Wind Speed Forecasting via Structured Output Learning

Annalisa Appice[1,2](✉), Antonietta Lanza[1], and Donato Malerba[1,2]

[1] Dipartimento di Informatica, Università degli Studi di Bari Aldo Moro via Orabona, 4 - 70126 Bari - Italy
[2] Consorzio Interuniversitario Nazionale per l'Informatica - CINI
annalisa.appice@uniba.it, antonietta.lanza@uniba.it, donato.malerba@uniba.it

Abstract. In the context of the wind energy management, the study of time series data by means of a predictability analysis can be very helpful. For example, accurate wind speed forecasts are necessary to schedule dispatchable generation and tariffs in the day-ahead electricity market. This paper examines the use of structured output learning, in order to model historical wind speed data and yield accurate forecasts of the wind speed on the day-ahead (24 h) horizon. The proposed method is based on a multi-resolution analysis of the historical data, which are represented at multiple scales in both space and time. Handling multi-resolution wind speed data allows us to leverage the knowledge hidden in both the spatial and temporal variability of the shared information, in order to identify spatio-temporal aided patterns that contribute to yield accurate wind speed forecasts. In an assessment, using benchmark data, we show that the multi-resolution structured output learning is able to determine more accurate forecasts than the state-of-the-art structured output models.

1 Introduction

Nowadays power and energy systems with wind energy being as integral system have been successful. The benefit of clean wind energy also brings the challenge of forecasting wind power for optimal management of electricity grids. However, the variable nature of wind speed [17] poses operational challenges for wind power integration into modern power systems. As the wind variability occurs in time, as well as in space scales, the profiles of the available power of wind sources depend on the geographic location, the season (or time of the year), the time of the day and other physical parameters.

In this paper, the wind speed forecasting task is addressed by considering wind speed data measured every 10 minutes along day-ahead time horizons. We propose a specific time series approach that applies artificial intelligence, in order to learn a forecasting model from the historical data only. A peculiar contribution is the consideration of multi-resolution representations of historical data, in order to handle the variability of the wind speed information at the space and time scales. Our purpose is the investigation of the implications of learning multi-resolution data on the accuracy of the forecasting operation. Specifically, we formulate the day-ahead forecasting task as a structured output predictive learning problem [4]. A multi-target model is considered, in order to learn a single model that predicts multiple output variables at the same time – one variable for each time point over the 24-ahead horizon. The decision of learning an output structured model is supported by various studies which have repeatedly proved that multi-target models are typically easier to interpret, perform better, and overfit less than single-target predictions [4, 29, 7].

Neighboring and windowing mechanisms are adopted, in order to represent the historical data at various scales in both space and time, respectively. These mechanisms are combined with the standard deviation operator that is used to quantify the spatial and/or temporal variability of the data. This multi-resolution representation of the wind speed variability contributes to define new input variables, as well as new output variables. In this way, we are able to learn a multi-target model that accounts for the variability of measurements at different sites and times. The viability of the proposed method is assessed in the structured output learning by comparing the accuracy of traditional multi-target models to the accuracy of multi-resolution multi-target models in a benchmark scenario. The sensitivity of the accuracy of the proposed forecasting model learned is evaluated along the size of the scale. Finally, the accuracy gained in by taking into account appropriate patterns of the data variability is explored.

The paper is organized as follows. In the next Section, we briefly report the state-of-the-art of the time series analysis for the problem of wind speed forecasting. In Section 3, we describe the basics of this study. In Section 4, we present the multi-resolution structured output learning phase proposed here. In Section 5 we describe the benchmark dataset considered for the empirical evaluation and illustrate the relevant results. Finally, Section 6 draws some conclusions and outlines some future work.

2 Related work

In the literature, different forecasting horizons have been investigated: long-term (from one day to one week ahead), medium-term (from 6 h to one day ahead), short-term (from 30 min to 6 h ahead) and very short-term (few seconds to 30 min ahead) [28, 11]. On the other hand, various approaches have been developed for wind speed forecasting in renewable energy systems. In particular, three main predictive categories are described in the literature [6, 30, 3]: the physical [1, 15], the time series [27, 8, 20, 13, 18, 22, 21, 2, 5] and the hybrid [16, 10, 12] approaches. The physical approach describes a physical relationship between wind speed, atmospheric conditions, local topography and the output from the wind power turbine. The time series approach consists of time series forecasts, which are based on the historical data (the wind speed collected at a specific site), while neglect commonly the meteorological data. Finally, the hybrid approach applies a combination of physical and time series models.

