Article

Improved Spatio-Temporal Linear Models for Very Short-Term Wind Speed Forecasting

Tansu Filik

Department of Electrical and Electronics Engineering, Anadolu University, Eskisehir 26555, Turkey; tansufilik@anadolu.edu.tr; Tel.: +90-222-3350580-6460

Abstract: In this paper, the spatio-temporal (multi-channel) linear models, which use temporal and the neighbouring wind speed measurements around the target location, for the best short-term wind speed forecasting are investigated. Multi-channel autoregressive moving average (MARMA) models are formulated in matrix form and efficient linear prediction coefficient estimation techniques are first used and revised. It is shown in detail how to apply these MARMA models to the spatially distributed wind speed measurements. The proposed MARMA models are tested using real wind speed measurements which are collected from the five stations around Canakkale region of Turkey. According to the test results, considerable improvements are observed over the well known persistence, autoregressive (AR) and multi-channel/vector autoregressive (VAR) models. It is also shown that the model can predict wind speed very fast (in milliseconds) which is suitable for the immediate short-term forecasting.

---

1. Introduction

Electricity consumption of the developing countries increases annually [1,2]. However, the authorities are aiming to reduce the greenhouse gas emission and also the electricity consumption by increasing the amount of renewable energy and improving the energy efficiency respectively [3]. Since wind energy is sustainable, emission-free and cost-effective, it is very attractive and a good candidate to achieve the above ambitious aims. In order to use these energy sources reliably in the future's optimum economic power system operations, it is critically important to accurately forecast wind power generation [4–6]. Since wind power is a function of the cube of wind speed, accurate wind power output prediction depends on wind speed prediction [7].

Wind speed prediction problem is widely investigated in literature and various methods are presented [5,7–13]. The available methods are generally separated as physical and statistical methods. However, for very short-term wind speed forecasting, physical model-based methods such as numerical whether prediction (NWP) have high computational complexity and lower accuracy [7,8]. Therefore, some hybrid of physical (NWP) and statistical methods are proposed in literature as in [8,9]. Computationally efficient but accurate and reliable statistical methods for very short-term wind speed forecasting are required especially for the electricity market-wind forecasting control [14]. The statistical methods can be classified as point and probabilistic forecasting approaches [8]. In point forecasting approach, future wind speed is given as a single value. However, in probabilistic forecasting case, the future wind speed value is modelled as random variable and its probability density function (pdf) is given as a result.

Recently, spatial correlation models, which also known as "spatio-temporal" methods, are appeared as a new trend in short-term wind speed forecasting [5]. These methods use measurements from neighbourhood of target location (wind farm) for more accurate wind speed forecasting with a modest processing overhead [15–21]. Since wind is a horizontal movement in atmosphere, its spatial correlation carries important information for such spatial models. However, the spatial correlation of low level wind directly depends on the complexity of the terrain. In [15], space-time forecasting model is proposed which promises more accurate results according to conventional time series models. However, this model which is called as calibrated probabilistic forecasting is designed only for the selected region. This region specific forecasting model is improved in [16], so it does not require any prior geographic information for the target region. In [17] a graph-learning based spatio-temporal analysis techniques are used to characterize probabilistic models for short-term forecasting. In [5], a methodology is proposed for optimum probabilistic forecasting of geographically dispersed information. In [19], multichannel adaptive filtering technique is applied for short-term prediction which promise lower complexity, improved robustness and ability to track seasonal variations. Most of the above methods are based on the statistical analysis and interpretation of the location specific multi-channel data collected in years.

On the other hand, the conventional linear time series models are easy to implement and requires no preliminary analysis for model development. Hence these models are widely preferred for short-term wind speed forecasting [12,22–24]. However, the multi-channel (spatio-temporal) linear methods, which uses the measurements from neighbourhood of target location, have not been addressed sufficiently for short term wind speed forecasting. The vector autoregressive (VAR) method is applied to geographically dispersed (multi-channel) wind speed data in [25]. There are also some other hybrid artificial neural network (ANN) based methods [26–28].

The multi-channel autoregressive moving average (ARMA) models are commonly used for blind identification of single input multi output (SIMO) systems in communications, source localization and medical imaging [29–31]. These multi-channel blind linear system models can also be applied for multi-channel wind speed prediction problem for more accurate results on target location [32,33].

In this paper, the multi-channel linear prediction models for short-term wind speed forecasting using neighbouring wind speed measurements around the target location which is sketched in Figure 1 are investigated and reviewed. These multi-channel linear prediction models can also be called as multi-channel ARMA or MARMA. The problem formulation, compact matrix forms and efficient multi-channel coefficient estimation approaches are presented and tested using hourly averaged real wind speed/direction values. These values are collected from the five synchronized measurements station of the Turkish State Meteorological Service. These stations are selected around the Canakkale Canel of the Turkey, namely Bozcaada (BOZ), Ipsala (IPS), Gonen (GON), Bandirma (BAN), and Sile (SIL). The root mean square error (RMSE) and mean absolute error (MAE) are used as the performance measurements of the prediction models. It is shown that MARMA model's prediction performance is better than uni-variable AR and multi-variable vector AR (VAR). It is also observed that the performance's of the MARMA increases when the forecast lead time is increased compared to other methods.

