Minute Ahead Wind Speed Forecasting Using a Gaussian Process and Fuzzy Assimilation

Miltiadis Alamaniotis
Dept. of Electrical and Computer Engineering
University of Texas at San Antonio
San Antonio, TX, USA
miltos.alamaniotis@utsa.edu

Georgios Karagiannis
Dept. of Mathematical Sciences
Durham University
Durham, UK
georgios.karagiannis@durham.ac.uk

Abstract—This paper presents an intelligent data driven method for forecasting minute ahead wind speed, which is essential in predicting the power output coming from wind generators. The proposed methodology, is based on the principle that "the most recent past should be used to predict the near future", and implements a two-stage forecasting method. In the first stage a Gaussian Process Regression model is trained multiple times on different length time window, and forecasts a set of next minute wind speed values. In the second stage, a fuzzy inference system collects the forecasts, rejects some of them and then provides a mean and a variance of a single forecast value. The proposed method is applied to a dataset of real-world data, and benchmarked against the autoregression (AR) model. Results exhibit the superiority of the proposed method over AR as well as over GPR which uses a single train set.

Index Terms—Wind speed forecasting, GPR, Fuzzy Inference, Renewable Energy.

I. INTRODUCTION

The vision of an intelligent power grid entails the coupling of information and intelligent systems with the power system infrastructure. The overarching goal is via his coupling to ensure the reliable, safe and non-stop delivery of electrical power from the generation areas to the load centers [1]. The cornerstone for implementing an intelligent grid exhibiting the above features is forecasting; forecasting of the values of grid operational parameters will contribute in attaining effective action-taking [2].

Machine learning offers a variety of tools with different capabilities in learning from data that may provide forecasts over monitored variables. In particular, machine learning allows pure data driven models that learn from observed data, and subsequently conduct forecasting [3]. Implicitly, utilization of data driven tools aims at capturing data patterns that may be observed in the near future. Thus, from a data driven perspective forecasting may be seen as the attempt of identifying past patterns the reoccur in the future [4]. Overall, utilization of data-driven approaches is a convenient way for modeling complex physical processes without explicitly considering every of the factors that affect the process.

Effective integration of renewable sources of energy in the power grid has been identified as one of the main characteristics of an intelligent power grid [5]. Forecasting of renewable power and especially of wind power is challenging due to the stochastic nature of renewable sources. Wind speed is the driving force behind wind power, and thus, its accurate forecasting may accommodate the efficient utilization of wind power [6]. Furthermore, it allows market operators to provide financial incentives to consumers increasing their consumption at time points where wind power is available. Furthermore, availability of wind power may also affect the electricity prices in a competitive electricity market given that there is excess supply that should preferably be consumed rather than wasted [7].

Wind speed forecasting has been extensively studied and several methodologies have been presented. Most of the presented methodologies utilize machine learning tools as the backbone of wind forecasting. A set of three different type of neural networks have been applied for wind speed forecasting in [8] and [9]. Neural networks have also been utilized for developing hybrid forecasting methods: neural networks with Bayesian statistics is presented in [10], neuro-fuzzy approaches [10], and synergism of neural networks with wavelet processing is introduced in [11]. Furthermore, an adaptive neuro fuzzy inference system (ANFIS) for wind speed forecasting has been discussed in [12], while empirical mode decomposition is integrated with neural networks in [13]. Other methods include kernel density estimators as presented in [14], and dynamic regression in [15]. Wind speed forecasting using support vector machines is introduced in [16], and fuzzy logic in a fuzzy logic approach in [17]. Time series methods such as autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) have been studied and tested in wind speed forecasting in [18] and [19] respectively, while the integration of Gaussian processes and particle swarm optimization has been proposed in [20]. More recently, several methods have been introduced in wind forecasting for time horizon of some minutes ahead of time: deep learning has been utilized for 10-minute forecasting in [21] and [22], and extreme learning machines in [23]. The majority of the methodologies have been utilized for hourly speed prediction and only a few have been used for minute ahead forecasting. In addition, most methodologies adopt large historical datasets to train the machine learning models that inherently introduce high uncertainty in forecasting.