By focusing the attention on the time series approach (which is the most popular in practice and also the subject of this paper), a wide plethora of time-series methods employ a general class of statistical models, that is, the Auto-Regressive Moving Average (ARMA) or Auto-Regressive Integrated Moving Average (ARIMA), in order to estimate future observations of a wind farm through a linear combination of the past data. The recent literature [26, 13, 19] has shown that these auto-regressive models are very well suited to capture short range correlations. Hence, they have been used extensively in a variety of (very) short-term forecasting applications (less than 6 hours). Recent studies have also proved that auto-regressive models can be profitably extended, in order to account for spatial characteristics of time series data and gain in accuracy [23–25]. In alternative, the time series approach also involves the use of artificial intelligence techniques, which are commonly well suited to produce accurate prediction in medium-term and long-term forecasting applications. Examples of artificial intelligence wind speed forecasting methods apply Neural Networks [9], Support Vector Machine [31], Regression Trees [18], K-Nearest Neighborhood [5] and Cluster analysis [22, 21].

By investing in artificial intelligence, this paper explores the use of the structured output learning in a medium-term wind speed forecasting application (24 h ahead). We note that the benefits of structured output learning in time series forecasting have been recently assessed in [7] considering the problem of deriving 24 h ahead solar radiation forecasts. Differently from this seminal study, that has modeled the spatial "correlation" of the solar radiation, in order to define new "input" variables only, we leverage here the power of a model of both the spatial and temporal "variability" of the wind speed, in order to define new "input" and "output" variables, which contribute to yield accurate 24 h ahead forecasts of the wind speed.

3 Basics

Premises Without loss of generality, the applicative scenario we consider in this work is described by the following four premises. First, the spatial location of a wind farm is modeled by means of 2-D point coordinates (e.g. latitude and longitude). Second, the spatial locations of the wind farms are known, distinct and invariant. Third, wind farms transmit measurements of the wind speed and they are synchronized in the transmission time. Finally, transmission time points are equally spaced in time.

Learning task Based upon these premises, the task we intend to perform is to forecast wind speed at each farm of the grid. The forecasting model is that learned from the input historical data of the wind speed, as they are collected

---

Multi-dimensional representations of geographic space can be equally dealt.

from a grid of wind farms, every 10 minutes, over $m+1$ consecutive days. We also consider additional input information, which models the wind speed variability at the spatial, temporal and spatio-temporal scales. The output of the learning phase is a structured output predictive model that allows us to yield fine-grained forecasts for the next day (24 hours) at 10 minutes intervals, based on the input historical wind speed data as they are measured at 10 minutes over the past m days.

Input and output variables Formally, let k_i be the i-th farm, (X_i, Y_i) are the geographic coordinates of k_i. Let us consider the historical wind speed data, measured from k_i, over days $1, \ldots, m, m+1$. They are transformed into a training example, that is represented by vectors $\mathbf{x_i}$, $\mathbf{x_i^S}$, $\mathbf{x_i^T}$ and $\mathbf{x_i^{ST}}$, which cover the role of independent input variables, and vectors $\mathbf{y_i}$, $\mathbf{y_i^S}$, $\mathbf{y_i^T}$ and $\mathbf{y_i^{ST}}$, which cover the role of dependent output (or target) variables, respectively. We note that the input variables are calculated over days $1, \ldots, m$, while the output variables are calculated over day $m+1$. These input and output variable vectors are formally described in the following.