The paper is structured as follows: (1) Multi-channel linear prediction models and their compact matrix forms for short-term wind speed forecasting is presented and reviewed. (2) Computationally efficient and accurate linear solution techniques with a new linear channel selection approach for multichannel coefficients are proposed and discussed. (3) MARMA forecasting models are tested using the real wind speed data which are collected from three different locations from the Canakkale region of Turkey. The RMSE and MAE performances are compared for various cases. The section organization of the paper is as follows. In Section 2, problem formulation of the multi-channel linear prediction models and the coefficient estimation techniques are presented. In Section 3, the selected region where the real multi-channel wind data collected is introduced and prediction performances

of the models in Section 2 are tested and compared with other methods. We conclude the results in Section 4.
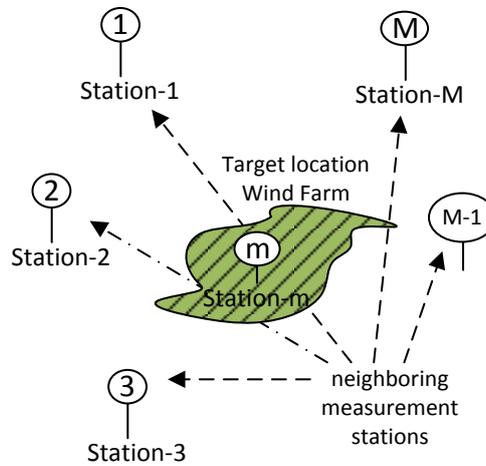

Figure 1. Multi-channel wind data measurement stations around the target location.

2. Multi-Channel Wind Data and Linear Prediction Models

2.1. Multi-Channel Wind Data

We consider M spatially distinct (geographically separated) measurement stations with known positions as shown in Figure 1. At each mth station (channel), discrete measurements are assumed to be collected as:

$$y_m[n] = \frac{1}{\Delta t} \int_{t_n - \Delta t}^{t_n} y_m(t) dt \tag{1}$$

where $y_m[n]$ is averaged wind speed values at discrete time index n respectively. Δt is the averaging time duration and can be chosen as a minute or a hour.

The problem is to forecast short-term wind speed value at m^{th} station, using M spatially distributed (multi-channel) averaged wind data measurements as in Figure 1. Since wind directions are spread to all directions, wind measurement stations should surround the target location for the best result.

2.2. Multi-Channel Linear Prediction Models

In this part, multi-channel ARMA model which is used for blind identification of SIMO systems in [31] are modified and implemented for the multichannel wind speed prediction model. AR model is applied to multi-channel real wind speed data which is called as vector autoregressive model (VAR) in [25]. The VAR predictor's Δ hour ahead output for the mth channel (target location) is given as:

$$y_m[n + \Delta] = \sum_{i=1}^{M} \sum_{p=1}^{P} a_{i,p}^m y_i[n - p] + w_m[n], \ m = 1, ..., M \tag{2}$$

where M is the number of channels, P is the number of coefficients and $w_m[n]$ is the additive noise (model error) terms at each channel and assumed as temporally and spatially white random process with variance σ_w^2.

Two different multi-channel ARMA models are proposed for short-term wind speed prediction. First model is called as MARMA-1 and Δ hour forecast lead time output at m^{th} location is defines as:

$$y_m[n + \Delta] = \sum_{i=1}^{M} \sum_{p=1}^{P} a_{i,p}^m y_i[n - p] + \sum_{q=1}^{Q} b_q^m s[n - q] + w_m[n], \ m = 1, ..., M \tag{3}$$

where $s[n]$ is the common input signal which is white noise random process with constant power spectrum and statistically independent from the additive channel noises $w_m[n]$ with variance σ_s^2. Second model is called as MARMA-2 and Δ hour forecast lead time output at m^{th} channel which differently using multi-channel spatially and temporally white noise inputs, $s_k[n]$, as follows:

$$y_m[n + \Delta] = \sum_{i=1}^{M} \sum_{p=1}^{P} a_{i,p}^m y_i[n - p] + \sum_{k=1}^{M} \sum_{q=1}^{Q} b_{k,q}^m s_k[n - q] + w_m[n], \ m = 1, ..., M \tag{4}$$

It is possible to put M channel wind data for the above linear prediction models in matrix form as:

$$\mathbf{y}[n + \Delta] = \mathbf{A}\mathbf{x}[n] + \mathbf{w}[n] \tag{5}$$

where $\mathbf{y}[n + \Delta] = [y_1[n + \Delta] \ldots y_M[n + \Delta]]^T$ is a $M \times 1$ vector and this vector (also known as snapshot) includes M channel wind values from different locations at the same time. $\mathbf{x}[n]$ is the input data for the multichannel linear prediction models and defined for MARMA-1 in Equation (6) and MARMA-2 in Equation (7) respectively as:

$$\mathbf{x}[n] = [\underline{\mathbf{y}}_1^T[n - 1] \ldots \underline{\mathbf{y}}_M^T[n - 1] \ \underline{\mathbf{s}}^T[n - 1]]^T \tag{6}$$

$$\mathbf{x}[n] = [\underline{\mathbf{y}}_1^T[n - 1] \ldots \underline{\mathbf{y}}_M^T[n - 1] \ \underline{\mathbf{s}}_1^T[n - 1] \ldots \underline{\mathbf{s}}_M^T[n - 1]]^T \tag{7}$$

where $\underline{\mathbf{y}}_m[n - 1] = [y_m[n - 1] \ldots y_m[n - P]]^T$ for $m = 1 \ldots M$ and $\underline{\mathbf{s}}[n - 1] = [s[n - 1] \ldots s[n - Q]]^T$. Similarly multi-channel white noise process in Equation (7) is defined as $\underline{\mathbf{s}}_m[n - 1] = [s_m[n - 1] \ldots s_m[n - Q]]^T$ for $m = 1 \ldots M$. $\mathbf{x}[n]$ in Equation (6) is a $(MP + Q) \times 1$ vector and $\mathbf{x}[n]$ in in Equation (7) is a $(M(P + Q)) \times 1$ vector. $\mathbf{w}[n] = [w_1[n] \ldots w_M[n]]^T$ is $M \times 1$ additive channel noise vector. Finally the multi-channel prediction filter coefficient matrix ($\mathbf{A}$) for MARMA-1 is defined as:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^1 & \cdots & \mathbf{a}_M^1 & \mathbf{b}^1 \\ \mathbf{a}_1^2 & \cdots & \mathbf{a}_M^2 & \mathbf{b}^2 \\ \vdots & & \vdots & \vdots \\ \mathbf{a}_1^M & \cdots & \mathbf{a}_M^M & \mathbf{b}^M \end{bmatrix} \tag{8}$$

