VECTOR AUTO REGRESSIVE MODELS WITH EXOGENOUS VARIABLES FOR PREDICTING WIND SPEED

Ergin Erdem¹, Ying She², Jing Shi¹, Steve Hsueh Ming Wang³
North Dakota State University¹, Nanchang Hangkong University², University of Alaska Anchorage³
1410 14th Avenue N., Fargo, ND 58108, USA¹
Nanchang, Jiangxi Province 330063, China²
3211 Providence Dr., Anchorage, AK 99508, USA³
ergin.erdem@my.ndsu.edu, sheying1980@hotmail.com, jing.shi@ndsu.edu, hswang@alaska.edu

ABSTRACT
In this study, we develop a vector auto regressive process with exogenous variables for forecasting the wind speed for very short term (i.e., 10 minutes ahead) and short term (i.e., 1 hour ahead) horizons. Meteorological variables such as air pressure, air temperature, wind shear, and relative humidity are incorporated into the vector auto regressive model. The classification of the wind direction and wind speed pairs into the corresponding groups are conducted by the $k$-means clustering approach and incorporated into the model as exogenous variables represented by categorical variables. Vector auto regressive models with exogenous variables are compared against the traditional auto regressive moving average based approach that is commonly used for forecasting the wind speed using the selected performance measures. The analysis reveals that the vector based auto regressive approach outperforms the traditional auto regressive moving average for both forecasting horizons of 1-hour and 10-minutes. It also indicates that on the average, the vector autoregressive approach performs 4.415% and 9.98% better in terms of the mean absolute and mean square error for 1-hour ahead forecasting. For very short term forecasting (10-min), the vector autoregressive models with exogenous variables still perform constantly better as compared to the traditional auto regressive moving average models albeit with smaller difference in terms of performance measures. The analysis reveals that for 10-min ahead forecasting, the relative differences in terms of mean absolute and mean square errors between the two methods are 4.415% and 1.385% respectively. In conclusion, it can be seen that incorporating the meteorological variables into model and developing classification scheme for the contribution of the wind speed to the wind direction helps improving the forecasts for wind speed especially for very short and short term periods. Further research is called for developing new approaches for incorporating meteorological variables for improving the accuracy of wind speed forecasting.

Keywords: Wind speed, wind direction, vector autoregressive model with exogenous variables, meteorological variables.

INTRODUCTION
Wind is one of the most promising energy sources for generating electricity. It is a renewable, clean, and plentiful energy source. The share of electricity generated from wind based resources is increasing over the years. Wind turbines that are used for generating electricity can be constructed both on land (on-shore) and on the platforms located at lake, seas and oceans (off-shore). Wind turbines might be used for providing electricity for the isolated locations, or might be used for feeding national power grid. Usually, offshore wind is stronger and more persistent, but construction and maintenance costs of the wind turbines for those locations are higher as compared to onshore counterparts (Spera, 1994).

The role of wind energy in overall energy markets is increasing. In the world, 83 countries are providing electricity from wind based sources, and 25% of the electricity of Denmark is produced from wind based resources (Dvice, 2014). As of 2010, in overall 2.5% of electricity worldwide is generated from wind based resources. It grows rapidly more than 10% per annum (GWEC, 2012).

Wind farms are the locations that are used for converting the wind energy into electric electricity. Individual wind turbines might produce outputs in up to 8 MW (Wind Power Monthly, 2014), whereas the output for wind farms consisting of several hundred wind turbines might reach to 1 GW or more (Terra Gen Press Release, 2012).

The amount of power that might be generated from wind based resources is pretty much consistent over the years, but might fluctuate significantly over shorter time periods. It might differ with respect to the time of the day and the season of the year along with other factors. Since wind is an intermittent energy source, and the proportion of electricity generated from wind based resources is increasing, the accurate forecasting of wind power is an important aspect.
The wind power is strongly associated with the wind speed. Usually, for certain wind speed below the threshold, no power is obtained. For wind speeds above a threshold level, in order to prevent damage to the turbines, the wind turbines are shut down. These threshold wind speeds (i.e., minimum and maximum wind speeds that a turbine operates) are also known as cut-in and cut-out speeds respectively. The power generated from the wind turbine increases cubically with the increasing wind speed up until rated output power and stays constant until the cut-out speed. In that regard, in order to predict the wind power, accurate forecasting of the wind speed carries an important concern.

