ECAI 2012
Luc De Raedt et al. (Eds.)

doi:10.3233/978-1-61499-098-7-975

Wind speed forecasting using spatio-temporal indicators

Orlando Ohashi [1,2] and **Luís Torgo** [2]

Abstract.
From small farms to electricity markets the interest and importance of wind power production is continuously increasing. This interest is mainly caused by the fact that wind is a continuous resource of clean energy. To take full advantage of the potential of wind power production it is crucial to have tools that accurately forecast the expected wind speed. However, forecasting the wind speed is not a trivial task. Wind speed is characterised by a random behaviour as well as several other intermittent characteristics. This paper proposes a new approach to the task of wind speed forecasting. The main distinguishing feature of this proposal is its reliance on both temporal and spatial characteristics to produce a forecast of the future wind speed. We have experimentally tested the proposed method with historical data concerning wind speed on the eastern region of the US. Nevertheless, the methodology that is described in the paper can be seen as a general approach to spatio-temporal prediction. We have compared our proposal to other standard approaches in the task of forecasting 2 hours ahead wind speed. Our extensive experiments show that our proposal has clear advantages in most setups.

1 Introduction

The importance of wind power production is continuously increasing, as countries are looking for more sustainable alternatives for their power grid. Wind power generation is an excellent option given that it is a continuous resource of clean energy. The main drawback of this technology is the large variability in production, which makes almost impossible to rely only in the wind energy. Generally, wind energy is used in conjunction with other types of technologies, like thermal, hydraulic, natural-gas, and so on. Wind power generation is also crucial in small remote autonomous locations, where it can be used as a fuel saver to reduce the operational costs. Some countries like the US [10], China [24] and UK [2] have electricity markets, which work similarly to an auction. Market participants rely on the expected future power production and on the market price to decide their bidding strategy. These expectations are usually considered for a short period, from a couple of hours to a day ahead. All these factors contribute to the crucial importance of having accurate prediction models of future power production. For wind energy this is even more relevant given its dependency on other sources of energy if wind speed is low. Having an accurate forecast of the wind speed in the next hours is of key importance to estimate wind power production and define the best bidding strategy that maximizes the profit and avoids the penalties from missing delivering energy.

According to Alexiadis et al. [1] wind power production is a function of the wind speed. This means that the accurate forecast of wind speed allows a better estimate of future wind power production. The wind is considered one of the most difficult meteorological parameters to forecast [20]. The wind speed behavior is influenced by several factors like: the topographical properties of the land, the rotation of the earth, temperature, pressure, obstacles, the height of the anemometer, etc. [12, 20]. Lei et al. [13] classify wind speed prediction models in four classes: physical models, conventional statistical models, spatial correlation models and artificial intelligence models. The physical models consider only characteristics like: terrain, obstacles, pressure and temperature to estimate the future wind speed. They generally have poor results in short term prediction. Conventional statistical models are based on time series techniques (ARMA, ARIMA, etc.) to forecast the future wind speed. Spatial models use the neighbourhood information as predictors of the wind speed, usually applied to locations where the wind speed measurement is not available. Artificial intelligent models use historical data to obtain machine learning models that can be used to forecast the future wind speed. The method proposed in this paper is an artificial intelligence approach that incorporates spatio-temporal predictors to forecast the future wind speed on any location.

The development of prediction tools for wind speed forecasting is not a new subject, and there is a considerable number of important contributions on this research field. Kavasseri and Seetharaman [11] use time series models to forecast the hourly average wind speed for up to 2 days ahead in North Dakota, US. Kusiak et al. [12] applied several machine learning models to forecast the next wind speed data, using the historical information of each site. Mohandes et al. [17] compare support vector machines against neural networks in the task of forecasting the average daily wind speed in Madina city, Saudi Arabia. In this study, support vector machines outperformed the neural network models. Sfetsos [20] compares machine learning models against time series models for forecasting the average hourly speed value in Greece. The study provided evidence in favour of the machine learning models. Damousis et al. [8] proposed a fuzzy model adjusted by a genetic algorithm for the prediction of wind speed 2 hours ahead, in Greece. Zhao et al. [24] proposed a hybrid approach including numerical weather prediction with a neural network and kalman filter to forecast the next day ahead wind power, in China. Li and Shi [14] compared three neural networks in the task of forecasting the next hour wind speed in North Dakota, US. Bilgili et al. [3] use a neural network model to forecast the mean monthly wind speed in Turkey. Among the inputs used in this work are the mean monthly value of neighbouring sites. Alexiadis et al. [1] proposed to forecast the speed of the next site in the wind direction based on the wind