where $\mathbf{a}_i^m = [a_{i,1}^m \ldots a_{i,P}^m]$ and $\mathbf{b}^m = [b_1^m \ldots b_Q^m]$ for $i = 1, \ldots, M$, $m = 1, \ldots, M$. $\mathbf{A}$ is a $M \times (MP + Q)$ matrix and it includes all the unknown coefficients in Equation (3). Similarly for MARMA-2, $\mathbf{A}$ matrix is defined as:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^1 & \cdots & \mathbf{a}_M^1 & \mathbf{b}_1^1 & \cdots & \mathbf{b}_M^1 \\ \mathbf{a}_1^2 & \cdots & \mathbf{a}_M^2 & \mathbf{b}_1^2 & \cdots & \mathbf{b}_M^2 \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{a}_1^M & \cdots & \mathbf{a}_M^M & \mathbf{b}_1^M & \cdots & \mathbf{b}_M^M \end{bmatrix} \tag{9}$$

where $\mathbf{a}_i^m = [a_{i,1}^m \ldots a_{i,P}^m]$ and $\mathbf{b}_k^m = [b_{k,1}^m \ldots b_{k,Q}^m]$ for $i = 1, \ldots, M, k = 1, \ldots, M, m = 1, \ldots, M$. In this case, $\mathbf{A}$ is a $M \times M(P + Q)$ matrix and it includes all the unknown coefficients in Equation (4).

It is required to efficiently solve linear prediction model coefficients in Equations (8) and (9). The matrix form of the multi-channel linear prediction models, which is given in Equation (5), is similar to well known array signal model in array theory [34]. Array signal processing area deals with the space-time signals which are collected by an array of sensors. It is possible to solve these

coefficients using the subspace methods in [30]. Another computationally efficient way of solving these coefficients is given in [31].

In the next section, computationally efficient and accurate linear solution technique with a new linear channel selection approach for multichannel coefficient estimation is presented.

2.3. Multi-Channel Linear Prediction Coefficient Estimation

In order to find multi-channel linear prediction coefficients for more accurate results, N snapshot measurements are collected and the data in Equation (5) is extended as:

$$\begin{bmatrix} \mathbf{y}[n+\Delta] \\ \mathbf{y}[n+\Delta-1] \\ \dots \\ \mathbf{y}[n+\Delta-N] \end{bmatrix} = \begin{bmatrix} \mathbf{A} & & & \\ & \mathbf{A} & & \\ & & \ddots & \\ & & & \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{x}[n] \\ \mathbf{x}[n-1] \\ \dots \\ \mathbf{x}[n-N] \end{bmatrix} + \begin{bmatrix} \mathbf{w}[n] \\ \mathbf{w}[n-1] \\ \dots \\ \mathbf{w}[n-N] \end{bmatrix} \tag{10}$$

which can be rewritten as:

$$\mathbf{Y}[n+\Delta] = \bar{\mathbf{A}}\mathbf{X}[n] + \mathbf{W}[n] \tag{11}$$

where $\mathbf{Y}[n+\Delta]$ is extended multi-channel linear prediction output vector with size $MN \times 1$ and $\mathbf{X}[n]$ is extended prediction input vector. $\mathbf{W}[n]$ is extended model error vector with size $MN \times 1$ and $\bar{\mathbf{A}}$ is the extended coefficient matrix. In order to solve MARMA-1 and MARMA-2 coefficients in Equations (8) and (9) respectively, it is possible to apply a selection matrix for the specified m^{th} target location as:

$$\mathbf{S}_m\mathbf{Y}[n+\Delta] = \mathbf{S}_m(\bar{\mathbf{A}}\mathbf{X}[n] + \mathbf{W}[n]) \tag{12}$$

where the selection matrix for the mth location is defined as:

$$\mathbf{S}_m = [\mathbf{e}_1 \ \mathbf{e}_2 \ \dots \ \mathbf{e}_M]^T. \tag{13}$$

$\mathbf{e}_k$ is a $1 \times MN$ row vector as:

$$\mathbf{e}_k = [\ \underbrace{0, \ \dots \ ,0}_{M(k-1)+m-1} ,1, \ \underbrace{0, \ \dots \ ,0}_{M(N-k+1)-m} \ ] \tag{14}$$

If we multiply N multi-channel data in Equation (10) with m^{th} selection matrix $\mathbf{S}_m$ as in Equation (12) we get linear set of equations for the mth location as:

$$\mathbf{Y}_m = \mathbf{H}\bar{\mathbf{a}}_m^T + \mathbf{w}_m \tag{15}$$

where $\mathbf{Y}_m = [y_m[n+\Delta]\dots y_m[n+\Delta-N]]^T$ is a $N \times 1$, m^{th} channel output data vector. $\mathbf{H}$ matrix is equivalent to the $\mathbf{X}[n]$ in Equation (11) and it is the measurement data matrix which consist from the previous multichannel wind data and white noise signal. $\bar{\mathbf{a}}_m$ is the mth row of the $\mathbf{A}$ matrix in Equations (8) or (9) which is the prediction coefficients of MARMA-1 and MARMA-2 respectively for the mth target location. This model is the well known linear model in classical estimation theory [35] and it is possible to apply linear least squares (LS) techniques to find the optimum prediction coefficients. In this case, it is required to minimize the following cost function:

$$\min_{\bar{\mathbf{a}}_m} J(\bar{\mathbf{a}}_m) = (\mathbf{Y}_m - \mathbf{H}\bar{\mathbf{a}}_m^T)^T(\mathbf{Y}_m - \mathbf{H}\bar{\mathbf{a}}_m^T) \tag{16}$$

where $()^T$ is for transpose operation and the optimum LS solution for the unknown prediction coefficients is:

$$\hat{\mathbf{a}}_m^T = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{Y}_m \tag{17}$$

There are some computationally efficient ways to solve the above matrix pseudoinverse solutions as in [36,37].

