Very Short-Term Photovoltaic Power Forecasting Using Stochastic Factors

Kriangkamon Khumma[1] and **Kreangsak Tamee**[2],

ABSTRACT

This paper proposes a photovoltaic (PV) power forecasting model, using the application of a Gaussian blur algorithm filtering technique to estimate power output and the creation of a stochastic forecasting model. As a result, affected power can be forecasted from stochastic factors with machine learning and an artificial neural network. This model focuses on very short-term forecasting over a five minute period. As it uses only endogenous data, no exogenous data is needed.

To evaluate the model, results were compared to the persistence model, which has good short-term forecasting accuracy. This proposed PV forecasting model gained higher accuracy than the persistence model using stochastic factors.

Keywords: Neural Network, Solar PV Generation Power Forecast, Stochastic Factor

1. INTRODUCTION

Energy is an essential factor in the development of a countrys economy, especially electrical energy. Fossil fuels have generated traditional electricity, causing environmental problems such as carbon dioxide and other greenhouse gasses (GHG), global warming, and climate change[1]. A worldwide effort to mitigate these issues resulted from the 2015 United Nations Climate Change Conference (COP21), known as the Paris Agreement. Many countries formed a consensus to the Paris Agreement by agreeing to reduce the release of GHG and CO2 to zero in the second half of 21st century. This is needed to limit the global warming problem and make sure the world remains livable in the future[2].

To replace conventional power generation by fossil fuels, energy must be gathered from a renewable source which can naturally replenish itself. Solar energy is a promising resource for power generation for residential, commercial, and industrial use[3]. Solar PV systems convert solar radiation into electric power using PV cells. This technology was exclusively used by satellites in space until solar PV came down to earth and became widely used around the world in the energy field [4].

Solar PV has gained exponential growth. Installation for power generation as can be seen Fig. 1, which shows Solar PV global capacity from solar PV installations from 2007 to 2017. Initially, power increased from 8 to 40 GW from 2007 to 2010 and then again from 40 to 402 GW from 2010 to 2017, thus demonstrating significant growth in the PV global capacity. The future growth rate will increase largely due to the increasing competitiveness of PV power generation, rising electricity demand in developing countries, and the potential of solar PV to mitigate pollution [5].

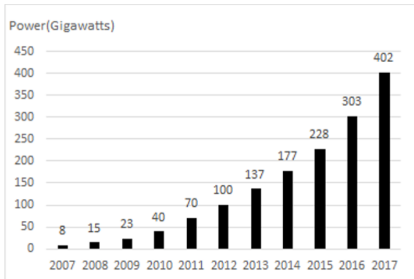


Fig.1: *PV power output generated by solar PV installations between 2007 and 2017.*

Solar PV power generation output can be uncertain due to many stochastic environmental factors, such as cloud movement, solar irradiance, and atmospheric conditions. In the power grid system, it is crucial to have an accurate forecast of future output from PV power plants, so authorities can manage grids to provide enough power for consumption. Many researchers have used various techniques and inputs to forecast power output with a high degree of accuracy. The origin of inputs, the forecasting types, and forecasting horizons are important considerations for this purpose.

There are two forecasting methods based on differing input origins. The first method uses only endogenous data. This includes current or lagged time series generated by PV power plants[6, 7]. The second method uses exogenous data. This includes meteorological measurements[8], satellite data[9], numerical weather prediction (NWP)[10], and neighboring PV

Manuscript received on June 27, 2019 ; revised on October 30, 2019.

Final manuscript received on November, 2019.

[1]The author is with Department of Computer Science and Information Technology, Faculty of Science Naresuan University Phitsanulok, 65000, Thailand., E-mail:kriangkamonk59@email.nu.ac.th

[2]The author is with Department of Computer Science and Information Technology, Research Center for Academic Excellence in Nonlinear Analysis and Optimization, Faculty of Science Naresuan University Phitsanulok, 65000, Thailand, E-mail:kreangsakt@nu.ac.th

plants[11].

There are two types of PV power generation forecasting. The first one is the direct forecasting of power that a PV should generate. The second one is indirect forecasting, which forecasts the solar irradiance and then uses simulation software to calculate the power output of the PV. Most PV power forecasting research has focused on direct forecasting with various techniques including the persistence model[12], statistical approaches[13], machine learning approaches[14], and hybrid techniques[15]. Mitsuru et al. [16] already studied both direct and indirect methods to compare their efficiencies. It has been found that direct forecasting yields higher accuracy than indirect forecasting.

