$

Solving the normal equations we have equation (6).:

$$\alpha = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i y_i \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \qquad (6)$$

Solving we have the linear regression model equation (7).:

$$\alpha = y - \beta x \qquad (7)$$

The OLS-type linear regression model draws a line as close as possible to the trained data, thus estimating a relationship between the dependent variable and one or more independent variables, this relationship is made by minimizing the sum of the squares in the difference, this model This gives us the ability to predict the values of the dependent variable based on the values of the independent variable. For the present study, two PV systems were modeled with data collected independently. The statistical description of the data for each system is provided in Table 1 and Table 2.

Table 1. Description of DC-DC PV data set.

	AC CURRENT	AC VOLTAJE	AC POWER	AC FRECUENCY	AC APARENT POWER	AC REACTIVE POWER	DC CURRENT	DC VOLTAGE	DC POWER
count	5041	5041	5041	5041	5041	5041	5041	5041	5041
mean	7,40	218,90	1623,63	59,99	1636,89	149,48	4,19	368,46	1648,20
std	4,86	4,17	1095,58	0,038	1086,64	41,95	2,70	78,67	1112,07
min	0	203,1	0	59,861	0	0	0	0	0
25%	2,83	216,6	600,1	59,969	614	135,07	1,647	370	609,2
50%	7,03	219,7	1535,6	59,996	1543	154,1	4,21	370,1	1558,4
75%	12,98	221,6	2872,8	60,023	2877,6	176,31	6,994	370,3	2914
max	13,84	228,7	3009	60,205	3016	216,71	8,206	445,7	3055

Table 2. Description of String PV data set.

	AC CURRENT	AC VOLTAJE	AC POWER	AC FRECUENCY	AC APARENT POWER	AC REACTIVE POWER	DC CURRENT	DC VOLTAGE	DC POWER
count	4893	4893	4893	4893	4893	4893	4893	4893	4893
mean	4,52	349,38	1553,73	6,77	220,55	1491,31	59,99	288,79	1499,69
std	3,02	22,49	1000,88	4,31	3,27	981,84	0,03	325,34	980,54
min	0,11	139,9	34	0,23	209,43	0	59,8	0	11
25%	1,79	335,3	653	2,87	218,67	615	59,96	0	618
50%	4,08	349,5	1449	6,39	220,82	1406	59,99	0,04	1413
75%	7,45	364,4	2519	10,98	222,55	2437	60,02	655,33	2445
max	10	412,3	3180	13,7	229,72	3028	60,18	655,35	3065

2.4. Cross Validation.

Although it is true that validation is looking for numerical results that quantify the hypothetical relationships between variables in acceptable ranges as descriptions of the data; This technique is generally used to quantify prediction models. Among the most used validation techniques we have: simple validation, which consists of randomly dividing the values into two groups: one to train the model and the other to test it. The second technique is cross validation, which starts as the previous method, but performs this process "n" times adjusting the model and thus reduces the variability caused by randomization in the process of choosing the data. The third technique is K-double cross validation, which consists of the same iterative process but works with k-1 groups to train the model against the only group to validate it, but this process is repeated k times using different groups for validation. , has the following advantages over the previous technique: it generates a balance between bias and variance, allowing a better estimation of the error, as well as reducing the computational cost, since it is recommended to take a k between 5 and a10, which generates a good result. The fourth technique is iterative double-K cross-validation, which is the same technique as above, but it is repeated “n” times to fit the model. As a fifth technique, we have Bootstrapping, which consists of taking a bootstrap sample, which is a sample obtained from the original sample by random sampling with replacement, and of the same size as the original sample, so that some observations appear multiple times in the bootstrap sample. and others none. These observations not taken are called out of bag. For each iteration, a new sample of the same size as the original is generated and evaluated with the samples out of bag.

If we make a comparison between these techniques, we notice that each one has advantages and disadvantages according to the type of data to be treated: some are used to validate models with small data sizes or samples, others for large sample sizes and others are focused on the comparison of models for more accurate estimation of metrics. Therefore, this feature of the validation techniques was used to compare the two models of grid-connected photovoltaic systems studied: traditional and with DC amplifiers.

