Received: 27 May 2022 | Revised: 15 August 2022 | Accepted: 17 August 2022

DOI: 10.1002/ese3.1295

ORIGINAL ARTICLE

Power production forecast for distributed wind energy systems using support vector regression

Ghali Yakoub | **Sathyajith Mathew** | **Joao Leal**

Department of Engineering Sciences, University of Agder, Grimstad, Norway

Correspondence
Ghali Yakoub, Department of Engineering Sciences, University of Agder, Jon Lilletunsvei 9, Grimstad 4879, Norway.
Email: Ghali.yakoub@uia.no

Funding information
Research Council of Norway (RFF Midt-Norge), Grant/Award Number: 299452

Abstract

Due to the inherent intermittency in wind power production, reliable short-term wind power production forecasting has become essential for the efficient grid and market integration of wind energy. The current wind power production forecasting schemes are predominantly developed for wind farms. With the rapid growth in the microgrid sector and the increasing number of wind turbines integrated with these local grids, power production forecasting schemes are becoming essential for distributed wind energy systems as well. This paper proposes a power production forecasting scheme developed explicitly for distributed wind energy projects. The proposed system integrates two submodels based on support vector regression: one for downscaling the wind speed predictions to the hub coordinates of the turbine and the other for predicting the site-specific performance of the turbine under this wind condition. The forecasting horizons considered are the hour ahead ($t + 1$) and the day ahead ($t + 36$), which align with the Nord pool's energy market requirements. For the day-ahead scheme, a multistep recursive approach is adopted. The accuracy and adaptability of the proposed forecasting scheme are demonstrated in the case of a distributed wind turbine.

KEYWORDS

distributed, wind energy, power management

1 | INTRODUCTION

With the record installations of 93 GW in 2020, the global wind power capacity has crossed over 742 GW.[1] Considering the present growth rate, the cumulative installations of wind energy systems are expected to reach a capacity of 1212 GW by 2025. One of the major challenges in the grid integration of wind energy is its temporal variability. Due to the stochastic nature of the wind, power output from the wind turbines can fluctuate significantly, even within short time intervals.[2] To efficiently manage such grids, variations in energy contributed by the turbines must be quantified in different time scales. Accurate wind power forecasts must understand these power fluctuations and manage the resulting uncertainties.

Wind power production forecasting can generally be classified as physical methods, traditional statistical methods, and, more recently, the so-called learning methods (e.g., machine learning [ML] approaches).[2–6] ML methods are considered an alternative to conventional methods as they have shown their ability to accurately predict wind power production.[6] One or more of these approaches can be combined to develop hybrid forecasting methods.[3,4] Wind power forecasting can also be categorized as direct and indirect forecasting. The direct method uses historical data to forecast future wind power production,[7] usually called a time series analysis. In another direct scheme, weather forecasts are directly used to predict wind power, which is often termed the wind-to-power (W2P) approach. This approach proved to be more accurate than time series analysis, especially for horizons more than 6 h ahead. However, the accuracy of this method highly depends on the accuracy of the used weather forecasts.[8,9]

On the other hand, in indirect forecasting, wind speeds are initially forecasted which is then utilized for power predictions using the turbines' power curves.[10] In this approach, errors are mainly caused due to the error in forecasting the wind speed.[11] Therefore, many studies on indirect wind power forecasting are focused on improving the accuracy of wind forecast.[12–14] Accuracy of the indirect approach can also depend on the nonlinear relationship between wind speed and power output,[2,15,16] which is generally represented by the power curve of the wind turbine. Several techniques have been proposed to model the power curve of wind turbines. Generally, these techniques can be classified as parametric and nonparametric[17] methods. Parametric models are built upon mathematical formulation based on a family of functions with some variables that are fitted specifically to a wind turbine. In contrast, nonparametric techniques, which generally learn the velocity-power relationship from the data, do not require any prespecified conditions. A comprehensive review of these power curve modeling techniques can be found in Carrillo et al.[18] and Lydia et al.[19] Among these, support vector regression (SVR) has shown a good capability to estimate the nonlinear relationship between wind speed and power.[2,20,21] In addition, SVR base models have capabilities for fast convergence and easy integration. A detailed description of SVR-based models can be found at The MetCoOp Team[22] and Pandit and Kolios.[23]

The manufacturers' power curves, which are generally used to model this relationship between wind velocity and power, are based on measurements under ideal test conditions.[18] Hence, they may not be able to capture the sit-specific dynamics of the complex W2P conversion process. Therefore, it is advised to use site-specific performance models, which are based on the actual performance data of the turbine at a given site, for correlating the wind speed and corresponding power in the wind power forecasting models.[19]

Though several methods and approaches have been proposed for wind power production forecasting, most of these are developed for wind farms, where several turbines are clustered together. Along with the rapid growth in these "centralized" wind farm sectors, the distributed wind energy sector has also grown rapidly in recent years. For example, in the United States alone, distributed wind systems' cumulative capacity was 1145 MW in 2019,[17] which is expected to be enhanced by 300% by 2030.[20] With varying rated capacities from a few kW to MW, these systems are often connected to microgrids, which are isolated or integrated with the main grid.[17] Short-term power production forecasts from wind turbines coupled with these microgrids are essential for efficient management.

Wind turbines perform differently while operating as a single turbine or as clusters, like, wind farms. This is mainly due to the wake interactions between the neighboring upwind and downwind turbines. So, the power forecasting schemes developed for wind farms cannot be adopted for predicting the performance of distributed wind energy systems. Hence, as emphasized in the Distributed Wind Research Program report,[17,21] there is a need to develop exclusive models for estimating and forecasting the performance of distributed wind turbines to integrate and manage them within hybrid microgrid energy systems efficiently.

This paper proposes an intelligent power production forecasting method, exclusively developed for distributed wind energy systems, using ML methods. The proposed method integrates wind speed forecasts, which are downscaled from Numerical Weather Predictions (NWPs) to the hub coordinates of the turbine, and a site-specific wind turbine performance model which predicts the performance of the given turbine under the forecasted wind field. SVR was used to develop the speed and turbine performance models.

