World Academy of Science, Engineering and Technology
International Journal of Energy and Power Engineering
Vol:14, No:12, 2020

# Enhancing Temporal Extrapolation of Wind Speed Using a Hybrid Technique: A Case Study in West Coast of Denmark

B. Elshafei, X. Mao

***Abstract*—**The demand for renewable energy is significantly increasing, major investments are being supplied to the wind power generation industry as a leading source of clean energy. The wind energy sector is entirely dependable and driven by the prediction of wind speed, which by the nature of wind is very stochastic and widely random. This s0tudy employs deep multi-fidelity Gaussian process regression, used to predict wind speeds for medium term time horizons. Data of the RUNE experiment in the west coast of Denmark were provided by the Technical University of Denmark, which represent the wind speed across the study area from the period between December 2015 and March 2016. The study aims to investigate the effect of pre-processing the data by denoising the signal using empirical wavelet transform (EWT) and engaging the vector components of wind speed to increase the number of input data layers for data fusion using deep multi-fidelity Gaussian process regression (GPR). The outcomes were compared using root mean square error (RMSE) and the results demonstrated a significant increase in the accuracy of predictions which demonstrated that using vector components of the wind speed as additional predictors exhibits more accurate predictions than strategies that ignore them, reflecting the importance of the inclusion of all sub data and pre-processing signals for wind speed forecasting models.

***Keywords*—**Data fusion, Gaussian process regression, signal denoise, temporal extrapolation.



## I. Introduction

In recent years, there has been a rise in understanding of the environmental effects of greenhouse gases. Considerations have given rise to demand for green energy sources and have motivated a rapid expansion of all its industries, a large fraction of which is produced by wind power through wind turbines. As of October 2019, more than 20 GW of wind power has been constructed in the United Kingdom. In addition, according to the government's plan, it intends to further increase demand for wind energy and its capacity [1]. Nevertheless, the erratic existence of this form of energy remains the biggest obstacle in incorporating the strength into the electricity grid. Future forecasts help overcome this dilemma in order to render wind energy a stable and consistent source of electricity. As a result, managing the supply and demand of electrical energy, which would reduce the cost impact on power system operators, will help integrate wind energy into the electrical system [2].

As mentioned above, the intermittent nature of wind energy is outlined in challenges that arise, such as non-storability, immense variability and limited predictability. This defying volatile nature of wind energy categorized in the parameters: wind speed and occurrence, have given operators prominence in all forecasting methods and how the accuracy of the forecasting methods could be enhanced by combining several techniques in a single hybrid method [3]. The vast number of methods makes finding the most desirable combination that produces the most accurate predictions tricky and highly investigative. However most of the approaches in the literature review rely on point-time forecasts and are classified under: (1) Persistence Method; (2) Physical Approach: Numerical Climate Prediction (NWP); (3) Statistical Approaches: Classical Time Series Models and Artificial Neural Network (ANN); (4) Machine Learning Methods; and (5) Dynamic hybrid Approaches [4], [5].

There are many models for each of the categories listed, and several of the models are improved by the early ones. The time period of the assignment plays a significant role in recruiting a model, in addition to the drawbacks and consequences of each model. The time-scale definition of wind energy is rather unclear. It may, however, be divided into four-fold horizons of specific ranges and applications:

- Very short-term, ranging from few seconds to 30 minutes ahead.
- Short-term, ranging from 30 minutes to 6 hours ahead.
- Medium-term, ranging from 6 hours to 1 day ahead.
- Long-term, ranging from 1 day to 1 week and more ahead [6].

However, we are interested in the machine learning approach, which is based on training with data collected that further use the difference between predicted and actual data in the almost immediate past to optimize model parameters [7]. In this study, the deep multi-fidelity GPR model, a non-parametric, stochastic process that follows the Bayesian regression approach, infers a probabilistic distribution of all possible values. Following this, the algorithm works on combining two sets of data with different levels of fidelity. The first is a (few) high fidelity data with (several) low-fidelity data, which provides more reliable forecasts without total reliance on heavy numerical, costly and time-consuming methods to produce high fidelity data collection. We strive to increase the amount and sources of low-fidelity data, taking into account derivatives and decomposed vector components of low-fidelity data.