In this paper, a new methodology is proposed for wind speed forecasting in renewable integration. The proposed methodology utilizes the synergism of a kernel modeled Gaussian process (GP) [24] with a fuzzy inference system [25]. The goal of the methodology is to forecast the next minute wind speed by utilizing the most recent measurements. The novelty of the work lies in the selection of various forecasts and assimilation of them into a single one, a process that is conducted by a proposed fuzzy inference system and a sliding time window.

In the next step, a brief introduction to Gaussian process regression (GPR) is given, while section III presents the wind speed forecasting methodology. Section IV discusses the results obtained on a set of real-world data, and lastly, section V concludes the paper.

II. Gaussian Process Regression

In statistics, as a Gaussian distribution is a probability distribution function that is modeled as a function of two parameters, i.e., the mean and the variance. Furthermore, the set of Gaussian distributions that share a joint distribution are identified as a Gaussian process [24] that is also modeled as a function of two parameters, namely, the mean and the covariance functions, as shown below:

$$GP \sim N\left(m(x), C(x^T, x)\right) \tag{1}$$

where $m(x)$ is the mean, and $C(x^T,x)$ is the covariance function [24].

In the realm of machine learning, a Gaussian process is identified as kernel machine model. The reason is the ability of Gaussian processes to be expressed as a function of a kernel [24]. A kernel is any valid mathematical function that may be expressed as:

$$k\left(x_1, x_2\right) = f\left(x_1\right)^T \cdot f\left(x_2\right) \tag{2}$$

with f(x) called the “basis function”.

A GP is modeled as kernel function through the covariance function in (1). In particular, the function $C(x',x)$ is set equal to a kernel function, i.e., $C(x',x) = k(x',x)$, and therefore a Gaussian process is transformed to kernel machine. Concurrently, the mean function in (1) is set equal to zero, i.e., m(x)=0 that is a convenient choice for deriving the Gaussian process regression (GPR) model. GPR is the form of GP that is used in solving regression problems.

The starting point of GPR derivation is the simple linear regression model:

$$y = b_0 + \sum_{i=1}^{N} b_i x_i \tag{3}$$

where coefficient b's stand for the regression coefficients and N is the number of training datapoints. It should be noted that a training dataset consists of pairs of datapoints: a known output **t** for a known input **x**.

The underlying idea is that the N training datapoints are part of the same Gaussian process and therefore the GPR framework may use all pairs (x_n, t_n) to predicting the target value t_{N+1} of an unknown input x_{N+1}. Based on the above the joint distribution between the N available datapoints and the unknown input x_{N+1} follows a Gaussian distribution. Utilizing the joint Gaussian distribution, it has been proved in [24,26] that the GPR provides a predictive distribution over the unknown target $t_{N+1}$with mean and covariance given by:

$$m(x_{N+1}) = \mathbf{k}^T \mathbf{C}_N^{-1} \mathbf{t}_N \tag{4}$$

$$\sigma^2\left(x_{N+1}\right) = k - \mathbf{k}^T \mathbf{C}_N^{-1} \mathbf{t}_N \tag{5}$$

where $\mathbf{C}_N$ is the NxN matrix of kernel values among the N training datapoints, **k** is the vector of kernel values between the new $N+1$ and each of the N training datapoints, and k is a scalar value that represents the value $k(x_{N+1}, x_{N+1})$.

Overall, we observe from (4) and (5) that the predictive distribution strongly depends on the form of the kernel. Thus, the output of the predictive distribution can be controlled by the system modeler. There are several kernels that have been proposed in the literature; selection of a kernel depends on the requirements of the application at hand and the experience of the modeler [27].

III. MINUTE AHEAD WIND SPEED FORECASTING

A. Problem Statement

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the heading margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

Integration of renewable energy in electricity supplier portfolio is a key point for developing the future smart power grid. Generation of renewable energy is not controllable by humans with respect to amount and time of generation. Taking into consideration the wind dynamics, wind power is available whenever wind blows, while the speed of the wind determines the amount of power that can be generated.

