Contents list available at IJRED website

International Journal of Renewable Energy Development

Journal homepage: https://ijred.undip.ac.id

Research Article

Daily Solar Radiation Forecasting based on a Hybrid NARX-GRU Network in Dumaguete, Philippines

Al Diego Pega Fuselero, Hannah Mae San Agustin Portus, Bonifacio Tobias Doma Jr.*

School of Chemical, Biological, and Materials Engineering and Sciences, Mapúa University, Manila 1002, Philippines

Abstract. In recent years, solar radiation forecasting has become highly important worldwide as solar energy increases its contribution to electricity grids. However, due to the intermittent nature of solar radiation caused by meteorological parameters, forecasting errors arise, and fluctuations in the power output of photovoltaic (PV) systems become a severe issue. This paper aims to introduce a forecasting hybrid model of daily global solar radiation time series. Meteorological data and solar radiation samples from Dumaguete, Philippines, are used to assess the forecasting accuracy of the proposed nonlinear autoregressive network with exogenous inputs (NARX) – gated recurrent unit (GRU) hybrid model. Four different models were trained using the meteorological and solar radiation data, which are the Optimizable Gaussian Process Regression (GPR), Nonlinear Autoregressive Network (NAR), NARX, and the proposed Hybrid NARX-GRU Network. Results show that the hybrid NARX-GRU model has a root mean square error (RMSE) of ~0.05 and a training time of 33 seconds. The proposed hybrid model has better forecasting performance compared to the three models which obtained RMSE values of 27.741, 39.82, and 28.92, for the GPR, NAR, and NARX, respectively. The simulation results demonstrate that the NARX-GRU model significantly outperforms the regression and single models in terms of statistical metrics and training efficiency. Furthermore, this study shows that the hybridized NARX-GRU model is able to provide an effective estimation for daily global solar radiation, which is important in the operation of PV plants in the country, specifically for unit commitment purposes.

Keywords: Forecasting, NARX-GRU, Neural network, Photovoltaic system, Solar Radiation

Received: 14th Feb 2022; Revised: 28th May 2022; Accepted: 6th June 2022; Available online: 18th May 2022

1. Introduction

In the Philippines, where imported oil and coal are relatively high, renewable sources such as solar energy are technical and economically promising alternatives. Though the increase in attention for solar-based renewable energy through technological advancement, critics have specified its intermittent nature, which deprived solar energy's suitability in becoming the primary energy source in many applications (Li *et al.*, 2016; Puah *et al.*, 2021). Because solar energy is non-dispatchable and is dependent on variables such as the state of the environment like weather conditions, prediction machines can be vital in the energy supply decision making (Almonacid *et al.*, 2014). Many utility companies and power producers have acknowledged that this resource can only provide a dispatchable generation capacity in the market through advanced forecasting technology.

The application of solar energy is the most cost-effective in developing countries such as the Philippines, where reliable solar irradiation forecasts and solar radiation measurement are either absent or are only available for a very limited number of locations. There is still no study conducted yet for solar radiation forecasting using machine learning (ML) in the Philippines despite the country's good solar potential for its geography. The Philippines' average solar radiation ranges from 128-203 W/m^2, or an average of 161.7 W/m^2, based on sunlight duration (Deutsche GIZ GmbH *et al.,* 2013). A better forecasting technique can help the renewable energy industry in tapping this solar energy potential in the country.

There are four techniques used to forecast solar radiation: physical models (Pazikadin *et al.,* 2020), statistical methods (Diagne *et al.*, n.d.), artificial intelligence (AI) (Awad & Khanna, 2015a), and hybrid methods of AI, fuzzy methods, and evolutionary algorithms (Wang *et al.*, 2020a). In recent years, the forecasting of solar irradiance progressed towards increasing the accuracy of the models. ML-based models solve the uncertainties of fluctuating meteorological and topographical data by extracting the complex relationships among variables to forecast solar radiation accurately or photovoltaic (PV) energy output (Jung *et al.,* 2020b). Recurrent neural networks (RNNs), variants of artificial neural networks (ANNs), are robust and state-of-the-art machine learning algorithms applied to forecast time series data. RNNs consider the time correlation of the input variables in forecasting by establishing cyclic or

* Corresponding author
Email: btdoma@mapua.edu.ph (B.T. Doma Jr)

https://doi.org/10.14710/ijred.2022.44755
ISSN: 2252-4940/© 2022.The Author(s). Published by CBIORE

chain structures that enable the network to learn from a sequential dataset (Akhter *et al.,* 2021; Wang *et al*., 2020).

One of the most conventional RNNs applied in time series forecasting is the nonlinear autoregressive with exogenous input (NARX) (Ahmad *et al.,* 2015; Gonzaga Baca Ruiz *et al.,* 2016; Pisoni *et al.,* 2009; Yadav & Chandel, 2014) because of its good performance. While single models perform well independently, it is common to combine models to take advantage of their strengths (Antonanzas *et al.,* 2016). In (Husein & Chung, 2019), an RNN-based long-short term memory (LSTM) network and predicted day-ahead solar irradiance with an acceptable root mean square error (RMSE) value of 60.31 W/m^2. Four neural network (NN) models were hybridized in the work of Huang *et al.* (X. Huang *et al.,* 2021) to predict hour-ahead irradiance and generated an RMSE of 32.1 W/m^2 which significantly outperformed their individual models. Kumari and Toshniwal's work (Kumari & Toshniwal, 2021) also proved the several limitations of single deep learning models and obtained better forecasting accuracy using a hybrid convolutional neural network (CNN)-LSTM model. Most recently, a study of Arun Kumar et al. (Arun Kumar *et al*., 2021) forecasted COVID-19 confirmed and recovered cases through the gated recurrent unit (GRU)-based and long-short term memory (LSTM)-based RNNs accounting for various factors contributing to the virus spread. GRU is a relatively new forecasting model introduced by Cho (Cho *et al*., 2014), which addresses the typical vanishing gradient problem in RNNs (Liu *et al*., 2021). It features an architecture involving an update gate that determines how much of the past information will be retained and a reset gate that controls the past information to be forgotten (Lai *et al*., 2021). The gating mechanism of the GRU network has been realized in several studies in solar radiation forecasting. In Faisal et al.'s work, a GRU-based neural network performed best, with an RMSE value of 0.891, among the RNN and LSTM networks in forecasting solar radiation (Faisal et al., 2022). A study comparing five different RNN classes in Jinju City, South Korea found that deeper RNN architecture such as bidirectional-LSTM and bidirectional-GRU provide the lowest RMSE and R^2 values. The bidirectional-GRU model performed better with 46.1 RMSE and 0.958 R^2 at a low computational cost (Jaihuni *et al.,* 2021). Lai et al. proposed deep time-series clustering (DTC)-feature attention based deep forecasting (FADF) hybrid model, which consists of many GRU layers, achieved the smallest solar forecasting error in next hour global horizontal irradiance (GHI) forecasting among smart persistence models (Lai *et al.,* 2021). Liu *et al*.'s work highlighted the application of the variational Bayesian convolutional GRU model in forecasting solar radiation of a spatial region consisting of 200 sites (Liu *et al*., 2019). Narvaez et al. trained encoder-decoder GRU and LSTM networks to predict global horizontal irradiance in daily and weekly forecasting horizons, and their results show the LSTM networks outperformed the GRU in both horizons (Narvaez *et al*., 2021). The GRU network's strengths were also studied in several works involving PV energy forecasting (Abdel-Basset *et al*., 2021; Khan *et al*., 2021; Li *et al.,* 2020; Mellit *et al.,* 2021).

