International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

DenPow Model for Solar PV Power Generation Forecasting

Kaustubha H. Shedbalkar*[1,2] and D. S. More[3]

Submitted: 29/01/2024 **Revised**: 07/03/2024 **Accepted**: 15/03/2024

Abstract*:* Due to environment dependent variation in solar power generation, in a day or throughout year, there is need of forecasting. This can help for planning at the distribution centre for selection of feed from different power plants. This can help to reduce the generation losses at the hydro power plants by putting load on solar plants during their peak generation hours. The historic power generation records of entire year can be used for time ahead forecasting of the power generation. The data used contains fields such as power generated, temperature of module and environment, irradiation, wind speed and rainfall. This paper addresses the forecasting challenge of solar power generation with 'DenPow' model. The model shows maximum error of 3% in predicted value compared to ground truth. The model outperforms over artificial neural network (ANN), recurrent neural network (RNN), RNN model with the combination of Long Short Term Memory (LSTM) and Auto-GRU layers and with combination of CNN and LSTM methods.

Keywords: *Solar power plant, Forecasting, Regression, Deep neural network, Performance.*

1. Introduction

Electricity distribution management is based on forecasting of power generation from different plants. The solar power generation forecasting which is volatile in nature is important factor that affects in such prediction and may lead to mismanagement. The solar power generation is directly dependent on different environmental parameters in which irradiance is important one. Also, solar panels show variation in power generation with respect to temperature of the environment. The temperature is dependent on various weather conditions which include wind speed.

The forecasting of solar power generation is the main objective considered in the work presented in this paper. In this context, an overview of different forecasting methods of solar irradiation which make use of machine learning methods are addressed. Mostly, conventional methods are used by various researchers such as support vector machines. The variety in datasets, time series variations, changes in geographical locations and respective weather station data are the main complications in these methods. Overall, the error of prediction is quite equivalent. The consideration of different parameters in power output prediction such as wind speed, temperature, irradiance with one at a time approach effect on predicted output power is analysed. Some authors also used ensemble approach and hybrid approach for improving the performance of prediction. In this paper, regression with the use of fully connected dense network model is proposed for the solar power plant in Maharashtra state of India which is located near the tropic of Capricorn.

2. Literature Survey

Extreme artificial neural network training is carried out for solar power forecasting by (F. Wang et al., 2015). The yield of the logistic function exhibits the best performance when compared to other activation functions. Regarding class regression analysis, the logistic-based regression model is taken into account. Through thorough research and various combinations of experiments, the ideal architectural model is discovered. (Teo et al., 2016) employed wavelet neural network and residual markov chain to estimate electricity generation. Correlation coefficients are obtained and key climatic variables that have the biggest impacts on photovoltaic power generation are found. Hua et al. (2016) employed straightforward data-driven models to estimate photovoltaic power. The availability and communication costs of the meteorological data are differentiated. (Munshi & Mohamed, 2016) used a bio-inspired clustering algorithm for doing data clustering as part of the preprocessing stage. Six clustering techniques are used for the experimental analysis, and the best technique is determined. The optimum weight learning approach was employed in (Z. Wang & Koprinska, 2017). The data source weighed closest neighbour (DWkNN) approach is suggested. Numerous studies demonstrate the use of traditional neural network-based techniques. Support vector machine (SVM) regression models can also be used to perform the regression. (Eseye et al., 2018) used an SVM regression model to predict the power generation. The wavelet transform and particle swarm optimization are used to linearize and pre-process the data. Actual recorded data from SCADA systems and meteorological data from the weather service are both used to forecast the production of