Recently, the bulk of models in the literature have drawn

B. Elshafei and X. Mao are with the University of Nottingham, Nottingham, NG7 2RD UK (e-mail: basem.elshafei@nottingham.ac.uk).

focus exclusively to the issue of wind speed forecasting over time series. Studies have thoroughly researched machine learning and the impact of using GPR in the area of wind speed forecasts, and some hybrid models have provided impressive findings of successful short-and long-term forecasts. For example, [8] used EWT, PACF (partial auto-correlation function) with GPR. EWT was used to collect useful data from the wind speed sequence via a modified wavelet filter array, and PACF provided input parameters for the GPR to model dynamic features and internal uncertainties. The system was able to produce positive precise forecasts and calculated wind speed periods, and the findings were favorable relative to other models at the point of wind speed forecast. Some time series representations include ARMA (Auto-Regressive Moving Average) [9], improved ARIMA (Auto-Regressive Integrated Moving Average) [10], FARIMA (Fractional Auto-Regressive Integrated Moving Average) [11], Grey Predictors [12], and exponential smoothing [13].

The alternate combination was preceded by [14], joined by AR (auto regressive) and GPR, where AR was used to catch the form of the wind speed sequence and GPR was derived from the local structure. In addition, automated significance determination (ARD) was used to determine the value of various variables, and specific forms of covariance functions were merged to collect data characteristics. This method was compared to the support vector machine (SVM) and the ANN and the benchmark forecasting method, the persistence approach. The results showed that the combination improved the point forecast compared to other models and generated satisfactory predictive intervals. A research by Cadenas and Rivera [15] finds non-linearity to be a self-regressive exogenous paradigm where the algorithm developed forecasts for 1 hour ahead and was compared to both the persistence approach and the NAR (nonlinear auto-regressive) which had only one component. The findings revealed that the NARX medal was the most reliable of the three, indicating the function performed by introducing additional variables to improve precision.

## II. Methodology

### A. Gaussian Process Regression

Gaussian process regression is a non-parametric, stochastic method that implements the Bayesian regression approach, which results in a likelihood distribution of all possible values. The Bayesian method defines the prior distribution, p(w) on the equation function, w, and the conversion of possibilities on the basis of the data obtained by implementing the Bays law. Following, the rule provides a posterior distribution, p (w|y, X) includes information from the data set and the previously mentioned distribution. However, to estimate values at unobserved points of interest, x *, we calculate the predictive distribution by taking into account all possible predictions using their calculated posterior distribution. We consider that all terms of the equation are Gaussian in order to map the regression, using this, we solve for the predictive distribution to get a Gaussian distribution, where we get a point estimate using its mean and a quantification of the uncertainty by its variance. The Gaussian cycle is different in that it is non-parametric and thus not restricted to a specific functional form. For the case of this experiment, the training dataset provides the posterior while the predictive posterior distribution is computed on the points of interest [16].

GPR has many benefits, operating well on limited data sets and being able to include ambiguity tests on predictions. Simply placed, it is the set of points provided by random variables that are indexed by time or space. Each and every finite collection of these random variables has a normal multivariate distribution. GPR is only satisfied if and only if there is a Gaussian distribution for any finite collection of all data.

$$posterior\ \frac{(likelihood \times prior)}{marginallikelihood} \tag{1}$$

$$p(w|y,X) = \frac{(p(y|X,w)p(w))}{p(y|X)}$$

$$p(f^*|x^*,y,X) = \int p(\,(f^*|x^*,w)p(w|y,X)dw \tag{2}$$

$$f(x) \sim GP\ (m(x),k(x,x') \tag{3}$$

Hence, the prior is defined by two parameters: the mean function, m(x), representing the trend of the function, secondly, the co-variance function (kernel), k (x, x'), reflecting the dependence of the structure and it is defined by hyperparameters. Moreover, the co-variance function must always satisfy a definite positive [17].

### B. Multi fidelity GPR

For the field of geostatistics, a more sophisticated method is followed, multi-fidelity GPR, where multivariate functions at different code levels that represent functions with variable fidelities, the lower fidelity data sets are more viable and cheaper, hence provide data-rich information that is qualitatively less interesting than that of higher fidelities. Combining data sets with several fidelities allows for a correlation between the different levels of accuracy which enhances the learning efficiency, thus giving more accurate data without the complete dependence on expensive methods. Furthermore, the study employs a GPR model based on the Bayesian approaches [18]. Thus, inference in Gausian Process (GP) is simple, a joint GP prior is put on training to test latent values, based on more than one level of coding. Hence, the two levels of multi-fidelity data sets considered in this study $f_1(h)$ and $f_2(h)$ representing the low and high-fidelity data sets.