3. Data and Test Results

3.1. Data Set

The accuracy of the proposed multi-channel linear prediction models are tested with hourly averaged wind speed and direction data which were collected from five stations around the Canakkale region of Turkey. Data is available in [38]. The three years hourly averaged wind speed and direction values between the years 2008 and 2010 are used. These five stations (Bozcaada, Ipsala, Gonen, Bandirma and Sile) belong to the Turkish State Meteorological Service and the locations are shown in Figure 2.
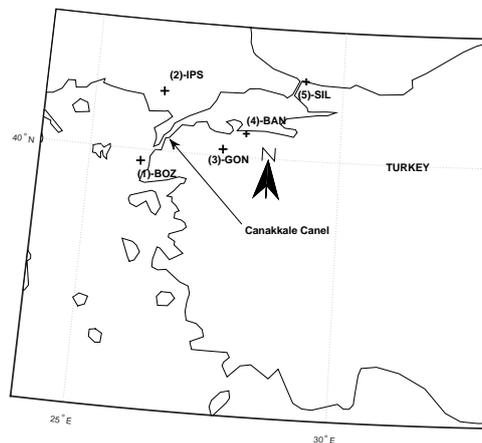

Figure 2. The measurement stations, Bozcaada (BOZ), Ipsala (IPS), Gonen (GON), Bandirma (BAN), and Sile (SIL) of Turkey where **N** indicates the North.

All wind measurements are taken from 10 meters height above ground. The region is known as having one of the highest wind energy potential in Turkey. These stations are selected arbitrary from the available measurement locations in that region. The topographic map of the region is shown in Figure 3. The topography is indicated by different colors; green colors indicated low altitude and white colors indicate hight altitude. As shown in Figure, these measurement stations are not close each other and the canal. BOZ is located at the highest point of an island. IPS is located in a valley. BAN is close to GON but it is separated from the canal. SIL is approximately 250 km far away from GON which is completely separated from the canal and other stations.
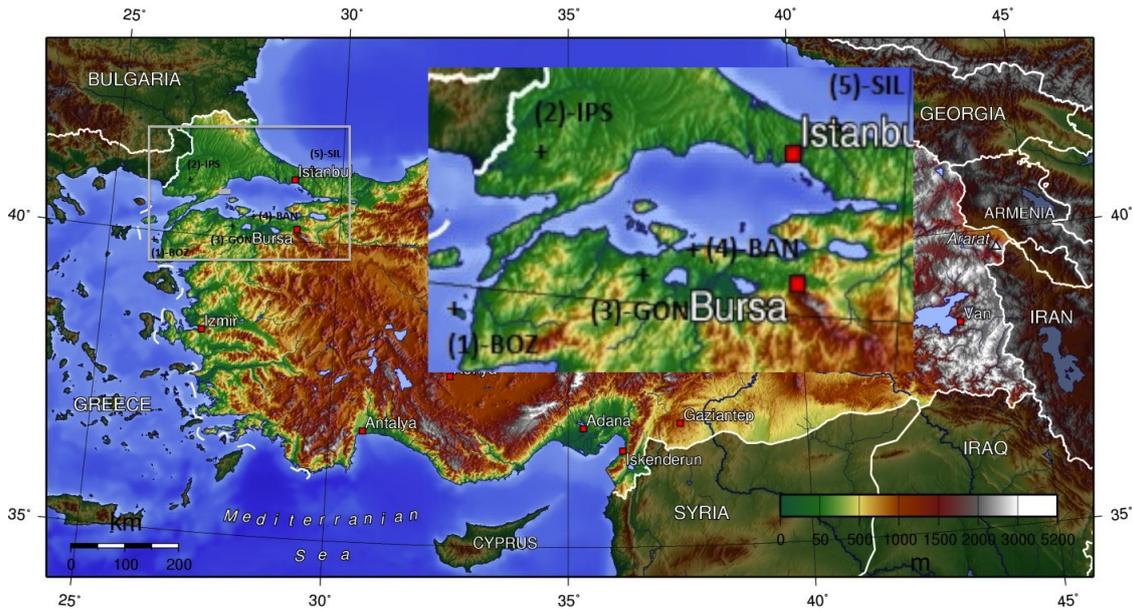
Figure 3. The topographic map of Turkey—A portion is zoomed for visual purposes. ("Turkey topo" by Captain Blood—Licensed under CC BY-SA 3.0)

Figure 4 shows the Auto and Cross-Correlation Coefficients of the stations with the target location GON for different time delays. All the correlations demonstrate a decline with time delay, except for maximum at diurnal periods (multiples of 24 h). It can be seen from Figure 4 that cross-correlation coefficient values of BOZ and IPS are higher than the other two (BAN and SIL) stations for short time delays, $1 \leq \Delta \leq 4$. SIL station has the lowest correlation values as expected. Since these stations are selected arbitrarily from the available stations their spatial dependencies are limited as shown. So it is not possible to apply a region specific space time method such as [15].
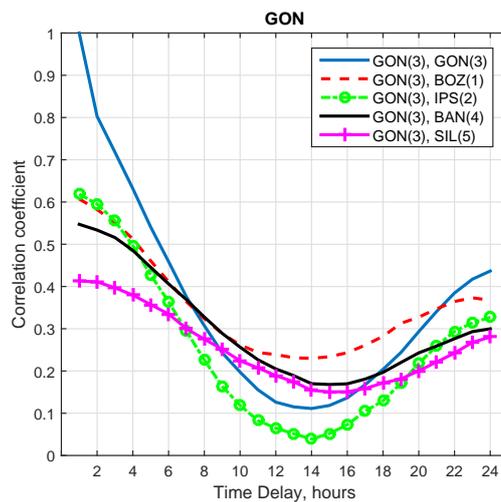

Figure 4. Autocorrelation and cross-correlation coefficients of wind measurements at Bozcaada (BOZ), Ipsala (IPS), Bandirma (BAN), and Sile (SIL) with Gonen (GON) in October–November 2008.