The paper is arranged as follows. After this introductory section, a brief description of the study case is presented, followed by an illustration of the framework for the proposed forecasting scheme. This is then followed by the details of various steps in developing the wind downscaling model. The performance of the downscaling approach at the hub coordinates of a 5 kW test turbine is then illustrated. The wind turbine performance model is then introduced with discussions on its development approaches and performances while applied to a 5 kW turbine. Finally, the performance of the integrated power production forecasting scheme is presented and discussed.

2 | STUDY CASE

To demonstrate the proposed SVR wind power production, the geospatial location of a fully instrumented 5 kW experimental wind turbine is considered in this study.

The wind turbine is located at Smøla island, on the west coast of Norway, as shown in Figure 1, within a wind farm composed of 67 turbines with an installed capacity of 148.4 MW. The uniqueness of its location introduces an interesting challenge for the proposed method.

For this study, data from January 1 to December 31, 2019 were retrieved from two sources. The first data set was from a met mast installed next to the experimental wind turbine at 20 m above ground level (AGL). These observed data were recorded at an interval of 5 min and consisted of the wind speed and direction. The second data set was extracted from the archived historical hourly raw weather forecast at the height of 10 m AGL from the regional NWP model METCoOp Ensemble Prediction System (MEPS). This model is a convection-permitting atmosphere ensemble model covering Scandinavia and the Nordic sea with a horizontal resolution of 2.5 km, 65 vertical levels, and 10 members.[22] Several weather parameters like wind speed, direction, air temperature, relative humidity, and air pressure were extracted from that model and considered in this study. To better understand the weather conditions at the point of interest, weather forecast data were extracted from the four grid cells nearest the wind mast location, as shown in Figure 1. The green icons show the locations of the centers of the four NWP grid cells, and the red icon shows the location of the target wind turbine at Smøla island.

It should be mentioned that the observed data were transformed to a lower resolution (hourly) to match the NWP's resolution, where the wind speed record was simply averaged while the wind directions were first transformed to radians, then the sin and cosine for the 5 min intervals were averaged and finally transformed back to a degree after using the inverse arctangent function, precisely the arctan2 function which is a four-quadrant inverse function giving out the angle between 0 and 2π radians.

3 | FORECASTING APPROACH

The proposed forecasting scheme is illustrated in Figure 2. It consists of two models, one for wind downscaling and the other for turbine performance, which is integrated to forecast the performance of the distributed wind energy system. In the downscaling model, the speed and direction of the wind, which is forecasted in low spatial resolution using the NWP models, are downscaled with enhanced resolution, corresponding to the hub coordinates of the turbine. The power responses of the turbine at these speeds are

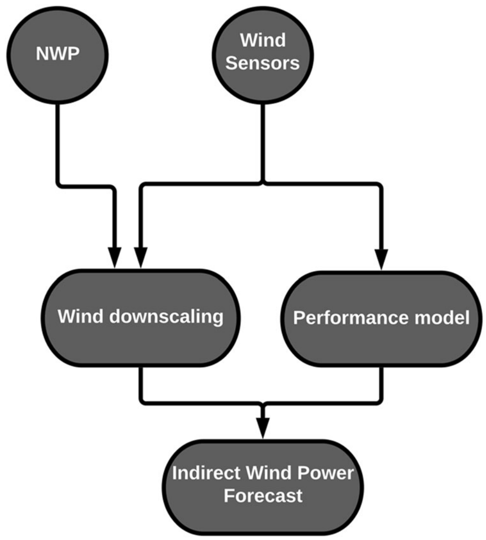

FIGURE 2 Forecasting approach proposed for distributed wind energy systems. NWP, Numerical Weather Prediction.

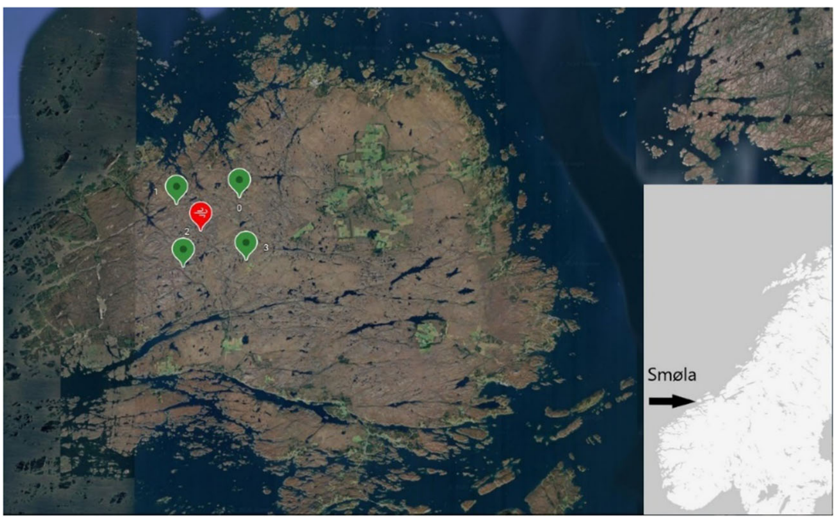

FIGURE 1 Wind turbine location and the four nearest NWP grid points. *Source*: Google earth. NWP, Numerical Weather Prediction.

then predicted using the wind turbine performance model. Algorithms based on SVR are used to develop the downscaling and turbine performance models. SVR-based models have proven to be an effective and valuable tool in actual value function estimation.[2,23–25] Furthermore, the lack of a long historical record limits the possibility of using more advanced neural networks and gradient boosting techniques. SVR models are trained using a symmetrical loss function, which penalizes high and low misestimates equally. Solving a quadratic optimization problem while minimizing Vapnik's ε insensitive loss function, the most widely used cost function, a flexible tube is formed symmetrically around the estimated function. Where points outside the tube are penalized, but those within the tube receive no penalty. One of the main advantages of SVR is that its computational complexity does not depend significantly on the dimensionality of the input space. In addition, it has excellent generalization capability, with high prediction accuracy.[26–29]

Nevertheless, achieving an optimal SVR architecture requires tuning several hyperparameters, namely, the kernel function, which maps lower-dimensional data into higher-dimensional data. The gamma parameter defines the inversed radius of the samples' influence selected by the model as support vectors.[30] The regularization parameter, which is inversely proportional to the strength of regularization. The epsilon parameter specifies the tube within which no penalty is associated in the training loss function with points predicted within a distance epsilon from the actual value.[31–33]

A 5 kW fully instrumented wind turbine is considered to demonstrate the viability of the proposed wind power forecasting method. Details of these models are discussed in the following sections.