The variability of wind speed with respect to time, accommodates various type of forecasting based on the length of the ahead-of-time forecasting horizon. The length of the horizon, may varies from a minute to month ahead, where each forecasting horizon may serve a different purpose in power grid management. The focus of the current manuscript is the extremely very-short-term forecasting of wind speed, as expressed of forecasting the wind speed in the next minute. Therefore, a single forecast must be provided every minute, in which case time also imposes a constraint; computation of forecasting should be significantly less than a minute in order to be of practical use to system operator.

B. Methodology

The proposed methodology integrates GP forecasting with fuzzy inference. The cornerstone of the proposed methodology is that “the future will be similar to the most recent past.” However, there is no clear guideline as to how much of the past should it be used in order to predict the future. The underlying idea in the current work is that to use a fuzzy inference system to determine the time window of the past that should be used for forecasting the future wind speed values. To that end, the fuzzy inference system will evaluate previous predictions and observations and will determine the amount of past observations that should be used for training the GPR model. The GPR model will be perform the forecasting of the minute ahead wind speed value. The overall block diagram of the proposed methodology is depicted in Fig. 1 where all the methodology steps are clearly given.

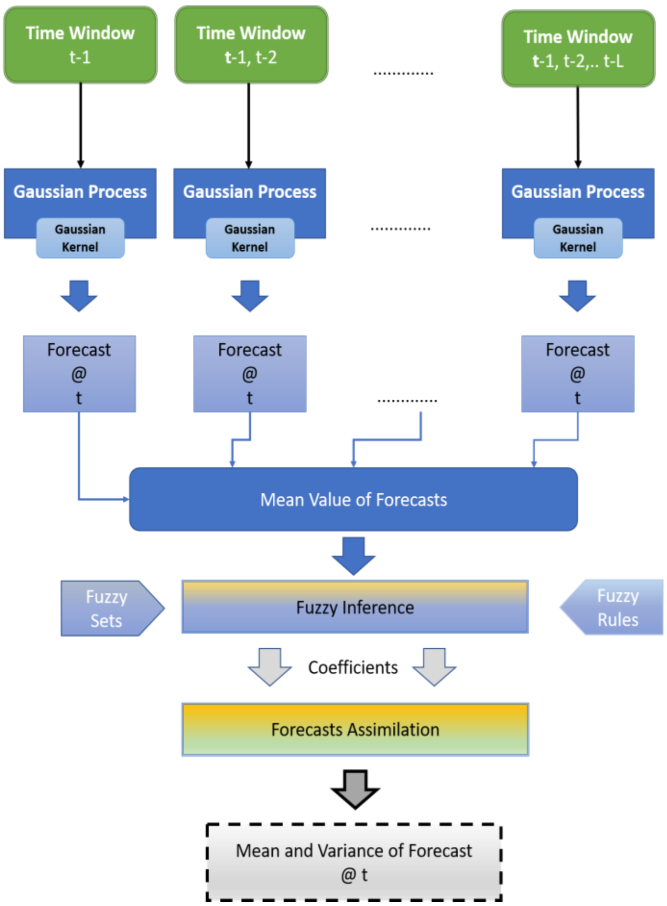

Figure 1. Block diagram of the proposed wind speed forecasting methodology for every minute.

Initially, a set of *N* Gaussian process regression models are created. Each of the GPR models is equipped with a kernel function and more specifically the Gaussian kernel whose analytical form can be found below:

$$k(x_1, x_2) = \exp\left(-\frac{\|x_1 - x_2\|^2}{2\sigma^2}\right) \tag{6}$$

where σ^2 is a hyperparameter evaluated using the training data, and x_1, x_2 are the input datapoints.

In the next step, each of the GPR models is provided with a set of training data. Each training dataset is comprised of various past observations; hence, we ensure that each GPR is trained based on different observations. Creation of the training datasets is conducted as follows. Assuming that we want to predict the wind speed at time *t*, we consider that all the wind speed values at past times are known. In other words, by assuming measured speed at minute intervals, then the wind speed values at minutes *t*-1, *t*-2,…,-*oo* are considered known and at our disposal. Given that there are *N* GPR models, then we create *N* training datasets. Each training dataset is created by a time window of length L where L=1,…,*N*+1. To make it clearer, each training dataset coincides with a time window that includes the observed values from *t*-1 to *t*-L. A visual representation of the training datasets is given in Fig. 2.