[1a]*Research Scholar, Electrical Engineering, Department, Walchand College of Engineering, Sangli,Maharashtra India*
[2]*Assistant Professor,Electrical Engineering, Department,Sharad Institute of Technology, Yadrav,Maharashtra India.*
[3]*Professor,Electrical Engineering Department, Walchand College of Engineering, Sangli,Maharashtra India*
** Corresponding Author Email: khshedbalkar@gmail.com*

power. By combining the results of neural network regression models, the difference between predicted values and actual values is reduced. Some researchers employed Gaussian regression models for predicting because the sparseness of the yearly dataset leads one to consider its Gaussian nature. (Semero et al., 2018) used a genetic algorithm (GA) in conjunction with a Gaussian regression analysis to identify key variables that influence electricity generation. Additionally, it is conceivable to consider some manually created fuzzy rules, which is doable with the aid of an adaptive neurofuzzy inference system (ANFIS). The ANFIS is used to forecast the data in (Semero et al., 2018) and (Manjili et al., 2018), pre-processed datasets for serial time domain analysis utilizing multivariate analysis. When the overall impact on prediction with respect to each individual variable is taken into account, the various variables used as input play a considerable role. Data can be normalised during pre-processing such that all values fall inside a given range. The normalisation has improved the ANN-based prediction. K-means, GRA, and Elman approaches are employed for prediction in (Lin et al., 2018). As an extra factor in power prediction, meteorological information is used. Wind speed and rainfall are two components of the meteorological data.

(Park & Park, 2018) employed relation detection inside clusters in addition to a hierarchical clustering technique. The clustering method results in the identification of the seasonal and flawed data. PV power generation is forecasted using a spatio-temporal model and probability density function (Agoua et al., 2019). Optimised SVM model was utilised by (VanDeventer et al., 2019) to predict power generation. The optimization is accomplished using the genetic method (GA). For estimating PV and wind power generation, (Sanjari et al., 2020) created a higher order multivariate Markov Chain-based approach. Pattern-based analysis employing window-based time-adaptive stochastic correlation is applied to PV and wind power. In the training state, the model needs both inputs at once. The performance indicators vary since solar irradiance and temperature have no effect on the production of wind power. On the other hand, the amount of wind has little effect on the production of solar power.

Irradiance is the most important factor in determining how much solar power will be produced by a power plant. Additionally, it has been observed that a lot of researchers also take other environmental characteristics obtained from weather station data into account. This paper's results and analysis section discusses the importance of each parameter for projecting solar power generation.

3. Proposed work

The block diagram in figure 1 depicts the intended work's stages. To improve neural network convergence, the input dataset is pre-processed to create more feature vectors. Next, the electricity generation on the same day the following year is predicted using the trained model created using this data.

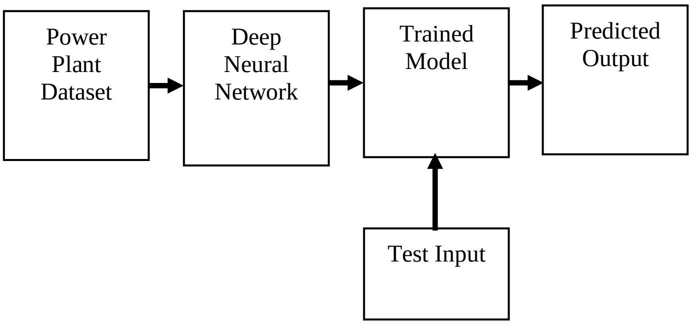

Fig 1: Proposed system Block diagram

Assumptions

The model consist of dense layers to perform regression task

The input dataset from actual power plant is used for training

The layers in the model are changed to observe the training results.

The proposed model consist of dense fully connected neural network layers which perform regression task on the input data vectors in X for target outputs Y. The number of layers in the model is concerned about complexity of the model as well as performance of prediction. As the prediction is regression task for proposed work, root mean square error metric is used to improve the training performance in number of iterations. The models are varied with increment in number of layers used during composition which on the variation shows change in performance during training of the model.

Figure 2 shows the multi-layer perceptron model with 11 layers. The increment in layers improves the performance of model. The increment after total 11 layers is stopped as performance of model after 11 layers decreases due to over fitting problem.