4 | WIND SPEED DOWNSCALING MODEL

The wind speed downscaling model used in this forecasting scheme is an extension of the method previously proposed by the authors.[12] The wind forecast from the regional NWP model MEPS has been used in this study, as it sufficiently covers the region where the experimental turbine is installed.[22] Wind speeds forecasted at the four grid cells nearest the wind turbine have been extracted.[12] They are then regressed with the corresponding wind speed and direction measured at the hub height of the experimental turbine for developing the SVR model.

The development of the SVR-based downscaling model had three major phases: preprocessing, model building, and postprocessing. Under the preprocessing, the data were cleaned for outliers and noise, relevant inputs for the model were identified, and the data were divided into two sets for training and testing the model. The inputs were selected by applying Pearson, Kendall, and Spearman correlations and Mutual Information Regression[34–36] to several weather parameters forecasted by the NWP and the corresponding wind measurements from the turbine hub level. Though several inputs were initially considered, wind speed and gust have been chosen as the model inputs through this feature selection, which has been conducted in previous research. For more details, see Yakoub et al.[12]

The data division was done by random resampling of the subsets (i.e., shuffling the lines). The data were divided into training and testing subsets using a supervised trial and error method with manual adjustments to get a satisfactory level of agreement between the statistical properties, in other words, maintaining a maximum relative difference of 4% for the statistical properties (mean, standard deviation, and coefficient of variation) of the subsets. This ensures that the seasonal and diurnal variations in the data are well reflected in both subsets.

The model building phase mainly involved developing and optimizing the SVR architecture. A stepwise constructive approach, using the Grid Search algorithm combined with the cross-validation method, was used to optimize different hyperparameters of the developed SVR architecture.

The optimal SVR model has a nonlinear kernel radial basis function (RBF) with a penalty parameter of 10, the gamma, and the epsilon parameters of 0.06 and 0.3, respectively. It should be noted that the optimization process used the calibration data integrated with the cross-validation method where $k = 10$ folds and the scoring function was chosen to be a mean-squared error to rank the models created.

Given its applications for power dispatch for distributed turbines integrated with microgrids and considering the energy market requirements, an hour-ahead $(t+1)$ and day-ahead $(t+36)$ forecasting schemes are considered. For the day-ahead $(t+36)$ forecast, which must be available before noon of each day, a multistep recursive forecasting strategy, as adopted in Yakoub et al.,[12] the strategy consists in running the downscaling models consecutively, starting by forecasting $(t+1)$ exactly as the intraday and then progressing to higher times. Note that the observed values in the previous hour are unknown for higher times and therefore replaced by forecasted ones. Also, for each forecast, the NWP values are maintained equal to the ones available at the beginning of the forecast period.

The model's performance has been evaluated in the postprocessing phase by considering several error metrics. Table 1 shows the performance of the wind speed downscaling model considering the forecasting schemes using data from January 2020 (note, these data were not used in the model development). In Table 1, $WS_{obs(t)}$ is the observed wind speed at the time (t), $avg\_WS_{Pi(t+1)}$ and $avg\_WG_{Pi(t+1)}$ are the average wind speed, and the wind gust of the four nearest points of the NWP at the time ($t+1$), respectively. It is clear from the results that the downscaled hour-ahead forecast improved the wind speed predictions by 63%, 67%, and 70%, considering the root mean-square error (RMSE), mean arctangent absolute percentage error, and the mean absolute error (MAE), respectively, compared with the latest updated NWP wind speed forecast. Similarly, the multistep recursive forecast (36 h ahead) improved the wind speed predictions by 35%. As expected, the multistep recursive forecast is relatively less accurate than the hour ahead. The relatively high error is due to the accumulation of errors from the successive hour-ahead predictions, which were used as additional inputs. Moreover, the recursive forecasting scheme (multistep) uses NWP values that are not updated throughout the forecast horizon. Therefore, the day-ahead forecasts showed higher error levels than the intraday ($t+1$) forecast, where NWP values are updated every 6 h.

Figure 3 shows the observed wind speeds on January 15 and 16, 2020 compared with the corresponding downscaled values for both the intraday and day-ahead forecasting schemes. The dashed pink lines define the day-ahead region of interest (ROI). The hour-ahead forecast is the one that follows best the actual measured values. It can be observed that the multistep recursive gives much better results compared with the NWP. This is confirmed by looking at the wind speed results within

TABLE 1 Evaluation of the forecasting strategies for January 2020

	Wind speed			
Properties	Intraday ($t+1$)		Day ahead[c] ($t+36$)	
Inputs	$WS_{obs(t)}$, $avg\_WS_{Pi(t+1)}$, $avg\_WG_{Pi(t+1)}$	NWP[a]	$WS_{obs(t)}$,[b] $avg\_WS_{Pi(t+1)}$, $avg\_WG_{Pi(t+1)}$	NWP[c]
RMSE (m/s)	1.31	3.56	2.22	3.42
NRMSE (%)	5.1	14.0	9.1	13.4
MAE (m/s)	0.94	3.20	1.71	3.02
MAAPE (%)	9.6	29.2	17.2	27.4
R^2	0.94	0.55	0.81	0.58

Abbreviations: MAAPE, mean arctangent absolute percentage error; MAE, mean absolute error; MSE, mean-squared error; NWP, Numerical Weather Prediction; RMSE, root mean-square error.

[a]Full record of NWP based on the latest forecast, that is, considering the four updates per day.

[b]Only using the observed recent value at the first step, that is, forecast ($t+1$), for the next steps using predicted downscaled ones.

[c]Considering the values based only on the region of interest.