[1] The work of Orlando Ohashi is supported by the PhD grant SFRH/BD/61795/2009 from FCT.
[2] LIAAD - INESC TEC / DCC - Faculdade de Ciências - Universidade do Porto - e-mails: {ohashijr, ltorgo}@dcc.fc.up.pt

speed in the previous site, using a neural network model. According to the authors, the main motivation behind the proposed technique was that the wind travels from one location to another and suffers disturbances during the propagation.

The most frequent approach used by machine learning models to predict the expected wind speed considers as predictors the previously observed values of this wind speed [12, 17, 20, 24]. Similar approaches are adopted by time series models [11]. All these approaches assume that the future wind speed depends on the recently observed wind speed on the same location. Given the fact that wind travels through the landscape this might be limiting for the models as they are being feed only with values from the same location for which a future prediction is required. These models ignore the spatial dependency that exists on this domain, where the wind speed at a certain location is clearly correlated with the wind speed at neighbouring locations. There are some attempts to use the spatial information of the domain. In the work of Bilgili et al. [3], they propose to use the monthly average wind speed at 4 neighbouring locations as inputs for a neural network model to forecast the monthly average at the target location. The work of Alexiadis et al. [1] tries to identify the temporal relationship of the wind speed between spatial locations. They try to identify a pattern of the wind speed measured in two different locations, based on the travel time of the wind from one location to the other. The authors use this relationship to forecast the wind speed in a sub-sequent location. The main drawback of this approach is that it limits the neighbors used in the analysis and requires the information of the wind direction between the locations. In situations where this information is not available or is unreliable we can not use this technique.

The main motivation for the approach we propose in this paper is the assumption that the future wind speed at any location depends not only on the recent wind speed on the same location, but also on this speed on neighbouring locations. In this context, our proposal may be succinctly described as an attempt to try to convey information on both the spatial and temporal historical wind speed values to the models, with the goal of improving their forecasting ability. Still, the approach is by no means dependent on this particular application and can actually be seen as a general approach to the problem of formalising prediction tasks in the context of spatio-temporal data. We have tested and compared our proposal against a series of alternative formalisations on a particular prediction task. As experimental benchmark we have used real world data collected in wind farms in the US (see Section 3.1). Our experiments show that our approach outperforms the standard formalisation that includes no information from the spatial neighbours, in the task of forecasting the wind speed 2 hours ahead. This result carries out for the majority of the machine learning models used in this experiment. We have also compared the machine learning models against two baseline models - a time series model (ARIMA) and a random walk approach.

In Section 2 we describe our proposed formalisation of the prediction problem that includes the definition of spatio-temporal indicators. Section 3 describes the experiments included in the paper, namely the data and the experimental methodology that were used. In Section 4 we present and discuss the results of our experiments, while on Section 5 we draw the conclusions of the work and describe our future research agenda.

2 Spatio-Temporal Indicators

The task being addressed in this paper consists on trying to forecast the future value of a time series variable on a certain geographical location, based on historical data of this variable collected on both this and other locations. The most common approach to time series forecasting using machine learning models consists in transforming the original problem into a multiple regression task, where the target variable is the future value of the series, while the predictors are previous past values of the series up to a certain p-length time window. This transformation technique is usually known as time delay embedding [21]. The idea is to provide the modelling techniques with information on the recent dynamics of the time series by means of the most recent values. An improvement over this simple strategy is frequently used within financial forecasting. In this field it is frequent to also use as predictors what are known as technical indicators. These variables are nothing more than summaries of certain properties of the time series. These properties include effects like tendency, acceleration, momentum and so on. Different indicators were developed to express these features of a time series. These indicators can be regarded as "sophisticated" descriptors of the recent dynamics of the time series we want to forecast.