The output of each dense layer in generalized way can be represented as,

$$Y^{in}=X*w+b$$

$$Y^{out}=ReLU(Y^{in}) \quad ..(1)$$

Where, Yin is output from neurons and Yout is output due to activation function rectified linear unit.

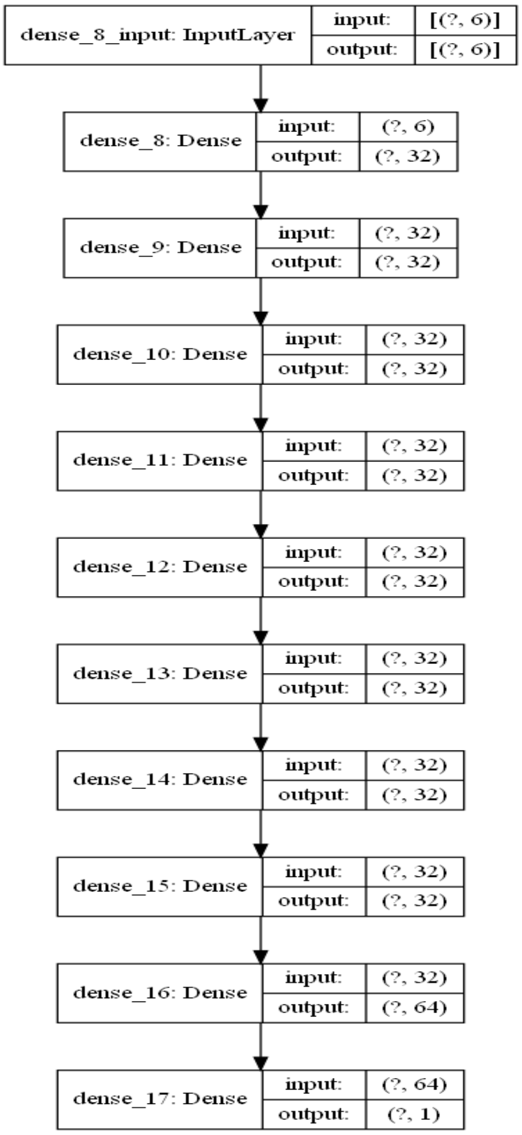

Fig 2: Proposed DenPow Model and other models used during development

4. Result Analysis

4.1 Data Pre-processing

For two years, all data were collected from two sources (2018-2019). The first dataset is power generation records from actual inverter. The power plant output in terms of generated power, module temperature, ambient temperature and solar irradiance are available in first dataset. In addition, second dataset consist of meteorological parameters like maximum and minimum temperature, maximum and minimum humidity, maximum and minimum wind speed and rainfall. There parametric records in second dataset are taken from Skymet Weather Services Pvt. Ltd. All the records are taken at 15 minutes time interval.

The additional features are extracted by estimating the mean of the total temperature of entire year, standard deviation on each day in ambient temperature and module temperature. In total 6 input features for each 15 minutes time interval are used for training with target power generation value.

The Mean of the temperature is obtained as,

$$Tm = \frac{\sum_{i=1}^{N} T(i)}{N}$$

(2)

Where, there are total N values of temperature in a year.

The standard deviation is thus obtained as,

$$\sigma\,(i) = Tm - T(i)$$

i=1,2,...,N

...(3)

Training Mean Square Error Analysis is performed for three models as shown in figure 3. These graphs show that final minimum value of root mean square achieved during training is similar in all the models but with increase in layers, this minimum value is seen to be achieved at early epochs. The performance shown in figure 3 is evaluated for all the dataset parameters as input in which solar irradiance, wind speed , temperature from weather station data are considered. The 11 layered model is named as Denpow model.