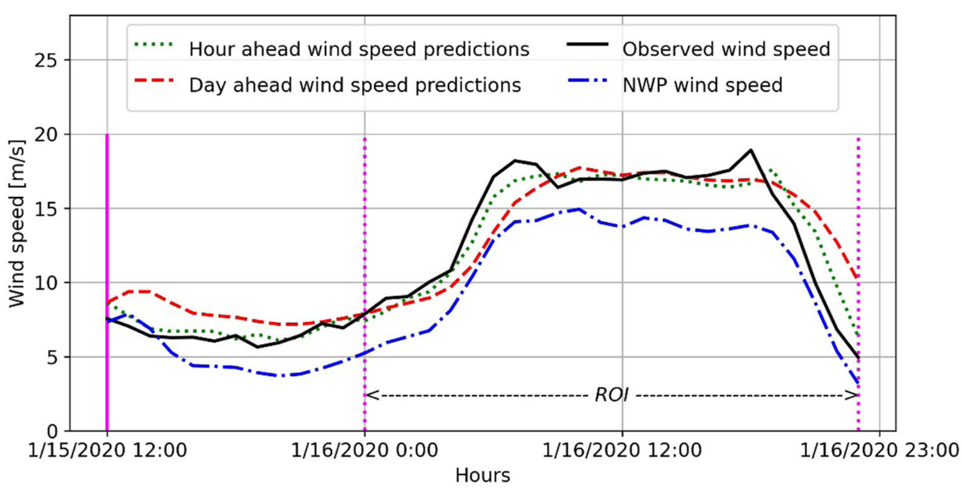

FIGURE 3 Wind speed forecasts for January 16, 2020. NWP, Numerical Weather Prediction; ROI, region of interest.

the ROI, where the recursive forecast effectively avoids the general underestimation of the NWP.

5 | WIND TURBINE PERFORMANCE MODEL

Once the wind speed at the turbine's hub coordinates is forecasted from the downscaling model at a given time of interest, this has further to be integrated with the turbine performance model for the proposed wind power forecasting scheme. The turbine performance model basically predicts the power produced by the turbine at different wind speeds. Conventionally, the turbine's wind speed–power response is estimated using the power curve provided by the manufacturer, which is developed under ideal test conditions following the specifications in IEC 61400 12. However, the manufacturers' power curve may not be able to capture the dynamics of the wind flow expected under real site environments.[15,18,19,37] Hence, a site-specific nonparametric modeling approach, using the speed power measurements from the turbine, is proposed for the wind turbine performance. In view of its capabilities in estimating the nonlinear relationship between wind speed and power, SVR-based algorithms are used to develop the turbine performance model.

The SVR-based turbine performance model was developed using the 5 kW experimental turbine data. The turbine's cut-in, rated, and cut-out speeds were 3, 12, and up to 60 m/s, respectively. Figure 4 shows the manufacturer power curve of the turbine compared with its measured performance. Differences in the expected and actual performances of the turbine are evident in Figure 4, highlighting the need to develop a site-specific performance model.

The wind speed and the corresponding power produced by the turbine, measured at 5-min intervals, have been aggregated on an hourly basis to match it with the temporal scale of the wind speed forecasts from the downscaling model. The hot-deck imputation approach has replaced some of the missing power data corresponding to the turbine's rated to cut-out wind speed.[38] The data were preprocessed, following a similar approach adopted for the downscaling model reported in Section 4. Though wind directions at the point of interest were also predicted using the downscaling model, they were not used as an input to the wind turbine performance model as it was verified that it creates more noise in the forecasting scheme, thereby lowering the forecast accuracy.

For optimizing the SVR architecture, the same approach used for the downscaling model is adopted. The optimal SVR wind turbine performance model has a nonlinear kernel RBF with a penalty parameter of 1.3, and gamma and epsilon parameters of 0.6 and 0.02, respectively.

Table 2 summarizes the SVR-based model's performances under calibration and testing. It can be seen that,

TABLE 2 SVR wind turbine performance evaluation

Performance metrics	Calibration set	Test set	MPC
RMSE (W)	165.02	166.28	807.62
NRMSE (%)	3.32	3.40	16.30
MAE (W)	96.19	98.25	659.85
R^2	0.99	0.99	0.71
Overfitting indicator		0.99	0.20

Abbreviations: MAE, mean absolute error; MPC, manufacturer power curve; NRMSE, normalized root mean-square error; RMSE, root mean-square error; SVR, support vector regression.

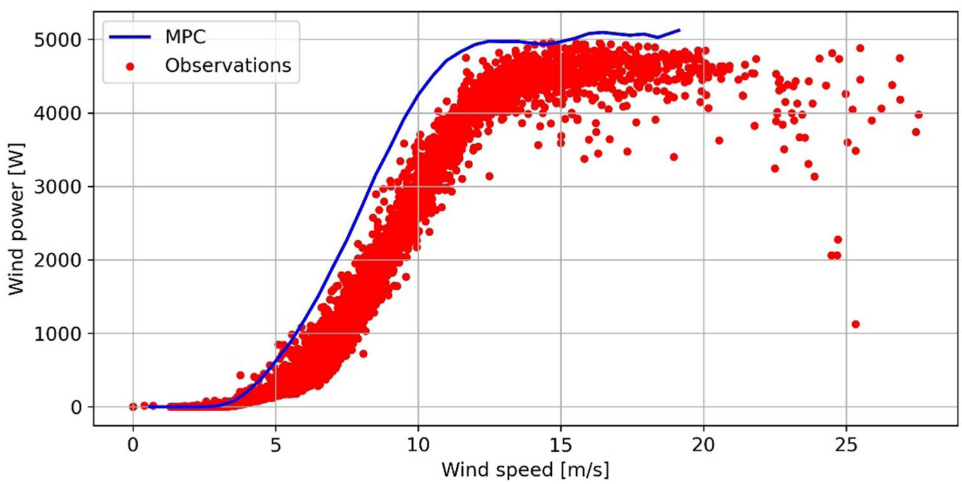

FIGURE 4 Manufacturer power curve (MPC) versus observations.

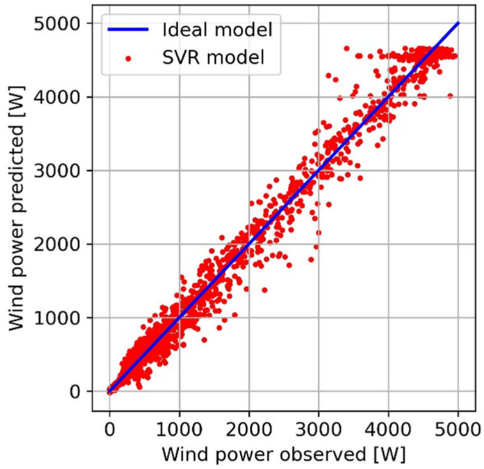

FIGURE 5 Comparison between the observed and predicted power using SVR (test set). SVR, support vector regression.