Figure 5 shows the frequency of the wind directions at the measurement stations as polar histograms. These polar plots show that prevailing wind directions at the stations are similar and along the Canakkale Canal from North East (NE) to South West (SW) and vice versa due to the large-scale circulation in that region.
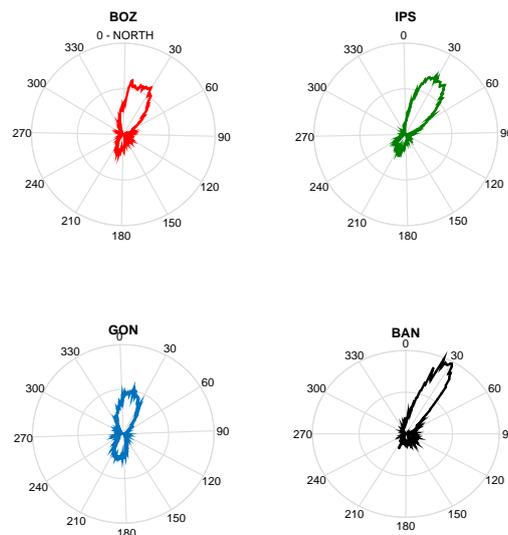
Figure 5. The frequency of the wind directions at the measurement stations for three year hourly averaged data where North is zero degree.

Some of the basic statistics (annual maximum, mean and variance values) of the used multi-channel data set are summarized in Table 1.

Table 1. Some basic statistics of the used multi-channel data set.

Year	(1)-BOZ (m/s)			(2)-IPS (m/s)			(3)-GON (m/s)			(4)-BAN (m/s)			(5)-SIL (m/s)		
	max	mean	var	max	mean	var	max	mean	var	max	mean	var	max	mean	var
2008	27.7	5.7	13.4	16.5	2.8	3.3	12.4	2.1	2.5	17.7	3.7	6.9	12.7	2.1	2.6
2009	22.8	5.6	12.9	14.7	2.7	3.1	11.5	1.9	2.1	17.3	3.7	7.1	11.0	2.2	1.9
2010	39.7	6.1	20.8	22.5	3.2	6.0	17.6	2.2	3.9	37.7	4.04	11.1	15.5	2.3	2.2

3.2. Test Results

In this section, real wind speed forecasting performances of the proposed multi-channel models are compared with the persistence, AR, VAR models. In order to compare and show the performances of the forecasting models, RMSE and MAE are calculated as,

$$RMSE_m(\Delta) = \sqrt{\frac{1}{K} \sum_k (\hat{y}_m[k+\Delta] - y_m[k+\Delta])^2}$$

$$MAE_m(\Delta) = \frac{1}{K} \sum_k |\hat{y}_m[k+\Delta] - y_m[k+\Delta]| \tag{18}$$

where $\hat{y}_m$ is the predicted value and y_m is the actual value. Δ is for forecast lead time in hours and m indicates the index of target location. K is the number of total predictions to calculate the RMSE and MAE in Equation (18). In this study, K is selected to cover the whole data between the years 2009 and 2010. In the following calculations of RMSE and MAE results total $K = 17280$ prediction values are used as in Equation (18) respectively. The persistence forecasting method in [25,39] is used as a benchmark to compare all the results. In persistence forecasting, the Δ ahead future value is taken as the current value. The prevailing wind directions are along the NE to SW and vice versa as shown in Figure 5. GON station is in the midst of the prevailing wind directions according to other stations. Therefore in the following case study, third station ($m = 3$) is selected as the target station which

is also surrounded by other stations. However, it is also possible to select the other stations as the target station.

3.2.1. Model Order

The linear prediction model orders of P and Q in Equations (2), (3) and (4) can be selected using the information criteria in [40] or the minimum description length in [41]. Figures 6 and 7 show the RMSE and MAE performances of the AR, VAR and MARMA models according to the model order for the 3rd station, GONEN ($m = 3$) and for the forecast lead time $\Delta = 2$, respectively. It is observed that the AR has minimum error for $P = 2$ and VAR and MARMA-1 gives minimum error when $P = 1$. On the other hand, MARMA-2's RMSE and MAE values are reducing when the filter order increased. Therefore, MARMA-2 model gives the best performance when $P = 4$ compared with other models.
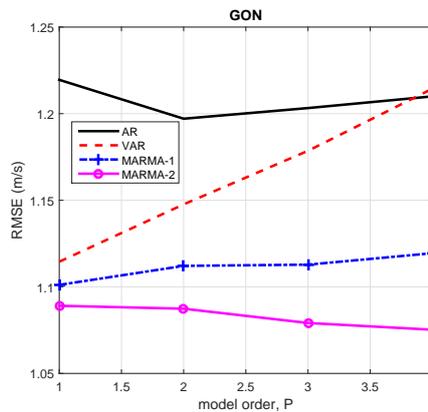

Figure 6. Root mean square error (RMSE) performances of the autoregressive (AR), vector autoregressive (VAR) and multi-channel autoregressive moving average (MARMA) models with respect to model order P when m = 3 (GONEN) and $\Delta = 2$ h.
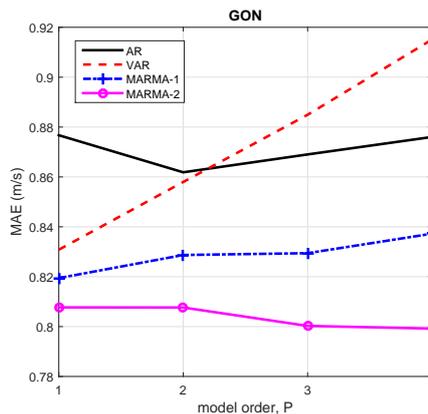

Figure 7. MAE performances of the AR, VAR and MARMA models with respect to model order P when m = 3 (GONEN) and $\Delta = 2$ h.