In this paper, we develop a mixed regression model based on the Vector Autoregressive model with exogenous variables (VARX) for forecasting the wind speed. The model incorporates meteorological factors factors (such as pressure, temperature, humidity) as well as wind related characteristics (i.e., wind speed and direction). In order to incorporate the wind direction, a classification scheme is utilized, and with respect to the classification, dummy variables representing category of the wind direction in terms of the wind speed is included in the regression model.

LITERATURE REVIEW

There are numerous studies in the literature regarding forecasting wind power output and direction. Since the market proliferation rate of the electricity obtained from wind based resources are increasing, forecasting the wind power output accurately contributes to planning and operations of the power grid in a more effective and efficient manner. Forecasting the wind power can effectively contribute to the solution of many problems regarding the grid operations such as standards of interconnection, competitive market design, real time grid operations, ancillary service requirements and costs, quality of power, capacity of transmission system, stability and reliability of power system, optimal reductions in green house, and optimal reductions in green house gas emissions (Wu and Hong, 2007; Lund, 2005). Additionally, improved forecasts help with functioning day ahead and hour ahead markets (Soman et al., 2010; Negevitsky et al., 2010). It has been indicated that the forecasting errors might contribute up to the 10% of the costs for selling wind energy (Pinson et al, 2010).

The forecasting of wind power output can be classified according to the specified time horizon. It has been indicated that the length of time horizon might span from few seconds (very short term) to 1 week or more (long term). Generally very short-term forecasts are used for the electricity market clearing and regulation purposes whereas long term forecasts are used for unit commitment decisions, reserve requirement decisions, and maintenance scheduling. In between those categories, short term and medium term forecasting is also used (Soman et al., 2010). There is a variety of methods that have been used for predicting wind speed. The most basic one is the persistence method (i.e., naïve predictor) which basically is providing forecast as the value of the wind speed experienced in previous period. Albeit the simplicity of the approach, it has been still used by the industry, and for very short term forecasting, the approach provides fairly good results (Potter and Negevitsky, 2006). On the other hand, numerical weather prediction methods rely on the physical approach which involves solving complex mathematical model that involve meteorological variables such as temperature, humidity, etc. The approach provides fairly good results for the long term forecasting under steady weather conditions (Candy and English, 2009, Potter and Negevitsky, 2006)

Additionally, along with the physical based models, the statistical approaches based on neural network and auto regressive moving average (i.e., ARMA) based approaches are also popular. There are various variations of neural network such as feed forward, multi-layer perceptrons, recurrent neural network, radial basis functions, and Adaline networks, etc. (Soman et al., 2010). Among ARMA models, fractional ARMA models (Kavasseri and Seetharaman, 2009), ARMA models with exogenous variables (i.e., ARMAX) (Yang and Huang, 1996), grey predictors (El Fouly et al., 2006), exponential smoothing (Cadenas et al, 2010), and linear prediction (Riahy and Abedy, 2008) based approaches are popular. Usually, statistical based methods are useful and accurate for short to short-medium term prediction of wind power and wind speed (Soman et al., 2010).

Along with the statistical and physical based methods, non-conventional approaches based on information science, control theory, artificial intelligence, and data mining approaches are also beginning to emerge. Some of these approaches also combine the spatial information by incorporating the correlation between different wind sites for forecasting purposes (Damousis et al., 2004) based on the assumption that the patterns in wind speed in some location is to some degree repeated at another wind site by some time lag. Among the unconventional models, wavelet transforms (Fichaut and Ranchin, 2002), ensemble forecasting (Sloughter et al., 2010), entropy based training (Bessa et al., 2009), fuzzy logic (Damousis and Dakapoulos, 2004) are proposed in the literature. These methods can be combined with other methods thus leading to hybrid approaches. Hybrid approaches such as combining the artificial neural network and the fuzzy logic is good at
very short term forecasts whereas combining numerical weather prediction with the neural network provides good estimates for the medium and long term forecasts (Soman et al., 2010).