From the above fundamental equations, the use of different co-variance functions such as: constant, linear, squared exponential, Mater and Rational quadratic, defines the prediction algorithm for the Gaussian process. Moreover, adjusting the hyperparameters defines the co-variance function and yields to a more accurate estimation. The additional levels used in this experiment are 1$^{st}$, 2$^{nd}$ derivatives and the decomposed east and north components of the low-fidelity wind speed data, the additional data possess a very similar trend to the original data, this invests more for the GPR and increases the efficiency of learning and locating the trends [19].

$$\begin{pmatrix} f_1 h \\ f_2 h \end{pmatrix} \sim GP\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} k_1(h,h^1) & \rho k_1(h,h^1) \\ \rho k_1(h,h^1) & \rho^2 k_1(h,h^1) + k_2(h,h^1) \end{pmatrix}\right) \quad (4)$$

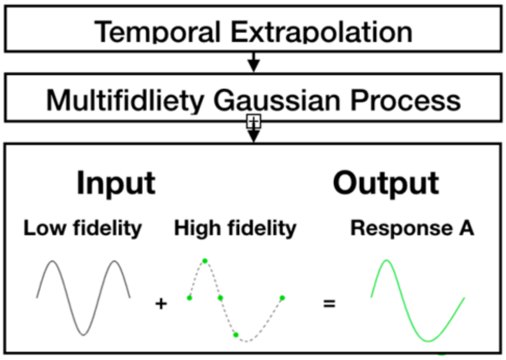


Fig. 1 Flow chart for temporal extrapolation with expected outcome

## III. Case Study

### A. Location Description

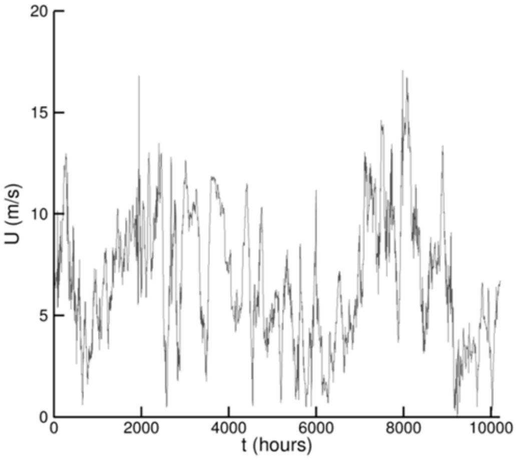


Fig. 2 Low-fidelity data from weather research forecast simulations

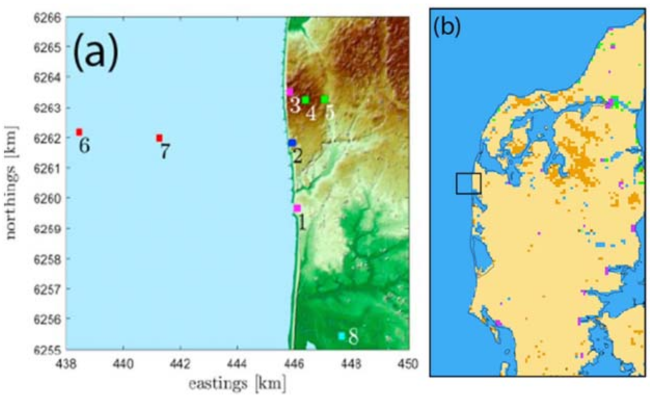


Fig. 3 (a) Location of all 7 lidars and (b) The RUNE experiment location at the west coast of Denmark

The experiment took place at a 9 km x 5 km area near the west coast of Denmark in the Lemvig municipality; this can be shown in Fig. 3 (b). In Fig. 3 (a), we observe the locations of the Lidar scanners used to generate the high-fidelity data discussed in Section III *C,* whereas the low-fidelity data were generated using weather research forecast simulations discussed in Section III *B*. This area is a strong location of interest for offshore wind development and studies. The site is well-serviced with 4G network, useful for device monitoring and data retrieval, needed for synchronization of the wind scanners. In addition, the location is considered very close to the test station of wind turbines at Hovsore. The topography of the area is described as almost flat coastal farmland, where the North Sea and grasslands are separated by a sand embankment. Moreover, towards the north side of the area, the aforementioned embankment transforms into cliffs that are covered by grass. Furthermore, position 1, position 2 and position 3 have terrains that are about 43 m above sea level. This is a feature of crucial importance for the lidars as it provides a clear line of sight to measure wind speeds accurately [2].