Vector $\mathbf{x_i}$ is defined as follows:

$$\mathbf{x_i} = (x_{i_1}, \ldots, x_{i_{144}}, x_{i_{145}}, \ldots, x_{i_{288}}, \ldots, x_{i_{144m}}), \tag{1}$$

where x_{i_t} denotes the wind speed measured from k_i at time t with $t = 1, \ldots, 144m$ (i.e. every day is divided into 144, ten minutes spaced, time points so that t denotes the time point that occurs every 10 minutes at days $1, \ldots, m$). Similarly, vector $\mathbf{y_i}$ is defined as follows:

$$\mathbf{y_i} = (y_{i_{144m+1}}, \ldots, y_{i_{144(m+1)}}), \tag{2}$$

where y_{i_t} represents the wind speed measured from k_i at time t with $t = 144m + 1, \ldots, 144(m+1)$ (every 10 minutes at day $m+1$).

By applying the standard deviation operator in combination with the neighboring and/or windowing mechanisms, we are able to define new data vectors that represent the variability of the multi-resolution wind speed data considered at spatial, temporal and spatio-temporal scales. In particular, the spatial scale is defined by the neighboring mechanism, the temporal scale is defined by the windowing mechanism, while the spatio-temporal scale is define by combining the neighboring and windowing mechanisms.

Given radius R, applying the neighboring mechanism to k_i, a circular neighborhood of k_i is constructed. This is a set of wind farms k_j so that $d(k_i, k_j) \leq R$ where $d(\cdot, \cdot)$ denotes the geographic distance. Considering the spatial scale defined by this neighboring mechanism, we define vectors $\mathbf{x_i^S}$ and $\mathbf{y_i^S}$, which represent farm k_i at the space scale with radius R over days $1 \ldots, m$ and day $m+1$, respectively. Procedurally,

$$\begin{aligned}
\mathbf{x_i^S} &= (x^S_{i_1}, \ldots, x^S_{i_{144}}, x^S_{i_{145}}, \ldots, x^S_{i_{288}}, \ldots, x^S_{i_{144m}}), \\
\mathbf{y_i^S} &= (y^S_{i_{144m+1}}, \ldots, y^S_{i_{144(m+1)}})
\end{aligned} \tag{3}$$

where:

$$x^S_{i_t} = stdev(\{x_{j_t}|d(k_i,k_j) \leq R\})\ with\ t = 1\ldots,144m,$$
$$y^S_{i_t} = stdev(\{y_{j_t}|d(k_i,k_j) \leq R\})\ with\ t = 144m+1\ldots,144(m+1). \quad (4)$$

Given length L so that L is a factor of 144, the windowing mechanism transforms the sequence of consecutive time points $t_1,\ldots,t_{144(m+1)}$ into the sequence of $\frac{144}{L}(m+1)$ consecutive time windows so that:

$$windowing[1\ldots144(m+1)] =$$
$$= \underbrace{W_1[1 \to L], W_2[L+1 \to 2L], \ldots, W_{\frac{144}{L}}[(\frac{144}{L}-1)L+1 \to 144]}_{day\ 1}, \underbrace{\ldots}_{day\ 2}, \ldots, \underbrace{\ldots}_{day\ m},$$
$$\underbrace{W_{\frac{144}{L}m+1}[t_{\frac{144}{L}m}+1 \to t_{\frac{144}{L}m+L}], \ldots, W_{\frac{144}{L}(m+1)}[t_{\frac{144}{L}m}+(\frac{144}{L}-1)L+1 \to t_{\frac{144}{L}(m+1)}]}_{day\ m+1}, \quad (5)$$

where each each window covers L consecutive time points. Considering the temporal scale defined by this windowing mechanism, we can define vectors $\mathbf{x_i^T}$ and $\mathbf{y_i^T}$, which represent farm k_i at the time scale with length L over days $1\ldots,m$ and day $m+1$, respectively. Procedurally,

$$\mathbf{x_i^T} = (x^T_{i_1},\ldots,x^T_{i_{\frac{144}{L}}},x^T_{i_{\frac{144}{L}+1}},\ldots,x^T_{i_{\frac{144}{L}2}}\ldots x^T_{i_{\frac{144}{L}(m-1)+1}},\ldots,x^T_{i_{\frac{144}{L}m}}),$$
$$\mathbf{y_i^T} = (y^T_{i_{\frac{144}{L}m+1}},\ldots,y^T_{i_{\frac{144}{L}(m+1)}}) \quad (6)$$

where:

$$x^T_{i_t} = stdev(\{x_{i_r}|r \in W_t\})\ with\ t = 1,\ldots,\frac{144}{L}m,$$
$$y^T_{i_t} = stdev(\{y_{i_r}|r \in W_t\})\ with\ t = \frac{144}{L}m+1,\ldots,\frac{144}{L}(m+1). \quad (7)$$