The studies featuring incorporating wind direction and other related variables in the prediction of wind speed starts in 1980’s. Geets (1984) indicates wind direction and air related variables can be integrated into Kalman filter approach for improving the results. In that line, Damousis et al. (2004) incorporates wind direction for prediction of wind speed by developing a membership function for the wind speed, and using the genetic algorithm based training, they develop the fuzzy logic based approach for predicting the wind speed. Sideratos and Hatziargyriou (2007), combining the fuzzy logic and neural networks, develop an advanced statistical methods based on the self-organized mapping. The model uses the forecasts for wind direction and wind speed as meteorological variables as well as the past wind power outputs for predicting the wind power. Watson (1992) discusses the numerical weather prediction model for forecasting wind speed and wind direction for reducing the fossil fuel related costs for start/stop operations of the wind turbine along with the spinning reserve. Lin et al. (1996) develop neural network models that combine wind speed and wind direction for prediction purposes. Marti et al. (2004) develop the power curve and power output model by incorporating two different approaches involving the power direction. The first approach features density corrected wind speed that is based on predictions of sea level pressure and 2 meters level and the second approach involves downscaling based on principal components for selecting regressor for multiple regressions. These approaches use the adaptive local polynomial regression with the dependence on the wind direction. Nielsen et al. (2004) develop a model to estimate the quantities of the prediction error based on the regression model. The regression model incorporate forecasted power from wind power prediction tool, wind power prediction tools and forecasts from High Resolution Limited Area model on air density, friction velocity, wind speed, and wind direction. Bao et al. (2009) use the Bayesian model averaging for producing the density function from ensemble forecasts incorporating wind direction as well as wind speed.

Erdem and Shi (2010) employ different types of ARMA based model for developing joint models for forecasting wind speed and direction simultaneously. Palomares and de Castro (2003) apply statistical downscaling method which is adjusted based on seasons and wind direction to predict the wind speed. Gallardo et al. (2008) launch a CASANDRA project. The project features a power curve model based on the multivariate regression over wind farm data (wind speed, direction, pressure and temperature), and model output statistics corrected winds, and other variables obtained from the mesoscale model.

Other than the wind speed, various approaches incorporate the other meteorological variables (pressure, temperature, etc.). In that regard, Akylas (1997) develop the multi-variable regression over the time series obtained from the meteorological masts. Those meteorological masts capture meteorological variables such as wind speed, temperature, pressure, and pressure tendency. As previously mentioned, CASANDRA project developed by Gallardo et al. (2008) features multivariate regression incorporating different weather related attributes. Ramirez-Rosado et al. (2009) use the meteorological variables obtained from numerical weather prediction model to develop neural networks. Fan et al. (2009) develop two-stage hybrid network with Bayesian clustering by using the support vector regression, and use the model to predict the wind speed up to 48 hours. The proposed model take the wind pressure, temperature, elevation of wind turbines located in the wind farms into consideration. Yesilbudak et al. (2013) use the k-nearest neighbor classification model for forecasting wind speed for very short term forecasting horizon. The meteorological variables used in this study are wind direction, air temperature, atmospheric pressure and relative humidity parameters. Hocaoglu and Karanfil (2013) identify the bi-relational causal relationship between different meteorological variables (wind speed, pressure and temperature) by employing the Granger causality and impulse response analysis. The authors suggest that incorporating this relationship might help enhancing the models that are used for predicting the wind speed.