TABLE 3 One hour ahead of power production forecast performance evaluation

Performance metrics	Intraday forecast ($t+1$) January 2020 SVR (WS) + SVR (WP)	NWP (WS)[a] + SVR (WP)	Persistence
RMSE (W)	592.67	1221.95	645.79
NRMSE (%)	12.1	24.9	13.1
MAE (W)	382.01	847.18	396.22
MAAPE (%)	24.4	42	25.3
R^2	0.89	0.53	0.87

Abbreviations: MAAPE, mean arctangent absolute percentage error; MAE, mean absolute error; NRMSE, normalized root mean-square error; NWP, Numerical Weather Prediction; RMSE, root mean-square error; SVR, support vector regression; WP, wind power; WS, wind speed.

[a]Wind speed forecast extracted from the latest NWP forecast.

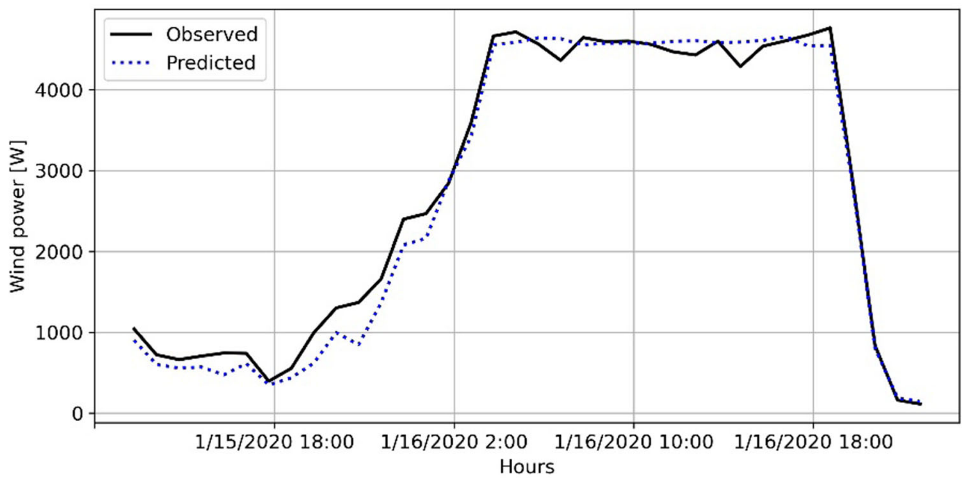

FIGURE 6 Performance model predictions over 1 day of January 16, 2020.

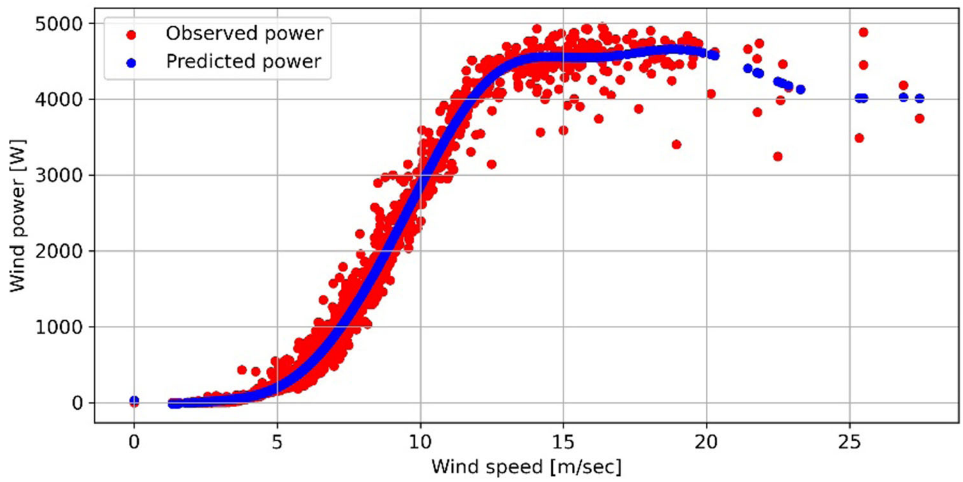

FIGURE 7 SVR-based site-specific performance turbine model. SVR, support vector regression.

with the SVR model, the overall normalized error in the power estimates could be reduced to 3.4% against the corresponding error of 16.3% while using the manufacturer's power curve. Figures 5 and 6 compare the actual power developed by the turbine with the power predicted by the proposed SVR model. Table 2 and both Figures 5 and 6 clearly demonstrate the proposed SVR model's ability to characterize the speed power response of the turbine at a given site. The results in Figure 5 show that the SVR model has no apparent tendency to either underpredict (points below the ideal model line) or overpredict (points above the ideal model line). On the basis of the SVR model, a site-specific performance curve for the turbine has been developed and is presented in Figure 7. The curve closely follows the actual performance of the turbine, especially between the cut-in to rated wind speeds under which the turbine predominantly operates.

6 | INTEGRATED WIND POWER PRODUCTION FORECASTING SYSTEM

The SVR-based wind speed downscaling and the wind turbine performance models were integrated to form the end-to-end wind power production forecasting system. The proposed system is made to temporally align with the Nord Pool's intraday and day-ahead markets.

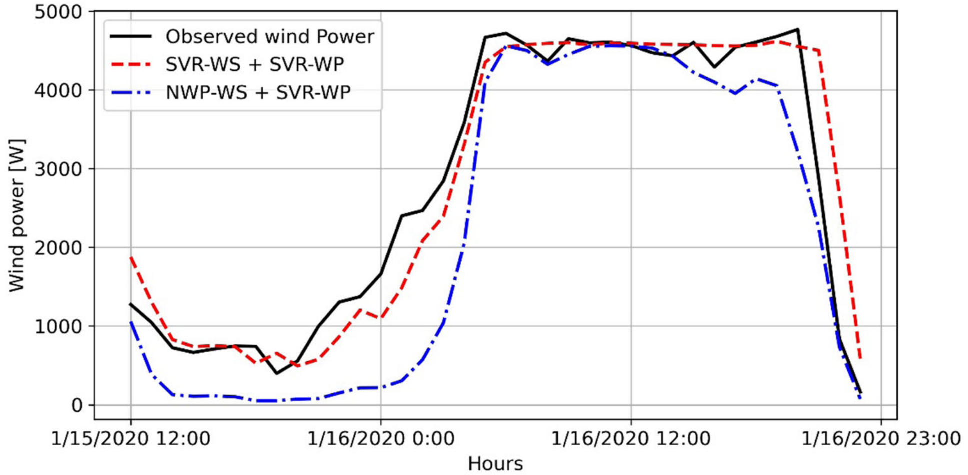

FIGURE 8 One hour-ahead wind power production forecast for January 16, 2020. NWP, Numerical Weather Prediction; SVR, support vector regression; WP, wind power; WS, wind speed.

