One-Day ahead Wind Speed/Power Prediction Based on Polynomial Autoregressive Model

Oktay Karakuş[1,*], Ercan E. Kuruoğlu[2], Mustafa A. Altınkaya[1]

[1]Department of Electrical-Electronics Engineering, Izmir Institute of Technology, İzmir, Turkey.

[*]oktaykarakus@iyte.edu.tr

[2]Institute of Science and Technology of Information, "A. Faedo", CNR, G. Moruzzi 1, 56124, Pisa, Italy.

Abstract: *Wind* has been one of the popular renewable energy generation methods in the last decades. Foreknowledge of power to be generated from wind, is crucial especially for planning and storing the power in a wind farm. It is evident in various experimental data that wind speed time series has nonlinear characteristics. It has been reported in the literature that nonlinear prediction methods such as ANN and ANFIS, perform better than linear AR and ARMA models. *Polynomial* AR (PAR) models, despite being nonlinear, are simpler to implement when compared to other nonlinear AR models due to their linear-in-the-parameters property. In this study, a PAR model is used for one-day ahead wind speed prediction by using the past hourly average wind speed measurements of Çeşme and Bandon and performance comparison studies between PAR model and ANN and ANFIS models are performed. In addition, wind power dataset which was published for Global Energy Forecasting Competition 2012 have been used to make power predictions up to 24 hours. Despite having lower number of model parameters, PAR models outperform all other models for both of the locations in speed predictions as well as in power predictions when the prediction horizon is longer than 12 hours.

1. Introduction

Wind speed or wind power prediction plays a key role in wind power management for governments and energy companies. A reliable and highly precise prediction provides better control and planning of power, e.g. daily and hourly scheduling, transmission, storage and balance of power and other managing related operations [1]. When it is not possible to measure and apply prediction operation directly to wind power, wind speed prediction becomes important since it is directly related to power. Short term wind speed prediction is crucial for preventing damage to wind farms due to sudden wind gust whereas long term prediction serves especially for planning of generated power. Moreover, predicting directly from power time series eliminates extra errors resulting from cubic conversion of erroneous wind speed estimates.

Several forecasting horizons can be used in power markets for both management and trading issues. For example, for management issues, from 2 to 3 days ahead predictions are required for systems like steam turbines, etc. Up to 6 hours predictions can be enough for systems with lower complexity like gas turbines. Predictions which lie within 48 hours are generally required for trading issues. However, longer or shorter horizons can be chosen according to the application and area in which the predictions will be used [2–4].

Wind speed/power forecasting models can be listed in two groups which are physical and statistical models. Physical models perform wind forecasting by using meteorological and atmospherical instances. *Numerical weather prediction* (NWP) is a popular approach in physical modelling which uses meteorological instances and performs predictions via solving complex mathematical models about physical equations of the atmosphere. Its performance is relatively well especially in stable weather conditions and is lower especially for short term because of the uncertainties on initial atmospheric conditions and parameterizations [5, 6].

There are several statistical models for wind speed/power prediction in the literature which can be listed as *the persistence model* [7], *linear time series based models* [8–10], *artificial neural networks* (ANN) [11, 12] and *adaptive neuro fuzzy inference systems* (ANFIS) [13, 14] . Persistence model is the simplest one among these models and can be defined as the naive predictor in meteorological studies [7] which makes predictions under the assumption that the future value (prediction) will be the same as the current measurement.

Linear time series based models have been frequently used in previous works on wind speed/power prediction. Those works which use *autoregressive* (AR), *autoregressive moving average* (ARMA) and *autoregressive integrated moving average* (ARIMA) or Seasonal-ARIMA (SARIMA) based models, have demonstrated a limited degree of success in predicting short-term (especially from 10 minutes up to 6 hours) wind speed/power. In particular, ARIMA(2,0,0) (equivalently AR(2)) model has performed better than *feed forward* ANN (FF-ANN) with different validation techniques at prediction horizon between 10 minutes and 4 hours [8]. Furthermore, ARMA models have been compared to ANN models in short term prediction [9] and achieved lower values of error measures than the ANN models in the simulations. In another study [10], short term prediction performance of ARMA models has been compared to that of the persistence model and these models performed better than persistence at different locations in Spain. These studies have clearly showed that linear time series models achieve considerably good performance and they are preferable models for short term over persistence model and several ANN.

When the prediction horizon is longer than 6 hours, linear time series methods fall short due to the nonlinear character of wind speed/power. In order to overcome this, various nonlinear methods such as ANN and ANFIS are used for both short and long term prediction in the literature. These methods outperform linear time series methods for various wind speed and wind power data from various locations in the world. Data mining approaches and hybrid systems have been proposed in studies [15–17] in order to predict wind speed/power up to 1-3 days and they outperform linear and some specific nonlinear methods. Furthermore, Gaussian processes [18], Random Forest and bagging trees [19, 20] are also utilized in prediction studies.