Some work has achieved high accuracy in one location, although the same set of variables might have resulted in lower accuracy at other PV power plant locations. They has led to gathering and testing of variables suitable for each area. To obtain the variables for meteorological data, sensors have been used at PV power plants. For accurate data, the sensors must be properly maintained, which is tedious work to do. Furthermore, some sensors have been affected by weather conditions, causing measurement error and lower forecasting efficiency. For example, pyranometers, which have been used to measure solar radiation flux density, have been affected by moisture. After precipitation occurs, the data from pyranometers should not be used for prediction until the moisture is removed or vaporized from the equipment[17].

Thus, some researchers have focused on PV power forecasting using only endogenous data without exogenous data to avoid the problems with data from meteorological measurement devices [6, 7, 13, 18, 19]. Some power plants have also had insufficient meteorological data due to negligence of sensor maintenance.

In PV power generation, the effect of stochastic factors on power output data has been shown. In solar power generation, as shown in Fig. 2, a dashed line shows the power that should be generated by PV without stochastic factors. The PV power would be generated when the sun rises and steadily increase to maximum power when the sun is at its peak in the daytime, and then gradually decrease until the sun sets, and then no more power is generated. The shape of the power output graph is a bell shaped curve. However, in the real world, power output is affected by stochastic factors. The factors causing the power output generation to become unsteady can be seen in Fig. 2, where the lower output line is not the same as the complete bell shaped curve. We hypothesize that prediction performance can be improved by considering stochastic factors in the forecasting model.

As a result, this work focused on creating a forecasting model using endogenous data to predict power output at 5 minute intervals as affected by stochastic factors. Intra-Hour forecasting can assist power plant operators in foreseeing ramping events that might occur, and allows for proper planning to handle the events to occur [20]. The forecasting model for power output uses two internal models: one without stochastic factors, and another model to predict power affected by stochastic factors by applying an artificial neural network (ANN) which has been used in previous various forecasting works [13]. All the experiments for this study were developed in Python with the aid of Keras and Sci-Py.

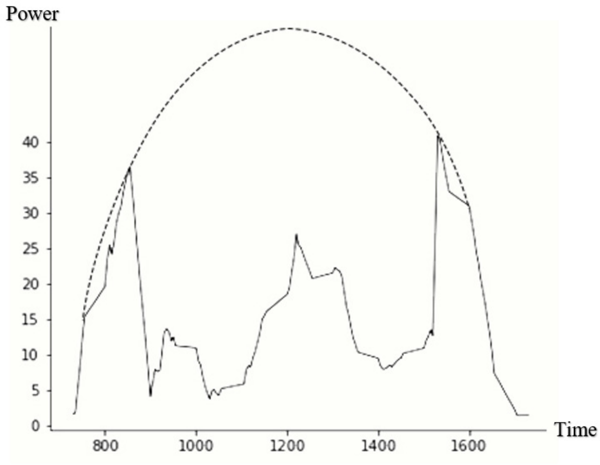


Fig.2: *PV power output graph. The dashed line represents expected power output, with no affect from stochastic factors. The lower line is measured power output as affected by stochastic factors.*

The rest of the paper is organized as follows. Section 2 contains the methodology used in the experiments and a dataset. Section 3 presents the results and discussion. Finally, the conclusion is drawn in section 4.

2. METHODOLOGY

2.1 Dataset

This work used a public dataset for validation of solar power model output forecasting from the National Renewable Energy Laboratory (NREL)[17], collected at five minute intervals from flat-plate PV modules installed in three different climatic locations in the United States of America: Colorado, Florida and Oregon. The selected PV data was from Eugene, Oregon from 2013 to 2014. Missing values were extrapolated using linear interpolation. Then the dataset size was reduced by choosing a subset of time between 9:00 and 17:00. The 2013 data was used for modeling, and 2014 data was used for model validation. All the data in the experiment was normalized with feature scaling.

2.2 Persistence model

The persistence model is the simplest model for forecasting time series data and is commonly used as a performance benchmark with other models[2]. As is stated in the name, the expected values in the future of the time series are calculated and this method assumes other environment variables remain unchanged. Thus, a time series under this study is stationary.