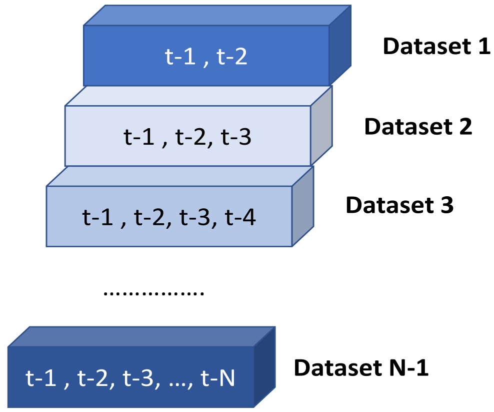

Figure 2. Visual representation of training datasets based on the time windows of part wind speed observations.

In the next step, each GPR utilizes the training dataset to provide a forecast over the wind speed at the next minute. The forecast is provided as a pair of values: the wind speed of the forecast and the variance around this value. The GPR computed values coincide with the values obtained by the predictive distribution in Eq. (4) and (5) respectively.

Once the forecasts have been obtained, then the mean value of the forecasts is calculated. The mean forecast is then forwarded to the fuzzy inference system, whose task is to assimilate the individual forecasts and provide a single one. To that end, the absolute difference between every individual forecast and the mean forecast. Denoting the mean forecast as μ_f and the individual forecasts as GPR_i, then we compute:

$$D_i = \left|\mu_f - GPR_i\right|, \quad i = 1, \ldots, N-1 \tag{7}$$

and thus, N-1 values are computed. These values are the inputs to the fuzzy inference system. The goal of the fuzzy inference engine is to evaluate a set of weight associated with each individual forecast. The idea is that the weight evaluation will be based on the difference of the individual forecast with respect to the forecast mean. By following this approach, we would like to give more weights to forecasts that are closer to the mean and eliminate the outliers (i.e., forecasts far away from the mean). By adopting the above approach, we anticipate that the majority of the GPR models will be able to capture the wind speed dynamics and provide forecasts that will be close to each other. Therefore, the forecasts that are clustered around a value will be evaluated with similar weights and therefore they will significantly contribute in the final assimilation. On the other hand, values that are far away from the mean will provide small or zero contribution to the assimilation process. Overall, the forecast assimilation process will provide a value that is close to the area with higher density of forecasts.

The fuzzy inference system has as an input the parameter D_i and as an output the weight value that lies in the interval [0 1]. The input fuzzy sets for the parameter "difference" D_i is depicted in Fig. 2, and the fuzzy sets for the output "weight" is depicted in Fig. 3 respectively. We observe that each variable is modeled by using 4 fuzzy sets. It should be noted that prior to any inference making the "difference" values are normalized (divide all values by the largest one).

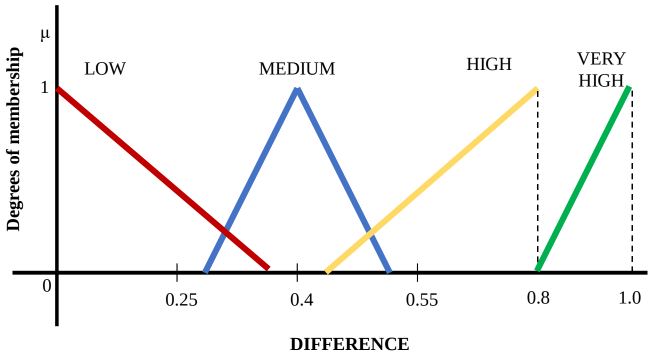

Figure 3. Fuzzy set modeling of the normalized parameter “difference” D_i.