Al-Ghezi *et al*. (Al-Ghezi *et al.,* 2022) studied the validity of two single and two polynomial linear regression models in measuring daily global horizontal solar radiation (GHSR) using data from different stations – Iraqi Meteorological Authority (IMA) and National Aeronautics and Space Administration (NASA). Among the different statistical tests performed in the study, such as RMSE, mean bias error (MBE), and mean percentage error (MPE), the formulated single linear regression model depending on the IMA data produced GHSR values that were closer to the actual GHSR values (RMSE: 0.4769 MJ/m^2/day, MBE: 0.0164 MJ/m^2/day, 0.2207) followed by the single linear regression model based on NASA (0.8641, 0.1773, -0.9680), the polynomial model based on NASA (0.6420, 0.3996, -1.1487), and lastly the polynomial based on IMA (0.9604, 0.218, -1.0225).

Solar radiation prediction has a specific corresponding application depending on time horizon – very short term, short term, medium term, and long term (Sharma & Kakkar, 2018). Very short-term forecasting horizon utilizes few seconds to minutes and is important in real time monitoring of PV plant operation. Short term forecasting horizon uses 48-72 hours of time steps for unit commitment purposes. Medium term employs one week ahead time steps and is mainly used for maintenance scheduling, while long term forecasting horizon utilizes months or years ahead of time steps for the purpose of PV plant design and network operations. Accurate solar radiation forecasting is essential in the design, planning, and operation of PV energy systems.

In the burgeoning solar energy industry in the Philippines, it is crucial to use Philippine data to predict solar radiation at different time horizons to help the industry design and operate. PV power output data scarcity is a key challenge in PV energy forecasting. PV energy still depends on solar irradiation. Hence, solar forecasting using meteorological variables and historical solar radiation data is still utilized to estimate the generation capacity of a PV plant. The use of machine learning strategies will significantly contribute to the solar industry in the Philippines.

The contributions of this paper include: (i) a hybrid NARX-GRU solar radiation forecasting model which aims to improve the performance of single models by taking advantage of the strengths of two individual models, (ii) examine the effectiveness of the proposed hybrid model to forecast daily global solar radiation using data from the Philippines, (iii) explore the performance of a hybrid NARX-GRU model using a relatively small dataset, as majority of past studies utilized huge datasets, (iv) establish the hybrid NARX-GRU model using the optimized hyperparameters through model training and validation, (v) a comprehensive hybrid NARX-GRU model demonstration through simulation results.

This study examines the effectiveness of machine learning strategies to forecast daily solar radiation using data from the Philippines, specifically in Dumaguete City, Negros Oriental (Latitude: 09^o 18' N Longitude: 123^o 18' E). The study is only limited to 1 year and 10 months of daily meteorological data provided by the Philippine Atmospheric, Geophysical, and Astronomical Services Administration (PAGASA) and National Solar Radiation Center (NSRC). The hybrid NARX-GRU and other models utilized were implemented in the MATLAB R2021a platform.

2. Materials and Methods

2.1 Dataset and pre-processing

2.1.1 Data gathering

Different meteorological data for Dumaguete, Philippines, were collected from the PAGASA and NSRC for the period of September 2016 to July 2018 (669 datapoints). The hourly average solar global radiation (W/m^2) data were converted to daily global solar radiation data. The data include six parameters:

i. T_{ave}: Daily average temperature (^oC)
ii. Rf: Daily rainfall amount (mm)
iii. Rh: Daily relative humidity (%)
iv. Wd: Daily wind direction (degree relative to North indicating where the wind is blowing from)
v. Ws: Daily wind speed (m/s)
vi. N/N_s: Daily relative sunshine duration (dimensionless quantity).

2.1.2 Data Smoothing and Normalization

The irregularities from the dataset were eliminated using the smooth function and min-max normalization function of MATLAB R2021a. The smooth function was carried out using the moving average filter formula:

$$y_{n,smooth} = \frac{1}{N}\sum_{i=0}^{N-1} x(n-i) \quad (1)$$

where y_n is the average value of the previous N input samples, x_n is the nth sample of an input parameter.

The smoothed data were normalized using the Min-Max method:

$$x_{i,normalized} = \frac{y_i - y_i^{min}}{y_i^{max} - y_i^{min}} \quad (2)$$

2.1.3 Input layer

The different types of meteorological data collected, namely average temperature, rainfall amount, relative humidity, wind direction, wind speed, and sunshine duration, were utilized as input variables.

$$x = [x_1, x_2, x_3, x_4, x_5, x_6] \quad (3)$$

Elements of the feature vector x are given as: x_1 is the average temperature, x_2 is rainfall amount, x_3 is relative humidity, x_4 is wind direction, x_5 is wind speed, and x_6 is sunshine duration.

2.1.4 Output layer

The given feature vector x was used to produce an output sequence of daily global solar radiation values. In MATLAB, a polynomial activation function was used for the linear hidden layer:

$$y = a \cdot w + b \quad (4)$$

where w is the weight and b is the bias.

2.2 Training, evaluation, and testing

2.2.1 Regression models

To further explore the relation of each input parameter (predictors) to the target data, different regression models from the Regression Learner application in MATLAB was used, including linear regression models, regression trees, support vector machines, Gaussian process regression models, ensembles of regression trees, and neural networks. All models were validated using cross-validation at different k-values of 2, 5, 10, and 20 folds.

Respectively, the best regression models were retrained at different feature selections with the k-value, which reported the best performance metric scores. Each training result varied as one feature from the feature selection was removed to assess which predictors significantly affect the regression model's performance and which predictors do not.

The best regression model was tuned with hyperparameter values by using an optimization scheme to minimize the model's root mean square error (RMSE).

2.2.2 Nonlinear Autoregressive with Exogeneous Input Network

The nonlinear autoregressive with exogenous input (NARX) network was modeled using MATLAB's Neural Net Time Series App. The smoothed and normalized input meteorological variables were utilized to determine the performance of the NARX network and nonlinear autoregressive (NAR) network.

Several iterations were performed using the NARX and NAR models to determine the optimum (a) dataset partition, (b) the number of hidden neurons, (c) the number of time delays, and (d) training algorithm. Various combinations of dataset partitions were utilized to determine the data division that yields the best model performance while holding the hidden neurons and delay constant at 10 and 2, respectively. The optimum number of hidden neurons and time delay were also determined. Lastly, Levenberg-Marquardt and Bayesian Regularization were used as training algorithms.

The schematic diagram and architecture of the NARX network are shown in Figs. 1 and 2, respectively. The schematic diagram of a NARX network normally consists of three layers: the input layer, hidden layer, and output layer. Other elements involved in the network are the neurons, weights, and activation functions. The direction of the information fed to the input layer flows to the output layer. Through the network neurons, each input vector is multiplied by the weight vector to give the scalar product in each layer.

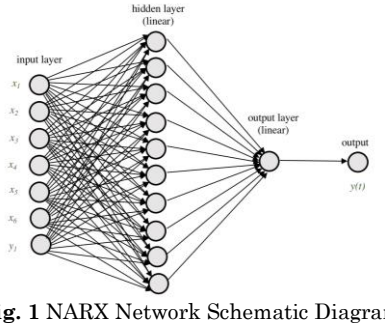


Fig. 1 NARX Network Schematic Diagram.