In our approach to wind speed forecasting we started with the assumption that the future values of the wind speed depend not only on the recent past values at the same location but also on nearby locations. This spatio-temporal dependency is not particular to wind speed forecasting. Several real world domains have similar forecasting problems with the same type of spatio-temporal data. In effect, with the profusion of mobile computing devices with GPS capabilities, the demand for the analysis of spatio-temporal data is increasing at a very high rate. The key idea behind our proposal is to try to develop predictors that are able to capture the spatio-temporal dynamics of the time series we aim to forecast. More precisely, we plan on mapping the concept of technical indicators used in financial forecasting to a spatio-temporal context. With this purpose we derive a series of spatio-temporal indicators that can be used as predictors in the task of developing forecasting models. Our assumption is that these extra predictors will provide the model with important information on the recent spatio-temporal dynamics of the time series, which in turn will improve the model prediction accuracy. In this context, we plan to formalise the prediction problem in such a way that the future values of the target time series are forecasted using not only previous values of the series and summaries of its temporal dynamics, but also with spatio-temporal indicators that summarise the dynamics of the series within the neighbourhood.

The first question we need to address is how to describe the behaviour of the time series within the neighborhood of the target location. Our proposal is based on the notion of spatio-temporal neighbourhood. In this context, we need to define a function to calculate the distance between any two points in the space-time dimension.

In this work a point in space-time is the value of a variable (in our application the wind speed) at a time t in a geographical location x, y, which we will denote as $w_{x,y}^{t}$. Let i and j be two points in space-time (i.e. two measurements $w_{x_i,y_i}^{t_i}$ and $w_{x_j,y_j}^{t_j}$). We define the spatio-temporal distance between i and j in a similar way to Mingyao et al. [16], namely,

$$D_{i,j} = d_{i,j} \times \alpha + t_{i,j} \times (1 - \alpha) \tag{1}$$

where $d_{i,j}$ is the spatial distance between the locations of the objects (x_i, y_i and x_j, y_j), $t_{i,j}$ is the time distance between the objects (t_i and t_j), and α is weighing factor between time and geographical distances that are assumed to be normalised. The spatial distance can be calculated using a standard metric, like for instance the Euclidean distance, or more sophisticated versions for geographical data like the great-circle distance [6]. In our experiments we have use this latter alternative given that our data is geographically indexed. The time distance is simply the absolute difference between the two time tags in some adequate time unit (e.g. hours).

Having defined the spatio-temporal distance between two objects we can define the spatio-temporal neighbourhood of a point o as the set of points within a certain spatio-temporal distance,

$$\mathcal{N}_o^\beta = \{k \in \mathcal{D} : D_{o,k} < \beta\} \tag{2}$$

where $\mathcal{D}$ is the available spatio-temporal data set.

Given the above definitions we can look at the spatio-temporal neighbourhood of a point as a kind of cone within space-time. Different settings for α and β lead to cones of difference sizes as shown in Figure 1.

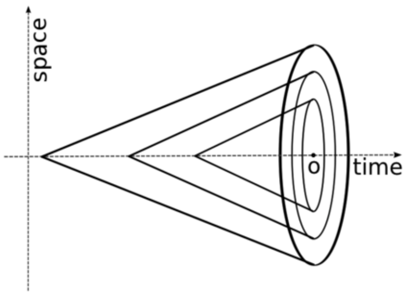

Figure 1. Defining spatio-temporal neighbourhoods with different sizes.

Each cone defines a neighbourhood around a central location. These cones represent which past values may influence the future value of the time series at that location. The cones can be regarded as the spatio-temporal equivalents of the idea of time-delay embedding. Increasing the size of the cone will increase the spatio-temporal embed size.

As we have mentioned before in finance it is common to summarise the dynamics of a time series by means of technical indicators, which may reflect different properties. For instance, the ratio between two moving averages calculated using two different embed sizes provides indications on the tendency of the series. If the value of the moving average with shorter embed surpasses the longer moving average we know that the time series is on an upwards tendency, while the opposite indicates a downwards direction. We have imported this idea into the spatio-temporal dimension. The ratio between two spatio-temporal averages provides us with information on how the time series values evolve in the space-time dimension. This ratio can be defined as follows,

$$\overline{W}_o^{\beta_1,\beta_2} = \frac{\overline{w}(\mathcal{N}_o^{\beta_1})}{\overline{w}(\mathcal{N}_o^{\beta_2})} \tag{3}$$

where β_1 and β_2 are two neighbourhood sizes and $\overline{w}()$ is the average of the target time series values for a set of points in the neighbourhood of o.

