ISSN 2714-7533

Support Vector Regression Approach for Wind Forecasting

Mohamad Yamin [a,1,*], Ahmad Fakhri Giyats [a]

[a] Department of Mechanical Engineering, Faculty of Industrial Technology, Gunadarma University, Depok 16424, Indonesia
[1] mohay@staff.gunadarma.ac.id
* corresponding author

ARTICLE INFO

Article history
Received May 14, 2022
Revised June 30, 2022
Accepted July 14, 2022

Keywords
Wind Forecast,
Support Vector Regression,
Sidrap

ABSTRACT

The government policies to fully support the G20 sustainability emphasize increased use of renewable energy. The high penetration of wind energy into power systems poses many challenges for energy system operators, primarily due to the unpredictability and variability of wind energy production. Wind power may not be provided, but accurate forecasting of wind speed and power generation helps grid operators reduce the risk of reduced electricity reliability. Accurately predicting wind speeds over 1 to 24 hours based on these conditions is important for predicting potential energy supply. These short-term forecasts are important to support wind power planning, so the required base load supply for the grid is always guaranteed (even if the wind power output fluctuates significantly). This task demonstrates that the relative forecasting performance of a support vector regression (SVR) wind forecasting system can be improved by systematically selecting and combining related input functions that affect wind speed. Shows the results of data collected in Sidrap, Indonesia, during the six months of 2019. This paper explained key methods of wind forecasting, based on the evaluation of wind speeds and wind speed prediction methods. The RMSE from the SVR shows an 8% - 9% improvement on the RMSE of the persistence forecast every 1 hour. Wind speed estimation using a support vector regression approach has the potential for further development, one of which is determining the potential location of wind-based renewable sources and Wind Energy Conversion System (WECS) can make more efficient.

1. Introduction

The government policies to fully support the G20 emphasize the increased use of renewable energy. Wind power is the most attractive renewable energy technology due to its high efficiency and low pollution level. However, because the power generated by the Wind Energy Conversion System (WECS) varies with atmospheric weather and wind speed, unexpected fluctuations in WECS power generation increase the operating cost of the power system by increasing the need for primary reserves, it can increase and pose a potential risk to power reliability [1].

By 2020, the global installed capacity of wind power will be 93 GW, a significant increase of 52.96% over the capacity of 2019. Of these, the newly installed capacities for onshore wind power and offshore wind power are 86.9 GW and 6.1 GW, respectively. In a wind farm with a large installed capacity, the amount of power generated fluctuates greatly as soon as the wind speed changes by 1 meter per second. This variation is due to the non-linear relationship between wind energy production and wind speed. Therefore, timely and effective wind energy forecasting is essential to optimize wind farm site selection and ensure sustainable energy development.[2]–[5]

However, due to the complexity and uncertainty of the causes of wind in nature, wind speed is affected by many factors, such as wind direction and atmospheric pressure [6], [7]. Additionally, wind power is essentially a process of converting air kinetic energy into electrical energy, but the air kinetic energy reaching the wind farm is affected by many aspects [8], [9]. Multiple factors contribute to the randomness, volatility, and intermittent nature of wind power generation and make wind energy prediction difficult [10], [11].

Grid operators need to anticipate changes in wind power to plan reserve power and control grid operation. Accurate forecast wind speed to reduce reserves and increase wind energy penetration. In addition, wind forecasts play an important role in controlling control power allocation. In addition, wind forecasts are used for traditional power plant event planning and spot market electricity trading. [1]

Although the prediction accuracy of wind power forecasting is lower than the prediction accuracy of load forecasting, wind forecasts still play a key role in addressing operational challenges in electricity supply. Recently, several methods have been employed the wind forecasting. Much literature has been devoted to improving wind forecasting approaches by researchers with wide experience in field tests. A few wind forecast methods have been developed and launched on wind sites. The wind forecasting methods can generally be categorized into six groups: persistence, physical, statistical, spatial correlation, artificial intelligence, and hybrids; provide a detailed wind forecast of existing tools used to predict wind speeds and wind speeds across timescales and identifies other possible future developments.