### B. Low-Fidelity Data

The low-fidelity data are usually a set of data generated using cheap and fast methods that are highly viable. They generate data-rich time-series in a short period of time; however, the data usually are high uncertainty. For this experiment, the low-fidelity data were generated using weather research forecast simulations. We gathered data from 12 different simulations combining various features such as: horizontal grid spacing; PBL schemes; atmospheric boundary conditions and SST sources. The simulation with the lowest accuracy was chosen to represent the low-fidelity data, in our case, YSU (0.5). The data set contained 10,298 values of east and north vector components each of the wind speed for an entire period from December 2015 to March 2016 with a prediction interval of 10 minutes represented in Fig. 2.

### C. High-Fidelity Data

This set was generated using expensive and highly accurate scanning lidars, the main problem with lidars is that they have to be physically present at each location. The allocation of lidars at extreme offshore locations results in destruction of equipment at early stages due to the extreme weather conditions, thus low availability of data. For this set of data, the equipment involved 3 PPI scans covering an azimuthal range of 60 degrees at an elevation angle of 1 degree located in locations (1, 2 & 3), additionally 2 dual setup lidars were configured to match their scans along 3 horizontal virtual lines at 50,100 and 150 m (ASL) at locations (1 & 2), 2 scanners were adjusted to perform RHI scans intersecting at location 7, hence a vertical profile with range gates with 36 different heights is achieved. Finally, vertical profiling lidars performed VAD scans. The high-fidelity data set had 113 values for the same period of time as the low-fidelity data and is shown in Fig. 4.

### D. Data Pre-Processing

The pre-processing of signals before using is always very beneficial, this step helps eliminate unwanted residuals and anomalous data in the signal by filtering data according to specific measures and trends of the entire signal. For this study we employed state-of-the-art EWT. This method works on constructing an appropriate wavelet filter bank to act as a filtering threshold for the provided data [20].

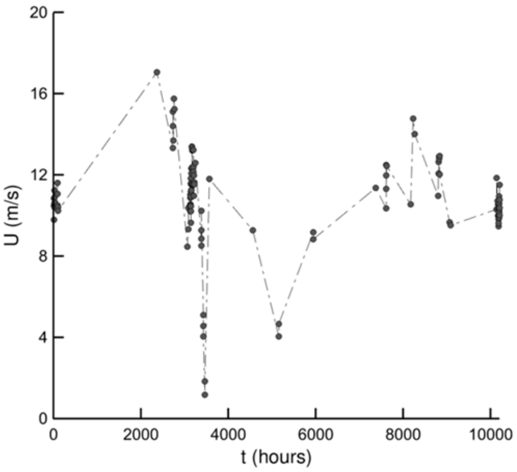


Fig. 4 High-fidelity data from wind scanning lidars

The advancement using this method from previous filtering methods relies in its capability of overcoming major issues with previous techniques such as; wavelet transform (WT) and empirical mode decomposition (EMD), where both mentioned process are known to remove sufficient amount of noise from the signal; however, the EMD model lacks any mathematical theories and is considered highly sensitive to both noise and sampling, while for the WT model, the model incapable of performing as a self-adaptive due to models requiring the parameters specified beforehand. The aforementioned issues are why EWT shows promising advantages as it eliminates those drawbacks. The approach can be summarized in the following main steps:

1. Extending the signal.
2. Fourier transforms.
3. Extracting boundaries.
4. Building a filter bank.
5. Extracting the sub band.

The EWT model identifies the time-series to extract different modes in it, the algorithm is then based on robust possessing for peaks detections where it performs a spectrum segmentation based on the maxima extracted, this is finally used to construct the wavelet filter bank [21]. In the figure EWT, we illustrate the difference between the filtered and original signals. As shown in Fig. 5, the filtered time-series (blue line) has fewer spikes and smoother curve form.

## IV. Results and Discussion

### A. Results of Each Method for the Three Models

Three different models of the multi-fidelity GPR were applied, where the low-fidelity data were employed to stretch the margin of values for the high-fidelity data. For the first method, only high- and low-fidelity data sets were employed, the extrapolation method produced data that follow the pattern of high-fidelity data, but the process shows a distorted curve in between the gaps of the high-fidelity data as shown in Fig. 6 (a). The RMSE of the experiment was 1.76.