A variation of this indicator can be easily obtained by using weighted averages of the values within the spatio-temporal neighbourhood. If we set the weights to the inverse of the spatio-temporal distance to the point o we have the effect that "closer" (in spatio-temporal terms) points are given more importance within the averages,

$$\widetilde{W}_o^{\beta_1,\beta_2} = \frac{\widetilde{w}(\mathcal{N}_o^{\beta_1})}{\widetilde{w}(\mathcal{N}_o^{\beta_2})} \tag{4}$$

where $\tilde{w}()$ is the weighed average of target time series for a set of points in the neighbourhood of o.

The spatio-temporal averages themselves can be seen as interesting indicators that provide information on the typical value of the time series within a certain spatio-temporal vicinity. Similarly, spatio-temporal standard deviations can be calculated to provide information on the dispersion of values within the neighbourhood of o. All these indicators can be easily calculated with their standard formulae applied to the cases inside the spatio-temporal neighbourhood of o.

Having defined a series of spatio-temporal indicators, our hypothesis is that they provide useful information for the target prediction task. In this context, given the goal of forecasting the value of the target time series for k time steps ahead at location o, we propose to tackle this problem using the following formalisation,

$$\begin{aligned} W_o^{t+k} = f(&W_o^t, W_o^{t-1}, \cdots, W_o^{t-m}, \\ &\overline{w}(\mathcal{N}_o^{k_1}), \overline{w}(\mathcal{N}_o^{k_2}), \overline{w}(\mathcal{N}_o^{k_3}), \overline{W}_o^{k_1,k_2}, \overline{W}_o^{k_2,k_3}, \\ &\widetilde{w}(\mathcal{N}_o^{k_1}), \widetilde{w}(\mathcal{N}_o^{k_2}), \widetilde{w}(\mathcal{N}_o^{k_3}), \widetilde{W}_o^{k_1,k_2}, \widetilde{W}_o^{k_2,k_3}, \\ &\sigma_w(\mathcal{N}_o^{k_1}), \sigma_w(\mathcal{N}_o^{k_2}), \sigma_w(\mathcal{N}_o^{k_3})) \end{aligned} \tag{5}$$

where $f()$ is the unknown regression function we are trying to model using a set of training data $\mathcal{D}$, m is the size of a temporal embed, k_1, k_2 and k_3 (with $k_1 < k_2 < k_3$) are spatio-temporal neighbourhood sizes, and $\sigma_w()$ is the standard deviation of the target time series calculated with the set of points in a neighbourhood of o.

We should note that this is simply one among many possible setups including spatio-temporal indicators as predictors. The decision of using 3 spatio-temporal neighbourhood sizes was arbitrary and other setups could make more sense depending on the application. Still, this was the setup used in our experiments with wind speed forecasting.

3 Experimental Evaluation

The main goal of our experiments is to test the hypothesis that motivates our work: using information on the wind speed of nearby locations in recent time will improve the predictive accuracy of our models when forecasting the future wind speed at a certain location. With the goal of collecting experimental evidence towards this hypothesis we have designed an experiment where we have compared different models that tackle this prediction task using different predictors. Namely, we have compared our approach that includes spatio-temporal indicators as shown in Equation 5, with other approaches where the predictors do not include data from this spatio-temporal vicinity. In order to exclude eventual dependencies of the outcome of the experiments on the used modelling tools, we have repeated the comparisons using several learning algorithms with different parameter settings.

3.1 Data Description

In this paper all the experiments were carried out using real world data publicly provided by the DOE/NREL/ALLIANCE[3]. The data

[3] http://www.nrel.gov/

consist in 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 is able to produce 580 GW, and each site produces between 100 MW and 600 MW. For our experiments we have selected two locations as our targets in terms of forecasting the future wind speed. This selection was guided by the availability of a larger number of neighbouring sites at these places. Figure 2 shows the geographical location of the data collection sites.

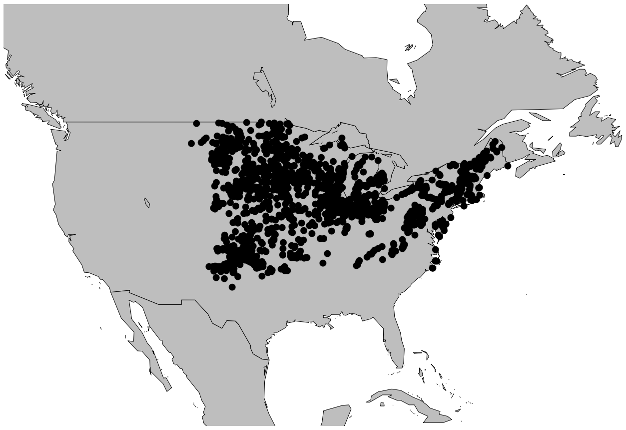

Figure 2. Wind Farm at Eastern US.