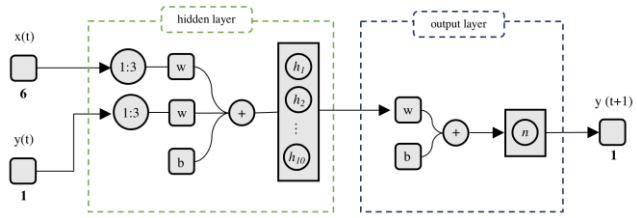


Fig. 2 NARX Network Architecture.

The architecture shown in Fig. 2 is a series-parallel (open-loop) NARX architecture, where the future value of the time series $\boldsymbol{y(t+1)}$ is predicted from both present and past values of $\boldsymbol{x(t)}$ and the true past values of the time series $\boldsymbol{y(t)}$. This architecture type shows a more precise forecast as input of the feedforward network uses true values as input.

2.2.3 Gated Recurrent Unit Network

The gated recurrent unit network is a relatively new variant of RNN which simplifies the architecture of the LSTM network which allows the GRU to be trained faster (Khan et al., 2021). The GRU neuron reduces the four gates of the LSTM into two gates as shown in Fig. 3. The improved network performance of the GRU network is due to the gating mechanism which allows the neurons to learn long-term dependencies of the input and output data.

The hidden state of the GRU, h_t, at a present time t was defined in Equation 5.

$$h_t = z_t\tilde{h}_t + (1 - z_t)h_{t-1} \quad (5)$$

The update gate of the GRU, z_t, was represented in Equation 6 which decides how much of the past information to retain. Consequently, the reset gate of the GRU, r_t, was defined in Equation 7 which determines how much of the past information to forget. The W_Z and W_R are the weights of the update gate and reset gate while the σ is a sigmoid function.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (6)$$

$$r_t = \sigma(W_R \cdot [h_{t-1}, x_t]) \quad (7)$$

The candidate activation state, $\tilde{\mathbf{h}}_\mathbf{t}$, is represented in Equation 8.

$$\tilde{h}_t = tanh(W \cdot [r_t \cdot h_{t-1}, x_t]) \quad (8)$$

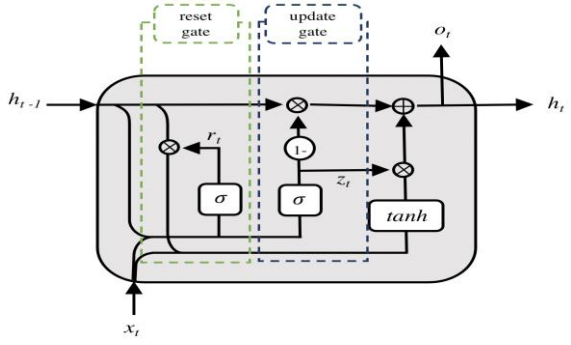


Fig. 3 GRU Cell Architecture.

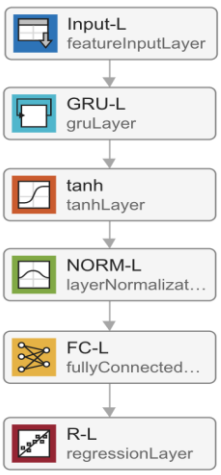


Fig. 4 Deep Network Design.

The deep network model which employs a GRU network was created using MATLAB's Deep Network Designer App. This neural network is shown in Fig. 4.

The deep network design was comprised of 6 layers. The Input-L is the feature input layer that expects the six variables stored using a datastore function. Next, the GRU-L is the GRU layer set to the default number of hidden units (128) with a tanh state activation function and sigmoid gate activation function. Consequently, a tanh layer was added as an activation function followed by a NORM-L layer to normalize the output variables. The normalized output variables were sent to an FC-L or fully connected layer, which is a linear layer. The output of the fully connected layer was set to 1, which was the predicted global solar radiation. It was sent to the R-L or regression layer, which determines the performance of the entire deep network.

The solver employed for the deep network training options was adam (adaptive moment estimation) optimizer with a default initial learning rate of 0.005. The validation frequency, maximum number of epochs, and mini-batch size were set to 50, 300, and 150, respectively.

2.2.4 Hybrid NARX-GRU network

The predicted global solar radiation, y', from the NARX network was utilized as an input variable in the GRU Network, hence the hybrid NARX-GRU model. The schematic diagram of the hybrid model is depicted in Fig. 5.

The NARX-GRU hybrid model was compared with other forecasting models such as the Gaussian Process Regression (GPR), NAR, and NARX models using different performance metrices. The dataset partition was set to 90% training and 10% validation of the hybrid, GPR, NAR, and NARX models. After the models were trained using the 90% training dataset (600 daily datapoints), the validation dataset (66 daily datapoints) was utilized to forecast day-ahead DGSR. The DGSR vs time (actual vs predicted) plot of each model was also generated to determine which model performed best. The trained hybrid model was also utilized to forecast DGSR starting July 21, 2018 to December 31, 2019 (529 daily datapoints) to visualize and explore the DGSR trend over such period to justify the hybrid model's forecasting accuracy.

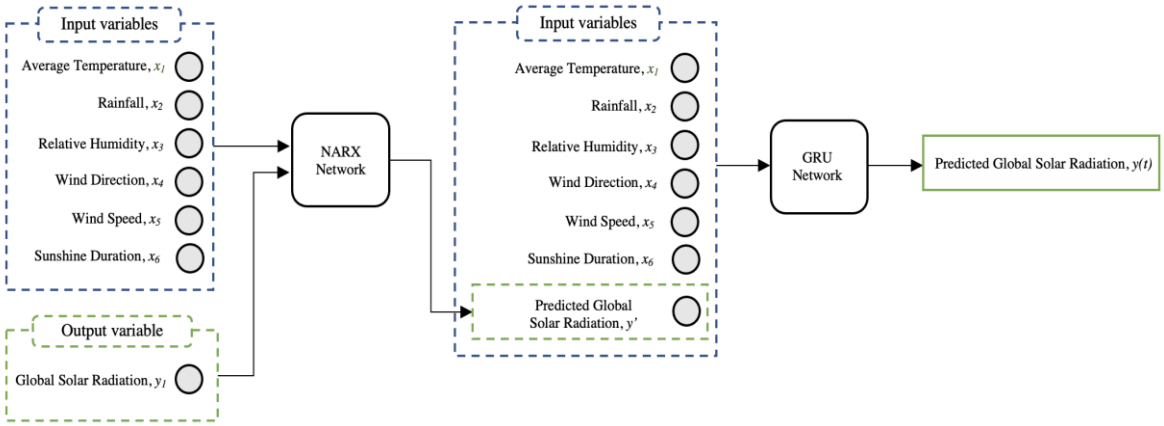


Fig. 5 NARX-GRU hybrid network diagram.