The persistence model assumes that the forecasted power will be the same as the previously measured value. For example, for 5 minutes ahead of the horizon, the power at 9:05 will be the same as power generation output at 9:00 as in Equation 1.

$$\hat{P}(t + \triangle t) = P(t) \tag{1}$$

$\hat{P}(t+\triangle t)$ is the predicted power in a time interval of $\triangle t$ and $P(t)$ is the measured power at t.

2.3 Proposed model

Coimbra et al.[21] explained that the PV power output is not stationary as in the persistence model. They proposed that the output power is the sum of expected power generation output under clear sky conditions and power influenced by stochastic factors as shown in Equation 2.

$$P(t) = P_{cs}(t) + P_{st}(t) \tag{2}$$

$P_{cs}(t)$ is the expected power generation output under clear sky conditions and $P_{st}(t)$ is affected power from stochastic factors.

Our work used the idea as shown in (2), which has both the expected power forecasting model in clear sky conditions and the affected power forecasting model including stochastic factors, but we applied a Gaussian blur filtering technique to Clear-sky model modeling. This was different from the method of Coimbra et al., which manually smoothed the surface of the solar power output graph. This manual smoothing technique was unspecified, so their work cannot be replicated in other locations.

The proposed model framework is shown in Fig. 3. The measured power (P) was transformed with Gaussian blur filtering technique. The expected power P_{cs} under clear sky conditions was generated from a previous process. Then, a Clear-sky model for P_{cs} was created. After P_{cs} data was obtained, it was used to calculate P_{st} by decomposing P_{cs} from P. Consequently, an ANN was applied to create a Stochastic forecasting model. Therefore, there were two models, a Clear-sky model and a Stochastic forecasting model. Both models were combined to forecast PV power output.

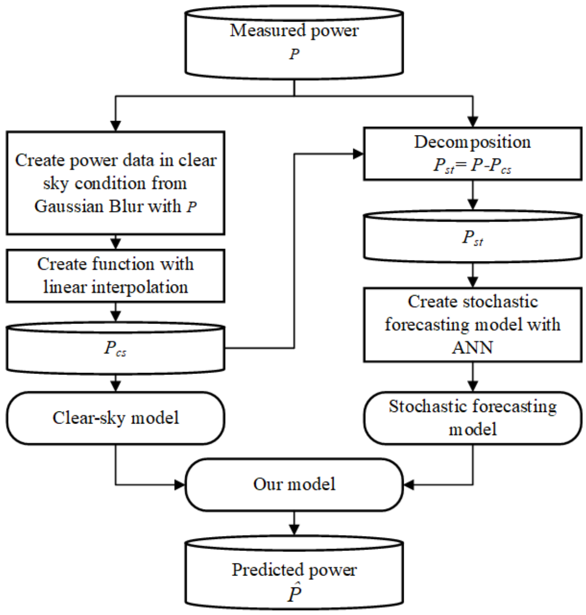


Fig.3: *Our modeling process included 2 previous models. One was the Clear-sky model to forecast the expected power output under clear sky conditions. The other was a Stochastic forecasting model to forecast power affected by stochastic factors.*