The bootstrapping algorithm used was:

1. Generate first sample of the same size as the original sample using random sampling with replacement.

2. Perform the model fit using the new sample generated in step 1.

3. Calculate the model error using those observations from the original sample that have not been included in the new sample. This error is known as a validation error.

4. Repeat steps 1, 2 and 3 “n” times and calculate the mean of the n validation errors.

5. Finally, and after the n repetitions, fit the final model using all the original training observations.

Although in the end a certain bias can be generated, this is reduced if the data to be processed are numerous, as is the case in this study.

3. Results.

3.1. Model Validation.

A. Linearity

This gives us the linear measurements with the data and the prediction of the proposed model; To verify this condition, a scatter diagram is plotted as shown in Fig. 8

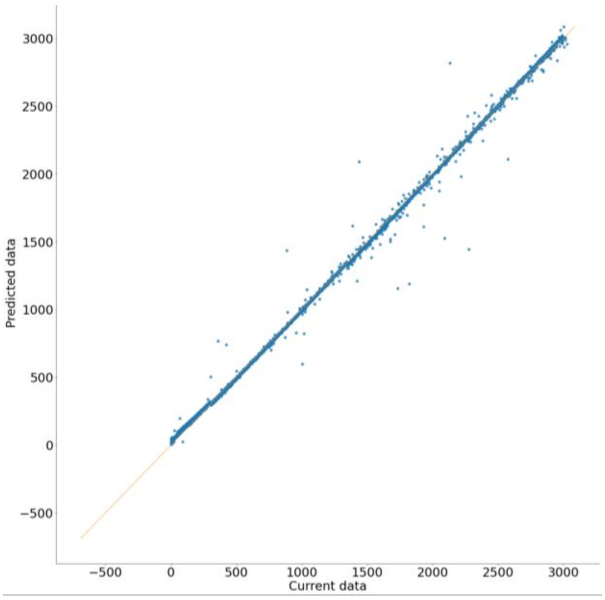

Fig. 8. OLS Linearity Current vs Predicted.

From the graph we note that there is a linear relationship between the real values and the values provided by the model, as well as the adjusted R-squared coefficient shown in Table 3, whose value is 0.99 and which, being greater than 0.7, indicates the existence of linearity.

B. Autocorrelation

For the analysis and detection of an autocorrelation in this regression model, the Durbin-Watson test was used to ensure that all the information was captured in order to avoid biases in the system. When the value of this test is from 0 to 2 the autocorrelation is considered positive, on the other hand if it has values from 2 to 4 it is considered negative.

The Durbin-Watson test for the present model provides a result of approximately 2 for both systems (CC-CC PV and String PV), so in both cases no or little autocorrelation is considered.

C. Normality of error terms

The normality of the error terms shows an interval estimate that follows a normal distribution, that is, it is distributed as a normal probability density function with zero mean and almost constant variance, this condition is indicated by the histogram shown in Fig 9.

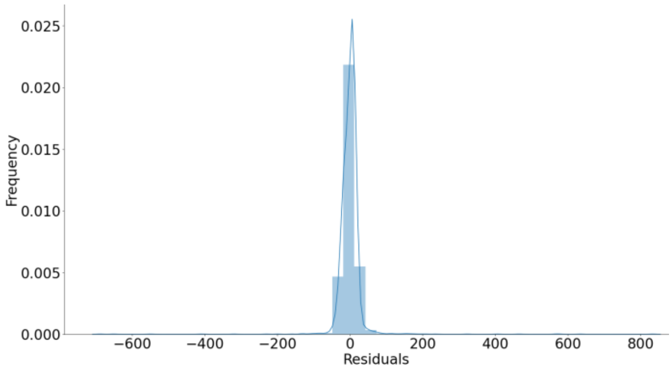

Fig. 9. OLS Distribution of Residuals

D. Correlation

The correlation between independent variables is avoided so as not to obtain a regression in bad conditions; this condition will help us minimize or eliminate some variables from our model. To show this condition, a heat map is plotted that is shown in Fig. 10