2.2.5 Performance metrics

Four performance parameters were used to evaluate the models' forecasting performance accuracy. These statistical parameters are defined as follows:

- Correlation

$$r = \frac{N\sum x_t^{Actual} x_t^{Predicted} - \left(\sum x_t^{Actual}\right)\left(\sum x_t^{Predicted}\right)}{\sqrt{\left[N\sum x_t^{(Actual)^2} - \left(\sum x_t^{Actual}\right)^2\right]\left[N\sum x_t^{(Predicted)^2} - \left(x_t^{Predicted}\right)^2\right]}} \quad (9)$$

when -1.0 ≤ r ≤ +1.0

- Coefficient of determination or r^2

$$r^2 = 1 - \frac{Unexplained\ variation}{Total\ variation} = 1 - \frac{\sum_i \left(x_t - x_t^{Predicted}\right)^2}{\sum_i \left(x_t - x_t^{Actual}\right)^2} \quad (10)$$

- Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(x_t^{Actual} - x_t^{Predicted}\right)^2} \quad (11)$$

- Mean Absolute Error (MAE)

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left|x_t^{Actual} - x_t^{Predicted}\right| \quad (12)$$

where x_t is the predicted target variable and T is the total sample size.

2.2.6 Software

MATLAB R2021a was used to set up both regression models and neural network design.

3. Results and Discussion

3.1 Data pre-processing

The raw daily meteorological data were plotted in two-dimensional plots as seen in Fig. 6 using MATLAB. These data were further processed using the Smooth Data task in MATLAB. Smoothing the data eliminated outlying data from the dataset and made trends or patterns more visible. Fig. 7 shows the data after pre-processing. After eliminating noise and outliers of the raw data through data smoothing and normalization, the two-dimensional plots of the meteorological data are shown in Fig. 7. The pre-processing of data improves the accuracy of the forecasting model.

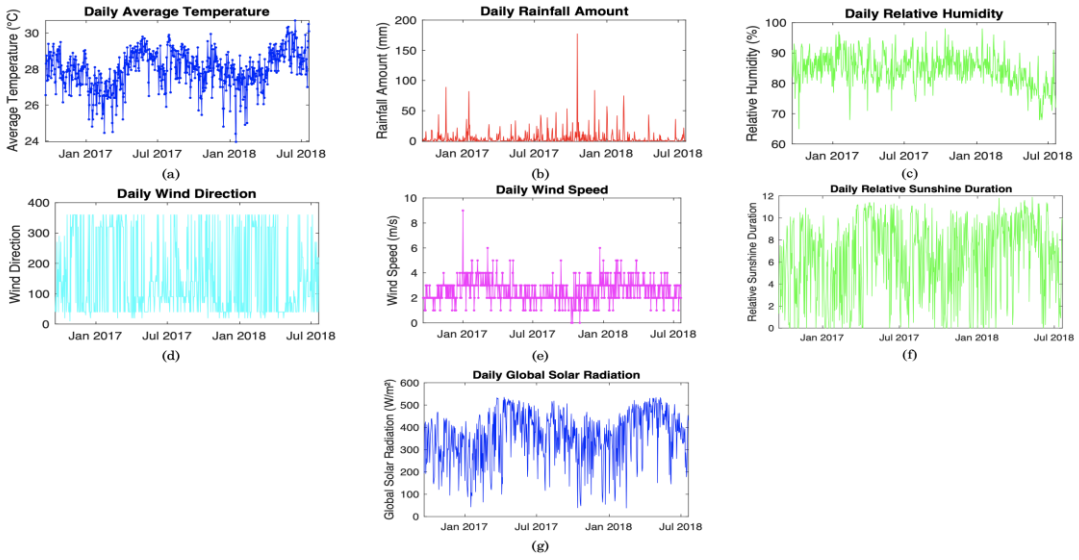


Fig. 6 Meteorological data from PAGASA and NSRC.

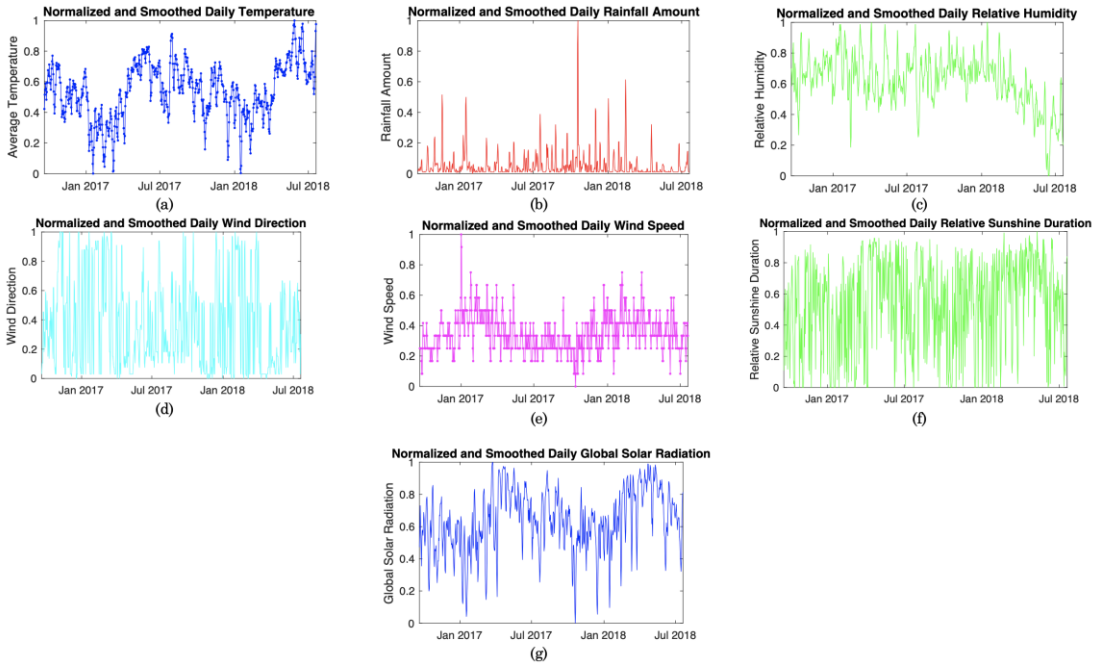


Fig. 7 Normalized and smoothed meteorological data.

Table 1
GPR model performance results at k=10 folds.

Model	RMSE (%)	R^2	MSE	MAE
Squared Exponential GPR	27.622	0.90	762.97	22.435
Matern 5/2 GPR	26.999	0.91	728.96	21.87
Exponential GPR	26.618	0.91	708.53	21.109
Rational Quadratic GPR	26.947	0.91	726.16	21.757

3.2 Regression models

After training a selection of different regression models with the normalized and smoothed data, results show that the Gaussian Process Regression (GPR) models had the best performance as evaluated from various performance indicators. From the comparison between different cross-validation folds, it can be found that ten folds can yield the best model performance. The extensive exploration on evaluating the influence and significance of each input variable on the prediction of daily solar irradiance reported higher metric scores when all available features were included. The performance of the trained GPR models is summarized in Table 1.

As shown in Table 1, the standard deviation of the above statistical parameters (RMSE, R^2, MSE, MAE) is positive. This non-zero value suggests that the selective different data inputs positively affect the GPR predictive model's performance. Regarding the regression framework, GPR works under the probabilistic framework. The GPR model takes N pairs of vector input and a noise scalar output as input to construct a model that delivers a well-distributed output at unseen input locations (Rohani *et al*., 2018).

$$y = f(x) + \varepsilon, \qquad \varepsilon \approx N\,(0, s^2_{noise}) \tag{13}$$

where s^2_{noise} is the variance of the noise output. In the Bayesian framework, unique indices follow consistent Gaussian distribution, limiting unnecessary function correlation. This then prioritizes the Gaussian process (GP) over functions. The consistency creates inference on function values that correspond to unseen inputs from the finite training data set.