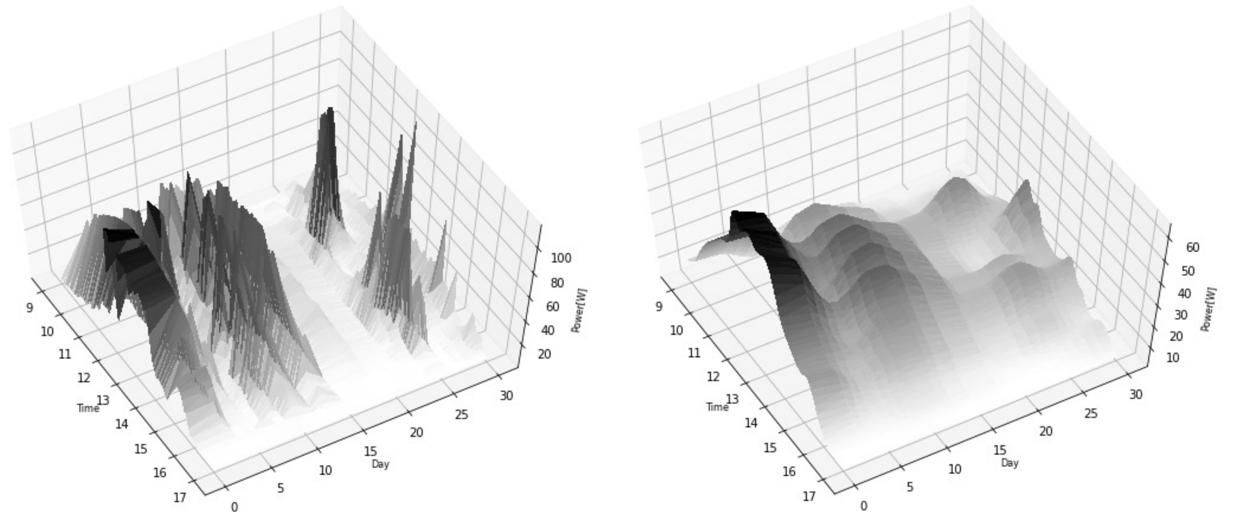
2.3.1 Clear-sky model

The PV power output is dependent on power plant location, technology, orientation of solar panels, and atmospheric conditions. All these factors can be used in forecasting models except atmospheric conditions, which occur randomly and thus cannot have predetermined exact values. Consequently, some researchers have assumed solar PV power output does not change with varying atmospheric conditions. They assume clear sky conditions. As a result, the Clear-sky model was created. Usually, Clear-sky models can be developed using the Radiative Transfer model[?] and the European Solar Radiation Atlas (ESRA) model[23].

Coimbra et al.[21] manually created a Clear-sky model from historical data, then created a function by linear interpolation. However, this study used a Clear-sky model and applied a filter technique. A Gaussian blur was used to remove noise from the PV power output and calculate new expected power output data from neighboring PV historical data.

To create the model we did the following steps:

- First, power output was represented as the pair of $P(Day, Time)$. Day was transformed to be day sequence number starting from 0 up to 365. $Time$ was the measured time in 24 hour clock time. For example, Fig.4(a) is the measured power output of subset January 2013, Day from 0 to 30 and Time from 9.00 to 17.00 represented in a surface graph, which fluctuated and was not smooth like the expected power under clear sky conditions.
- Secondly, a Gaussian blur was applied to measured



(a)Power output surface graph as a function of $P(Day, Time)$

(b)Power output surface graph smoothed by a Gaussian blur filtering technique for expected power output as function of $P_{cs}(Day, Time)$.

Fig.4: *Surface graph as function of* $(Day, Time)$

data, calculated with neighboring $P(Day, Time)$. The data from this calculation was used to create an interpolation function to be used as a Clear-sky model as a function $P_{cs}(Day, Time)$, which is represented in the graph shown in Fig. 4(b). Both Fig. 4(a) and Fig. 4(b) were created in the python program and function $P_{cs}(Day, Time)$ was created to be used for the later experiments.

2.3.2 Stochastic forecasting model

An ANN was used for stochastic forecasting modeling from P_{st} data. The input of the ANN was P_{st} data in a time-series format. For example, a one lagged time series $(t-1)$,$(t) \rightarrow (t+1)$ with a sequence of affected power $P_{st}1$, $P_{st}2$, $P_{st}3$, $P_{st}4$, $P_{st}5$ was transformed into five inputs for each P_{st} as shown in table 1 for five P_{st}. In table 1, input patterns No 1 and 2 were not complete, as there were no previous inputs for them and so they could not be used. The usable patterns with complete input interval steps were 3, 4 and 5.

Table 1: *Example of one lagged time series format*

No	input1	input2	output
1	-	-	$P_{st}1$
2	-	$P_{st}1$	$P_{st}2$
3	$P_{st}1$	$P_{st}2$	$P_{st}3$
4	$P_{st}2$	$P_{st}3$	$P_{st}4$
5	$P_{st}3$	$P_{st}4$	$P_{st}5$

To find suitable ANN parameters, input formats between two lagged and five lagged time series were tested with one hidden layer ANN containing different numbers of hidden nodes including three, five, and seven nodes, as shown in table 2.

The performance and accuracy of the model could be evaluated via several performance metrics. The metrics were used for performance comparison between different models. Each metric focused on a particular point distribution. Thus, there was not a unique metric that could use for all situations. Zhang et al[24] described various metrics for solar PV assessment. We used Root Mean Square Error ($RMSE$) as shown in (3). It penalizes significant errors in a square order.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{P}-P)^2} \tag{3}$$

$\hat{P}$ is the predicted power output and P is the measured power output.