Based on the literature review, it can be seen that various approaches have been developed for forecasting the wind speed. Some approaches also incorporate the other meteorological variables. Our approach combines classification algorithm and statistical based VAR (vector auto regressive) methodology for forecasting the wind speed for very short term (i.e., 10 minutes) and short to medium term (1 hour). In the next section, we present the methodology and provide the data sources that are used as the input for predicting the wind speed.

METHODOLOGY
In the model, we develop two approaches for forecasting the wind speed for very short term (i.e., 10 minutes) and short term (1 hour). The first approach involves the traditional auto regressive moving average for forecasting the wind speed. The second approach incorporates the meteorological variables and the wind direction as well in a mixed model that incorporate not only past observations of the wind speed, but also past
observations of the meteorological variables as well as the classification of the forecasts for the wind direction. The second methodology is comprised of four sub-stages which can be expressed as follows:

- **Classification of wind speed with respect to wind direction:** Using the k-means algorithms, the classification of the wind directions with respect to the associated wind speed is conducted. For this purpose, three different clusters are formed. Number of clusters is decided based on the trade-off between capturing the variation among wind directions with respect to wind speed, and the model coherence. Based on the trial and error, number of clusters to be formed is decided as 3. Using the k-means clustering algorithm, corresponding past wind speed-direction observations are classified into three different clusters.

- **Forecasting the wind direction:** For this purpose, the linked ARMA based models is used for forecasting the wind direction. Using the link function, the circular variable (i.e., the wind direction) is converted to the linear variables for the prediction purposes. Using the ARMA methodology, after the wind directions in linearized forms are obtained, using the inverse link functions, these variables are converted back to the actual wind directions.

- **Forecasting the wind speed:** In this step, a mixed linear regression model that incorporates the past meteorological variables as well as the wind direction clusters are used. The mixed regression model is composed of two parts. The first part is the Vector Auto Regressive part (VAR) that incorporates the meteorological variables. These variables are:
  - Wind speed (m/s)
  - Pressure (mBar)
  - Temperature (°C)
  - Wind shear (s⁻¹)
  - Relative Humidity (%)

Apart from the VAR part, the second part incorporates the categorical variables that are used for representing the clusters associated with the wind direction. This mixed model leads to Vector Autoregressive with exogenous variables (i.e., VARX) model. In the next subsections, more information regarding the established methodology is provided.

**ARMA Model**

An ARMA model which can be denoted by ARMA (p,q) can be expressed as follows (Montgomery et al., 2008):

\[ y_t = \delta + \sum_{i=1}^{p} \varphi_i y_{t-i} + \sum_{j=1}^{q} \phi_j e_{t-j} + e_t, \]  

where \( \delta \) is the mean of the underlying process, \( \varphi_i \) is the \( i^{th} \) autoregressive coefficient, and \( \phi_j \) is the \( j^{th} \) moving average coefficient, and \( e_t \) is the residual term at time period \( t \). Note that the based on the assumptions, error term is normally distributed with mean 0, and \( y_t \) is the wind speed observed or forecasted at time period \( t \).

**Linked ARMA model**

A linked ARMA model incorporates the link function to convert the circular variables to the linear variables. After the circular variables are transformed into the linear variables, using the ARMA based approach, the forecasts are provided. Various link functions exist in the literature; we adopt the cumulative distribution function based on standardized normal distribution following Erdem and Shi (2011). To do so, we make use of the linked function which is expressed in Eq. (2) as (Fisher and Lee, 1994),

\[ g(u) = 2\pi(\Phi(u) - \frac{1}{2}) + \mu \]  

where \( \Phi(u) \) is the standardized normal cumulative distribution function, and \( \mu \) is the mean wind direction parameter. Accordingly, the inverse of the probit link function is presented in Eq. (3),

\[ g^{-1}(b) = \Phi^{-1}((b - \mu)/2\pi + \frac{1}{2}) \]  

where \( \Phi^{-1} \) is the probit function (i.e., inverse of the standardized normal distribution function. Note that Eq. (3) is the inverse link function that is used for converting the circular variables back to the linear variables. Based on the ARMA approach, forecasting of the linearized forms of the circular variables are conducted. Using Eq. (4), the forecasted variables are converted back to the circular variables to obtain the forecasts of the wind speed.
where $S$ is the sum of the sine values of all the angles with respect to the north axis (i.e., $S = \sum_{i=1}^{n} \sin \theta_i$), $C$ is the sum of the cosine values with respect to the north axis (i.e., $C = \sum_{i=1}^{n} \cos \theta_i$).