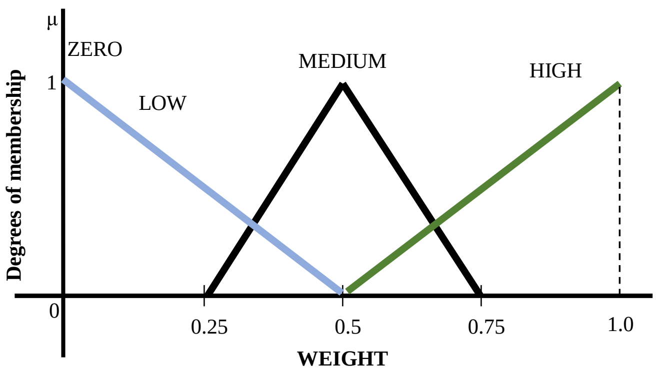

Figure 4. Fuzzy set modeling of the parameter “weight”.

In addition to the fuzzy sets, the inference system contains a small set of fuzzy rules. The rules that are of the IF…THEN form and are the following ones:

IF Difference is LOW, THEN Weight is HIGH,

IF Difference is MEDIUM, Then Weight is MEDIUM,

IF Difference is HIGH, Then Weight is LOW,

IF Difference is VERY HIGH, Then Weight is ZERO.

Evaluation of the above rules is done using the Mamdani Min operator [25], while the defuzzifying method is the centroid method [25]. Notably, the last rule assigns a zero weight to those predictions that are far away from the mean forecast. The latter assignment is conducted by modeling the fuzzy set *ZERO* as a singleton of the form (1,0) [24].

Overall, N-1 weights are being computed by the fuzzy inference system, with each weight being assigned to the respective GPR forecast. The final assimilation process is performed via two formulas. The first formula computed the final forecasted speed value and is given by:

$$FF = \frac{\sum_{i=1}^{N-1} w_i GPR_i}{\sum_{i=1}^{N-1} w_i} \tag{8}$$

while the second formula computed the forecast variance and is taken by:

$$FV = \frac{\sum_{i=1}^{N-1} w_i^2 V_{GPR_i}}{\sum_{i=1}^{N-1} w_i^2} \qquad (9)$$

where V_{GPRi} is the variance computed by the GPR model *i*.

Overall, we observe that the presented methodology allows the assimilation of various forecasts in order to get a single forecast value and the associated variance. With the adoption of several models, we anticipate that the majority of them will be able to capture the wind speed dynamics to various degrees, and their forecasts will cluster together.

IV. Forecasting Results

A. *Setup*

In this section, the proposed methodology is applied on a set of real-world wind speed data that have measured by the National Renewable Energy Laboratory (NREL) Observed Atmospheric and Solar Information System (OASIS) [28]. The available data contain measurements that have been obtained in minute intervals in the time period from April 1, 2017 to April 15, 2017. The wind speed has been recorded as the average speed with measurement unit in terms of m/s.

Furthermore, the presented methodology is benchmarked against a single GPR model that uses N measurements for training as well as against the statistical tool of autoregression (AR) of order 8 (p=8). The obtained results are recorded with respect to mean square error (MSE) and the variance, while are grouped with respect to daily performance. It should be noted that MSE is appropriate for this work because it handles cases where the wind speed is zero (as opposed to widely used measures like the mean average percentage error MAPE that fails if there are zero values) [29].

B. *Test Results*

The presented methodology is applied on minute ahead wind speed forecasting for various time intervals of the available data. More specifically, our tests contain three cases that contain measurements from three different cases in the time interval April 1-April 15.

In the first case, we forecast the minute ahead wind speed at day April 1, 2017 starting from 4.08am and finishing at 8.00pm (overall, there are 953 measurements). In this case, we present the detailed process for obtaining the first measurement (forecast at 4.08am) in order to present in a clear way how our methodology works; subsequently we provide the results of the whole-time interval. It should be noted that in all cases we considered that N=8, i.e., we had 8 GPR models providing predictions.