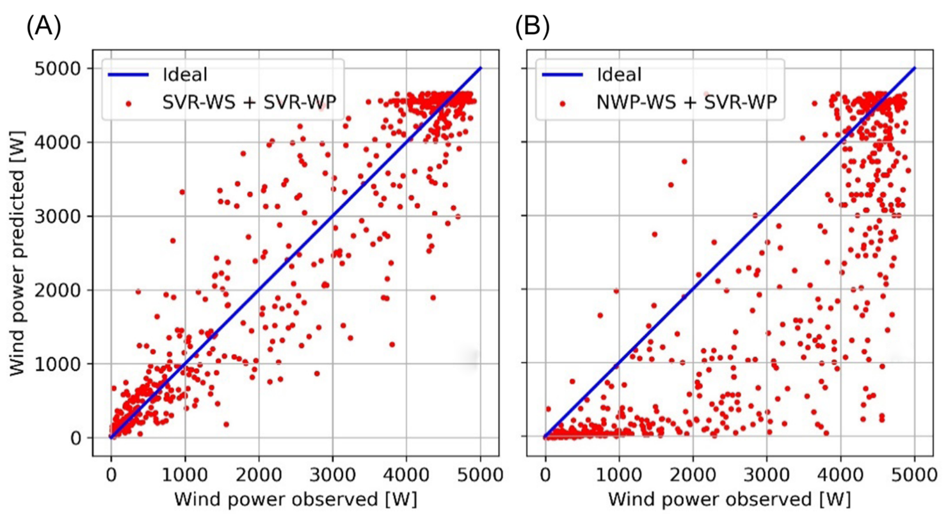

FIGURE 9 One hour-ahead wind power observed versus predicted: (A) results from the proposed indirect forecast and (B) results from using the NWP forecast directly with the developed SVR performance model (January 2020). NWP, Numerical Weather Prediction; SVR, support vector regression; WP, wind power; WS, wind speed.

The performance of the integrated system in the hour-ahead wind power production forecasting is shown in Table 3. Forecasting schemes combining the turbine performance model with both the NWP wind predictions and the SVR-based downscaled wind forecasts are compared and benchmarked with the persistence approach. This is further illustrated in Figure 8. The forecasting system, which combines the wind downscaling model with the turbine performance model, outperforms the other options. For example, compared with the approach with NWP wind predictions, the forecasting errors in terms of RMSE and MAE could be reduced by 52% and 55% by downscaling the wind speeds to the turbines' hub coordinates. In comparison with the persistence approach, corresponding error reductions were 8% and 3.5%, respectively.

Figure 9 shows a comparison of 1 h ahead wind power observed and the predicted using (A) the proposed indirect approach and (B) the NWP wind forecast with the developed SVR performance model. The proposed forecast is robust near the two ends of the ideal line Figure 9A, that is, within 20% of the cut-in speed and the rated speed, while the errors are magnified in the region between cut-in and rated speeds. Using the NWP results directly in a generalized underprediction of the wind power Figure 9B.

TABLE 4 Day-ahead recursive wind power production forecast performance evaluation

Performance metrics	Day-ahead forecast ($t + 36$) January 2020 "ROI": SVR (WS) + SVR (WP)	NWP (WS)[a] + SVR(WP)
NRMSE (%)	21.5	26.4
MAE (W)	656.76	839.71
MAAPE (%)	36.7	41
R^2	0.65	0.47

Abbreviations: MAAPE, mean arctangent absolute percentage error; MAE, mean absolute error; NRMSE, normalized root mean-square error; NWP, Numerical Weather Prediction; RMSE, root mean-square error; ROI, region of interest; SVR, support vector regression; WP, wind power; WS, wind speed.

[a]Wind speed forecast extracted from NWP forecast at 12 noon each day covering the next 36 h.

The recursive forecasting scheme, developed for the day-ahead wind power production forecasting, also showed similar trends as in the case of the intraday forecasts discussed above. Table 4 shows the error evaluation of this recursive forecasting approach. By downscaling the NWP wind predictions to the hub coordinates of the turbine, errors in the day-ahead recursive forecasting could significantly be reduced. For example, RMSE and MAE were reduced by 8% and 22%, respectively, compared with the forecasting approach using NWP wind predictions directly.

Performances of both approaches in the recursive day-ahead wind power production forecasting are compared in Figure 10 for January 16, 2020. Due to market regulations, the time at which the forecasting must be made is at 12:00 each day, and the forecasting ROI, between 00:00 and 23:00 the next day (dotted vertical pink lines), is also indicated in Figure 10. It should be noted that the NWP data, which are used for the downscaling, are extracted at the forecasting time 12:00, which could have affected the accuracy of the proposed day-ahead recursive forecasts. Nevertheless, in

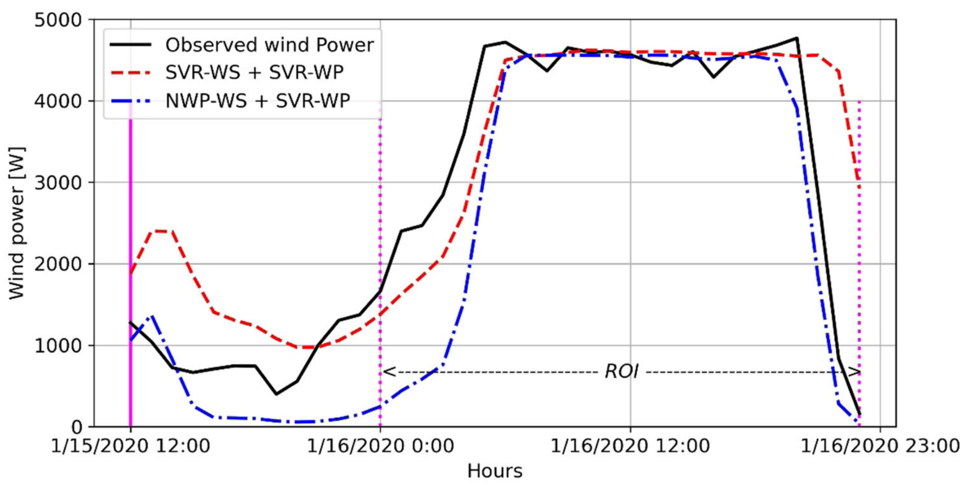

FIGURE 10 Day-ahead wind power production forecast for January 16, 2020. NWP, Numerical Weather Prediction; SVR, support vector regression; WP, wind power; WS, wind speed.