**K-means Algorithm**

K-means algorithm is one of the basic classification algorithms. The algorithm is used for categorizing the wind directions with respect to the corresponding wind speed. Associated wind directions are assigned to the corresponding clusters based on the proximity of the wind speed-wind direction to the particular cluster center (i.e., centroid). Based on the inclusion of those pairs, the center for those clusters is updated. The procedure is iterative in nature and consists of mainly two steps. In the first step, each observation pair is assigned to the particular cluster based on the partitioning according to the Voronoi Diagrams. In that instance, each observation $(x_p)$ is assigned to only one cluster $S_{(w)}$ (i.e., cluster), where $p$ represents the index of the observation, $b$ represents the index of the cluster, and $w$ represents the index of step, and $S_{b(w)}$ represents the center location of the cluster. After the clusters are formed, the subsequent step is updating the location of the centroids for the new steps. These two steps follow each other in iterative manner until the total distance of the pairs to the center of the clusters is minimized. Mathematically, it can be expressed as follows (Lloyd, 1982);

- **Assignment step**

  
  
  $$
  S_{b(w)} = \left\{ x_p \right| \| x_p - m_{b(w)}^{(w)} \| \leq \| x_p - m_{i(w)}^{(w)} \| \ \forall 1 \leq c \leq k \}
  $$

  
  where $m_{c(w)}$ is the mean of cluster $c$ at step $w$.

- **Update step**

  
  $$
  m_{c(w+1)} = \frac{1}{|S_c|} \sum_{x_p \in S_c} x_p
  $$

  
  The problem is NP-hard in nature and there are heuristic approaches for solving the classification algorithm. A heuristic version of the k-means algorithm is developed and implemented. Three different clusters (that are low, medium, and high) are formed and wind direction-wind speed pairs are assigned with respect to those clusters.

**VARX Model**

After forming the clusters and forecasting the wind direction next step is developing the VAR model with exogenous variables (i.e., VARX). VARX model incorporates six types of meteorological variables namely wind direction, wind speed, temperature, pressure, relative humidity, and the wind shear. Using the classification algorithm described in the previous section, the wind direction speed pairs are grouped in the corresponding bins, and represented by the dummy variables indicating the categories that particular wind speed and direction is assigned to. Three different clusters are represented by two dummy variables. If the wind direction and speed belongs to the first category, both of the variables (i.e., $z_1$ and $z_2$) take the value of 0, if the pair belongs to the second category, the $z_1$ variable takes the value of 1, and $z_2$ takes the value of 0, if the wind speed-pair direction belongs to the third category then the $z_2$ variable takes the value 1, and $z_1$ variable takes the value of 0. Using the classification algorithm that is described in the previous section, clusters are formed and represented by the dummy variables. Those dummy variables represent corresponding clusters for forecasting wind direction. The VARX model, which is comprised of autoregressive terms belonging to five different meteorological variables and the categorical variable representing the wind direction, can be represented as;

$$
\begin{align*}
  y_t &= \delta + \sum_{i=1}^{\tau} \beta_i y_{t-i} - \sum_{i=1}^{\eta} \gamma_{ki} w_{k(t-i)} + z_1 q_1 + z_2 q_2
\end{align*}
$$
where \( q_1 \) and \( q_2 \) are the coefficients associated with the categorical variables, \( \delta \) is the constant term, \( \varphi_i \) are the autoregressive components belonging to the past observations of the wind speed, \( y_{(i-1)} \) is the value of the wind speed observed at \( i \)-th period prior to the current period, and \( m_i \) and \( m_q \) are the values of the corresponding dummy variables, \( \eta_k \) is the coefficient of the meteorological variable of type \( k \) for time period \( i \), and \( w_{k(i-1)} \) is the actual value of the meteorological variable that is observed \( i \) periods prior to the current period. Note that the first meteorological variable is temperature, the second one is the relative humidity, the third variable is the pressure, and the fourth variable is wind shear and categorical variables represent the wind directions. For scaling purposes relative humidity and pressure is divided into 100, and incorporated into mathematical model.