3.2 Used Machine Learning Models

We have tried to select a wide range of modeling approaches to test our hypothesis. The idea is to confirm its validity independently of the technique used to forecast. All used tools are freely available in the R software environment [19], which ensures easy replication of our work. The following is a list of the methods used in our experiments as well as the considered parameter variants:

Random Walk - a simple baseline method that uses the last wind speed measurement as prediction for the 2 hours ahead wind speed;

Arima - a time series Box-Jenkins model [18] based on the R package `forecast` [9]. The function `auto.arima` automatically selects the best parameters for the algorithm;

Regression Trees (RT) - a regression tree (e.g. [4]) based on the R package `rpart` [22]. In our experiments we have used an interface to the `rpart` function provided in package `DMwR` [23] and have tried 4 different variants by using the parameter `se` that controls the level of pruning with values: 0, 0.5, 1 and 1.5.

Support Vector Machines (SVM) - an implementation of SVMs (e.g. [7]) available in the R package `e1071`. Six variants were tried by using the parameter *cost* with the values 10 and 100, and the parameter *epsilon* with the values 0.1, 0.3 and 0.5.

Random Forest (RF) - an implementation of random forests [5] available in the R package `randomForest` [15]. We have used 3 variants of the parameter *ntree* with the values 500, 1000 and 1500.

3.3 Experimental Methodology

Each model variant that we have considered in our experiments (a combination of a learning algorithm plus parameter settings), was applied to 6 different prediction tasks. These tasks have exactly the same target variable (the wind speed at time $t + 2h$), but differ in the way they use the available past data to obtain the predictors used to forecast the target variable. One of these 6 tasks only uses data from the same spatial location, i.e. it only uses information from the past values of the wind speed measured on the site for which we want a forecast. The other 5 variants use the formalisation we have proposed in Equation 5, with different configurations of the 3 neighbourhoods. As we have seen these neighbourhoods are cones defined by Equation 2. The problem formalisation proposed in Equation 5 uses three of these cones. An alternative way of defining a cone is by its maximum radius and its height from the base. This equivalent specification of the neighbourhood is more intuitive in our application. For instance, the cone with maximum radius of 10km and height of 10 days, defines a neighbourhood that for the current time uses points that are at most 10km away from the target location, and goes back in time at most 10 days. Using this alternative specification of neighbourhoods we can describe the remaining five variants of the problem specification as follows: i) [50km, 10 days], [100km, 20 days] and [150km, 30 days]; ii) [140km, 10 days], [350km, 20 days] and [730km, 30 days]; iii) [75km, 10 days], [150km, 20 days] and [300km, 30 days]; iv) [100km, 10 days], [500km, 20 days] and [900km, 30 days]; and v) [150km, 10 days], [675km, 20 days] and [1200km, 30 days]. Regards the first variant using only data from the same location we have used exactly the same predictors as in Equation 5. However, all indicators are calculated using only the wind speed values of the same location, i.e. the spatial neighbours are ignored. It is like we were using a cylinder of spatial radius near zero, instead of the cones.

The predictions of the different trials were evaluated using the mean absolute error (MAE),

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |\hat{y}_i - y_i| \tag{6}$$

where $\hat{y_i}$ is the predicted wind speed value for a true value of y_i.

With the goal of obtaining statistically reliable estimates of this error measure we have used a Monte Carlo simulation. The simulation was designed to provide estimates of the MAE at predicting the wind speed for two hours ahead of the different alternatives considered in our experiments. To increase the statistical reliability of the experiments we have repeated the process 10 times at randomly selected time points within the available data interval (10 minutes measurements throughout all 2004). For each of these 10 randomly selected time points, and for each of the two sites, the alternatives were evaluated by means of their predictions during the next day (144 predictions given that the periodicity of the data is 10 minutes and the test window is a full day). The predictions for the next day were obtained using a sliding window approach. Each model obtained with this approach was learnt using data from the same past window, although using it to build different predictors as we have seen. For instance, at time t and site A we use the available training data to obtain a model that is used to forecast the wind speed at time $t + 2h$. After this prediction is obtained, the training window is slided one time step (i.e. 10mins) and another model is obtained to forecast the value of wind speed at time $t + 2h + 10mins$. This sliding window process is repeated until we have predictions for all time points in the next day. All model variants are evaluated using the same data.