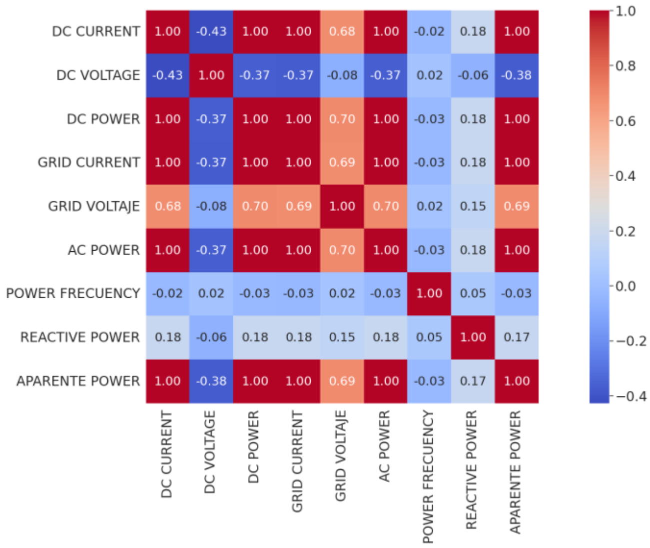

Fig. 10. heat map between independent variables

E. Homoskedasticity

In statistics, a prediction model is said to be homoscedastic when the variance of the error conditional on the explanatory variables is constant throughout the observations, a statistical model relates the value of a variable to be predicted to that of others; if the model is unbiased, the predicted value is the mean of the variable to be predicted.

In any case, the model gives an idea of the value that the variable to be predicted will take. To avoid that a subset of data is assigned an inappropriate weight, it is plotted with the residuals and thus it is determined whether the variance is uniform as shown in Fig.11.

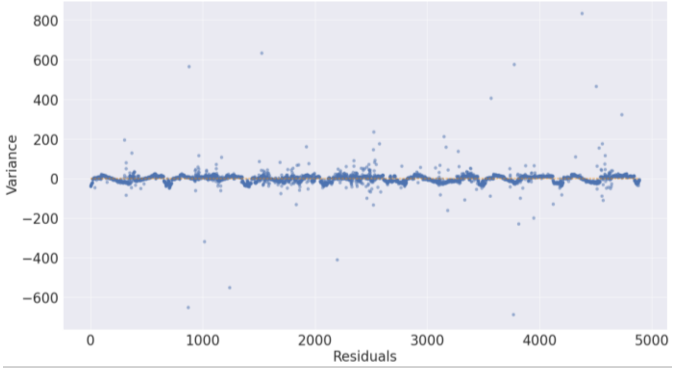

Fig. 11. OLS Homoscedasticity

3.2. Model Performance.

From Table 3, the score for both models is higher than 99.9%, which indicates that the percentage of correctly predicted values with respect to the total values is adequate. Likewise, obtaining an MAE of 4.21 for the CC-CC PV system and 12.52 for the String PV system indicates the proximity of the predictions with respect to the real results. Likewise, the MSE value of 8.77 for the CC-CC PV model and 24.10 for the String PV model creates a single value that summarizes the error in the model and is generally taken as the square root of this value. Finally, the coefficients of determination and adjusted coefficients of determination, which in all four cases are greater than 99.9%, indicate that the models performed effectively represent the systems analyzed.

Table 3. Model performance.

	OLS	
	DC-DC PV	**String PV**
Score	0.99993534671124	0.999391262336387
MAE	4.21155133395433	12.5204268446744
MSE	8.77227535829964	24.1003997219785
Determination coefficient	0.99993534671124	0.999391262336387
Adjusted coefficient of determination	0.99993524392381	0.999390265223097

3.3. Cross Validation

The equations obtained that model the systems to be compared by means of OLS regression are:

A. DC-DC PV:

Value of the slopes or coefficients "a":
[4.56883182e+01 1.57976130e+00 -1.15721484e+00 2.24060833e-01 4.52685101e-02 -3.33755215e+00 -9.36013521e-02 5.71671771e-01]

Intercept value or coefficient "b":
-258.0901368608595

B. String PV:

Value of the slopes or coefficients "a":
[2.43497292e+01 1.04241200e-01 5.02161023e-01 8.33985448e+00 1.07256736e+00 4.11125860e+00 -7.44605696e-04 3.75203843e-01]

Intercept value or coefficient "b":
-537.8103635478974

Cross-validation was performed from these two models, obtaining predictions from one model with the data from the other, as shown in Table 4 and Table 5

Table 4: String PV data in the DC-DC PV model.