There are many ways to predict wind, organized according to timescale and method. The chronological classification of wind prediction methods differs in describing various kinds of literature. However, combined with some references, wind forecasting methods can be divided into four categories according to the time scale. [12]

- Ultra-short-term forecasting: From a few minutes to an hour ago.
- Short-term forecasting: From an hour to a few hours ago.
- Medium-term forecasting: From a few hours to a week ago.
- Long-term forecasting: From a week to over a year ago.

Support Vector Regression is a regression algorithm in which the cost function ignores all training data within the adjustable margin (determined by the hyperparameter epsilon) around the model prediction. In addition, based on some advantages of SVR, SVR has been successfully applied to Biology, medicine, environmental protection, information technology, engineering technology, and other fields.[13]

2. Materials and Method

The SVR forecast is compared with the endurance forecast. The moving time window is used in SVR where, for each forward prediction step in time, the training data window is also changed. Initial training of the SVR algorithm is performed using the training dataset, with predictions made on the validation set. The training dataset includes past data of the predicted parameter. It can add additional features like wind speed at different center heights, wind speed, and air temperature to provide more information for the algorithm to make better predictions. Adding different features to the training data to improve the prediction accuracy is called feature selection or customization.

A. Support Vector Machine Theory

The non-linear epsilon insensitive SVR attempts to fit a function,

$$(\bar{x}) = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) k(\bar{x_\iota}, \bar{x}) + b,$$

Through the training samples, where all training data points are within epsilon (ε) from the function. ε is the hyperparameter that represents the band where there would be no penalty to the function. K (xi, x) is the kernel function. In this case, the RBF kernel, defined by

$$k(\bar{x_\iota}, \bar{x}) = \exp(-\gamma || \bar{x_\iota} - \bar{x} ||^2).$$

α_i and α_i^* are the Lagrange multipliers that solve the objective function. [1]

$$Maximize = \begin{cases} -1/2 \sum_{i,j=1}^{l} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) k(\bar{x_\iota}, \bar{x}) \\ \epsilon \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) + \sum_{i,j=1}^{l} (\alpha_i - \alpha_i^*), \end{cases}$$

Subject to,

$$\sum_{i,j=1}^{l} (\alpha_i - \alpha_i^*) = 0 \; and \; \alpha_i, \alpha_i^* \in [0, C].$$

l is the number of support vectors, and b is the bias that can be solved using the Karush-Kuhn-Tucker conditions.

$$\max\{-\epsilon + \gamma_i - \langle \omega, \bar{x_\iota} \rangle | \alpha_i < C \; or \; \alpha_{i^*} > 0\} \leq b \; \leq \min\{-\epsilon + \gamma_i - \langle \omega, \bar{x_\iota} \rangle | \alpha_i > C \; or \; \alpha_{i^*} < 0\}$$

From the formulation, the hyper parameter C determines the penalty assigned to target values outside the epsilon band and thus controls the degree of regularization. In contrast, the hyper parameter gamma (γ) is inversely proportional to the width of the RBF kernel, which is placed over each support-vector.

These hyper parameters (ε, γ, and C) used in the SVR are selected through a rough grid search using the validation set prediction, varying the hyper parameters to optimize the root mean square error (RMSE). The RMSE is calculated by

$$RMSE = \sqrt{\frac{1}{m} \sum_{i}^{m} (\bar{x_\iota} - \bar{x})^2}$$

This process is repeated for each model with different training features. The final prediction is made on the testing data set. The training and validation data sets are combined to form a new training set with the optimized hyper parameters. The hyper parameter selection is made using the validation set to lessen the chances of the SVR overfitting on the testing data set. [14]

B. Data Description

This paper uses the core of windy. The following information needed to complete this study is the wind speed in the Sidrap area. The map clearly shows that the wind speed around the Sidrap area is between 7.0 and 7.5 m/s at 100 meters.