Finally, given radius R and length L, we define vectors $\mathbf{x_i^{ST}}$ and $\mathbf{y_i^{ST}}$, which represent farm k_i at the space scale with radius R and the time scale with length L over days $1,\ldots,m$ and day $m+1$, respectively. Procedurally,

$$\mathbf{x_i^{ST}} = (x^{ST}_{i_1},\ldots,x^{ST}_{i_{\frac{144}{L}}},x^{ST}_{i_{\frac{144}{L}+1}},\ldots,x^{ST}_{i_{\frac{144}{L}2}}\ldots x^{ST}_{i_{\frac{144}{L}(m-1)+1}},\ldots,x^{ST}_{i_{\frac{144}{L}m}})$$
$$\mathbf{y_i^{ST}} = (y^{ST}_{i_{\frac{144}{L}m+1}},\ldots,y^{ST}_{i_{\frac{144}{L}(m+1)}}) \quad (8)$$

where:

$$x^{ST}_{i_t} = stdev(\{x_{j_r}|d(k_i,k_j) \leq R\ and\ r \in W_t\})\ with\ t = 1,\ldots,\frac{L}{144}m,$$
$$y^{ST}_{i_t} = stdev(\{y_{j_r}|d(k_i,k_j) \leq R\ and\ r \in W_t\})\ with\ t = \frac{144}{L}m+1,\ldots,\frac{144}{L}(m+1). \quad (9)$$

4 Multi-Resolution Structured learning - MuReS

Let us consider a wind farm grid $\mathcal{K}$, which is composed of N wind farms $k_1, k_2, \ldots, k_N$, and a historical dataset $\mathcal{D}$, which comprises wind speed measurements collected from $\mathcal{K}$ over $m+1$ days. Adopting the notation introduced in Section 3, $\mathcal{D}$ is spanned over an independent input space $\mathbf{X} \times \mathbf{X^S} \times \mathbf{X^T} \times \mathbf{X^{TS}}$ and a dependent output space $\mathbf{Y} \times \mathbf{Y^S} \times \mathbf{Y^T} \times \mathbf{Y^{TS}}$. The structured predictive model f^{+ST} can be learned from $\mathcal{D}$ so that:

$$f^{+ST} \colon \mathbf{X} \times \mathbf{X^S} \times \mathbf{X^T} \times \mathbf{X^{TS}} \rightarrow \mathbf{Y} \times \mathbf{Y^S} \times \mathbf{Y^T} \times \mathbf{Y^{TS}}, \tag{10}$$

This predictive model is a "multi-resolution" upgrade of the traditional structured output predictive model [4, 29]. We note that the output space of $f^{+ST}(\cdot)$ yields 24 h forecasts of the fine-grained wind speed ($\mathbf{Y}$), as well as 24 h ahead forecasts of the winds speed variability at the space and time scales considered ($\mathbf{Y^S}$, $\mathbf{Y^T}$ and $\mathbf{Y^{TS}}$). However, this study aims at yielding accurate fine-grained forecasts of wind speed; hence the empirical study will explore the accuracy of model $f^{+ST}(\cdot)$ along $\mathbf{Y}$ only.

In this study, predictive model $f^{+ST}(\cdot)$ is learned as a tree, i.e. a hierarchy of clusters (Predictive Clustering Trees (PCTs)): the top node corresponds to one cluster containing all the data, which is recursively partitioned into smaller clusters while moving down the tree. CLUS, including PCTs for multi-target regression [14], is available at clus.sourceforge.net.