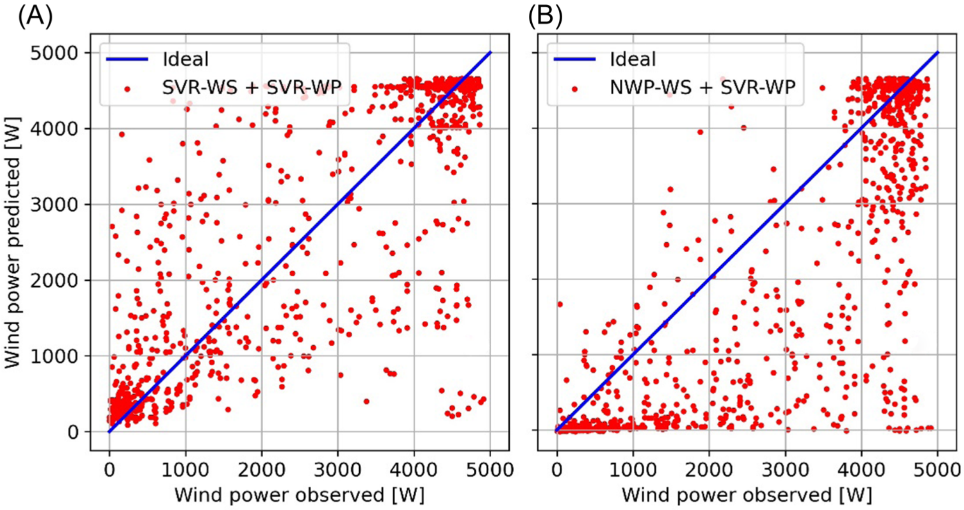

FIGURE 11 Day-ahead wind power observed versus predicted: (A) results from the proposed indirect forecast and (B) results from using the NWP forecast directly with the developed SVR performance model (January 2020). NWP, Numerical Weather Prediction; SVR, support vector regression; WP, wind power; WS, wind speed.

general, the forecasts based on the downscaling approach outperform the NWP-based option.

Similarly to Figure 9, Figure 11 shows the comparison for day-ahead power forecasts. The performance of the proposed forecast degrades as the time horizon increase, which can be observed by the increased variations shown in Figure 11A. Despite that, the model continues to not have a clear tendency to overpredict or underpredict the observations. Interestingly, the direct use of NWP wind forecast shows consistency in a larger forecast horizon by preserving a generalized underprediction pattern Figure 11B.

7 | CONCLUSIONS

In this study, we have presented an SVR-based wind power forecasting, explicitly developed for distributed wind energy systems integrating the NWP forecasts from the nearest grid cells to the studied system. The proposed method consisted of two models, one for downscaling the wind speed predictions from NWP to the hub coordinates of the wind turbine and the other for estimating the power produced by the turbine under these downscaled wind conditions. The models were then integrated to form an end-to-end wind power production forecasting scheme, where both hour-ahead and day-ahead timeframes were considered possible applications in grid and market integration. A multistep forecasting approach (recursive) was adopted for the day-ahead timeframe ($t + 36$), which implied running the model several times, considering the latest prediction as an input to the next hour's prediction. Both timeframes were tested individually with the case of a 5 kW turbine and showed a high accuracy level compared with relying only on pure NWP forecast.

The proposed wind power forecasting scheme is unique as it specifically targets stand-alone distributed wind turbines. Such forecasts are significant in managing microgrids integrated with distributed wind systems. The SVR-based downscaling of wind speeds, especially in conjunction with the multistep recursive approach in the day-ahead forecasting, has significantly contributed to enhancing the wind speed prediction accuracies. Similarly, the SVR-based "site-specific" approach could minimize the normalized error of the W2P modeling by 12.9%, in comparison with the conventional method of using the manufacturers' power curve. With these improvements, the proposed wind power forecasting schemes could minimize the normalized error of hour-ahead and day-ahead forecasts by 12.1% and 18.5%, which is significantly lower than the conventional indirect wind power forecasting options which combine the NWPs with the manufacturers' power curves.

AUTHOR CONTRIBUTIONS

Ghali Yakoub: Conceptualization, methodology, software, visualization, investigation, validation, and writing the original draft. **Sathyajith Mathew**: Supervision, conceptualization, validation, and writing review and editing. **Joao Leal**: Cosupervision, conceptualization, validation, and writing review and editing.

ACKNOWLEDGMENTS

The authors would like to thank the National Wind Energy Center in Norway (NVES) for their cooperation in providing the observed data needed for this study. This study was partly funded by the Research Council of Norway (RFF Midt-Norge) through project number 299452. The first author also acknowledges the Ministry of Education in Norway for the Ph.D. fellowship at the University of Agder.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Ghali Yakoub http://orcid.org/0000-0003-3228-7360