In this study, we consider the mean absolute error (MAE) and mean square error (MSE) for assessing the quality of the forecast, which can be expressed with the following formulas by Eqs. (8) and (9) respectively;

\[
MAE = \frac{\sum_{t=1}^{N} \left| y_t - \hat{y}_t \right|}{N}
\]  

\[
MSE = \frac{\sum_{t=1}^{N} (y_t - \hat{y}_t)^2}{N}
\]

where \( y_t \) is the realized value at time period \( t \), and \( \hat{y}_t \) is the forecast pertaining to that period, and \( N \) is the number of data points.

**WIND DATA**

The data is obtained from meteorological mast from the wind site located in Colorado, USA. As previously mentioned, relative humidity, temperature, wind shear, wind direction information are captured by using the instruments placed in the meteorological mast. The data are recorded continuously and averaged over every 10 minutes and one hour to obtain the wind attributes as well as the meteorological variables. The data belongs to the period of January 1, 2012 and December 31, 2012.

For short term forecasting (with a forecasting horizon of 1 hour), the yearly data averaged on hourly basis is used. In total, there are 8,784 data points in which 7,000 data points are used for the model building, and the remaining 1784 data points are used for testing the model based on MAE and MSE measures.

For very short term forecasting (with a forecasting horizon of 10 minutes), the data pertaining two months (i.e., January and February, 2012) is used. For this purpose, the data is averaged over 10 minutes. Out of 8,640 data points, in order to establish a common platform between the efficacy of the model for very short term and short term forecasting, similar number of data points are used for model building and model testing correspondingly. In short, 7,000 data points are used for model building purpose, while the remaining 1,640 points are used for testing purpose. Based on the performance measures, the developed VARX models are tested against their ARMA counterparts. In the next section, more information is provided in the underlying model structure. Along with the model coefficients, the comparison between the models based on the MAE and MSE measures are provided.

**RESULTS**

Basically two different model structures with respect to the forecasting horizon (very short term and short term) are constructed. As previously suggested, the very short term planning horizon incorporates 10-min ahead forecasts, while the short term forecasts provides the 60-min ahead forecasts. Both are conducted in rolling one-step ahead forecast manner. For the short term forecasting, based on the Bayesian Information Criterion (i.e., BIC), corresponding values for different model structures are compared. Note that BIC can be approximated for large sample sizes based on the following formula (Wit et al., 2012);

\[
BIC = -2 \ln \hat{L} + d \ln(n)
\]

where \( \hat{L} \) is the maximized value of the likelihood function of the underlying ARMA process, \( n \) is the sample size, and \( d \) is the number of estimated parameters. Based on the BIC values, ARMA(2,2) is chosen. Additionally, an analysis is performed to measure the correlation of wind speed with respect to other meteorological variables. As a result, the following correlation coefficients are obtained for 10-min and 1-hour data.
Table 1. Correlation coefficients between wind speed and other meteorological variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Averaged over 10-min</th>
<th>Averaged over 1-hour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind speed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Air temperature</td>
<td>0.369332</td>
<td>-0.02918</td>
</tr>
<tr>
<td>Relative humidity</td>
<td>-0.48431</td>
<td>-0.31152</td>
</tr>
<tr>
<td>Air pressure</td>
<td>-0.28198</td>
<td>-0.35734</td>
</tr>
<tr>
<td>Wind shear</td>
<td>0.096679</td>
<td>0.88309</td>
</tr>
<tr>
<td>Wind direction</td>
<td>0.468147</td>
<td>0.365679</td>
</tr>
</tbody>
</table>

Based on Table 1, it is concluded that the correlation between air temperature and wind speed is significant based on 10-min averaged data. However, for the data averaged over 1 hour, the correlation is not significant. Similarly, wind shear is strongly correlated with wind speed based on the data averaged over 1 hour, but the correlation between the wind shear and wind speed is not significant when data is averaged over 10 minutes.