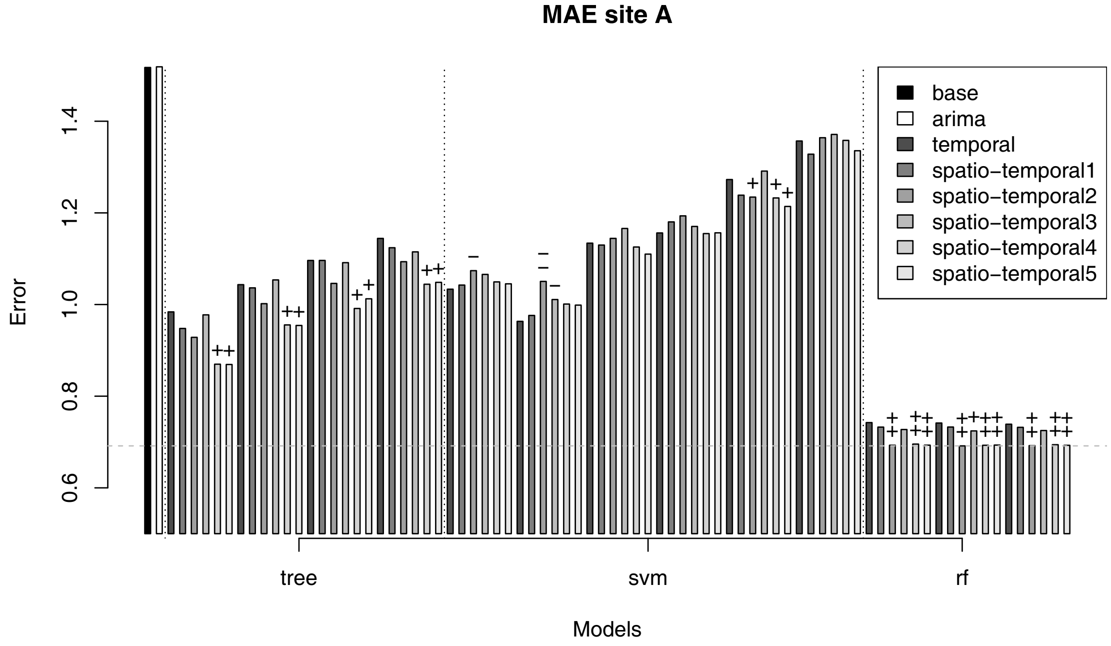

Figure 3. Results for site A.

4 Experimental Results and Discussion

Figures 3 and 4 summarise the results of all experiments. They present the Monte Carlo estimates of the MAE of all considered variants for the sites A and B, respectively. Each bar is the MAE estimate of a variant. There are four groups of model variants. The first group includes the baseline approaches: the random walk and the arima model. Then we have all variants of the regression trees, SVMs and random forests. For each of the parameter settings we have considered (c.f. Section 3.2) we show 6 bars, corresponding to each of the 6 alternative problem formulations we have described in Section 3.3. Recall that the main goal of our experiments is to compare the use of the spatio-temporal indicators as predictors against the use of indicators built with data from the same location only. This means we want to compare the 5 last bars of each variant against the first bar (darkest bar of the six). On top of the last five bars we may have one or two symbols ($+$ or $-$). They represent the statistical significance of the difference in performance against the first bar according to a paired t-test. A single $+$ ($-$) means that the respective bar is better (worse) than the first bar with 95% confidence. Two symbols increase the confidence to 99%.

In general, with the exception of some SVM variants, we can say that these experiments confirm our hypothesis that the use of predictors based on data from a spatio-temporal neighbourhood is advantageous in terms of predictive performance. Moreover, for the best models in the set we have considered (Random Forests), this advantage is even more marked. As shown in the graphs the best overall predictive performance is always obtained by some random forest variant using our spatio-temporal indicators. Regression trees have achieved a performance surprisingly competitive with SVMs, and they have also taken advantage of the use of our indicators. The results with SVMs are a bit contradictory and their generally poor performance may provide indications that further parameter tuning may be required for improving their performance.