In literature, many studies have determined that meteorological data such as temperature, cloud covers, sunshine duration, and solar irradiance affect the solar output (Duffie *et al*., 1994). The best result from the GPR regression feature selection was obtained when all features were included. Compared to other predictor variables, which did not affect the solar output significantly, the performance of the GPR model drastically dropped when the duration of sunshine predictor was not included. Sunshine duration has the greatest correlation coefficient value to the target data among the six predictors. A deliberate selection of appropriate features balances complexity and bias in a prediction model (Najibi *et al.,* 2021). Determining the proper features through feature selection prior to developing the final hybrid model impedes overfitting and adjusts its sensitiveness.

Table 2
Optimizable GPR model training hyperparameter results.

Parameter	Result
RMSE	27.741
R-squared	0.91
MSE	715.07
MAE	21.342

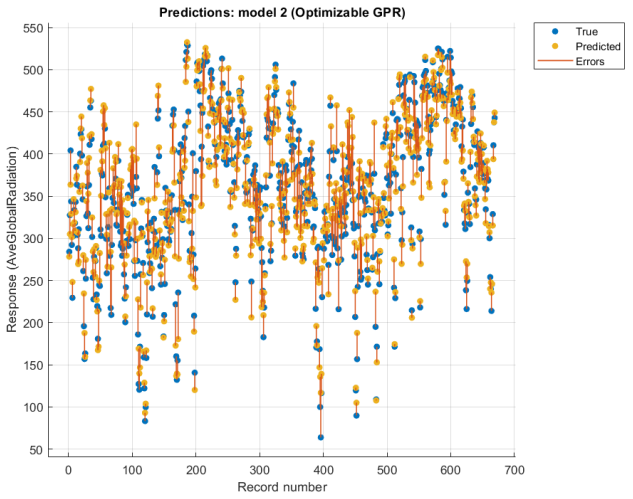


Fig. 8 Optimizable GPR model Response Plot.

In the Regression Learner app, after choosing the best model to train, the optimizable hyperparameter option automates the selection of hyperparameter values. Results from the training show that the simulation for optimizable GPR yields a better model performance compared to other GPR models. The results from the optimizable training are summarized in Table 2. Additionally, a response plot is shown in Fig. 8 in reflection of the validated optimized model results.

3.3. Parameter estimation

Results show that the NARX Network using the Bayesian Regularization (NARX-BR) training algorithm performed better compared to the NAR Network. The dataset partition was determined best at 85%, 10%, and 5% for the training, validation, and testing data, respectively. There were 569, 67, and 33 daily timesteps for the training, testing, and validation. The lowest training and testing mean square error was obtained using 10 hidden neurons and 3 delays for the NARX-BR Network. The NARX-BR Network with 50 hidden neurons and 3 delays performed well in the training dataset but has a significantly poor performance in the testing data, hence, the observance of overfitting. The correlation coefficient is 0.94542 and 0.91119 for the training and testing data, respectively, which indicates that the NARX-BR model is adequately fit for the data. The regression plots for the training, testing, and overall data for the NARX-BR Model are shown in Fig. 9.

As shown in Fig. 10a, the lowest validation error is obtained by the NARX-BR model at the 108th epoch, which indicates its best training performance. Further training the model at higher epochs reduced the error but overfit the training data. The error histogram in Fig. 10b shows that the errors range from -13.32 to 12.5.

All tests were performed both for the NARX and NAR models. After all the training procedures, the NARX model showed the best performance using the Bayesian Regularization algorithm with 10 hidden neurons, 3 delays, and a dataset partition of 85% for training, 10% for validation, and 5% for testing. The said model at these test parameters generated training, and testing mean square error of 836.49 and 1502.00, respectively. These parameters of the NARX-BR model were utilized to build the Hybrid NARX-GRU Network.

3.4. Hybrid NARX-GRU model

3.4.1. Training Progress

According to the training progress plots of the NARX-GRU hybrid model in Fig. 11, the RMSE recorded after multiple iterations reached an improved and better performance. The distribution of RMSE values was highly consistent after the 50th epoch, and the total training time was shortened (33 secs.) while at ~0 loss. This indicates that the proposed NARX-GRU model can effectively utilize the meteorological variables and the solar radiation information from the NARX network to predict solar radiation in parallel.

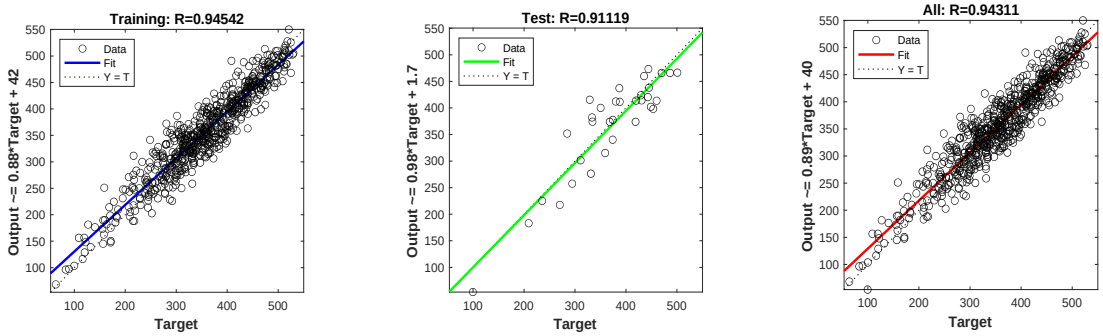
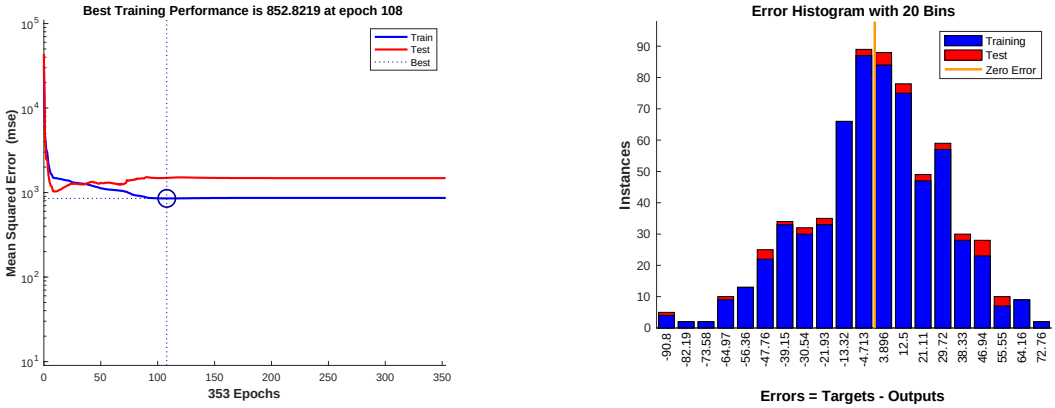


Fig. 9 Regression plots of the NARX-BR model (10 hidden neurons, 3 delays, 10% validation data, 5% testing data).



(a) (b)

Fig. 10 (a) Training performance and (b) Error histogram of the NARX-BR model (10 hidden neurons, 3 delays, 10% validation data, 5% testing data).