The model coefficients are estimated using the maximum likelihood method by using SAS 9.2™. The results are provided in Table 2. In a similar fashion, the corresponding BIC values are calculated for different VARX model based on the order of auto-regression terms. The model that provides lowest BIC value is chosen as the underlying model. It turns out that VARX model that incorporates the autoregressive term up to third order provides the lowest BIC value. The model parameters are estimated in two stages. In the first stage, the unrestricted VARX model is run in SAS 9.2™. In the second stage, the model coefficients that have associated p values greater than 0.1 are restricted to be equal to 0. Based on this approach, the following results that are presented in Table 3 are obtained.

It has been indicated that the wind speed up to autoregressive order of 3 plays a significant role in forecasting the wind speed. In a similar fashion, temperature and relative humidity up to autoregressive order of 2 are significant, for pressure and wind shear; only first term autoregressive coefficients (i.e., the coefficients associated with one period prior) are significant. The categorical variables are significant indicating that wind directions pertaining to the second and third group contributes positively to the wind speed, whereas the contribution of the second group is higher than the contribution of the third group. In terms of MAE and MSE values, the comparison is provided in Table 4.

Table 2. ARMA model coefficient for 1 hour ahead prediction

<table>
<thead>
<tr>
<th>Short Term Forecasting Horizon (1-hour ahead)</th>
<th>ARMA (2,2)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Model parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>1.475515***</td>
<td></td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>-0.50477***</td>
<td></td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.55436***</td>
<td></td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>0.16613***</td>
<td></td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.143161***</td>
<td></td>
</tr>
</tbody>
</table>

**: Significant at 1%

Table 3. VARX model coefficient for 1 hour ahead prediction

<table>
<thead>
<tr>
<th>Wind speed</th>
<th>Coefficient value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind speed</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.91402***</td>
</tr>
<tr>
<td>2</td>
<td>-0.16322***</td>
</tr>
<tr>
<td>3</td>
<td>0.09311***</td>
</tr>
<tr>
<td>Temperature (C°)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.08039***</td>
</tr>
<tr>
<td>2</td>
<td>-0.08159***</td>
</tr>
<tr>
<td>Relative humidity/100 (%)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1.5967***</td>
</tr>
<tr>
<td>2</td>
<td>1.07457*</td>
</tr>
<tr>
<td>Pressure/100 mBar (mm)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-3.6674***</td>
</tr>
<tr>
<td>Wind shear</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-3.54262***</td>
</tr>
<tr>
<td>Wind direction coefficients</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.61215***</td>
</tr>
<tr>
<td>2</td>
<td>0.06621**</td>
</tr>
<tr>
<td>Constant term</td>
<td>30.90954***</td>
</tr>
</tbody>
</table>

**: Significant at 1%
**: Significant at 5%
*: Significant at 10%
Table 4. Corresponding MAE and MSE values for the models

<table>
<thead>
<tr>
<th></th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>VARX</td>
<td>3.154861</td>
<td>1.293287</td>
</tr>
<tr>
<td>ARMA(2,2)</td>
<td>3.469839</td>
<td>1.341397</td>
</tr>
<tr>
<td>Relative difference</td>
<td>9.983957%</td>
<td>3.719959%</td>
</tr>
</tbody>
</table>

The MAE value pertaining to the traditional ARMA model on the average is 3.72% worse as compared to the VARX model. When the MSE values are compared, the difference is even more, that reaches to approximately 10%. This significant difference in terms of MSE values are important in terms of predicting the wind power output, since the generated power is directly proportional to the cube of the wind speed in the given range. In terms of the very short term forecasting (i.e., 10-min ahead), a similar approach is followed. Table 5 provides the results for the ARMA model for one step 10-min ahead forecasting. The analysis on the VARX model reveals the coefficients presented in Table 6. Note that similar to the previous approach, the variables that has significance level above 10% is restricted to have the value of 0. The MAE and MSE values associated with the ARMA and VARX based models are provided in Table 7.