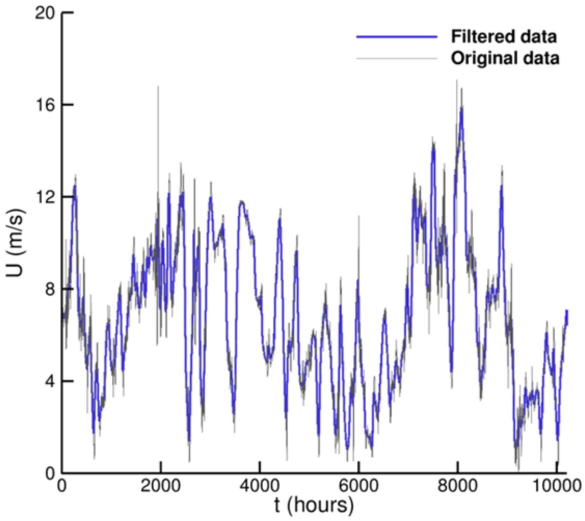


Fig. 5 Original dataset (Black) and filtered response (Blue)

The second model, which involves derivatives of the low-fidelity data, showed a drop in the RMSE to 1.43, and a response curve that had a similar pattern to the low-fidelity data and also covered the high-fidelity data appropriately represented in Fig. 6 (b). Furthermore, for the introduced model, where we pre-process the low-fidelity data to denoise the signal and introduce the vector components of the wind speed signal in the east and north directions, the response had a satisfying drop in the RMSE to 0.82, reflecting the improvement in accuracy of predictions. Additional confidence bounds (90%) were used to reflect the low uncertainty in modelling the high-fidelity data using the response of the temporal extrapolation process shown in Fig. 6 (c). In Fig. 6 (d), a closer look on the representation of the response to the high-fidelity data is considered, we show the data for hours between 3000 and 3600, the time with the richest high-fidelity values.

Finally, Fig. 6 (e) compares the error for all three methods by plotting the difference between all high-fidelity data and their corresponding in the response curve. In this figure we see that few points, the first and second model, perform better predictions, despite the overall superiority to the third model. This is mainly due to the presence of these points in positions where the pattern observed by the GPR model was more accurate at this trial.

### B. Results of RMSE of All Models across all the Data Points

In addition to comparing the predictions for a single point, as shown in Fig. 6, further analysis of the model was carried out. The three models were executed for a straight line of points starting with the farthest offshore point 6 to the most onshore point 2, a total of 36 points. For the first model, the generated RMSE's ranged between 1.1284 and 2.106, with the highest at the first onshore point, and a majority of the highest errors at the onshore area. This is due to the complex terrain of the area's geometry. On the other side, offshore points showed lower average error due to the calm geometrical conditions at sea, although going further offshore shows an increase in error as the weather conditions become harsher.