REFERENCES

1. Lee L, Zhao F, Dutton A, et al. Global wind report 2021. In: G.W.E.C., ed. *Global Wind Report*. Global Wind Energy Council; 2021.
2. Veena R, Mathew S, Petra MI. Artificially intelligent models for the site-specific performance of wind turbines. *Int J Renew Energy Environ Eng*. 2020;11(3):289-297.
3. Pockey JR, Song EY, Sutfin EL, et al. Current methods and advances in forecasting of wind power generation. *Renewable Energy*. 2012;37(1):1-8.
4. Mana M, Burlando M, Meissner C. Evaluation of two ANN approaches for the wind power forecast in a mountainous site. *Int J Renewable Energy Res*. 2017;7:1629-1638.
5. Okumus I, Dinler A. Current status of wind energy forecasting and a hybrid method for hourly predictions. *Energy Convers Manage*. 2016;123:362-371.
6. Sfetsos T. A comparison of various forecasting techniques applied to mean hourly wind speed time series. *Renewable Energy*. 2000;21:23-35.
7. Wang J, Zhou Q, Zhang X. Wind power forecasting based on time series ARMA model. In: *IOP Conference Series: Earth and Environmental Science*. Vol 199. IOP Publishing; 2018:022015.
8. De Giorgi MG, Ficarella A, Tarantino M. Assessment of the benefits of numerical weather predictions in wind power forecasting based on statistical methods. *Energy*. 2011;36:3968-3978.
9. Browell J, Gilbert C, McMillan D. Use of turbine-level data for improved wind power forecasting. In: *2017 IEEE Manchester PowerTech*. IEEE; 2017.
10. Ran L, Yue W. *Short-term wind speed forecasting for wind farm based on empirical mode decomposition*. In: *2008 International Conference on Electrical Machines and Systems*. IEEE; 2008.
11. Wang Y, Hu Q, Srinivasan D, Wang Z. Wind power curve modeling and wind power forecasting with inconsistent data. *IEEE Trans Sustain Energy*. 2019;10(1):16-25.
12. Yakoub G, Mathew S, Leal J. Downscaling and improving the wind forecasts from NWP for wind energy applications using support vector regression. *J Phys: Conf Ser*. 2020;1618:062034.
13. Liu H, Mi X, Li Y. An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and ELM algorithm. *Renewable Energy*. 2018;123(C):694-705.
14. Yu M, Zhou W, Wang B, Jin J. The short-term forecasting of wind speed based on EMD and ARMA. In: *12th IEEE Conference on Industrial Electronics and Applications (ICIEA)*; 2017.
15. Gottschall J, Peinke J. How to improve the estimation of power curves for wind turbines. *Environ Res Lett*. 2008;3:15005-15007.
16. Wan Y-h, Ela E, Orwig KD. Development of an Equivalent Wind Plant Power-Curve: Preprint. National Renewable Energy Lab. (NREL); National Renewable Energy Laboratory (NREL); 2010.
17. Foster N, Orrell A, Homer J, Tagestad J. The “perfect storm” for distributed wind markets. *Renewable Energy*. 2020;145:1033-1039.
18. Carrillo C, Obando Montaño AF, Cidrás J, Díaz-Dorado E. Review of power curve modelling for wind turbines. *Renewable Sustainable Energy Rev*. 2013;21:572-581.
19. Lydia M, Kumar SS, Selvakumar AI, Prem Kumar GE. A comprehensive review on wind turbine power curve modeling techniques. *Renewable Sustainable Energy Rev*. 2014;30(C):452-460.
20. Lantz E, Sigrin B, Gleason M, Preus R, Baring-Gould I. *Assessing the Future of Distributed Wind: Opportunities for Behind-the-Meter Projects*. National Renewable Energy Laboratory (NREL); 2016.
21. Orrell AC, Kazimierczuk K, Sheridan LM. *Distributed Wind Market Report: 2021 Edition*. Pacific Northwest National Laboratory (PNNL); Pacific Northwest National Laboratory (PNNL); 2021.
22. The MetCoOp Team. C.a.U.A. *ALADIN-HIRLAM Newsletter No. 8 (Corresponding Author: Ulf Andrae)*; 2017:98-103. https://www.umr-cnrm.fr/aladin/spip.php?article260
23. Pandit R, Kolios A. SCADA data-based support vector machine wind turbine power curve uncertainty estimation and its comparative studies. *Appl Sci*. 2020;10(23):8685.
24. Díaz S, Carta JA, Matías JM. Performance assessment of five MCP models proposed for the estimation of long-term wind turbine power outputs at a target site using three machine learning techniques. *Appl Energy*. 2018;209:455-477.
25. Pandit R, Infield D, Kolios A. Comparison of advanced non-parametric models for wind turbine power curves. In: *Renewable Power Generation*. Vol 13. IET; 2019:1503-1510.
26. Khanna R, Awad M. *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. Apress; 2015.
27. Li C, Lin S, Xu F, Liu D, Liu J. Short-term wind power prediction based on data mining technology and improved support vector machine method: a case study in Northwest China. *J Cleaner Prod*. 2018;205:909-922.
28. Li L, Cen ZY, Tseng ML, Shen Q, Ali MH. Improving short-term wind power prediction using hybrid improved cuckoo search arithmetic—support vector regression machine. *J Cleaner Prod*. 2021;279:123739.
29. Wu C-H, Ho J-M, Lee D. *Travel-time prediction with support vector regression*. *IEEE Trans Intell Transp Syst*. 2005;5:276-281.
30. Pedregosa F, Varoquaux, G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2012;12:2825-2830.
31. Chapelle O, Vapnik V. *Model Selection for Support Vector Machines*; 2000.
32. Vladimir Cherkassky FM. In: Sons JW, ed. *Learning from Data: Concepts, Theory, and Methods*. Vol 2, 2nd ed. John Wiley & Sons; 2007:404-466.

33. Cherkassky V, Ma Y. Selecting of the loss function for robust linear regression. *Neural Comput*. 2002.
34. Guyon I, Elisseeff A. An introduction to variable and feature selection. *J Mach Learn Res*. 2003;3:1157-1182.
35. Ross BC. Mutual information between discrete and continuous data sets. *PLoS One*. 2014;9(2):e87357.
36. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. *Phys Rev E*. 2004;69(6):066138.
37. Sohoni V, Gupta S, Nema R. A critical review on wind turbine power curve modelling techniques and their applications in wind based energy systems. *J Energy*. 2016;2016:1-18.
38. Andridge RR, Little RJA. A review of hot deck imputation for survey non-response. *Int Stat Rev*. 2010;78(1):40-64.

How to cite this article: Yakoub G, Mathew S, Leal J. Power production forecast for distributed wind energy systems using support vector regression. *Energy Sci Eng*. 2022;10:4662-4673. doi:10.1002/ese3.1295