Table 5. ARMA model coefficients for 10-min ahead prediction

<table>
<thead>
<tr>
<th>Model parameters</th>
<th>ARMA (5,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_1$</td>
<td>1.00989***</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>-0.04041***</td>
</tr>
<tr>
<td>$\phi_3$</td>
<td>-0.06432**</td>
</tr>
<tr>
<td>$\phi_4$</td>
<td>0.06626***</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.159133**</td>
</tr>
</tbody>
</table>

***: Significant at 1%
**: Significant at 5%

Examining Table 6, it can be seen that compared to the 1-hour ahead forecasting, fewer variables that are associated with the meteorological variables are significant for the prediction of wind speed at 10-min ahead. Comparing Tables 4 and 7, it can be observed that the forecasting power of VARX model over the ARMA counterpart diminishes, but there is still some significant difference between the two approaches both in MAE and MSE values. The ARMA based approach performs approximately 4.4% worse as compared to VARX counterpart, and for the MSE values, the difference is reduced to 1.38%. Note that for one hour forecast, ARMA (2,2) model provides the best results; while for the 10-min forecast, ARMA(4,0) performs the best in terms of the BIC value.

Table 6. VARX model coefficient for 1 hour ahead prediction

<table>
<thead>
<tr>
<th>Wind Speed</th>
<th>Coefficient Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind Speed</td>
<td>0.20243**</td>
</tr>
<tr>
<td>Relative humidity/100 (%)</td>
<td>-76.0149**</td>
</tr>
<tr>
<td>Wind direction coefficients</td>
<td>75.1798**</td>
</tr>
<tr>
<td>Constant term</td>
<td>0.54673***</td>
</tr>
</tbody>
</table>

***: Significant at 1%
**: Significant at 5%

Table 7. Corresponding MAE and MSE values for the models

<table>
<thead>
<tr>
<th></th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>VARX</td>
<td>2.946603</td>
<td>1.169026</td>
</tr>
<tr>
<td>ARMA(4,0)</td>
<td>2.987438</td>
<td>1.220642</td>
</tr>
<tr>
<td>Relative difference</td>
<td>1.385813%</td>
<td>4.415336%</td>
</tr>
</tbody>
</table>

The forecasted values provided by the VARX and ARMA models and the actual values are compared for 120 data points both for 1-hour and 10-min forecasts. The results are provided in Figures 1 and 2 for 10-min ahead and 1 hour ahead forecasting, respectively.
CONCLUSIONS

In this paper, we develop the vector autoregressive with exogenous variables approach for forecasting wind speed 10-min and 1 hour ahead. Rather than incorporating the wind direction related variables into the vector autoregressive approach directly, using the classification scheme, the clusters that these wind direction speed pairs are classified into, are represented as separate groups in the vector autoregressive models as exogenous variables. For this purpose, along with the wind direction, four different types of meteorological variables (i.e., pressure, relative humidity, temperature, and wind shear) are incorporated into vector autoregressive equation. Different orders of autoregressive terms are compared using the BIC values to decide on the best model.

Our approach, as compared to the benchmark ARMA based approach, provides significantly better results especially for the 1 hour ahead forecasting. It can be seen that for the short term (10-min forecasting), the performance of both models are similar. Especially, the difference between the performance measures in terms of the MSE values reaches 10%, which is quite significant especially for the wind power prediction based on the fact that for certain intervals, the wind power is directly proportional to the cube of the wind speed. As a future research direction, the development of the Kalman filter based approaches involving meteorological variables would be considered. In line with this approach, different constructs for building the state space models would be examined to improve forecasting of the wind speed.
REFERENCES