Fig. 11 displays an animated plot of the hybrid model's mini-batch loss and accuracy, and additional information on the training progress. Considering the study's small dataset, the best training performance for the hybrid model was small value options for the mini-batch size and validation frequency and loss. As monitored from the plot, the network accuracy quickly improved from an RMSE of ~2 to ~0.05 at the hundredth iteration. Simultaneously, keeping the loss at zero. The built GRU model was specifically applicable with adam solver algorithm. Since multiple input parameters are used, the adam optimizer adapts to its different learning rate (Chakrabarty *et al*., 2021). The adam algorithm solves the randomness of the system which leads to a local optimum result.

The trend in the RMSE plot shows a positive result, which suggests that the selective data inputs positively affect the overall performance of the hybrid model. In many literatures, GRU along with other deep learning models like LSTM, RNN, and hybrids are frequently applied. The GRU in this study's hybrid model was able to learn the data's short-term deficiencies, providing a more effective and compact overall performance. Since GRUs train less parameters and uses less memory, the training time of the model was executed with a fast elapsed time. The experimental results for the training process suggests that assembling the NARX network with the deep learning model GRU enhanced the model's predictive performance.

3.4.2. Model performance

This study uses evaluation metrics for a more comprehensive comparison between the previous predicting models and the recently proposed hybrid machine learning model. The comparison results are shown in Table 3, which indicates that the NARX-GRU hybrid model outperformed the Optimizable GPR, NAR, and NARX models. Likewise, it is noted that as the RMSE values decrease, forecasting preciseness is further achieved. It is evident that the NARX-GRU hybrid model has the best forecasting results with the most negligible losses and outperformed the Optimizable GPR, NAR, and NARX models.

Table 3
Comparison of models in terms of RMSE metric.

Model	RMSE (%)
Optimizable GPR	27.741
NAR	39.82
NARX	28.92
NARX-GRU Hybrid	~0.05

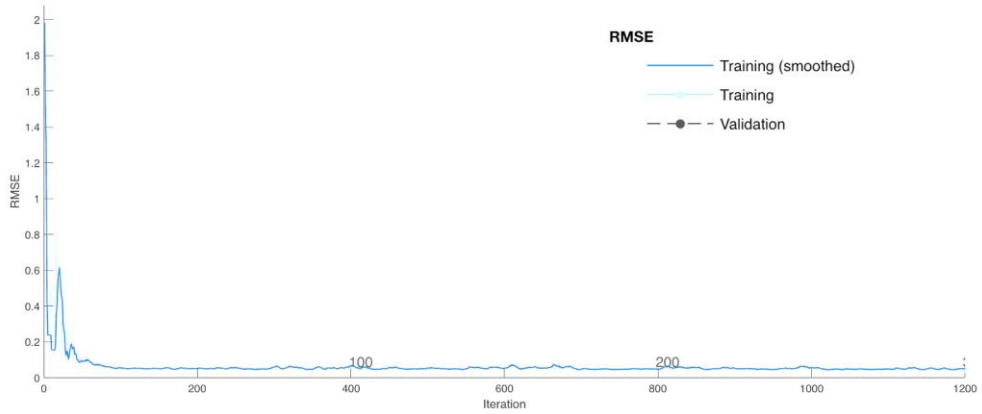
The actual DGSR values from NSRC and the forecasted DGSR values obtained from the hybrid NARX-GRU, NARX, and NAR are summarized in Table 3. The performance of the forecasting models is illustrated in Fig. 12. The forecast for day-ahead DGSR provided by the hybrid NARX-GRU, NARX, and NAR were plotted to show their accuracy. Fig. 12 shows that the hybrid model performed the best compared to the NARX and NAR. Statistically, the hybrid NARX-GRU model outperformed the models because of its significantly low RMSE value. The performance plots show that the other forecasting models still performed well.

The results presented in Fig. 12 highlight the fundamental role of the hybrid NARX-GRU model in forecasting daily global solar radiation. The hybridized network integrates the strengths of the individual forecasting models which generated better results compared to that of individual models being studied.

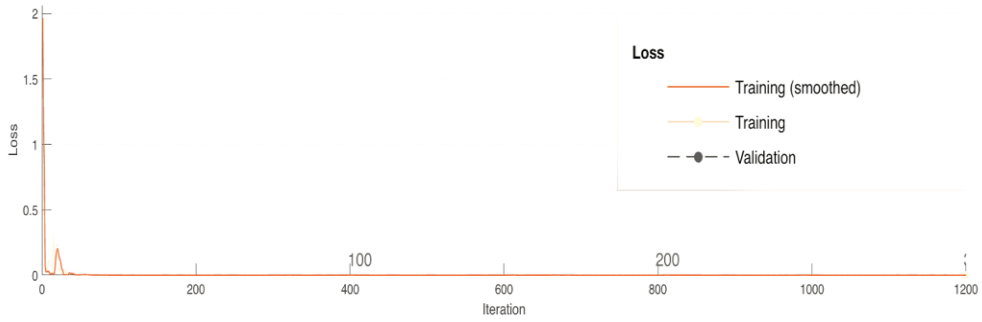
3.4.3. Forecast Results

The result for the daily global solar radiation forecast according to the trained and improved database obtained with the NARX-GRU hybrid model is shown in Fig. 13. The forecasted days are from July 2018 to December 2019 (529 datapoints). In this figure, the y-axis represents the DGSR in W/m^2, and the x-axis represents time in days. Actual DGSR data are represented by the red line, while the projected DGSR values with NARX-GRU are depicted by the blue lines. From Fig. 12, it was observed that the hybrid model fits quite well with low residuals. This suggests that the generated plot in Fig. 13 showed good results for day-ahead short-term forecasting. The forecasted DGSR were within the range of 32.3593 – 535.043 W/m^2.

The fluctuation trend on the forecasted plot can be explained by the variations in DGSR due to the changing weather conditions in Dumaguete, Philippines. It is evident from Fig. 13 that the highest recorded DGSR peaked during the warmest months in the country, which extends from March to June. Comparing the forecasted DGSR to the historical data, its fluctuation seemed larger. This is expected due to the increasing solar radiation potential and ultraviolet (UV) rays as a result of ozone layer depletion and climate change (Belmahdi *et al.*, 2020). The solar radiation time series plot shows the importance of precise adaptation models in improving the training data and ultimately develop accurate forecasting models.



(a)



(b)

Fig. 11 (a) RMSE and (b) loss NARX-GRU training process plot.

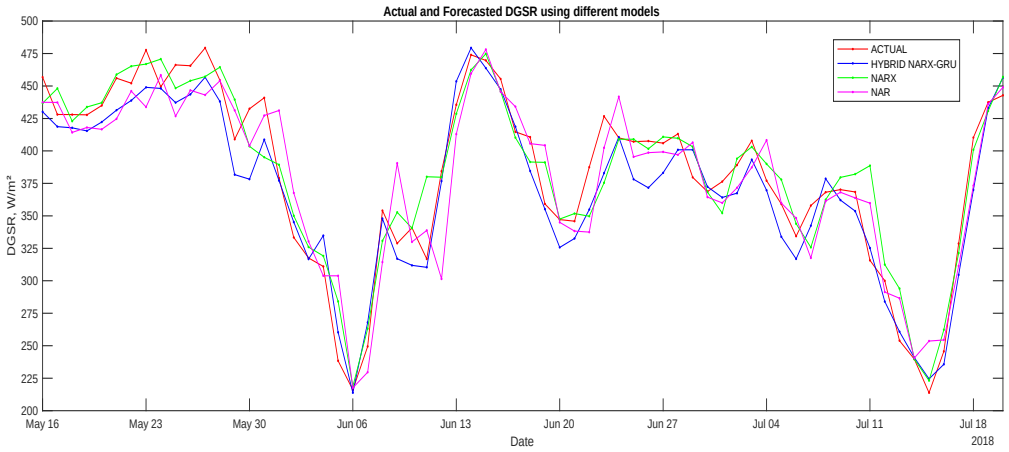


Fig. 12 Forecasting results of the NAR, NARX, and NARX-GRU models.