Fig. 5 presents the steps for obtaining the first forecast. In the first step the individual GPR models are utilized for obtaining the individual forecasts; hence, we get a set of 8 measurements. Next, we compute the mean value of the forecasts that is equal to 0.322m/s. Furthermore, we compute the differences between forecasts and the individuals and then we normalized those differences. The normalized differences are fed to the fuzzy inference system that provides a set of weights. We observe in Fig. 5 that there is one weight whose value is set equal to zero; this forecast is the farthest forecast from the mean of forecasts and thus the fuzzy system assigns to it a zero value according to the fourth rule. The final step assimilates the forecasts by providing a single value equal to 0.218, which is the closest value to real value among all forecasts.

Following are the minute ahead forecasting results of tested interval (April 1, 2017, 4.08am - 8.00pm). In particular, we obtained:

- *MSE* = 1.1136 for the GPR-fuzzy methodology
- *MSE* = 1.2262 for GPR with *N*=8
- *MSE* = 1.3266 for AR(8),

where the above results may also be found in Table I. Furthermore, we depict the computed forecasted values taken with the GPR-fuzzy methodology superimposed to the real wind speed values in Fig. 5.

There, we also see that the presented methodology was able to provide close forecasts despite the high volatility of the data.

Results for the next two tested cases are provided in Table I (together with the results of case 1 that are also provided there). Case 2 contains minute measurements taken from the day April 7, 2017 from 4.08am-8.00pm, and case 3 contains wind speed measurements from the day April 15, 2017 from 4.08am to 8.00pm.

Step 1: GP forecasts for various N							
N=1	N=2	N=3	N=4	N=5	N=6	N=7	N=8
0.067	0.121	0.205	0.581	0.726	0.643	0.191	0.037
Step 2: Compute the mean of forecasts							
Mean = 0.322							
Step 3: Find the differences							
N=1	N=2	N=3	N=4	N=5	N=6	N=7	N=8
0.254	0.200	0.116	0.259	0.404	0.322	0.130	0.284
Step 4: Find the Fuzzy Inference Values							
N=1	N=2	N=3	N=4	N=5	N=6	N=7	N=8
0.205	0.500	0.797	0.200	0	0.163	0.789	0.179
Step 6: Find the Mean and the variance forecast based steps 3 and 1 (Assimilation)							
Final Forecasted Value = 0.218 m/s				Real Value: 0.3480 m/s			
Final Forecast Variance = 0.0955 m/s							

Figure 5. Step by step forecasting of wind speed at 4.08am for case 1.

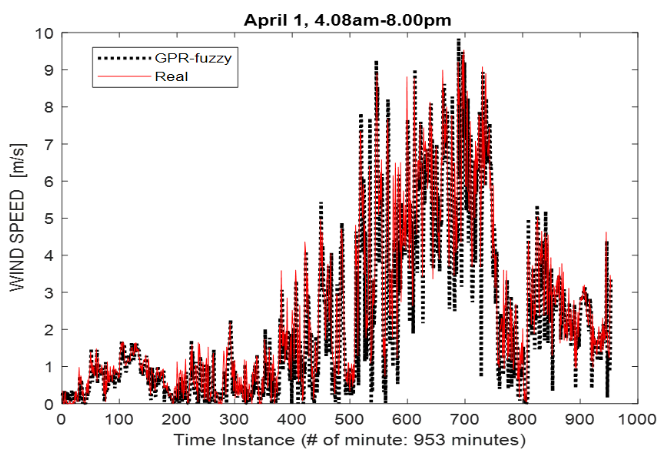

Figure 6. Wind speed forecasting using GPR-fuzzy against real values for case 1 (April 1, 2017, 4.08am-8.00pm).

TABLE I. RESULTS TAKEN FOR THE THREE TESTED CASES IN TERMS OF MEAN SQAURE ERROR (MSE)

Forecaster	GPR-Fuzzy	GPR with N=8	AR(8)
Case 1	1.1136	1.2262	1.3266
Case 2	0.5043	0.5308	2.44