5 Experimental study

The experiments are carried out using real world data publicly provided by the DOE/NREL/ALLIANCE3 (http://www.nrel.gov/). The data (see Figure 1) consist of wind speed measurements from 1326 different locations at 80m of height in the Eastern region of the US. The data were collected in 10 minutes intervals during the year of 2004. This wind farm grid was able to produce 580 GW, and each farm produces between 100 MW and 600 MW. For the evaluation of the results, we consider the root mean squared error (RMSE), computed over the grid at each time point, as an indicator of the predictive performance. We derive twelve (training and testing) datasets, which are constructed as follows: for every month, days 1-11 defines a training dataset that is processed to learn the forecasting model (with $m = 10$), while days 15-25 defines the testing set used to evaluate the performance of the forecasting model learned on the corresponding training dataset. The 24 h ahead forecasting errors, averaged on the twelve datasets, are analyzed. For this empirical evaluation, the multi-resolution information is modeled at the spatial scale with radius $R = 10$ km or $R = 50$ km, as

The traditional structured output predictive model $f \colon \mathbf{X} \rightarrow \mathbf{Y}$ can be simply learned in this scenario by neglecting the information on the data variability.

The information on the wind speed variability is included in the output learning setting as a constraint to improve the predictive ability of the forecasting model learned.

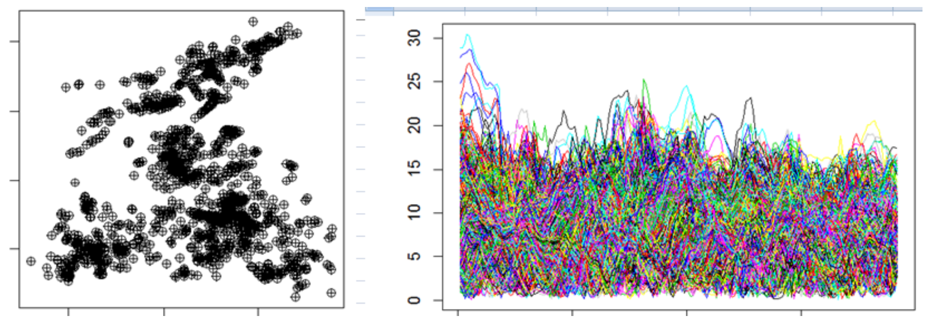

Fig. 1. Wind speed data

well as at the temporal scale with length $L = 1$ hour (6 consecutive time points) or $L = 3$ hours (18 consecutive time points). The performance of the standard deviation operator is compared to that of the sum and mean operators. Finally, the forecasting performance of the multi-resolution structured output predictive model ($f^{+ST}(\cdot)$) is compared to the performance of the baseline structured output predictive model ($f(\cdot)$).

Evaluating scale size We start by analyzing the performance of the multi-resolution structured output predictive models learned by MuReS along the size of the space (R) and time (L) scales. We run MuReS with the standard deviation operator computed over neighborhoods with radius $R = 10$km or $R = 50$km and windows with length $L = 1$h or $L = 3$h. Average RMSE results are plot in Figure 2. They show that the accuracy of the forecasting model is more sensitive to the time scale than to the spatial scale. In any case, the decrease in the forecasting accurcy performance, due to a large scale in the temporal resolution, starts being observed starting from forecasts produced 12 hours far from the current time point. Therefore, selecting the appropriate scale size is a crucial issue to yield accurate long-term forecasts. Based on these preliminary results, we select $R = 10$km and $T = 1$h for the remaining of this study.

Evaluating multi-resolution operator We proceed by exploring the performance of the multi-resolution structured output predictive models learned by MuReS along the selection of the multi-resolution operator used to model the wind speed variability. Considering $R = 10$km and $T = 1$h, we compare the forecasting accuracy achieved when the data variability model is computed through the standard deviation operator to the accuracy achieved when the variability model is computed through the sum or mean operators. Average RMSE results are plot in Figure 3. Results empirically support the effectiveness of our choice of resorting to the standard deviation as the most appropriate second order statistic to model the wind speed variability in both space and time. It actually contributes to gain in forecasting accuracy in this peculiar application.