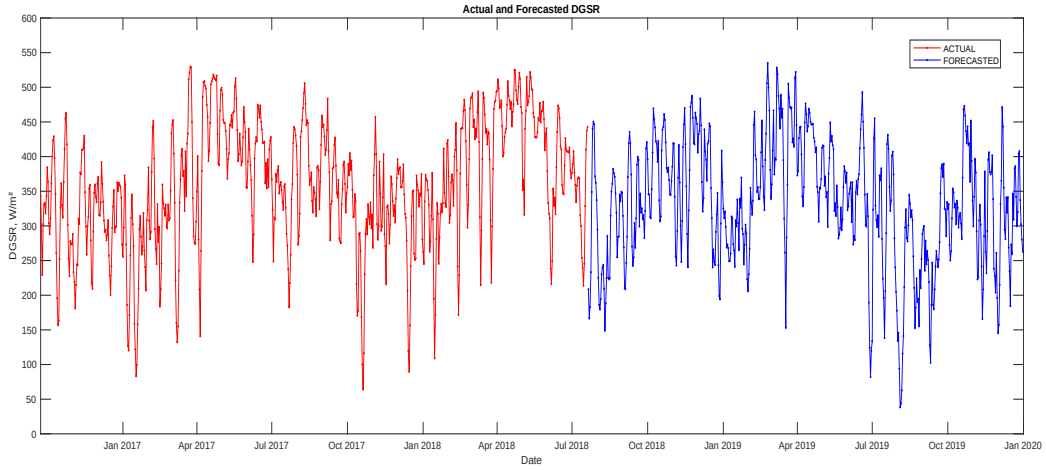


Fig. 13 NARX-GRU solar radiation time series.

4. Conclusion

As concerns on environmental protection and the global energy crisis continue to aggravate, PV power generation is becoming crucial on the energy portfolio of countries around the world. The Philippines' solar energy capacity has been exponentially increasing over the past decades, advantageous in PV power plant operations. In this paper, we developed a hybrid model for forecasting the global components of solar irradiance based on a machine learning algorithm – NARX-GRU.

The NARX-GRU model was equipped with hyperparameters to achieve the optimal model by minimizing the training loss. The data includes rainfall amount, relative humidity, duration of sunshine, average temperature, wind direction, and wind speed as input and solar radiation as the final output. The data were collected from PAGASA and NSRC with 669 data points. The NARX network was employed with the Bayesian Regularization algorithm to train at the optimal configuration of hyperparameters. A partition of 85% training, 10% validation, and 5% testing with 10 hidden neurons and 3 delays, showed the best NARX result of 0.91119 correlation and 1502.00 MSE (testing). The NARX output was then used for deep learning forecasting. We obtained optimal configuration with the hybrid model after arranging the provided hyperparameters such as solver type, learning rate, activation function, batch size, and batch normalization statistics.

The performance of the hybrid prediction model was measured through evaluation metrics. The proposed hybrid model performed well with minor errors and is comparable with other NNs. Compared with other state-of-the-art machine learning models, the results of this study showed that the NARX-GRU hybrid performed best. The hybrid model showed promising results for day-ahead forecasting. In addition, the training time of the NARX-GRU hybrid model is only at 33 secs., which shows that the proposed model has high training efficiency.

Acknowledgments

The authors would like to extend their gratitude to the Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA) and the National Solar Radiation Center (NSRC), for providing the needed historical data used in this study.

Author Contributions: A.F. and H.P.: Writing—original draft and editing, formal analysis, validation, methodology, B.D.: conceptualization, supervision, formal analysis, writing—review.

Funding: The author(s) received no financial support for the research, authorship, and/or publication of this article.

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

References

Abdel-Basset, M., Hawash, H., Chakrabortty, R. K., & Ryan, M. (2021). PV-Net: An innovative deep learning approach for efficient forecasting of short-term photovoltaic energy production. *Journal of Cleaner Production*, 303, 127037. https://doi.org/10.1016/J.JCLEPRO.2021.127037

Ahmad, A., Anderson, T. N., & Lie, T. T. (2015). Hourly global solar irradiation forecasting for New Zealand. Solar Energy, 122, 1398–1408. https://doi.org/10.1016/J.SOLENER.2015.10.055

Akhter, M. N., Mekhilef, S., Mokhlis, H., Ali, R., Usama, M., Muhammad, M. A., & Khairuddin, A. S. M. (2021). A hybrid deep learning method for an hour ahead power output forecasting of three different photovoltaic systems. *Applied Energy*, 118185. https://doi.org/10.1016/J.APENERGY.2021.118185

Al-Ghezi, M. K., Mahmoud, B. K., Alnasser, T. M., & Chaichan, M. T. (2022). A Comparative Study of Regression Models and Meteorological Parameters to Estimate the Global Solar Radiation on a Horizontal Surface for Baghdad City, Iraq. *International Journal of Renewable Energy Development*, 11(1), 71-81. https://doi.org/10.14710/ijred.2022.38493

Almonacid, F., Pérez-Higueras, P. J., Fernández, E. F., & Hontoria, L. (2014). A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator. *Energy Conversion and Management*, 85, 389–398. https://doi.org/10.1016/j.enconman.2014.05.090

Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., Martinez-de-Pison, F. J., & Antonanzas-Torres, F. (2016). Review of photovoltaic power forecasting. *Solar Energy*, 136, 78–111. https://doi.org/10.1016/J.SOLENER.2016.06.069

ArunKumar, K. E., Kalaga, D. v., Kumar, C. M. S., Kawaji, M., & Brenza, T. M. (2021). Forecasting of COVID-19 using deep

layer Recurrent Neural Networks (RNNs) with Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM) cells. *Chaos, Solitons & Fractals*, 146, 110861. https://doi.org/10.1016/J.CHAOS.2021.110861

Awad, M., & Khanna, R. (2015). *Machine Learning*. In Efficient Learning Machines (pp. 1–18). Apress. https://doi.org/10.1007/978-1-4302-5990-9_1

Belmahdi, B., Louzazni, M., & Bouardi, A. el. (2020). One month-ahead forecasting of mean daily global solar radiation using time series models. *Optik*, 219, 165207. https://doi.org/10.1016/J.IJLEO.2020.165207

Chakrabarty, A., Danielson, C., Bortoff, S. A., & Laughman, C. R. (2021). Accelerating self-optimization control of refrigerant cycles with Bayesian optimization and adaptive moment estimation. *Applied Thermal Engineering*, 197, 117335. https://doi.org/10.1016/J.APPLTHERMALENG.2021.117335

Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). *Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation*. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1724–1734. https://doi.org/10.3115/v1/D14-1179

Deutsche GIZ GmbH, Renewable Energy Developers Center, & WWF Philippines. (2013). IT'S MORE SUN IN THE PHILIPPINES Facts and Figures on Solar Energy in the Philippines Project Development Programme (PDP) Southeast-Asia. https://www.doe.gov.ph/sites/default/files/pdf/netmeter/policy-brief-its-more-sun-in-the-philippines-V3.pdf