The similar confirmation is repeated for the Q parameter's of the MARMA models. Figures 8 and 9 show the RMSE and MAE performances of the MARMA-1 and MARMA-2 models respectively according to Q. It is observed that the increasing the model order Q slightly reduces the performances. For the best performance, Q is selected as 1 for MARMA models.
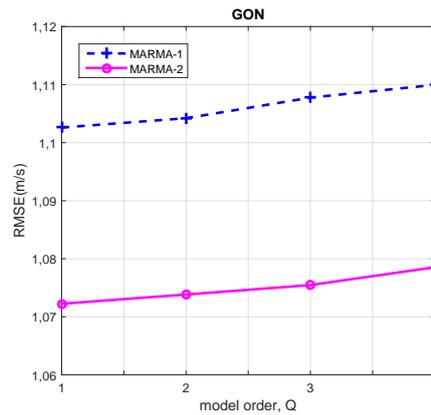
Figure 8. RMSE performances of the MARMA-1 and MARMA-2 models with respect to model order Q when m = 3 (GONEN) and Δ = 2 h.
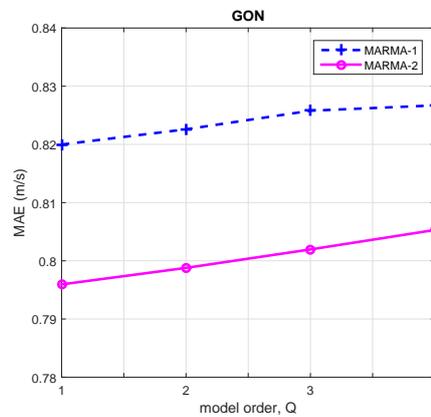

Figure 9. Mean absolute error (MAE) performances of the MARMA-1 and MARMA-2 models with respect to model order Q when m = 3 (GONEN) and Δ = 2 h.

3.2.2. Number of Samples

In order to solve the multichannel linear prediction filter coefficients, the selection of the number of previous samples, *N*, in Equation (10) is another critical parameter. It is observed that increasing the number of *N* after certain value do not improve the forecasting performances of the AR and VAR models as shown in Figure 10. On the other hand, MARMA-2's forecasting performance is better than the other models when relatively large number of previous samples are used. MARMA-2 uses different random noise processes for each channels and if the large number of previous samples are used, this model gives statistically efficient results.
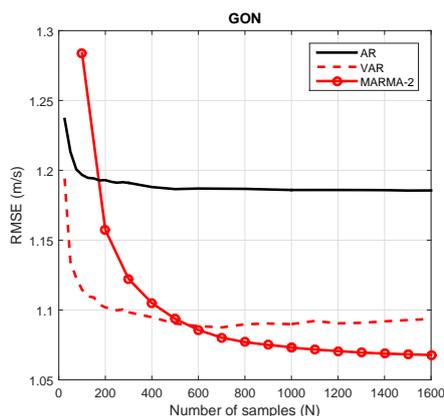
Figure 10. RMSE performances of the AR, VAR and MARMA models with respect to number of previous samples N when m = 3 (GONEN) and $\Delta = 2$ h.

In the following section to make a fair comparison, *N* is selected as 1000 h for all models.

3.2.3. Number of Channels

In this part the effect of number of channels, *M*, in Equations (2), (3) and (4) is investigated. Table 2 shows RMSE performances of important channel selections. For $M = 4$, if we exclude BOZ or IPS from data set, RMSE increases which indicates the significance of these measurements for GON. However, if we exclude SIL, which has minimum correlation value with GON, RMSE value almost unchanged which verifies the correlation values in Figure 4.

Table 2. RMSE and MAE performances of MARMA-2 for different M when $m = 3$ (GON), $\Delta = 2$ and $P = 4$.

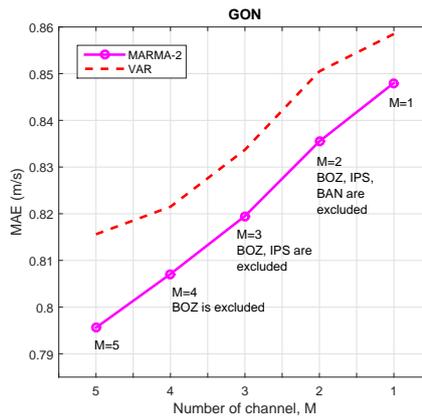| Number of channel, M	MAE	RMSE
M = 5	0.7956	1.0721
M = 4 BOZ(1) is excluded	0.8071	1.0909
M = 4 IPS(2) is excluded	0.8035	1.0898
M = 4 BAN(4) is excluded	0.7994	1.0765
M = 4 SIL(5) is excluded	0.7968	1.0730
M = 3 BOZ(1), IPS(2) are excluded	0.8195	1.1146
M = 3 IPS(2), SIL(5) are excluded	0.8040	1.0891
M = 3 BAN(4), SIL(5) are excluded	0.8003	1.0763
M = 2 BOZ(1), IPS(2), BAN(4) are excluded	0.8355	1.1403
M = 2 IPS(2), BAN(4), SIL(5) are excluded	0.8142	1.1017
M = 1 All other channels are excluded	0.8480	1.1699

Figure 11. MAE performances of the VAR and MARMA-2 models with respect to number of channel M when m = 3 (GONEN) and Δ = 2 h.

Figure 11 shows the RMSE performances with respect to the channel number, *M*. As it is seen, if the used channel is decreased the forecasting error demonstrate a steady increase.

3.2.4. Forecasting Results

Table 3 shows the RMSE and MAE of the target station ($m = 3$) according the forecast lead time (Δ). The multi-channel MARMA-2 has better RMSE and MAE performance than the persistence, AR and VAR models. Table 3 also show that when the lead time period is increased the MARMA models have much better performance than the others.