Another metric used for assessment is the Coefficient of Determination (R^2), also know as Pearsons coefficient. It shows how correlated the forecasted and real values are, as shown in (4) where Var is Variance. Lower R^2 means lower accuracy and more error from predicted data, which reflects model forecast performance.

$$R^2 = 1 - \frac{Var(\hat{P}-P)}{Var(\hat{P})} \tag{4}$$

In table 2 various ANN parameters were tested to find a suitable parameter. The two lagged time series input format and five hidden nodes gained better performance compared to the other parameters. Thus the two lagged time series format and five hidden nodes were determined to be suitable parameters to be used in the stochastic forecasting model. The ANN structure in our stochastic forecasting model has three layers: an input layer with two nodes, a hidden layer with five hidden nodes, and one node output layer. For the activation function in ANN, Rectified Linear units (ReLu) was used.

Table 2: *ANN experiment parameters*

Hidden nodes	2 lagged		5 lagged	
	$RMSE$	R^2	$RMSE$	R^2
3 nodes	8.50	0.84	8.50	0.84
5 nodes	**8.15**	**0.86**	8.26	0.85
7 nodes	8.07	0.86	8.16	0.86

3. RESULT AND DISCCUSSION

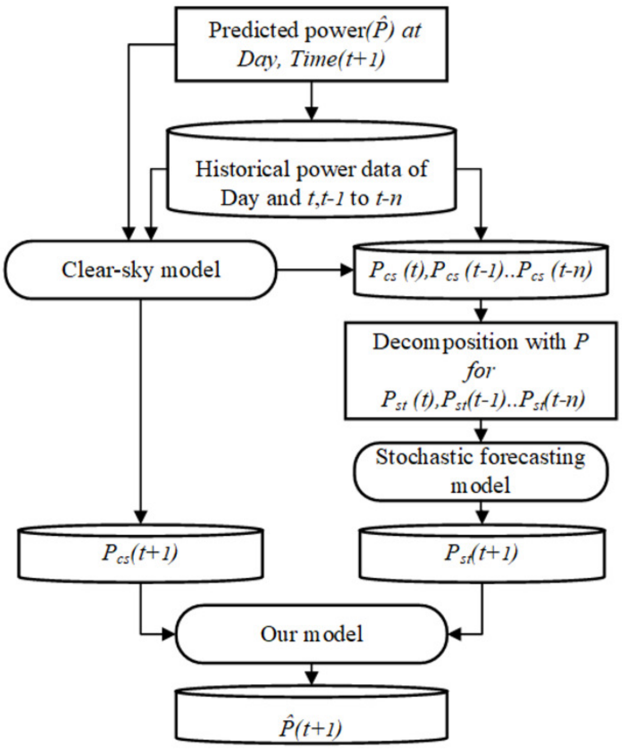


Fig.5: *Forecasting process to get predicted power $\hat{P}$ at Day and Time(t+1)*

The developed forecasting model was applied to a very-short time power generation forecast and its performance was compared against the persistence model.

The process for power prediction $\hat{P}$ of Day and $Time(t+1)$ is show in Fig. 5. Our model consisted of two components. First, the power from clear sky conditions $P_{cs}(t+1)$ was generated with the Clear-sky model. Next, the power affected by stochastic factors $P_{st}(t+1)$ used historical power at a 2 lagged time series format back from Time (t) format $(t-2, t-1, t \rightarrow t+1\ )$ to be the model input and got $P_{cs}(t)$, $P_{cs}(t-1)$ and $P_{cs}(t-2)$. Then we decomposed $P(t)$, $P(t-1)$ and $P(t-2)$ with $P_{cs}(t)$,$P_{cs}(t-1)$ and $P_{cs}(t-2)$ to get $P_{st}(t)$,$P_{st}(t-1)$ and $P_{st}(t-2)$ to predict $P_{st}(t+1)$. Finally, both $P_{st}(t+1)$ and $P_{cs}(t+1)$ were used to calculate the predicted power $\hat{P}$ at Day and $Time(t+1)$.