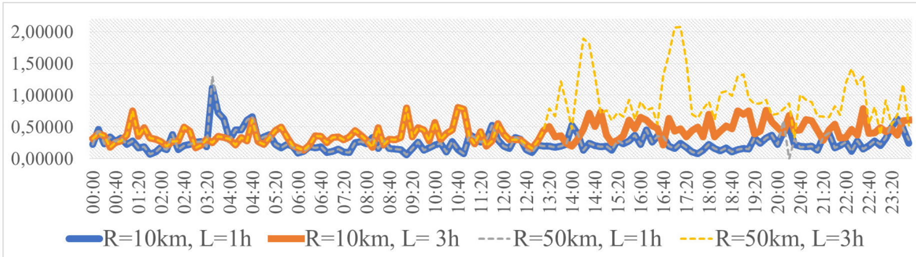

Fig. 2. Multi-resolution structured output predictive models ($R = 10, 50$km, $L = 1, 3$h): testing RMSE (averaged on twelve testing datasets - axis Y) plotted with respect to 10 minutes-spaced time points (over 24 h ahead horizon - axis X).

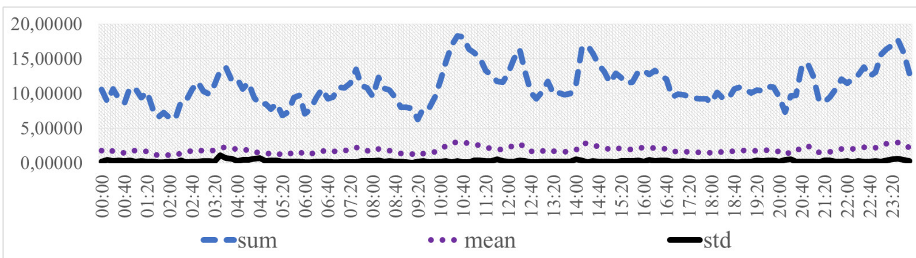

Fig. 3. Multi-resolution structured output predictive models (multi-resolution operator analysis - sum, mean and standard deviation): testing RMSE (averaged on twelve testing datasets - axis Y) plotted with respect to 10 minutes-spaced time points (over 24 h ahead horizon - axis X).

Evaluating multi-resolution learning schema We complete this study by comparing the performance of the multi-resolution structured output predictive models learned by MuReS with $R = 10$km, $L = 1$h and the standard deviation as multi-resolution operator to the performance of the baseline structured output predictive models learned neglecting data variability at space and time scales. Average RMSE results are plot in Figure 4. These results show empirically the viability of the main idea inspiring this study: the accuracy of the structured output predictive learning in the wind speed forecasting can be greatly improved by augmenting both the input and output spaces of the learning problem with multi-resolution information modeling the data variability of the wind speed observed at both space and time scales.

6 Conclusion

This paper studies the problem of the medium-term (24 h ahead) wind speed forecasting by considering different dimensions of analysis: data variability, spa-

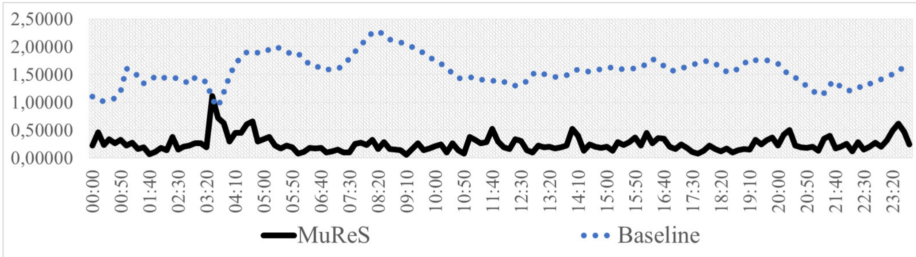

Fig. 4. Multi-resolution structured output predictive models (multi-resolution analysis - multi-resolution structured output model (MuReS) vs structured output model (Baseline): testing RMSE (averaged on twelve testing datasets - axis Y) plotted with respect to 10 minutes-spaced time points (over 24 h ahead horizon - axis X).

tial and temporal resolution, structured output learning, aiming to investigate the relevant implications for the dealing with the multi-resolution representation of the data variability for the problem at hand. Results in a benchmark dataset clearly show that accounting for the data variability at space and time scales allows us to learn output prediction models, which are much more accurate than traditional models that neglect the multi-resolution information. Moreover, experimental results confirm that the standard deviation is an appropriate multi-resolution operator of the data variability to be taken into account in this specific application. Finally, defining the size of the scale (particularly in the time scale) can be crucial issue to guarantee accurate long-term forecasts. As future work, we intend to explore the implications of these models of the data variability in statistical time series models (e.g Arima or Var). We also plan to investigate more sophisticated incremental learning methods that are able to update the forecasting model as new historical data are collected, in order to fit the learned models to drifting data.