Table 3. RMSE and MAE performances of the persistent, AR, VAR and MARMA models according the forecast lead time (Δ).

Model	$\Delta = 1$ h		$\Delta = 2$ h		$\Delta = 3$ h		$\Delta = 4$ h	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Persistent	0.8998	1.2443	1.0304	1.4148	1.1686	1.5842	1.2851	1.725
AR	0.7151	0.9947	0.8582	1.1909	0.9893	1.3578	1.1073	1.5081
VAR	0.6962	0.9438	0.8156	1.0984	0.9215	1.2252	1.0127	1.3348
MARMA-2	0.6854	0.9301	0.7956	1.0721	0.8879	1.1834	0.9588	1.2700

Table 4 shows the percentage improvements of the AR, VAR and MARMA-2 methods over persistence method. It is observed that the proposed MARMA-2 method has the best performance and approximately 2.6% more improvements on the average than the multichannel VAR method. It is also seen in Table 4, the multichannel (spatio-temporal) models (VAR and MARMA) which using the neighbouring measurements have significant improvements according the only temporal AR model.

Table 4. The percentage improvements in MAE of the models with respect to persistence model.

Model	$\Delta = 1$	$\Delta = 2$	$\Delta = 3$	$\Delta = 4$	**Average**
AR	20.52%	16.71%	15.34%	13.83%	16.60%
VAR	22.62%	20.84%	21.14%	21.19%	21.44%
MARMA-2	23.83%	22.79%	24.02%	25.40%	24.01%

The average execution times of the used and the proposed methods are given in Table 5 for a single Δ hour ahead forecasting. The used desktop computer has Intel Core(TM) i7-3770K CPU @ 3.50 GHz Processor and 16 GB RAM. Since all the single and multi-channel models are linear and

uses efficient linear least square techniques, the observed execution times in table are less than one second with an ordinary desktop computer. It is possible to forecast very short term (in seconds) wind speed values with the proposed spatio-temporal linear MARMA model.

Table 5. The execution times of the used and the proposed models in milliseconds (ms).

Model	Prediction Time
Persistence	0.021 ms
AR	0.712 ms
VAR	1.500 ms
MARMA-2	31.20 ms

4. Conclusions

In this study, the spatio-temporal (multi-channel) linear models, which use the neighbouring measurements around the target location, are investigated for short-term wind speed forecasting problem. The problem formulation of the multi-channel ARMA models (called as MARMA) are presented and efficient multi-channel prediction coefficient estimation techniques are revised. The proposed MARMA models and solution techniques are tested using hourly averaged real wind values from the five station around Canakkale region of Turkey. The forecasting RMSE and MAE's of the MARMA-2 model is compared with the persistence, AR and multi-channel AR (VAR) methods. As a result, considerable improvements are observed compared to well known temporal persistence (24.01% improvement) and AR (7.41% improvement) methods. The proposed MARMA-2 model gives 2.6% better results than the spatio-temporal VAR method. It is shown that MARMA-2's performance is continuously improved when number of previous samples (N) and filter order are increased unlike the other models. It should be also noted that since the proposed MARMA model moves on the data set using the N previous available samples, it can also adapt the seasonal variations. It is also shown that the proposed multi-channel linear model can predict Δ hour wind speed value using an ordinary desktop computer in milliseconds which is suitable for very short term (in seconds) wind speed forecasting.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Filik, Ü.B.; Filik, T.; Gerek, Ö.N. New electric transmission systems—Experiences from Turkey. In *The Handbook of Clean Energy Systems*; Yan, J.Y., Ed.; John Wiley and Sons Ltd: New York, NY, USA, 2015; Volume 4, pp. 1981–1994.
2. Filik, Ü.B.; Gerek, Ö.N.; Kurban, M. A novel modeling approach for hourly forecasting of long-term electric energy demand. *Energy Convers. Manag.* **2011**, *52*, 199–211.
3. Moldan, B.; Janouskova, S.; Hak, T. How to understand and measure environmental sustainability: Indicators and targets. *Ecol. Indic.* **2012**, *17*, 4–13.
4. Farhangi, H. The path of a smart grid. *IEEE Power Energy Mag.* **2010**, *8*, 18–28.
5. Tastu, J.; Pinson, P.; Trombe, P.J.; Madsen, H. Probabilistic forecasts of wind power generation accounting for geographically dispersed information. *IEEE Trans. Smart Grid* **2014**, *5*, 480–489.
6. Bracale, A.; De Falco, P. An Advanced Bayesian Method for Short-Term Probabilistic Forecasting of the Generation of Wind Power. *Energies* **2015**, *8*, 10293–10314.
7. Lei, M.; Shiyan, L.; Jiang, C.W.; Liu, H.L.; Zhang, Y. A review on the forecasting of wind speed and generated power. *Renew. Sustain. Energy Rev.* **2009**, *13*, 915–920.

8.　Zhu, X.X.; Genton, M.G. Short Term Wind Speed Forecasting for Power System Operations. *Int. Stat. Rev.* **2012**, *80*, 2–23.

9.　De Giorgi, M.G.; Campilongo, S.; Ficarella, A.; Congedo, P.M. Comparison Between Wind Power Prediction Models Based on Wavelet Decomposition with Least-Squares Support Vector Machine (LS-SVM) and Artificial Neural Network (ANN). *Energies* **2014**, *7*, 5251–5272.

10.　Jung, J.; Broadwater, R.P. Current status and future advances for wind speed and power forecasting. *Renew. Sustain. Energy Rev.* **2014**, *31*, 762–777.

11.　Zhang, Y.; Wang, J.X.; Wang, X.F. Review on probabilistic forecasting of wind power generation. *Renew. Sustain. Energy Rev.* **2014**, *32*, 255–270.