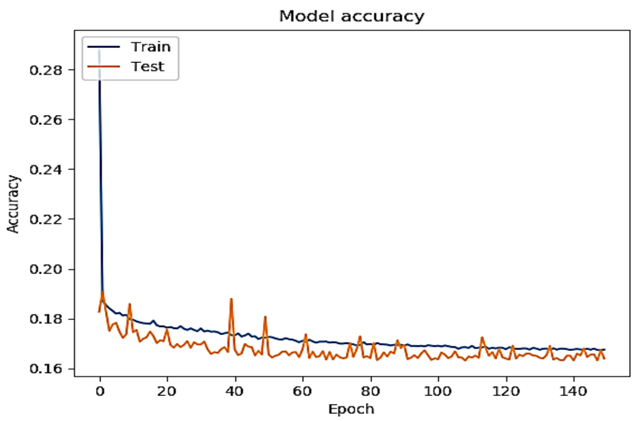

Fig 3: RMSE analysis of DenPow Model

The RMSE is estimated as,

$$\boldsymbol{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} \|y(i) - y'(i)\|^2}{N}} \quad (4)$$

The Mean Absolute Error (MAE) is estimated as,

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N} (y(i) - y'(i))}{N} \quad (5)$$

Where, y' is prediction for ith sample and y is actual value in total N samples.

Analysis of Power output prediction with respect to different parameters, in which one at a time and then all together taken as input to DenPow model, is shown figure 4. The graph shows the effective parameter is irradiance.

A comparative of model with 9 layers and 10 layers is shown in figure 5. The 11 layered model shows better performance and prediction very near to actual.

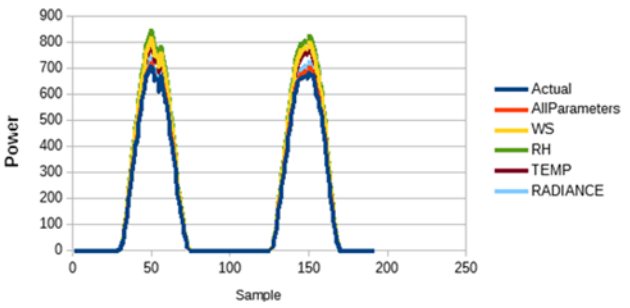

Fig 4: Individual parameter based analysis compared with actual Power Generation

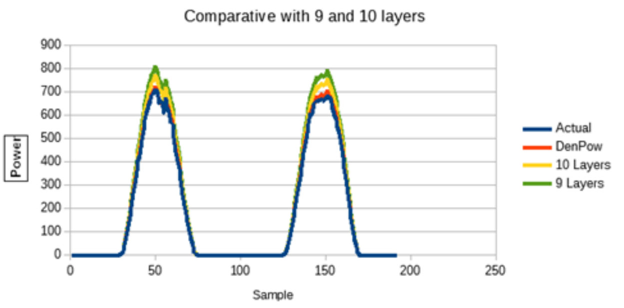

Fig 5: Comparative of prediction of models with 9 layers and 10 layers with DenPow

The performance is evaluated for prediction of power generation in which the recorded data of entire year of 2018 is used for training and data for 2019 is predicted. The per day error analysis is performed for DenPow Model. The performance of prediction for 5 Days and it's percent error is shown in figure 6 and 7 respectively.