String PV data								**String PV data in the DC-DC PV model**			
AC current	AC voltage	AC frecuency	AC aparent power	AC reactive power	DC current	DC voltage	DC power	AC power string PV model	AC power String PV data	Difference	Prediction of original data
0.75	219.82	59.96	104	655.32	0.44	381.8	170	166.9989975	101	65.998998	125.14114
0.77	220.53	60.05	110	655.31	0.46	388.8	181	175.8405863	110	65.840586	135.431158
0.79	220.06	59.97	117	655.34	0.48	392.5	192	183.5495394	118	65.549539	143.78781
0.84	219.93	59.95	129	655.32	0.51	393.4	203	194.443577	130	64.443577	154.833685
0.86	219.96	59.99	139	655.34	0.53	394.4	212	202.5846548	141	61.584655	164.059818
0.89	219.7	59.9	153	0	0.56	396.3	224	183.7013941	150	33.701394	176.356439
0.93	220.04	59.91	165	0.01	0.59	396.4	235	194.922559	165	29.922559	187.862945

Table 5: DC-DC PV data in the String PV model.

DC-DC PV Data								**DC-DC PV data in the String PV model**			
AC current	AC voltage	AC frecuency	AC aparent power	AC reactive power	DC current	DC voltage	DC power	AC power DC DC PV model	AC power DC DC PV data	Difference	Prediction of original data
0.65	218.2	59.98	143.86	118.23	0.247	370.1	91.47	92.65765203	81.97	10.687652	101.3103508
0.66	217.7	59.97	140.5	112.46	0.231	370.2	85.41	87.5475483	88.12	-0.5724517	96.60828707
0.68	217.8	60.01	152.04	118.7	0.261	369.9	96.46	98.47213987	95.01	3.46213987	106.7000532
0.74	217.5	59.97	162.53	105.37	0.303	369.8	112.21	111.3534251	123.74	-12.386574	119.6336947
0.75	217.7	59.95	164.38	110.81	0.333	369.8	122.87	118.3427168	121.41	-3.0672832	127.0843422
0.78	217.2	59.98	169.93	103.49	0.383	370.2	141.8	131.041112	134.78	-3.738888	139.1576808
0.85	217.4	59.96	182.32	107.79	0.403	370	149.28	140.6332972	147.04	-6.4067028	149.8914876

From table 4 the final validation percentage is 1.8951927, which results from running the values of the second system: String PV in the first regression model: CC-CC PV. From Table 5 the validation percentage is -4.214518 which results from running the values of the first system: CC-CC PV in the second regression model: String PV. From what we deduce that the CC-CC PV system has a better performance in 6.1097727 for these systems at 3800 m.a.s.l.

No literature directly related to the cross-validation of the CC-CC PV and String PV systems at 3800 m.a.s.l. was found.

4. Conclusion

Two factors are taken into account to improve energy harvesting through solar panels: type and location of the DC and AC converters and how they are integrated into the electrical grid. Therefore, in this study, two models of photovoltaic energy generation were implemented: the first, an SFCR with DC-DC optimizers with 10 monocrystalline photovoltaic modules of 370W of the ERA SOLAR brand, model ESPSC370 with 10 Edge P370 solar DC-DC converters and a single-phase inverter with HD-Wave Solar Edge SE3000H technology. The second: a SFCR with a single-phase String inverter, with 12 polycrystalline photovoltaic modules of 270 W TALESUN model TP660P- and a String inverter of 3 kW SUNNY BOY 3.0. For the collection, IEC 61724-2017 regulations were complied with. Both systems were implemented in the city of Juliaca, which is located at 3,800 meters above sea level. For the comparison of both models, the data provided by the CC-CC PV and String PV systems were analyzed, then the regression model was trained separately, obtaining in both cases scores higher than 99.9%, in the same way the determination coefficients and adjusted determination coefficients in all cases are higher than 99.9%. Subsequently, both systems were validated by means of linearity, normality of error terms, correlation, autocorrelation and homoscedasticity to check the usefulness of both models. Bootstrapping cross-validation was used, whose characteristic is that it is used to compare models with small data sizes or samples, as in the present study, and thus establish which model has the best performance. It was determined that the DC-DC PV model has a better performance of 6.09% over the String PV model, so we conclude that the DC-DC PV converter performs better at 3800 meters above sea level. For future work we will perform hyperparameter tuning, variable elimination by means of recursive models such as RFE and also use new boosting type models to perform the regression models such as XGBoost.