12.　Tascikaraoglu, A.; Uzunoglu, M. A review of combined approaches for prediction of short-term wind speed and power. *Renew. Sustain. Energy Rev.* **2014**, *34*, 243–254.

13.　Sun, W.; Liu, M.; Liang, Y. Wind Speed Forecasting Based on FEEMD and LSSVM Optimized by the Bat Algorithm. *Energies* **2015**, *8*, 6585–6607.

14.　Potter, C.W.; Negnevitsky, M. Very short-term wind forecasting for Tasmanian power generation. *IEEE Trans. Power Syst.* **2006**, *21*, 965–972.

15.　Gneiting, T. Larson, K.; Westrick, K.; Genton, M.G.; Aldrich, E. Calibrated probabilistic forecasting at the stateline wind energy center: The regime-switching space time method. *J. Am. Stat. Assoc.* **2006**, *101*, 968–979.

16.　Hering, A.S.; Genton, M.G. Powering up with space-time wind forecasting. *J. Am. Stat. Assoc.* **2010**, *105*, 92–104.

17.　He, M.; Yang, L.; Zhang, J.S.; Vittal, V. A spatio-temporal analysis approach for short-term forecast of wind farm generation. *IEEE Trans. Power Syst.* **2014**, *29*, 1–12.

18.　Xie, L.; Gu, Y.Z.; Zhu, X.X.; Genton, M.G. Short-term spatio-temporal wind power forecast in robust look-ahead power system dispatch. *IEEE Trans. Smart Grid* **2014**, *5*, 511–520.

19.　Dowell, J.; Weiss, S.; Hill, D.; Infield, D. Short term spatio temporal prediction of wind speed and direction. *Wind Energy* **2014**, *17*, 1945–1955.

20.　Mohandes, M.A.; Rehman, S.; Rahman, S.M. Spatial estimation of wind speed. *Int. J. Energy Res.* **2012**, *36*, 545–552.

21.　Pourhabib, A.; Huang, J.Z.; Ding, Y. Short-term Wind Speed Forecast Using Measurements from Multiple Turbines in a Wind Farm. *Technometrics* **2015**, *58*, 138–147.

22.　Sfetsos, A. A comparison of various forecasting techniques applied to mean hourly wind speed time series. *Renew. Energy* **2000**, *21*, 23–35.

23.　Riahy, G.H.; Abedi, M. Short term wind speed forecasting for wind turbine applications using linear prediction method. *Renew. energy* **2008**, *33*, 35–41.

24.　Brown, B.G.; Katz, R.W.; Murphy, A.H. Time series models to simulate and forecast wind speed and wind power. *J. Clim. Appl. Meteorol.* **1984**, *23*, 1184–1195.

25.　Hill, D.C.; McMillan, D.; Bell K.R.W.; Infield, D. Application of auto-regressive models to UK wind speed data for power system impact studies. *IEEE Trans. Sustain. Energy* **2012**, *3*, 134–141.

26.　Shukur, O.B.; Lee, M.H. Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. *Renew. Energy* **2015**, *76*, 637–647.

27.　Cadenas, E.; Rivera, W. Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA ANN model. *Renew. Energy* **2010**, *35*, 2732–2738.

28.　Hernandez, L.; Baladron, C.; Aguiar, J.M.; Calavia, L.; Carro, B.; Sanchez-Esguevillas, A.; Perez, F.; Fernandez, A.; Lloret, J. Artificial Neural Network for Short-Term Load Forecasting in Distribution Systems. *Energies* **2014**, *7*, 1576–1598.

29.　Swami, A.; Giannakis, G.; Shamsunder, S. Multichannel ARMA processes. *IEEE Trans. Signal Process.* **1994**, *42*, 898–913.

30.　Moulines, E.; Duhamel, P.; Cardoso, J.F.; Mayrargue, S. Subspace methods for the blind identification of multichannel FIR filters. *IEEE Trans. Signal Process.* **1995**, *43*, 516–525.

31.　Yu, C.P.; Zhang, C.S.; Xie, L.H. Blind identification of multi-channel ARMA models based on second-order statistics. *IEEE Trans. Signal Process.* **2012**, *60*, 4415–4420.

32. Filik, T. A Multidimensional Short-Term Wind Speed/Directional Analysis Based on Spatial Covariance Matrix Model. In Proceedings of the Turkish-German Conference on Energy Technologies, METU, Ankara, Turkey, 13–15 October 2014.

33. Filik, T. An Improved Multichannel Spatial Correlation Matrix Based ARMA Model for Short-Term Wind Forecasting. In Proceedings of the 5th International 100% Renewable Energy Conference (IRENEC-2015), Istanbul, Turkey, 28–30 May 2015.

34. Van Trees, H.L. *Detection, Estimation, and Modulation Theory, Optimum Array Processing*; John Wiley and Sons Ltd: New York, NY, USA, 2004.

35. Kay, S.M. *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*; Prentice Hall: Upper Saddle River, NJ, USA, 1993.

36. Bjorck, A.; Elfving, T. Accelerated projection methods for computing pseudoinverse solutions of systems of linear equations. *BIT Numer. Math.* **1979**, *19*, 145–163.

37. Golub, G.H.; Van Loan, C.F. *Matrix Computations*; The Johns Hopkins University Press (JHU Press): Baltimore, MD, USA, 2012; Volume 3.

38. Filik, T. Wind Forecasting, Multichannel Wind Speed and Direction Data. Available online: http://www.eem.anadolu.edu.tr/tansufilik/EEM%20547/icerik/wind_forecasting.htm (accessed on 29 February 2016).

39. Parkes, J.; Tindal, A.; Works, S.V.S.; Lane, S. Forecasting short term wind farm production in complex terrain. In Proceedings of the 2004 European Wind Energy Conference, London, UK, 22–25 November 2004.

40. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723.

41. Rissanen, J. Modeling by shortest data description. *Automatica* **1978**, *14*, 465–471.