Table 1 summarises the results on the number of significant differences between the spatio-temporal neighbourhood variants and the strategy of using only the temporal information. Each row shows the number of significant wins ($+$'s) and losses ($-$'s) of the spatio-temporal variants for the different experimental configurations.

	Site A												Site B											
	tree				svm				rf				tree				svm				rf			
	++	+	−	−−	++	+	−	−−	++	+	−	−−	++	+	−	−−	++	+	−	−−	++	+	−	−−
sp1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	1	3	0	0	0
sp2	0	0	0	0	1	0	1	1	0	3	0	0	3	0	0	0	3	1	1	0	0	3	0	0
sp3	0	0	0	0	0	0	1	0	1	0	0	0	4	0	0	0	0	0	3	1	0	3	0	0
sp4	4	0	0	0	1	0	0	0	0	3	0	0	4	0	0	0	1	0	2	0	0	3	0	0
sp5	4	0	0	0	1	0	0	0	0	3	0	0	0	4	0	0	1	0	1	1	0	3	0	0
Tot.	8	0	0	0	3	0	2	1	1	9	0	0	11	4	0	0	7	1	9	3	3	12	0	0

Table 1. Number of significant wins and losses.

5 Conclusion and Future Work

This paper has described a new methodology for short-term wind speed prediction, a class of problems with extreme relevance for electricity markets and wind power production. We proposed a new formalisation of this spatio-temporal prediction problem, which includes the definition of spatio-temporal indicators. These predictors provide information on the spatio-temporal dynamics of the target time series. Our proposal is general and can be applied to any spatio-temporal prediction task. These type of prediction problems are becoming more and more relevant with the prevalence of mobile computing devices with localisation features. In this paper we have tested our proposal on the task of forecasting the wind speed for a two hours

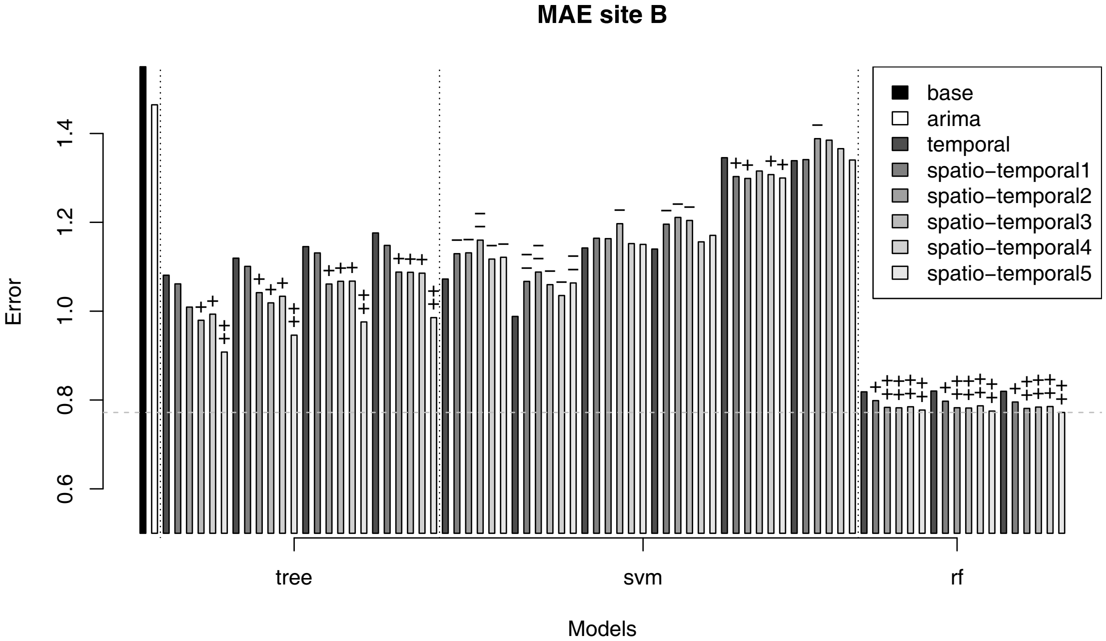

Figure 4. Results for site B.

ahead horizon in the eastern region of the US. Our experimental results confirm the advantages of the use of spatio-temporal information on this prediction task. Models using our spatio-temporal indicators have generally obtained superior performance.