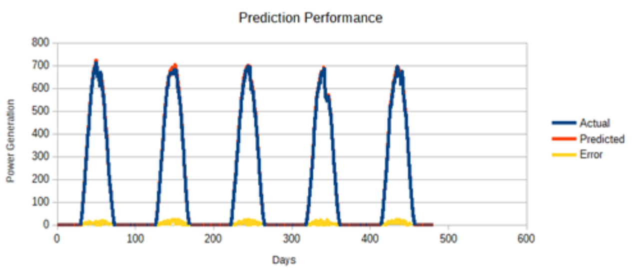

Fig 6: Prediction performance

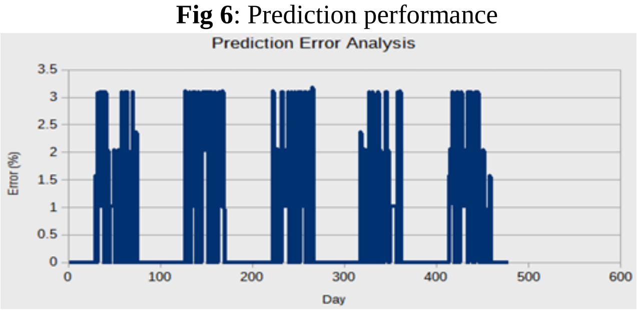

Fig 7: Percent prediction error analysis

The analysis with the use of K-fold approach is performed in which 5 folds are done for training the model. The performance in terms of RMSE is seen improved in 5th fold of the training. There are 35040 samples in the dataset of entire year recording. In the 5-fold analysis, the dataset is split in 80%-20% parts for training and validation sets. The validation set as gets selected randomly in each fold, the average RMSE value goes on decreasing in each fold and found minimum in 5th fold

Table 1: 5 Fold Analysis

Fold	**Total Sample Size**	**Training Sample Size**	**Training Samples Set Indexes**	**Testing sample size**	**Testing Samples Set Indexes**
1	35040	28032	1,2,3,4	7008	5
2	35040	28032	2,3,4,5	7008	1
3	35040	28032	3,4,5,1	7008	2
4	35040	28032	4,5,1,2	7008	3
5	35040	28032	5,1,2,3	7008	4

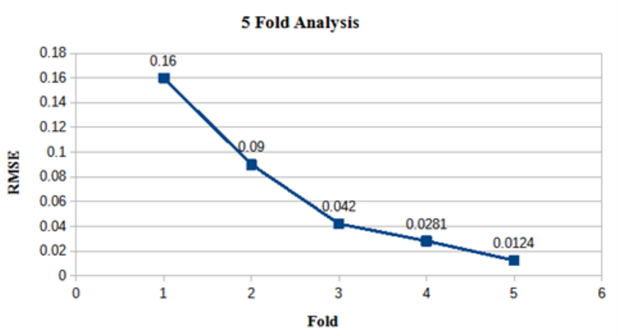

Fig 8: 5-fold analysis

4.2 Comparative Analysis

(González Ordiano et al., 2016), have focused on development of system for forecasting PV output on data driven models and without considering weather data. The models used show experimentation with artificial neural network (ANN) with 6 neurons. The model shows less RMSE compared to other models in their paper. This model is compared with DenPow model for finding out the better model.

(G. Li et al., 2019) used recurrent neural network (RNN) for forecasting the PV output. The intra-day data input based comparative analysis in this paper shows LSTM as better model for forecasting the PV output. The performance evaluation strategy from authors include analysis of mean absolute percent error (MAPE) estimation at different time horizon. We consider the data with 15 minutes time interval on the same model for comparative analysis.

(AlKandari et al., 2022) developed an RNN model with the combination of LSTM and Auto-GRU layers. The ensemble approach used during composition of hybrid model of machine learning, the weighted averaging with linear and nonlinear approach and simple averaging are combined. The performance of this model is compared with our DenPow Model.

(Lim et al., 2023) developed a hybrid model with combination of CNN and LSTM. The 1D CNN model focuses on forecasting of weather data while the predicted vector is fed to LSTM layer which then predicts the power.

The mean absolute percent error is given as,

$$MAPE = MAE * 100 \quad \text{... (6)}$$

Where MAE is estimated by using equation (5).