Table 3: *Predicted power in Fig 6 area A.*

Time	Measured	Proposed	Persistence
11.00	83.526	82.7372	82.3744
11.05	84.5517	83.9049	83.526
11.10	85.4103	84.8544	84.5517
11.15	86.1621	85.5379	85.4103
11.20	86.6192	86.3027	86.1621
11.25	87.6981	86.5591	86.6192
11.30	88.6026	88.088	87.6981
11.35	89.0831	88.8868	88.6026
11.40	89.6491	89.0939	89.0831
11.45	89.7733	89.734	89.6491
11.50	90.0941	89.4349	89.7733
11.55	89.7658	89.9324	90.0941
12.00	89.7816	89.03	89.7658

Table 4: *Predicted power in Fig 6 area B.*

Time	Measured	Proposed	Persistence
14.10	72.872	73.0582	74.3088
14.15	71.1757	71.1457	72.872
14.20	69.6444	69.3096	71.1757
14.25	68.3058	67.8221	69.6444
14.30	65.8827	66.7656	68.3058
14.35	63.9132	63.5887	65.8827
14.40	61.0715	61.9846	63.9132
14.45	56.9015	58.601	61.0715
14.50	52.9941	53.7329	56.9015
14.55	42.7255	50.0806	52.9941
15.00	28.4719	38.323	42.7255
15.05	32.1711	27.8727	28.4719
15.10	18.7843	31.8015	32.1711

Fig. 6. shows the solar power output graph including measured power shown as a continuous line, the power from our model as the dotted line, and the predicted power from the persistence model as a dashed line. The graph shows power generation for a typical day, where the power graph is not smooth like the bell-shaped curve mentioned for clear sky conditions. Area A (table 3), which includes both our model and the persistence model from 11.00 to 13.00, shows almost identical power output yields, so the plots nearly overlap into one line. This implies that under clear sky conditions with no effect from severe stochastic factors, both models give a satisfactory prediction. However, when sky conditions are not normal, as can be seen in area B (table 4) from 14.00 to 15.00, the power initially slowly decreased until, at 14.40, power output rapidly decreased. The persistence model could not yield accurate data compared to our model. As can be seen, the gap between measured data and the persistence model is more significant than the difference between our model and measured data. Our model provided better power

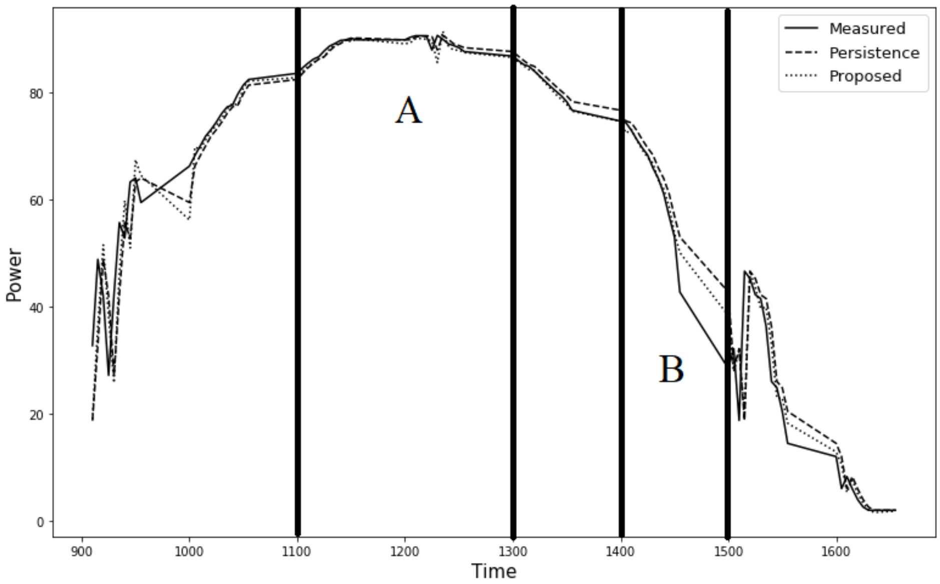


Fig.6: *Typical result in forecasted power output in Januaray 2014. This graph shows measured power, forecasted power by the proposed model, and forecasted power by persistence model*

output prediction with lower error, even with stochastic factors included, when compared to the persistence model.