7 Acknowledgments

Authors thank Enrico Laboragine for his support in developing the algorithm presented and running the experiments. This work was partially supported by the EU-funded project TOREADOR (ICT-16-2015, Grant Agreement A. no. 688797), as well as by the University of Bari Aldo Moro - funded ATENEO Project 2014 on "Mining of network data" and ATENEO Project 2015 on "Models and Methods to Mine Complex and Large Data".

References

1. W. Albert, S. Chi, and C. J.H. An improved grey-based approach for electricity demand forecasting. *Electric Power Systems Research*, 67:217–224, 2003.

2. V. Almeida and J. Gama. Collaborative wind power forecast. In *3rd International Conference on Adaptive and Intelligent Systems ICAIS 2014*, volume 8779 of *LNCS*, pages 162–171. Springer, 2014.
3. D. Ambach and P. Vetter. Wind speed and power forecasting - a review and incorporating asymmetric loss. In *2016 Second International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management (SMRLO)*, pages 115–123, 2016.
4. A. Appice and S. Dzeroski. Stepwise induction of multi-target model trees. In J. N. Kok, J. Koronacki, R. L. de Mántaras, S. Matwin, D. Mladenic, and A. Skowron, editors, *Machine Learning: ECML 2007, 18th European Conference on Machine Learning, Warsaw, Poland, September 17-21, 2007, Proceedings*, volume 4701 of *Lecture Notes in Computer Science*, pages 502–509. Springer, 2007.
5. A. Appice, S. Pravilovic, A. Lanza, and D. Malerba. Very short-term wind speed forecasting using spatio-temporal lazy learning. In N. Japkowicz and S. Matwin, editors, *Discovery Science - 18th International Conference, DS 2015, Banff, AB, Canada, October 4-6, 2015, Proceedings*, volume 9356 of *Lecture Notes in Computer Science*, pages 9–16. Springer, 2015.
6. M. Bhaskar, A. Jain, and N. V. Srinath. Wind speed forecasting: Present status. In *2010 International Conference on Power System Technology*, pages 1–6, 2010.
7. M. Ceci, R. Corizzo, F. Fumarola, D. Malerba, and A. Rashkovska. Predictive modeling of PV energy production: How to set up the learning task for a better prediction? *IEEE Trans. Industrial Informatics*, 13(3):956–966, 2017.
8. S. Fan, J. Liao, R. Yokoyama, L. Chen, and W.-J. Lee. Forecasting the wind generation using a two-stage network based on meteorological information. *IEEE Transactions on Energy Conversion*, 24(2):474–482, 2009.
9. U. B. Filik and T. Filik. Wind speed prediction using artificial neural networks based on multiple local measurements in eskisehir. *Energy Procedia*, 107:264 – 269, 2017. 3rd International Conference on Energy and Environment Research, ICEER 2016, 7-11 September 2016, Barcelona, Spain.
10. G. Giebel, J. Badger, I. M. Perez, P. Louka, and G. Kallos. Short-term forecasting using advanced physical modelling-the results of the anemos project. results from mesoscale, microscale and cfd modelling. In *Proc. of the European Wind Energy Conference 2006*, 2006.
11. G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl. *The State-Of-The-Art in Short-Term Prediction of Wind Power: A Literature Overview, 2nd edition*. ANEMOS.plus, 2011. Project funded by the European Commission under the 6th Framework Program, Priority 6.1: Sustainable Energy Systems.
12. Z. Hui, L. Bin, and Z. Zhuo-qun. Short-term wind speed forecasting simulation research based on arima-lssvm combination method. In *ICMREE) 2011*, volume 1, pages 583–586, 2011.
13. R. G. Kavasseri and K. Seetharaman. Day-ahead wind speed forecasting using f-arima models. *Renewable Energy*, 34(5):1388 – 1393, 2009.
14. D. Kocev, C. Vens, J. Struyf, and S. Dzeroski. Tree ensembles for predicting structured outputs. *Pattern Recognition*, 46(3):817–833, 2013.
15. M. Lange and U. Focken. New developments in wind energy forecasting. In *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*, pages 1–8, 2008.
16. M. Negnevitsky, P. Johnson, and S. Santoso. Short term wind power forecasting using hybrid intelligent systems. In *Power Engineering Society General Meeting, 2007. IEEE*, pages 1–4, 2007.