Vol. 4, No. 2, August 2022, pp. 95-101

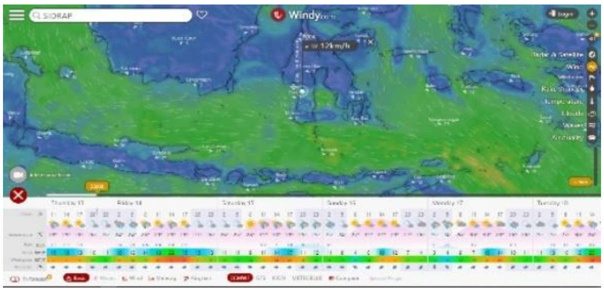

Figure 1. A sample coordinate take point forecast Sidrap.

The average wind speed in the Sidrap region for a year in 2019 has also been reported by local research. The report indicates that the average wind speed is between 5.0 and 7.5 m/s. Strong winds were observed occurring from April to October, while weaker wind speeds occurred from November to March. Also, during the feasibility study conducted by the developer on 14 different wind monitoring stations in the area using direct observation and remote sensing, the average wind speed varies from 5.2 m/s to 10.6 m/s (Renewable Energy Project Report) and by a simple average of mean wind speed of 7.7 m/s can be determined for the area. It is important to know that the wind turbine specifications indicate that electricity can be generated at a minimum wind speed of 3 m/s. Looking at the "Weibull distribution," wind speeds in the region, including wind speeds between 0 and 3 m/s, turbines may stop working and produce no electricity due to this low wind speed.[15]

C. Data Pre-Processing

Wind speed data from windy is composed of measures from January 1, 2019, to August 28, 2019. Data is measured every 1 hour. For missing values, data was interpolated between previous and next. The resulting 1,600 data points were then split into the training and testing sets. The training set contains 1,200 (80%) consecutive data points from January 1, 2019. The testing set contains the remaining 400 (20%) data points from July 30, 2019. Additionally, when the proposed model of this paper was used, data were scaled between 0 and 1 to enhance SVR training time requirements.

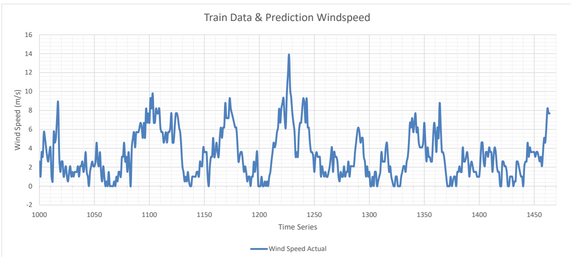

Figure 2. A sample line graph data time series wind speed (an hour).

D. Iteration Parameter EMD

Empirical Mode Decomposition (EMD) is an adaptive time series decomposition technique proposed by Norden E. Huang. [12] The principle of this signal processing method is to decompose the original time series with various fluctuations into a stationary one with different characteristics. Each series that is obtained after decomposition is treated as an intrinsic mode function (IMF), which satisfies the following two conditions: (1) in the whole-time range, the number of local external points and over zero must be equal, or the maximum difference is one; and (2) the mean value of the two envelopes formed by the local maxima and local minima, respectively, is zero at any point. [3]

This method was applied in this research and used to get results in Fig.4 comparison of actual and prediction data.

E. Quality Metrics

Performance evaluation of persistence, autoregressive, and our proposed methods was done in terms of root mean square. In doing so, we provide a completer overview of the performance of the evaluated methods and their utility for wind farms. [16]

3. Result and Discussion

The RMSE of an SVR model with feature set wind actual is compared to a persistence forecast to benchmark. The persistence forecast and SVR predicts the wind speed at a 100 m hub height, as shown in Figure 1. The persistence forecast is improved at every step by the SVR. The RMSE from the SVR shows an 8% - 9% improvement on the RMSE of the persistence forecast every 1 hour.

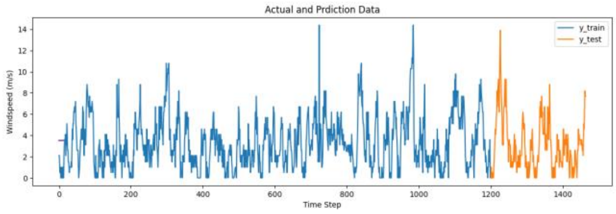

Figure 3. A sample coordinate take point forecast Sidrap.