Acknowledgment

This work was financed by CONCYTEC-FONDECYT call E041-01 [contract number N°180-2018-FONDECYT-BM-IADT-AV

References

[1]S. R. P. Chitturi, E. Sharma, and W. Elmenreich, Efficiency of photovoltaic systems in mountainous areas', 2018 IEEE Int. Energy Conf. ENERGYCON 2018, pp. 1–6, 2018, doi: 10.1109/ENERGYCON.2018.8398766.

[2] S Javed, F., & Raza, A. (2022). Impact of temperature & illumination for improvement in photovoltaic system efficiency. Ijree.Net, 6(1). https://ijree.net/wp-content/uploads/2021/04/2021-1-1-2.pdf

[3]Verma, S., Yadav, D. K., & Sengar, N. (2021). Performance Evaluation of Solar Photovoltaic Power Plants of Semi-Arid Region and Suggestions for Efficiency Improvement. International Journal of Renewable Energy Research, 11(2), 762–775. https://doi.org/10.20508/ijrer.v11i2.11957.g821

[4] J. Solis, A. Råberg, J. André and M. Nilsson, "Analyzing the effect of snow in PV regulator response in a PV solar park," 2021 9th International Conference on Smart Grid (icSmartGrid), 2021, pp. 72-75, doi: 10.1109/icSmartGrid52357.2021.9551233.

[5] S. Huaquipaco, N. Beltran, V. Sarmiento, H. Pizarro, J. Cruz, R. Condori, J. Ramos, C. Romero, N. A Achahuanco, 'Solar library', in Proceedings of the ISES Solar World Congress 2019 and IEA SHC International Conference on Solar Heating and Cooling for Buildings and Industry 2019, 2020, pp. 2507–2513, doi: 10.18086/swc.2019.52.01.

[6] Q. Lv, J. Zhang, K. Ding, Z. Zhang, H. Zhu and R. Hou, "The Output Power Smoothing Method and Its Performance Analysis of Hybrid Energy Storage System for Photovoltaic Power Plant," 2021 10th International Conference on Renewable Energy Research and Application (ICRERA), 2021, pp. 36-39, doi: 10.1109/ICRERA52334.2021.9598777

[7] K.N. Nwaigwe, P. Mutabilwa, E. Dintwa. An overview of solar power (PV systems) integration into electricity grids. Materials Science for Energy Technologies. Volume 2, Issue 3. 2019. Pages 629-633, ISSN 2589-2991, https://doi.org/10.1016/j.mset.2019.07.002.

[8] D. Siswantoro, S. Zainal. Is Grid Solar Power Still Attractive Amid Decreasing Fuel Prices? The Case of Indonesian Electrical Power. International Journal of Smart Grid-ijSmartGrid, 2021, vol. 5, no 2, p. 83-87.

[9] A. Belkaid, I. Colak, K. Kayisli, R. Bayindir. Design and implementation of a cuk converter controlled by a direct duty cycle INC-MPPT in PV battery system. International Journal of Smart Grid-ijSmartGrid, 2019, vol. 3, no 1, p. 19-25.

[10] D. Icaza, F. Córdova and M. Avila, "Modeling and Simulation of a Solar System in the Quingeo Church in Ecuador," 2018 International Conference on Smart Grid (icSmartGrid), 2018, pp. 158-163, doi: 10.1109/ISGWCP.2018.8634526.