Diagne, H. M., David, M., Lauret, P., Boland, J., Schmutz, N., Review, N. S., & Ma¨ımouna Diagne, M. (n.d.). Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renewable and Sustainable Energy Reviews*, 27, 65–76. https://doi.org/10.1016/j.rser.2013.06.042

Duffie, J. A., Beckman, W. A., and Worek, W. M. (1994). Solar Engineering of Thermal Processes, 2nd ed.. ASME. *J. Sol. Energy Eng*. February 1994;. https://doi.org/10.1115/1.2930068

Faisal, A. N. M. F., Rahman, A., Habib, M. T. M., Siddique, A. H., Hasan, M., & ; Khan, M. M. (2022). Neural networks based multivariate time series forecasting of solar radiation using meteorological data of different cities of Bangladesh. *Results in Engineering*, 13, 100365. https://doi.org/10.1016/J.RINENG.2022.100365

Gonzaga Baca Ruiz, L., Pegalajar Cuéllar, M., Delgado Calvo-Flores, M., del Carmen Pegalajar Jiménez, M., Riquelme, J. C., Troncoso, A., & Martínez-Álvarez, F. (2016). An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings. *Energies,* 9(9). https://doi.org/10.3390/en9090684

Huang, X., Li, Q., Tai, Y., Chen, Z., Zhang, J., Shi, J., Gao, B., & Liu, W. (2021). Hybrid deep neural model for hourly solar irradiance forecasting. *Renewable Energy*, 171, 1041–1060. https://doi.org/10.1016/J.RENENE.2021.02.161

Husein, M., & Chung, I.-Y. (2019). Day-Ahead Solar Irradiance Forecasting for Microgrids Using a Long Short-Term Memory Recurrent Neural Network: A Deep Learning Approach. *Energie*s, 12(10). https://doi.org/10.3390/en12101856

Jaihuni, M., Basak, J. K., Khan, F., Okyere, F. G., Sihalath, T., Bhujel, A., Park, J., Lee, D. H., & Kim, H. T. (2021). A novel recurrent neural network approach in forecasting short term solar irradiance. *ISA Transactions*. https://doi.org/10.1016/j.isatra.2021.03.043

Jung, Y., Jung, J., Kim, B., & Han, S. U. (2020). Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: Case study of South Korea. *Journal of Cleaner Production*, 250, 119476. https://doi.org/10.1016/J.JCLEPRO.2019.119476

Khan, A. T., Khan, A. R., Li, S., Bakhsh, S., Mehmood, A., & Zaib, J. (2021). Optimally configured Gated Recurrent Unit using Hyperband for the long-term forecasting of photovoltaic plant. *Renewable Energy Focus*, 39, 49–58. https://doi.org/10.1016/J.REF.2021.07.002

Kumari, P., & Toshniwal, D. (2021). Deep learning models for solar irradiance forecasting: A comprehensive review. *Journal of Cleaner Production*, 318, 128566. https://doi.org/10.1016/J.JCLEPRO.2021.128566

Lai, C. S., Zhong, C., Pan, K., Ng, W. W. Y., & Lai, L. L. (2021). A deep learning based hybrid method for hourly solar radiation forecasting. *Expert Systems with Applications*, 177, 114941. https://doi.org/10.1016/j.eswa.2021.114941

Li, J., Ward, J. K., Tong, J., Collins, L., & Platt, G. (2016). Machine learning for solar irradiance forecasting of photovoltaic system. *Renewable Energy*, 90, 542–553. https://doi.org/10.1016/j.renene.2015.12.069

Li, P., Zhou, K., Lu, X., & Yang, S. (2020). A hybrid deep learning model for short-term PV power forecasting. *Applied Energy*, 259, 114216. https://doi.org/10.1016/J.APENERGY.2019.114216

Liu, X., Lin, Z., & Feng, Z. (2021). Short-term offshore wind speed forecast by seasonal ARIMA - A comparison against GRU and LSTM. *Energy*, 227, 120492. https://doi.org/10.1016/J.ENERGY.2021.120492

Liu, Y., Qin, H., Zhang, Z., Pei, S., Wang, C., Yu, X., Jiang, Z., & Zhou, J. (2019). Ensemble spatiotemporal forecasting of solar irradiation using variational Bayesian convolutional gate recurrent unit network. *Applied Energy*, 253, 113596. https://doi.org/10.1016/J.APENERGY.2019.113596

Mellit, A., Pavan, A. M., & Lughi, V. (2021). Deep learning neural networks for short-term photovoltaic power forecasting. *Renewable Energy*, 172, 276–288. https://doi.org/10.1016/J.RENENE.2021.02.166

Najibi, F., Apostolopoulou, D., & Alonso, E. (2021). Enhanced performance Gaussian process regression for probabilistic short-term solar output forecast. *International Journal of Electrical Power & Energy Systems*, 130, 106916. https://doi.org/10.1016/J.IJEPES.2021.106916

Narvaez, G., Giraldo, L. F., Bressan, M., & Pantoja, A. (2021). Machine learning for site-adaptation and solar radiation forecasting. *Renewable Energy*, 167, 333–342. https://doi.org/10.1016/j.renene.2020.11.089

Pazikadin, A. R., Rifai, D., Ali, K., Malik, M. Z., Abdalla, A. N., & Faraj, M. A. (2020). Solar irradiance measurement instrumentation and power solar generation forecasting based on Artificial Neural Networks (ANN): A review of five years research trend. *Science of the Total Environment*, 715, 136848. https://doi.org/10.1016/j.scitotenv.2020.136848

Pisoni, E., Farina, M., Carnevale, C., & Piroddi, L. (2009). Forecasting peak air pollution levels using NARX models. *Engineering Applications of Artificial Intelligence*, 22(4), 593–602. https://doi.org/10.1016/j.engappai.2009.04.002

Puah, B. K., Chong, L. W., Wong, Y. W., Begam, K. M., Khan, N., Juman, M. A., & Rajkumar, R. K. (2021). A regression unsupervised incremental learning algorithm for solar irradiance prediction. *Renewable Energy*, 164, 908–925. https://doi.org/10.1016/j.renene.2020.09.080

Rohani, A., Taki, M., & Abdollahpour, M. (2018). A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (Part: I). *Renewable Energy*, 115, 411–422. https://doi.org/10.1016/J.RENENE.2017.08.061

Sharma, A., & Kakkar, A. (2018). Forecasting daily global solar irradiance generation using machine learning. *In Renewable and Sustainable Energy Reviews* (Vol. 82, pp. 2254–2269). Elsevier Ltd. https://doi.org/10.1016/j.rser.2017.08.066

Wang, F., Xuan, Z., Zhen, Z., Li, K., Wang, T., & Shi, M. (2020). A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework. *Energy Conversion and Management,* 212, 112766. https://doi.org/10.1016/J.ENCONMAN.2020.112766

Wang, H., Cai, R., Zhou, B., Aziz, S., Qin, B., Voropai, N., Gan, L., & Barakhtenko, E. (2020). Solar irradiance forecasting based on direct explainable neural network. *Energy Conversion and Management,* 226, 113487. https://doi.org/10.1016/j.enconman.2020.113487

Yadav, A. K., & Chandel, S. S. (2014). Solar radiation prediction using Artificial Neural Network techniques: A review. *Renewable and Sustainable Energy Reviews,* 33, 772–781. https://doi.org/10.1016/J.RSER.2013.08.055