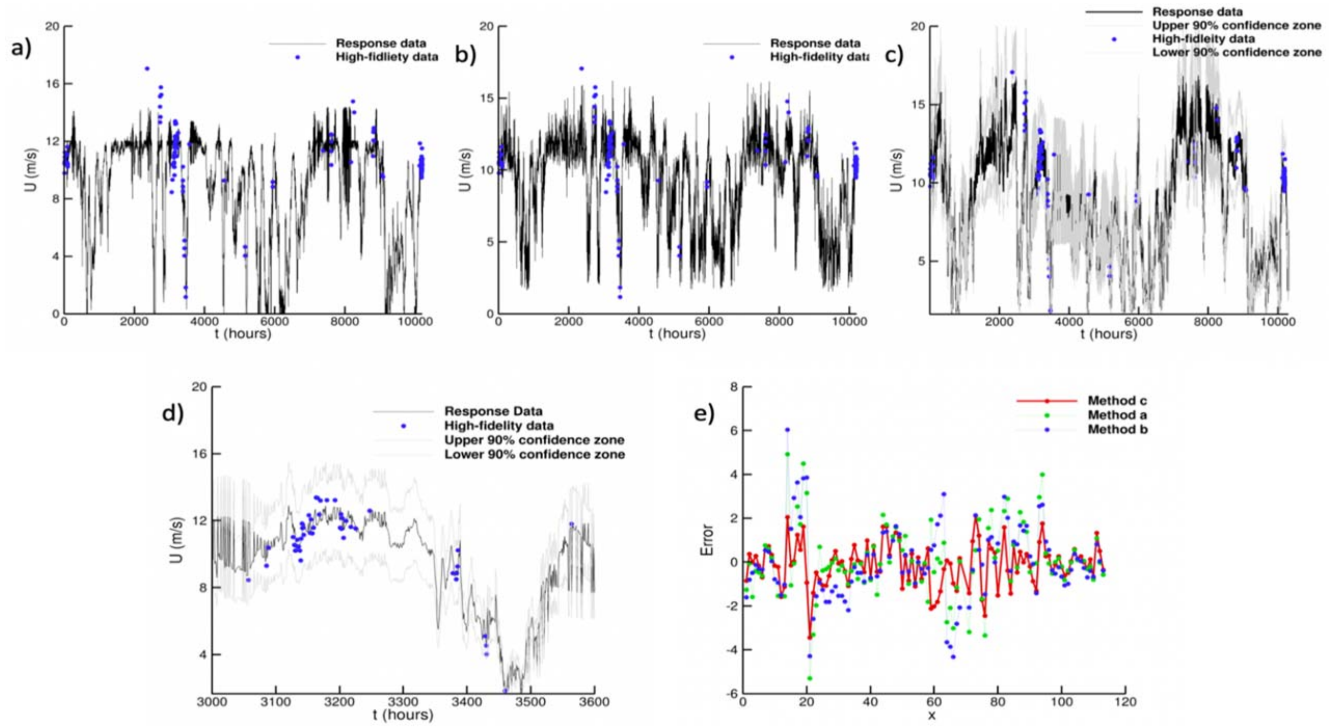

Fig. 6 (a) Original GPR method, (b) GPR with derivatives, (c) GPR with derivatives, filtering and vector components, (d) Close up for the third method with confidence bounds and (e) Error between response and high-fidelity points for all three methods

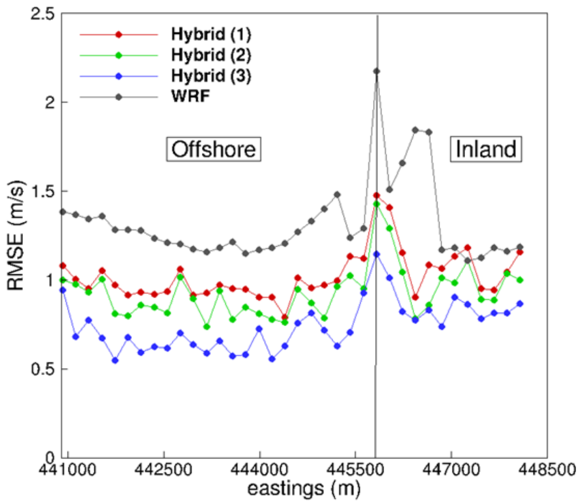


Fig. 7 Results of RMSE across all Easting points for all three Multi-fidelity GPRs

Secondly, the test which showed pleasant improvements decreasing the highest and lowest error values to 2.037 and 1.0538 at roughly the same positions raised a question about the number of derivatives used, we found that increasing the number of derivatives will in fact decrease the error, but with more additional layers introduced, the system becomes of high complexity and the process time increases. Finally, the third model showed a significant drop in RMSE across all points as expected. The highest and lowest errors were 1.6332 and 0.784, respectively. As shown in Fig. 6, the performance of the third model exceeds that of the other two and shows a clear superiority and improvement in prediction. Furthermore, the pattern of errors matches that of the literature provided by the Technical university of Denmark, where they used meteorological mist and the deviation was found higher at far offshore and onshore with the onshore area a little higher [2].

## V. Conclusion

The growing demand for wind energy incorporation into the power grid is rapidly increasing and, because the nature of wind energy is erratic and stochastic, a single-parameter or single-model method is not of a sufficient precision compared to industrial standards, so effective and robust hybrid models for forecasting wind speed in the short and long term are of high importance. In this paper, we showed how pre-processing the wind speed signal, hence eliminating the residuals, makes the predictions more accurate. In addition, increasing the number of layers for the deep multi-fidelity GPR also had a drastic effect on the precision of the prediction as it acted as additional layer of helpful information for the data fusion resulting in a positive direct correlation reflected in a significant reduction in the RMSE [3]. On the other hand, adding more layers of input data increases the complexity of the system, as more hyperparameters are required to be solved for. This results in more execution time to carry out the same process for an application where predictions could be for few minutes ahead such as; short and very short time horizons. In conclusion, achieving higher precision in wind speed forecasting through increasing the predictors is a helpful method, but the outrageous use might result in huge consumption of time that is of crucial importance in this field.

## ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon Research and Innovation Programme under the Marie Skaodwska-Curie Grant agreement no. 777717 and the Future and Emerging Technologies Programme with agreement no. 828799.

## REFERENCES

[1] Negnevitsky, M., Johnson, P., & Santoso, S. (2007). Short term wind power forecasting using hybrid intelligent systems. 2007 IEEE Power Engineering Society General Meeting. doi: 10.1109/pes.2007.385453.
[2] Hoolohan, V., Tomlin, A. S., & Cockerill, T. (2018). Improved near surface wind speed predictions using Gaussian process regression combined with numerical weather predictions and observed meteorological data. Renewable Energy, 126, 1043–1054. doi: 10.1016/j.renene.2018.04.019.
[3] Zhang, C., Wei, H., Zhao, X., Liu, T., & Zhang, K. (2016). A Gaussian process regression based hybrid approach for short-term wind speed prediction. Energy Conversion and Management, 126, 1084–1092. doi: 10.1016/j.enconman.2016.08.086.
[4] Tascikaraoglu, A. and Uzunoglu, M., 2014. A review of combined approaches for prediction of short-term wind speed and power. Renewable and Sustainable Energy Reviews, 34, pp.243-254.
[5] Zhu, X. and Genton, M., 2012. Short-Term Wind Speed Forecasting for Power System Operations. International Statistical Review, 80(1), pp.2-23.
[6] Shao, H., Wei, H., Deng, X. and Xing, S., 2017. Short-term wind speed forecasting using wavelet transformation and AdaBoosting neural networks in Yunnan wind farm. IET Renewable Power Generation, 11(4), pp.374-381.
[7] S. S. Soman, H. Zareipour, O. Malik and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," North American Power Symposium 2010, Arlington, TX, 2010, pp. 1-8, doi: 10.1109/NAPS.2010.5619586.
[8] Hu, J. and Wang, J., 2015. Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression. Energy, 93, pp.1456-1466.
[9] Erdem Ergin, Shi Jing. ARMA based approaches for forecasting the tuple of wind speed and direction. Appl Energy 2011;88:1405e14.
[10] Liua Hui, Tian Hong-qi, Li Yan-fei. An EMD-recursive ARIMA method to predict wind speed. J Wind End Ind Aerodyn 2015;141:27e38.
[11] Katris Christos, Daskalaki Sophia. Comparing forecasting approaches for Internet traffic. Expert Syst Appl 30 November 2015;42(21):8172e83.
[12] Li Yunhua, Ling Lina, Chen Jiantao. Combined grey prediction fuzzy control law with application to road tunnel ventilation system. J Appl Res Technol 2015;13:313e20.
[13] Hu Jianming, Wang Jianzhou, Zeng Guowei. A hybrid forecasting approach applied to wind speed time series. Renew Energy 2013;60:185e94.
[14] Zhang, C., Wei, H., Zhao, X., Liu, T. and Zhang, K., 2016. A Gaussian process regression based hybrid approach for short-term wind speed prediction. Energy Conversion and Management, 126, pp.1084-1092.
[15] Cadenas, E., Rivera, W., Campos-Amezcua, R. and Cadenas, R., 2015. Wind speed forecasting using the NARX model, case: La Mata, Oaxaca, México. Neural Computing and Applications, 27(8), pp.2417-2428.
[16] Petelin, D., n.d. Gaussian Processes For Machine Learning.
[17] Alamaniotis, M. and Karagiannis, G., 2017. 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, 5(3), pp.1-14.
[18] Parussini, L., Venturi, D., Perdikaris, P. and Karniadakis, G., 2017. Multi-fidelity Gaussian process regression for prediction of random fields. Journal of Computational Physics, 336, pp.36-50.
[19] Raissi, M. and Em Karniadakis, G., 2016. Deep Multi-fidelity Gaussian Processes. ArXiv,.W.-K. Chen, *Linear Networks and Systems* (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.
[20] Hu J, Wang J. Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression. Energy 2015;1456:1466-93.
[21] Catalao JPS, Pousinho HMI, Mendes VMF. Short-term wind power forecasting in portugal by neural networks and wavelet transforms. Renew Energy 2011;36:1245-51.