Table 5: *Model performance evaluation*

Model	$RMSE$	R^2
Persistence model	8.87	0.89
Proposed model	7.26	0.94

Furthermore, the results from table 5 showing the model performance evaluation of both models based on the 2014 dataset with evaluation metrics $RMSE$ and R^2 indicate that the proposed model outperformed the persistence model with $R^2 = 0.94$, while the persistence model gave $R^2 = 0.89$ and the RMSE indicated greater error. The proposed model had better performance because our model included stochastic factors and handled other conditions that affected the solar power generation better than the persistence model, which uses the assumption that future conditions will remain the same as those measured previously and thus yield the same output. In the case of severe fluctuation of stochastic factors, the assumption of the persistence model means it cannot perform as well as the proposed model, which aims to handle stochastic factors related to PV power generation.

4. CONCLUSION

This work proposes a PV power forecasting model which considers the stochastic factors relevant to PV power generation output using only endogenous data. The model forecasts expected power output under clear sky conditions and power output affected by stochastic factors. It then combines both of the forecast calculation results to predict power output 5 minutes into the future. In the experiment, the model obtained R^2= 0.94 and $RMSE$ = 7.26, which outperformed the persistence model. As can be seen from the results, Clear-sky model creation with Gaussian Blur and error correction from the Stochastic forecasting model with ANN gives acceptable performance with simple techniques which can be replicated. It also demonstrates that consideration of stochastic factors can increase forecasting performance when compared to existing model methods which do not include them, as described in the hypothesis.

This study used standard Gaussian Blur and ANN from the Python library without any special modifications. In the future, a proper estimation and approximation function will be investigated, and tests for other ANNs will be performed to gain higher forecasting accuracy.

ACKNOWLEDGMENTS

The work was supported by a Phd scholarship from the Division of Research Promotion, Naresuan University, Phitsanulok, Thailand. The authors are also thankful to NREL for providing the PV power dataset that was used in the experiment. Thanks are due to Mr. Thomas Elliott, Graduate School, Naresuan University, Phitsanulok, for English language editing services.

References

[1] U. K. Das, K. S. Tey, M. Seyedmahmoudian, S. Mekhilef, M. Y. I. Idris, W. Van Deventer, B. Horan, and A. Stojcevski, “Forecasting of pho-

tovoltaic power generation and model optimization: A review," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 912-928, 2018.

[2] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. Martinez-de Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Solar Energy*, vol. 136, pp. 78111, 2016.

[3] S. Sobri, S. Koohi-Kamali, and N. A. Rahim, "Solar photovoltaic generation forecasting methods: A review," *Energy conversion and management*, vol. 156, pp. 459497, 2018.

[4] C. Puvareka and N. Kannan, "Fundamentals of Photovoltaic Technology (PV):A Review," *Middle-East Journal of Scientific Research*, vol. 25, no. 3, pp. 536-542, 2017. [Online]. Available: `https://www.idosi.org/mejsr/mejsr25(3)17/15.pdf`

[5] A. Zervos. Renewables 2018 global status report. [Online]. Available: `http://www.ren21.net/gsr-2018/`

[6] M. Rana, I. Koprinska, and V. G. Agelidis, "Forecasting solar power generated by grid connected pv systems using ensembles of neural networks," in *2015 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2015, pp.1-8.

[7] H. Wang, H. Yi, J. Peng, G. Wang, Y. Liu, H. Jiang, and W. Liu, "Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network," *Energy Conversion and Management*, vol. 153, pp. 409-422, 2017.

[8] S. Leva, A. Dolara, F. Grimaccia, M. Mussetta, and E. Ogliari, "Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power," *Mathematics and computers in simulation*, vol. 131, pp. 88-100, 2017.

[9] S. Pfenninger and I. Staffell, "Long-term patterns of european pv output using 30 years of validated hourly reanalysis and satellite data," *Energy*, vol. 114, pp. 1251-1265, 2016.

[10] D. P. Larson, L. Nonnenmacher, and C. F. Coimbra, "Day-ahead forecasting of solar power output from photovoltaic plants in the american southwest," *Renewable Energy*, vol. 91, pp. 11-20, 2016.

[11] A. Tascikaraoglu, B. M. Sanandaji, G. Chicco, V. Cocina, F. Spertino, O. Erdinc, N. G. Paterakis, and J. P. Catalao, "A short-term spatiotemporal approach for photovoltaic power forecasting," in *2016 Power Systems Computation Conference (PSCC)*. IEEE, 2016, pp. 1-7.