Case 3	0.9595	1.0285	1.6355
Forecasted Variance	0.1917	0.3018	----

Obtained result exhibit that the highest accuracy among the three forecasted is attained by the presented methodology, i.e., GPR-fuzzy in Table I. In particular, we observe that the GPR-fuzzy provides the lowest MSE in all three cases, while the AR(8) model is the least accurate one in all tested cases. Furthermore, we observe that the simple GPR model accuracy is closer to that of the GPR-fuzzy; the latter observation designates that success of our methodology by integrating the GPR model with a fuzzy inference system. The work done by fuzzy in terms of rejecting some forecasts and assign weights to forecast assimilation has driven us to higher accuracy expressed with lower MSE compared to GPR. Regarding the variance of the forecast variance the GPR-fuzzy provides a lower mean variance compared to single GPR model. Lastly, with respect to execution time, the GPR-fuzzy methodology required about 2 sec to provide a forecast that is significantly lower than the upper limit of 60 sec.

V. CONCLUSION

In this paper, a new methodology for very short-term speed forecasting applicably to renewable power is presented. In particular, the methodology integrates a set of GPR models, which have been trained on various, with a fuzzy inference system. The fuzzy inference system allows the assimilation of the multiple GPR forecasts into a single one by weighting the forecasts according to their distance from the mean of forecasts. Results exhibit that the presented GPR-fuzzy methodology provided higher accuracy compared to single GPR and to AR(8) model for a set of three days comprised of minute ahead wind speed measurements taken from the April 2017.

Future work will focus on improving the fuzzy inference system, by adding a higher number of fuzzy sets and fuzzy rules. Furthermore, it will focus on adopting other kernel function beyond the Gaussian kernel and on testing on a higher volume of datasets taken from various time seasons.

REFERENCES

[1] L.H. Tsoukalas, and R. Gao, "From smart grids to an energy internet: Assumptions, architectures and requirements," In *Proc. 2008 IEEE Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, pp. 94-98, 2008.

[2] M. Alamaniotis, N. Gatsis, and L.H. Tsoukalas, "Virtual Budget: Integration of electricity load and price anticipation for load morphing in price-directed energy utilization," *Electric Power Systems Research*, vol. 158, pp. 284-296, 2018.

[3] S. Theodoridis, *Machine learning: a Bayesian and optimization perspective*, New York: Academic Press, 2015.

[4] M. Alamaniotis, D. Bargiotas, and L.H. Tsoukalas, "Towards smart energy systems: application of kernel machine regression for medium term electricity load forecasting," *SpringerPlus*, vol. 5(1), pp. 58, 2016.

[5] M. Liserre, T. Sauter, and J.Y. Hung, "Future energy systems: Integrating renewable energy sources into the smart power grid through industrial electronics," *IEEE industrial electronics magazine*, vol. 4(1), pp. 18-37, 2010.

[6] S.S. Soman, H. Zareipour, O. Malik, and P. Mandal, P., "A review of wind power and wind speed forecasting methods with different time horizons," In *Proc. North American power symposium,* pp. 1-8, 2010.

[7] A. Nasiakou, M. Alamaniotis, and L.H. Tsoukalas, "Power Distribution Network Partitioning in Big Data Environment Using K-Means and Fuzzy Logic," In *Proc. 10th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (Med Power 2016)*, November 6-9, 2016.

[8] G. Li, and J. Shi, "On comparing three artificial neural networks for wind speed forecasting," *Applied Energy*, vol. 87(7), pp. 2313-2320, July 2010.

[9] A. Daraeepour, and D.P. Echeverri, "Day-ahead wind speed prediction by a Neural Network-based model," In *Proc. 2014 IEEE PES Innovative Smart Grid Technologies Conference (ISGT),* pp. 1-5, February 2014.

[10] G. Li, J. Shi, and J. Zhou, "Bayesian adaptive combination of short-term wind speed forecasts from neural network models," *Renewable Energy*, vol. 36(1), pp. 352-359, 2011.

[11] H. Liu, T.Q. Tian, D.F. Pan, Y.F. Li, "Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks," *Applied Energy*, vol. 107, pp. 191-208, 2013.

[12] F. Castellanos, and N. James, "Average hourly wind speed forecasting with ANFIS," In *Proc. 11th American Conference on Wind Engineering*, pp. 1-8, June 2009.