Table 2 : Comparative Results

Method/Parameter	**RMSE**	**MAE**	**MAPE**
ANN6 (González Ordiano et al., 2016)	0.32	0.042	4.2%
RNN (G. Li et al., 2019)	0.231	0.036	3.6%
RNN-LSTM-AutoGRU(AlKandari et al.,	0.129	0.0877	3.4%
CNN-LSTM(Lim et al., 2023)	0.4387	0.34	7.21%
DenPow	0.012	0.029	2.9%

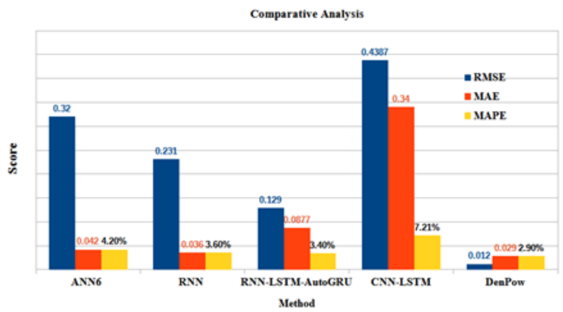

Fig 9: Comparative analysis of state-of-the-art methods and DenPow performance

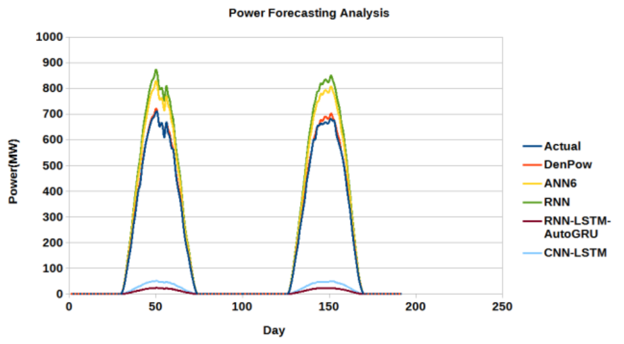

Fig 10: Comparative of Power Generation Prediction

Figure 9 ,10 and table 2 shows the comparative analysis of proposed DenPow model with other models. The proposed model shows good match with that of actual power with slight error in prediction.

5 Conclusion

This paper contributes for solar power generation forecasting using yearly data for training the deep neural network model for predicting next year power generation. The DenPow model is proposed which shows better results with maximum up to 3% error while predicting the power generation using yearly historic data. The data contains various parametric values of solar irradiation. Also, simultaneous of different weather parameters are also used during training of the proposed model. The comparative study is performed to identify the significance of each parameter with respect to error in prediction. Also, it is observed that all the parameters play important role for boosting the prediction accuracy. Based on experimental observations, the multiple parameters based approach is most suitable for better accuracy. The comparative analysis shows that the proposed model outperforms over state-of-the-art methods.

Acknowledgements

Authors would like to thank Galileo Solar India Pvt. Ltd. and Skymet Weather Services Pvt. Ltd. for providing real-time site data for experimentation.

Author contributions

The paper background work, conceptualization, methodology, Dataset collection, implementation, result analysis and comparison, preparing and editing draft, visualization, review of work and project administration, have been done by Kaustubha H Shedbalkar under the supervision of Dr D S More.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Wang, F., Zhen, Z., Mi, Z., Sun, H., Su, S., & Yang,

G. “Solar irradiance feature extraction and support vector machines based weather status pattern recognition model for short-term photovoltaic power forecasting”, *Energy and Buildings*, Vol. 86, pp 427–438, 2015. 10.1016/J.ENBUILD.2014.10.002

[2] Teo, T. T., Logenthiran, T., & Woo, W. L. “Forecasting of photovoltaic power using extreme learning machine”. *Proceedings of the 2015 IEEE Innovative Smart Grid Technologies - Asia, ISGT ASIA 2015*, 2015. 10.1109/ISGT-ASIA.2015.7387113

[3] Hua, X., Le, Y., Jian, W., & Agelidis, V. ,” Short-term power forecasting for photovoltaic generation based on wavelet neural network and residual correction of Markov chain”. *Asia-Pacific Power and Energy Engineering Conference, APPEEC*, 2016-January. 10.1109/APPEEC.2015.7381043