[11] M. N. Hamidi, D. Ishak, M. A. A. M. Zainuri, C. A. Ooi and T. Tarmizi, "Asymmetrical Multi-level DC-link Inverter for PV Energy System with Perturb and Observe Based Voltage Regulator and Capacitor Compensator," in Journal of Modern Power Systems and Clean Energy, vol. 9, no. 1, pp. 199-209, January 2021, doi: 10.35833/MPCE.2019.000147

[12] N. Amouzegar Ashtiani, S. M. Azizi and S. A. Khajehoddin, "Robust Control Design for High-Power Density PV Converters in Weak Grids," in IEEE Transactions on Control Systems Technology, vol. 27, no. 6, pp. 2361-2373, Nov. 2019, doi: 10.1109/TCST.2018.2867212

[13] S. Farhangi, A. Nazer, G. R. Moradi and E. Asadi, "Improved Performance of Solar Array Simulator Based on Constant Voltage/Constant Current Full-Bridge Converter," 2020 11th Power Electronics, Drive Systems, and Technologies Conference (PEDSTC), 2020, pp. 1-5, doi: 10.1109/PEDSTC49159.2020.9088485.

[14] F. Thomas and J. Mathew, "Transformer-Less Solar PV Inverter With Fuzzy Assisted MPPT for Standalone/ Grid Connected System," 2020 IEEE Recent Advances in Intelligent Computational Systems (RAICS), 2020, pp. 1-6, doi: 10.1109/RAICS51191.2020.9332489

[15]T. T. Guingane, D. Bonkoungou, e. Korsaga, e. Simonguy, Z. Koalaga and F. ZOUGMORE, "Modeling and Simulation of a Photovoltaic System Connected to The Electricity Grid with MALTAB/Simulink/Simpower Software," 2020 8th International Conference on Smart Grid (icSmartGrid), 2020, pp. 163-168, doi: 10.1109/icSmartGrid49881.2020.9144807.

[16]S. Huaquipaco ,N. Beltran,V. Sarmiento,L. Huanchi, R. Tito, R. Contreras, M. Reyes, A. Apaza, 'Photovoltaic charger system for mobile devices using quick charge 3.0 technology.', Proc. LACCEI Int. Multi-conference Eng. Educ. Technol., pp. 27–30, 2020, doi: 10.18687/LACCEI2020.1.1.341.

[17] U. K. Das, K. S. Tey, M. Y. Idna Idris and S. Mekhilef, "Maximum Power Flow Management for Stand-alone PV Based Battery Charging System," 2019 10th International Conference on Power Electronics and ECCE Asia (ICPE 2019 - ECCE Asia), 2019, pp. 1-6, doi: 10.23919/ICPE2019-ECCEAsia42246.2019.8797266.

[18] R. Wongsathan, A. Nuangnit and I. Seedadan, "Optimal Intelligent Controllers for Multi-Stage CCCV Charging of Solar PV Battery Charger System," 2018 International

Conference on Computing, Power and Communication Technologies (GUCON), 2018, pp. 105-110, doi: 10.1109/GUCON.2018.8675010.

[19] B. K. Verma, S. Devassy, S. K. Ram, A. Abhishek and A. K. Dhakar, "Performance Evaluation of PV Integrated DSTATCOM Based on Complex Variable Filter," 2018 4th International Conference on Electrical Energy Systems (ICEES), 2018, pp. 94-99, doi: 10.1109/ICEES.2018.8442402.

[20] S. Ganguly, S. Pal, I. Khan, D. Das, A. Ghosh and A. Bhattacharjee, "Design of an efficient Solar PV Battery Charge Controller under dynamic Motor load conditions," 2018 2nd IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), 2018, pp. 546-551, doi: 10.1109/ICPEICES.2018.8897360.

[21] Bharath, K. R., & Suresh, E. (2017). Design and implementation of improved fractional open circuit voltage based maximum power point tracking algorithm for photovoltaic applications. International Journal of Renewable Energy Research, 7(3), 1108–1113. https://doi.org/10.20508/ijrer.v7i3.5899.g7141

[22] M. Duan, D. Sun, J. Duan, L. Sun and Y. Liu, "Interleaved Modulation Scheme with Optimized Phase Shifting for Double-Switch Buck-Boost Converter," in IEEE Access, vol. 9, pp. 55422-55435, 2021, doi: 10.1109/ACCESS.2021.3071314.