In this paper, we explored support Vector Regression for forecasting accuracy of a 1 to 24-hour ahead wind speed prediction at a 100 m hub height by wind speed. Additional needed for more accuracy such as wind speed acceleration, change in wind speed acceleration, wind direction, air pressure, air temperature, and relative humidity.

A. Comparison Train and Test Data

The predictions that show the most noticeable improvement on the base case in each preceding Figure are shown in Figures 1 & 2 to provide an overall clear visual comparison of the most noted results. The SVR prediction of the wind speed at 60 m, using only the 60 m wind speed time series in the training data set, has good results Fig.4.

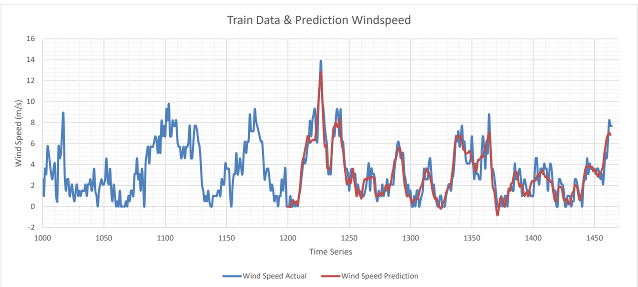

Figure 4. A sample coordinate take point forecast Sidrap.

In Figs 5 & 6, detailed images of the regression hyperparameter results which have values almost similar to the data that have been taken as validation data in the 1360 to 1380 times series, there are poor estimation results as a result of training there is previous data that has decreased the intensity of the wind speed.

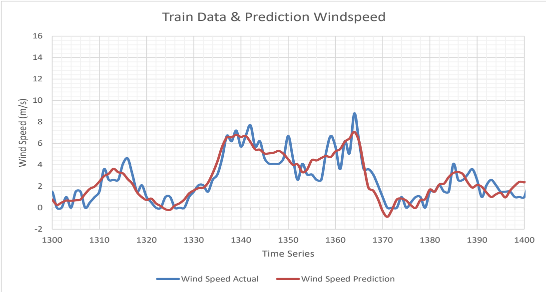

Figure 5. A sample coordinate take point forecast Sidrap.

Figure 5 is a timescale estimation magnification with data intervals multiples of 10; minor unfavorable results can be seen from the estimated air velocity in the 1370-1372 interval, where the estimation results (hyperplane) from breaking the limit of the value of the actual airspeed results.

4. Conclusion

Wind speed forecasting is a key component of wind energy production. Based on the simple approach in this research, it can be seen that the wind forecast for the development of Sidrap PLTB electricity production is still theoretically acceptable. Wind speed estimation using a support vector regression approach has the potential for further development, one of which is to determine the potential location of wind-based renewable sources. In addition, developments related to wind forecasting can be integrated so that it can estimate the power generated from the power plant in the next few days.

Acknowledgment

Universitas Gunadarma funds this research. We also thank the reviewers for the constructive comments in improving the quality of this paper.

References

[1] N. Botha dan C. M. van der Walt, “Forecasting wind speed using support vector regression and feature selection,” dalam *2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech)*, Bloemfontein, South Africa, Nov 2017, hlm. 181–186. doi: 10.1109/RoboMech.2017.8261144.

[2] A. Banik, C. Behera, Tirunagaru. V. Sarathkumar, dan A. K. Goswami, “Uncertain wind power forecasting using LSTM-based prediction interval,” *IET Renew. Power Gener.*, vol. 14, no. 14, hlm. 2657–2667, Okt 2020, doi: 10.1049/iet-rpg.2019.1238.

[3] T. Niu, J. Wang, K. Zhang, dan P. Du, “Multi-step-ahead wind speed forecasting based on optimal feature selection and a modified bat algorithm with the cognition strategy,” *Renew. Energy*, vol. 118, hlm. 213–229, Apr 2018, doi: 10.1016/j.renene.2017.10.075.