[13] Z. Guo, W. Zhao, H. Lu, and L. Wang, "Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model," *Renewable Energy*, vol. 37(1), 2012, pp. 241-249.

[14] J. Zhou, J. Hongjie, Z. Tian, and L. Hu, "Short-term wind speed forecasting based on non-parametric kernel density estimation," In *Proc. IEEE Innovative Smart Grid Technologies-Asia,* pp. 1-5, May 2012.

[15] B. Banerjee, D. Jayaweera, and S.M. Islam, "Wind power forecasting with dynamic regression models and ageing considerations," In *Proc. IEEE PES Innovative Smart Grid Technologies Asia,* pp. 1-7, Nov. 2011.

[16] J. Zhou, J. Shi, and G. Li, "Fine tuning support vector machines for short-term wind speed forecasting," *Energy Conversion and Management*, vol. 52(4), pp. 1990-1998, 2011.

[17] X. Wang, G. Sideratos, N. Hatziargyriou, and L.H. Tsoukalas, "Wind speed forecasting for power system operational planning," In *Proc. IEEE International Conference on Probabilistic Methods Applied to Power Systems*, pp. 470-474, September 2004.

[18] E. Erdem, and J. Shi, "ARMA based approaches for forecasting the tuple of wind speed and direction," *Applied Energy*, vol. 88(4), pp. 1405-1414, 2011.

[19] I. Colak, S. Sagiroglu, M. Yesilbudak, E. Kabalci, and H.I. Bulbul, "Multi-time series and-time scale modeling for wind speed and wind power forecasting part I: Statistical methods, very short-term and short-term applications," In *Proc. 2015 IEEE International Conference on Renewable Energy Research and Applications,* pp. 209-214, November 2015.

[20] M. Alamaniotis, M., and G. Karagiannis, "Integration of Gaussian Processes and Particle Swarm Optimization for Very-Short-Term Wind Speed Forecasting in Smart Power," *International Journal of Monitoring and Surveillance Technologies Research*, vol. 5(3), pp. 1-14, 2018.

[21] J. Chen, G.Q. Zeng, W. Zhou, W. Du, and K.D. Lu, "Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization," *Energy Conversion and Management*, vol. 165, pp. 681-695, 2018.

[22] M. Khodayar, and J. Wang, "Spatio-temporal Graph Deep Neural Network for Short-term Wind Speed Forecasting," *IEEE Transactions on Sustainable Energy*, pp. 670-681, 2018.

[23] X. Luo, J. Sun, L. Wang, W. Wang, W. Zhao, J. Wu, J.H. Wang, and Z. Zhang, "Short-term wind speed forecasting via stacked extreme learning machine with generalized correntropy," *IEEE Transactions on Industrial Informatics*, vol. 14(11), pp. 4963-4971, 2018.

[24] C.E. Rasmussen, and C. Williams, *Gaussian processes for machine learning*, MIT Press: Boston, 2006.

[25] L.H. Tsoukalas, and R.E. Uhrig, *Fuzzy and Neural Approaches in Engineering*, New York: Wiley and Sons Ltd., 1997.

[26] D.J.C. Mackay, "Introduction to Gaussian Processes," in *Neural Networks and Machine Learning*, C.M. Bishop, Ed. Berlin: Springer-Verlag, 1998.

[27] M. Alamaniotis, J. Mathew, A. Chroneos, M. Fitzpatrick, and L.H. Tsoukalas, "Probabilistic Kernel Machines for Predictive Monitoring of Weld Residual Stress in Energy Systems," *Engineering Applications of Artificial Intelligence*, vol. 71, pp. 138-154, May 2018.

[28] A. Andreas, and S. Wilcox, "Observed Atmospheric and Solar Information System (OASIS)," NREL, Tucson, Arizona (Data), NREL Report No. DA-5500-56494, 2010.

[29] M. Alamaniotis, A. Ikonomopoulos, and L.H. Tsoukalas, "Evolutionary Multiobjective Optimization of Kernel-based Very Short-Term Load Forecasting," *IEEE Transactions on Power Systems*, vol. 27(3), pp. 1477-1484, August 2012.