[4] Munshi, A. A., & Mohamed, Y. A. R. I.,”Photovoltaic power pattern clustering based on conventional and swarm clustering method”, *Solar Energy*, Vol. 124, pp. 39–56, 2016 .10.1016/J.SOLENER.2015.11.010

[5] Wang, Z., & Koprinska, I, “ Solar power prediction with data source weighted nearest neighbors”, *Proceedings of the International Joint Conference on Neural Networks*, pp. 1411–1418, May 2017. 10.1109/IJCNN.2017.7966018

[6] Eseye, A. T., Zhang, J., & Zheng, D.” Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and Meteorological information”, *Renewable Energy*, Vol. 118,pp. 357–367, 2018. 10.1016/J.RENENE.2017.11.011

[7] Semero, Y. K., Zhang, J., & Zheng, D., ”PV power forecasting using an integrated GA-PSO-ANFIS approach and Gaussian process regression based feature selection strategy”, *CSEE Journal of Power and Energy Systems*,Vol, 4, no.2, pp. 210–218, 2018. 10.17775/CSEEJPES.2016.01920

[8] Manjili, Y. S., Vega, R., & Jamshidi, M. M, “ Data-Analytic-Based Adaptive Solar Energy Forecasting Framework”, *IEEE Systems Journal*, Vol. 12, no. 1, pp. 285–296, 2018. 10.1109/JSYST.2017.2769483

[9] Lin, P., Peng, Z., Lai, Y., Cheng, S., Chen, Z., & Wu, L., “Short-term power prediction for photovoltaic power plants using a hybrid improved Kmeans-GRA-Elman model based on multivariate meteorological factors and historical power datasets”, *Energy Conversion and Management*, Vol. 177, pp. 704–717,2018. 10.1016/J.ENCONMAN.2018.10.015.

[10] Park, S., & Park, Y. B.,” Photovoltaic power data analysis using hierarchical clustering”, *International Conference on Information Networking*, pp. 727–731, January 2018. 10.1109/ICOIN.2018.8343214

[11] Agoua, X. G., Girard, R., & Kariniotakis, G., “Probabilistic Models for Spatio-Temporal Photovoltaic Power Forecasting”, *IEEE Transactions on Sustainable Energy*, Vol.10,no. 2, pp. 780–789, 2019. 10.1109/TSTE.2018.2847558

[12] VanDeventer, W., Jamei, E., Thirunavukkarasu, G. S., Seyedmahmoudian, M., Soon, T. K., Horan, B., Mekhilef, S., & Stojcevski, A., “Short-term PV power forecasting using hybrid GASVM technique”. *Renewable Energy*, Vol. 140, pp. 367–379, 2019. 10.1016/J.RENENE.2019.02.087

[13] Sanjari, M. J., Gooi, H. B., & Nair, N. K. C. , “Power generation forecast of hybrid PV-Wind system”, *IEEE Transactions on Sustainable Energy*, Vol. 1, no. 2, pp. 703–712, 2020. 10.1109/TSTE.2019.2903900

[14] González Ordiano, J. Á., Waczowicz, S., Reischl, M., Mikut, R., & Hagenmeyer, V. ,”Photovoltaic power forecasting using simple data-driven models without weather data”, *Computer Science - Research and Development 2016*, Vol. 32, no.1, pp. 237–246, 2016. 10.1007/S00450-016-0316-5

[15] Li, G., Wang, H., Zhang, S., Xin, J., & Liu, H. , “ Recurrent Neural Networks Based Photovoltaic Power Forecasting Approach”, *Energies* Vol. 12,no. 12, pp 2538, 2019. 10.3390/EN12132538

[16] AlKandari, Mariam, and Imtiaz Ahmad. “Solar Power Generation Forecasting Using Ensemble Approach Based on Deep Learning and Statistical Methods.” *Applied Computing and Informatics*, 2019.

[17] Lim, Su-Chang et al. “Solar Power Forecasting Using CNN-LSTM Hybrid Model.” *Energies*, Vol. 15, Page 8233, 2022.