[4] C. Yang, X. Yang, dan X. Xiao, “Data-driven projection method in fluid simulation: Data-driven projection method in fluid simulation,” *Comput. Animat. Virtual Worlds*, vol. 27, no. 3–4, hlm. 415–424, Mei 2016, doi: 10.1002/cav.1695.

[5] M. H. Do dan D. Söffker, “State-of-the-art in integrated prognostics and health management control for utility-scale wind turbines,” *Renew. Sustain. Energy Rev.*, vol. 145, hlm. 111102, Jul 2021, doi: 10.1016/j.rser.2021.111102.

[6] V. Nikolić, S. Sajjadi, D. Petković, S. Shamshirband, Ž. Ćojbašić, dan L. Y. Por, “Design and state of art of innovative wind turbine systems,” *Renew. Sustain. Energy Rev.*, vol. 61, hlm. 258–265, Agu 2016, doi: 10.1016/j.rser.2016.03.052.

[7] Z. Peng *dkk.*, "A novel deep learning ensemble model with data denoising for short-term wind speed forecasting," *Energy Convers. Manag.*, vol. 207, hlm. 112524, Mar 2020, doi: 10.1016/j.enconman.2020.112524.

[8] S. R. Moreno, V. C. Mariani, dan L. dos S. Coelho, "Hybrid multi-stage decomposition with parametric model applied to wind speed forecasting in Brazilian Northeast," *Renew. Energy*, vol. 164, hlm. 1508–1526, Feb 2021, doi: 10.1016/j.renene.2020.10.126.

[9] W. Fu, K. Wang, J. Tan, dan K. Zhang, "A composite framework coupling multiple feature selection, compound prediction models and novel hybrid swarm optimizer-based synchronization optimization strategy for multi-step ahead short-term wind speed forecasting," *Energy Convers. Manag.*, vol. 205, hlm. 112461, Feb 2020, doi: 10.1016/j.enconman.2019.112461.

[10] Ll. Lledó, V. Torralba, A. Soret, J. Ramon, dan F. J. Doblas-Reyes, "Seasonal forecasts of wind power generation," *Renew. Energy*, vol. 143, hlm. 91–100, Des 2019, doi: 10.1016/j.renene.2019.04.135.

[11] E. C. Malz, V. Verendel, dan S. Gros, "Computing the power profiles for an Airborne Wind Energy system based on large-scale wind data," *Renew. Energy*, vol. 162, hlm. 766–778, Des 2020, doi: 10.1016/j.renene.2020.06.056.

[12] N. E. Huang *dkk.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. Ser. Math. Phys. Eng. Sci.*, vol. 454, no. 1971, hlm. 903–995, Mar 1998, doi: 10.1098/rspa.1998.0193.

[13] S. Zhang, C. Liu, W. Wang, dan B. Chang, "Twin Least Square Support Vector Regression Model Based on Gauss-Laplace Mixed Noise Feature with Its Application in Wind Speed Prediction," *Entropy*, vol. 22, no. 10, hlm. 1102, Sep 2020, doi: 10.3390/e22101102.

[14] Y. Ren, P. N. Suganthan, dan N. Srikanth, "A Novel Empirical Mode Decomposition With Support Vector Regression for Wind Speed Forecasting," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 8, hlm. 1793–1798, Agu 2016, doi: 10.1109/TNNLS.2014.2351391.

[15] L. Lisapaly, "An academic review on the performance of the Sidrap wind turbine, Sulawesi – Indonesia," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 878, no. 1, hlm. 012058, Okt 2021, doi: 10.1088/1755-1315/878/1/012058.

[16] G. Santamaría-Bonfil, A. Reyes-Ballesteros, dan C. Gershenson, "Wind speed forecasting for wind farms: A method based on support vector regression," *Renew. Energy*, vol. 85, hlm. 790–809, Jan 2016, doi: 10.1016/j.renene.2015.07.004.