In the future we plan to extend our tests to other applications and also to explore the reasons for the sub-optimal performance of SVMs on the prediction task we have considered. Finally, we plan to study the use of other alternative spatio-temporal indicators to further improve the performance of the models.

REFERENCES

[1] MC Alexiadis, PS Dokopoulos, and HS Sahsamanoglou, 'Wind speed and power forecasting based on spatial correlation models', *Energy Conversion, IEEE Transactions on*, **14**(3), 836–842, (1999).

[2] RJ Barthelmie, F. Murray, and SC Pryor, 'The economic benefit of short-term forecasting for wind energy in the uk electricity market', *Energy Policy*, **36**(5), 1687–1696, (2008).

[3] M. Bilgili, B. Sahin, and A. Yasar, 'Application of artificial neural networks for the wind speed prediction of target station using reference stations data', *Renewable Energy*, **32**(14), 2350–2360, (2007).

[4] L. Breiman, *Classification and regression trees*, Chapman & Hall/CRC, 1984.

[5] L. Breiman, 'Random forests', *Machine learning*, **45**(1), 5–32, (2001).

[6] M.R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, volume 319, Springer Verlag, 1999.

[7] N. Cristianini and J. Shawe-Taylor, *An introduction to support Vector Machines: and other kernel-based learning methods*, Cambridge Univ Pr, 2000.

[8] I.G. Damousis, M.C. Alexiadis, J.B. Theocharis, and P.S. Dokopoulos, 'A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation', *Energy Conversion, IEEE Transactions on*, **19**(2), 352–361, (2004).

[9] Rob J Hyndman, *forecast: Forecasting functions for time series*, 2011. R package version 3.11.

[10] P. Joskow and E. Kahn, 'A quantitative analysis of pricing behavior in california's wholesale electricity market during summer 2000', Technical report, National Bureau of Economic Research, (2001).

[11] R.G. Kavasseri and K. Seetharaman, 'Day-ahead wind speed forecasting using f-arima models', *Renewable Energy*, **34**(5), 1388–1393, (2009).

[12] A. Kusiak, H. Zheng, and Z. Song, 'Short-term prediction of wind farm power: A data mining approach', *Energy Conversion, IEEE Transactions on*, **24**(1), 125–136, (2009).

[13] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan, 'A review on the forecasting of wind speed and generated power', *Renewable and Sustainable Energy Reviews*, **13**(4), 915–920, (2009).

[14] G. Li and J. Shi, 'On comparing three artificial neural networks for wind speed forecasting', *Applied Energy*, **87**(7), 2313–2320, (2010).

[15] A. Liaw and M. Wiener, 'Classification and regression by randomforest', *R News*, **2**(3), 18–22, (2002).

[16] Q. Ming-yao, M. Li-xin, and S. Jie, 'A spatio-temporal distance based two-phase heuristic algorithm for vehicle routing problem', in *Fifth International Conference on Natural Computation, ICNC'09*, pp. 352–357. IEEE, (2009).

[17] MA Mohandes, TO Halawani, S. Rehman, and A.A. Hussain, 'Support vector machines for wind speed prediction', *Renewable Energy*, **29**(6), 939–947, (2004).

[18] A. Pankratz, *Forecasting with univariate Box-Jenkins models*, volume 3, Wiley Online Library, 1983.

[19] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, 2010.

[20] A. Sfetsos, 'A comparison of various forecasting techniques applied to mean hourly wind speed time series', *Renewable Energy*, **21**(1), 23–35, (2000).

[21] F. Takens, 'Detecting strange attractors in turbulence', *Dynamical systems and turbulence Warwick 1980*, **898**(1), 366–381, (1981).

[22] T. M. Therneau and B. Atkinson. R port by B. Ripley., *rpart: Recursive Partitioning*, 2009. R package version 3.1-44.

[23] L. Torgo, *Data Mining with R, learning with case studies*, CRC Press, 2010.

[24] P. Zhao, J. Wang, J. Xia, Y. Dai, Y. Sheng, and J. Yue, 'Performance evaluation and accuracy enhancement of a day-ahead wind power forecasting system in china', *Renewable Energy*, (2011).