[23] Huaquipaco, S., Cruz, J., Beltran Castañon, N. J., Pineda, F., Romero, C., Chura Acero, J. F., & Mamani Machaca, W. (2021). Modeling And Prediction Of A Multivariate Photovoltaic System, Using The Multiparametric Regression Model With Shrinkage Regularization And Extreme Gradient Boosting. https://doi.org/10.18687/laccei2021.1.1.557

[24] COLAK, R. BAYINDIR and S. SAGIROGLU, "The Effects of the Smart Grid System on the National Grids," 2020 8th International Conference on Smart Grid (icSmartGrid), 2020, pp. 122-126, doi: 10.1109/icSmartGrid49881.2020.9144891.

[25] F. Ayadi, I. Colak and R. Bayindir, "Interoperability in Smart Grid," 2019 7th International Conference on Smart Grid (icSmartGrid), 2019, pp. 165-169, doi: 10.1109/icSmartGrid48354.2019.8990680.

[26] P. Tamilarasi and R. U. Rani, "Diagnosis of Crime Rate against Women using k-fold Cross Validation through Machine Learning," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 1034-1038, doi: e10.1109/ICCMC48092.2020.ICCMC-000193.

[27] K. Pal and B. V. Patel, "Data Classification with k-fold Cross Validation and Holdout Accuracy Estimation Methods with 5 Different Machine Learning Techniques," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 83-87, doi: 10.1109/ICCMC48092.2020.ICCMC-00016.

[28] N. Li, T. Zong and Z. Zhang, "Prediction of the Electronic Work Function by Regression Algorithm in Machine Learning," 2021 IEEE 6th International Conference on Big Data Analytics (ICBDA), 2021, pp. 87-91, doi: 10.1109/ICBDA51983.2021.9403202.

[29] M. M. S. Pangaliman, F. R. G. Cruz and T. M. Amado, "Machine Learning Predictive Models for Improved Acoustic Disdrometer," 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology,Communication and Control, Environment and Management (HNICEM), 2018, pp. 1-5, doi: 10.1109/HNICEM.2018.8666256.

[30] A. Tanaka and H. Imai, "Kernel Ridge Regression with Autocorrelation Prior: Optimal Model and Cross-Validation," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 3872-3876, doi: 10.1109/ICASSP40776.2020.9053423.

[31] J. Tong, J. Xi, Y. Yu and P. O. Ogunbona, "Cross-Validated Bandwidth Selection for Precision Matrix Estimation," 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018, pp. 4479-4483, doi: 10.1109/ICASSP.2018.8462469.

[32] D. Ö. Şahin, O. E. Kural, S. Akleylek and E. Kılıç, "Comparison of Regression Methods in Permission Based Android Malware Detection," 2020 28th Signal Processing and Communications Applications Conference (SIU), 2020, pp. 1-4, doi: 10.1109/SIU49456.2020.9302502.

[33] M. Kavita and P. Mathur, "Crop Yield Estimation in India Using Machine Learning," 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), 2020, pp. 220-224, doi: 10.1109/ICCCA49541.2020.9250915.

[34] R. K. Pachauri, "Investigation on SP and TCT Photovoltaic Array Configurations under Obscured Shading Conditions," 2020 IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), 2020, pp. 1-6, doi: 10.1109/UPCON50219.2020.9376510.

[35] H. Sangrody, N. Zhou and Z. Zhang, "Similarity-Based Models for Day-Ahead Solar PV Generation Forecasting," in IEEE Access, vol. 8, pp. 104469-104478, 2020, doi: 10.1109/ACCESS.2020.2999903.

[36] A. Lahouar, A. Mejri and J. Ben Hadj Slama, "Importance based selection method for day-ahead photovoltaic power forecast using random forests," 2017 International Conference on Green Energy Conversion Systems (GECS), 2017, pp. 1-7, doi: 10.1109/GECS.2017.8066171.

[37] B. Hashemi, A. -M. Cretu and S. Taheri, "Computational Intelligence Based Snow Cover Prediction for Photovoltaic Systems," 2021 IEEE Electrical Power and Energy Conference (EPEC), 2021, pp. 1-6, doi: 10.1109/EPEC52095.2021.9621702