[12] S. Dutta, Y. Li, A. Venkataraman, L. M. Costa, T. Jiang, R. Plana, P. Tordjman, F. H. Choo, C. F. Foo, and H. B. Puttgen, "Load and renewable energy forecasting for a microgrid using persistence technique," *Energy Procedia*, vol. 143, pp. 617-622, 2017.

[13] S. I. Vagropoulos, G. Chouliaras, E. G. Kardakos, C. K. Simoglou, and A. G. Bakirtzis, "Comparison of sarimax, sarima, modified sarima and ann-based models for shortterm pv generation forecasting," in *2016 IEEE International Energy Conference (ENERGYCON)*. IEEE, 2016, pp. 1-6.

[14] H. Zhu, X. Li, Q. Sun, L. Nie, J. Yao, and G. Zhao, "A power prediction method for photovoltaic power plant based on wavelet decomposition and artificial neural networks," *Energies*, vol. 9, no. 1, p. 11, 2016.

[15] E. Ogliari, A. Dolara, G. Manzolini, and S. Leva, "Physical and hybrid methods comparison for the day ahead pv output power forecast," *Renewable energy*, vol. 113, pp. 11-21, 2017.

[16] M. Kudo, A. Takeuchi, Y. Nozaki, H. Endo, and J. Sumita, "Forecasting electric power generation in a photovoltaic power system for an energy network," *Electrical Engineering in Japan*, vol. 167, no. 4, pp. 16-23, 2009.

[17] B. Marion, A. Anderberg, C. Deline, J. delCueto, M. Muller, G. Perrin, J. Rodriguez, S. Rummel, T. J. Silverman, F. Vignola et al., "New data set for validating pv module performance models," in *2014 IEEE 40th Photovoltaic Specialist Conference (PVSC)*. IEEE, 2014, pp.1362-1366.

[18] M. Rana, I. Koprinska, and V. G. Agelidis, "Univariate and multivariate methods for very shortterm solar photovoltaic power forecasting," *Energy Conversion and Management*, vol. 121, pp.380-390, 2016.

[19] F. Wang, Z. Zhen, B. Wang, and Z. Mi, "Comparative study on knn and svm based weather classification models for day ahead short term solar pv power forecasting," *Applied Sciences*, vol. 8, no. 1, p. 28, 2018.

[20] C. Voyant, G. Notton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, and A. Fouilloy, "Machine learning methods for solar radiation forecasting: A review," *Renewable Energy*, vol. 105, pp. 569-582, 2017.

[21] H. T. Pedro and C. F. Coimbra, "Assessment of forecasting techniques for solar power production with no exogenous inputs," *Solar Energy*, vol. 86, no. 7, pp. 2017-2028, 2012.

[22] R. Mueller, K.-F. Dagestad, P. Ineichen, M. Schroedter-Homscheidt, S. Cros, D. Du-mortier, R. Kuhlemann, J. Olseth, G. Pier-navieja, C. Reise et al., "Rethinking satellite-based solar irradiance modelling: The solis clear-sky module," *Remote sensing of Environment*, vol. 91, no. 2, pp. 160-174, 2004.

[23] C. Rigollier, O. Bauer, and L. Wald, "On the clear sky model of the esra–european solar radiation atlas–with respect to the heliosat method," *Solar energy*, vol. 68, no. 1, pp.33-48, 2000.

[24] J. Zhang, A. Florita, B.-M. Hodge, S. Lu,

H. F.Hamann, V. Banunarayanan, and A. M. Brock-way, "A suite of metrics for assessing the performance of solar power forecasting," *Solar Energy*, vol. 111, pp.157-175, 2015.

Kriangkamon Khumma received his B.Sc.degree (Computer Science) from Chiangmai University (CMU), Chiangmai, Thailand, and his M.Sci degree in Computer Science from Naresuan University, Phitsanulok, Thailand. He is a PhD candidate at the Department of Computer Science and Information Technology, Faculty of Science, Naresuan University, Pitsanulok, Thailand. His research interests include machine learning and data science.

Kreangsak Tamee received his B.Sc. degree (Physics) from Chiangmai University (CMU), Chiangmai, Thailand, in 1997. He received his M.Eng. and D.Eng. degrees in electrical engineering from King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand, in 2001, and 2011, respectively. He is a Lecturer at the Department of Computer Science and Information Technology, Faculty of Science, Naresuan University, Pitsanulok, Thailand. His research interests are in the areas of optical sensors, optical signal processing, machine learning, data analytics, and optimization.