17. M. Negnevitsky, P. Mandal, and A. K. Srivastava. An overview of forecasting problems and techniques in power systems. In *2009 IEEE Power Energy Society General Meeting*, pages 1–4, 2009.
18. O. Ohashi and L. Torgo. Wind speed forecasting using spatio-temporal indicators. In *ECAI 2012 - PAIS-2012) System Demonstrations Track*, volume 242 of *Frontiers in Artificial Intelligence and Applications*, pages 975–980. IOS Press, 2012.
19. J. Palomares-Salas, J. De la Rosa, J. Ramiro, J. Melgar, A. Aguera, and A. Moreno. Arima vs. neural networks for wind speed forecasting. In *IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, CIMSA 2009*, pages 129–133, 2009.
20. C. Potter and M. Negnevitsky. Very short-term wind forecasting for tasmanian power generation. *IEEE Transactions on Power Systems*, 21(2):965–972, 2006.
21. S. Pravilovic, A. Appice, A. Lanza, and D. Malerba. Mining cluster-based models of time series for wind power prediction. In S. Greco and A. Picariello, editors, *22nd Italian Symposium on Advanced Database Systems, SEBD 2014, Sorrento Coast, Italy, June 16-18, 2014.*, pages 9–20, 2014.
22. S. Pravilovic, A. Appice, A. Lanza, and D. Malerba. Wind power forecasting using time series cluster analysis. In *17th International Conference on Discovery Science, DS 2014*, volume 8777, pages 276–287. Springer, 2014.
23. S. Pravilovic, A. Appice, and D. Malerba. An intelligent technique for forecasting spatially correlated time series. In *13th International Conference of the Italian Association for Artificial Intelligence, AI*IA 2013*, volume 8249 of *LNCS*, pages 457–468. Springer, 2013.
24. S. Pravilovic, A. Appice, and D. Malerba. Integrating cluster analysis to the ARIMA model for forecasting geosensor data. In *21st International Symposium on Foundations of Intelligent Systems, ISMIS 2014*, volume 8502 of *LNCS*, pages 234–243. Springer, 2014.
25. S. Pravilovic, M. Bilancia, A. Appice, and D. Malerba. Using multiple time series analysis for geosensor data forecasting. *Inf. Sci.*, 380:31–52, 2017.
26. J. Shi, X. Qu, and S. Zeng. Short-term wind power generation forecasting: Direct versus indirect arima-based approaches. *International Journal of Green Energy*, 8(1):100–112, 2011.
27. G. Sideratos and N. Hatziargyriou. An advanced statistical method for wind power forecasting. *Power Systems, IEEE Transactions on*, 22(1):258–265, 2007.
28. S. S. Soman, H. Zareipour, O. Malik, and P. Mandal. A review of wind power and wind speed forecasting methods with different time horizons. In *North American Power Symposium 2010*, pages 1–8, 2010.
29. D. Stojanova, M. Ceci, A. Appice, D. Malerba, and S. Dzeroski. Dealing with spatial autocorrelation when learning predictive clustering trees. *Ecological Informatics*, 13:22–39, 2013.
30. X. Wang, P. Guo, and X. Huang. A review of wind power forecasting models. *Energy Procedia*, 12(0):770 – 778, 2011.
31. J. Zhou, J. Shi, and G. Li. Fine tuning support vector machines for short-term wind speed forecasting. *Energy Conversion and Management*, 52(4):1990 – 1998, 2011.