Nonlinear time series methods are also used for wind speed/power prediction studies. A Volterra model based prediction scheme trained by a *recursive* ANN (R-ANN) has been utilized in [21] and a pattern based hybrid wind speed forecasting scheme with a Hammerstein-AR model, has been proposed [1]. In addition, in other studies *nonlinear autoregressive moving average with exogenous inputs* (NARMAX) models have been proposed [22, 23].

PAR models are based on the Volterra series expansion (1), which has been applied successfully in areas such as biological systems [24], industrial plant control systems [25], communications [26] and seismology [27]. PAR processes differ from the other nonlinear AR processes in that they are *linear-in-the-parameters* and thus many mathematical applications developed for linear models can be employed without much difficulty [28].

In a preliminary work presented as conference communication [29], PAR models performance was compared to linear models. In this study, we extend our comparisons to state of the art nonlinear models such as ANN and ANFIS models whose number of parameters and computational complexities are relatively higher than the proposed model, PAR. We also enrich our simulations on various experimental data: in addition to hourly average wind speed data of Çeşme (İzmir, Turkey), data from another location in USA, Bandon (Oregon) are used for this study. Performance comparison is provided visually for two different periods (July-August-September and January-February-March) using error measures such as *normalized root mean square error* (NRMSE), *normalized mean absolute percentage error* (NMAPE) and *bias*.

In this paper, in addition to wind speed predictions, we have performed a wind power prediction study directly using wind power time series. For this purpose, globally reachable wind power measurements which was published for Global Energy Forecasting Competition 2012 [30] have been utilized for the prediction study and performances of models have been compared by error metrics NRMSE, NMAPE and bias.

Rest of the paper is organized as follows: PAR models and representative state of the art models, ANN and ANFIS models are examined in Section 2. Section 3 exhibits estimation methods and error measures used in prediction procedure. Wind data information, the simulation results and error analysis have been provided in Section 4. Section 5 concludes the paper with a discussion of experimental results.

2. Models

2.1. PAR Models

PAR models can be represented as [28]:

$$x(l) = \mu + \sum_{i}^{k} a_i^{(1)} x(l-i) + \sum_{i}^{k}\sum_{j}^{k} a_{i,j}^{(2)} x(l-i)x(l-j) + \ldots + \sum_{i,\ldots}^{k,\ldots} a_{i,\ldots}^{(p)} x(l-i)\ldots + \epsilon(l), \quad (1)$$

where μ is the intercept, $\epsilon(l)$ is an *independent and identically distributed* (i.i.d.) excitation sequence with distribution $\mathcal{N}(0, \sigma_\epsilon^2)$, $a_i^{(1)}$, $a_{i,j}^{(2)}$ and $a_{i,\ldots}^{(p)}$ are coefficients for 1^{st}, 2^{nd} and p^{th} order polynomials, respectively, p is the nonlinearity degree and k is the AR order. We represent a PAR model with the notation: $\text{P}^{(p)}\text{AR}(k)$.

A $\text{P}^{(p)}\text{AR}(k)$ model given by (1) can be expressed in matrix-vector form by using the *linear-in-the-parameters* property:

$$\mathbf{x} = \mu + \mathbf{X}\mathbf{a}^{(p,k)} + \boldsymbol{\epsilon}, \quad (2)$$

where $\mathbf{x}$ is an $n \times 1$ data vector, $\mathbf{X}$ is an $n \times w$ matrix of past samples and polynomial products of the data, $\mathbf{a}^{(p,k)}$ is a $w \times 1$ coefficient vector, $\boldsymbol{\epsilon}$ is an $n \times 1$ vector of excitation sequence, n is the data length and w is the number of model coefficients. The coefficient vector $\mathbf{a}^{(p,k)}$ and the matrix $\mathbf{X}$ can be defined as in equations (3) and (4), respectively.

$$\mathbf{a}^{(p,k)} = \left[a_1^{(1)}, a_2^{(1)}, \ldots, a_k^{(1)}, a_{1,1}^{(2)}, a_{1,2}^{(2)}, \ldots, a_{k,k}^{(2)}, \ldots, \underbrace{a_{k,k,\ldots,k}^{(p)}}_{\text{p}^{\text{th}}\text{ order}} \right]^T, \quad (3)$$

$$\mathbf{X} = \begin{bmatrix} x(0) & x(-1) & \ldots & x(1-k) & x^2(0) & x(0)x(-1) & \ldots & x^2(1-k) & \ldots & x^p(1-k) \\ x(1) & x(0) & \ldots & x(2-k) & x^2(1) & x(1)x(0) & \ldots & x^2(2-k) & \ldots & x^p(2-k) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x(n-1) & x(n-2) & \ldots & x(n-k) & x^2(n-1) & x(n-1)x(n-2) & \ldots & x^2(n-k) & \ldots & x^p(n-k) \end{bmatrix}. \quad (4)$$

AR models are subsets of PAR models. So, $\text{P}^{(p)}\text{AR}(k)$ models are equivalent to $\text{AR}(k)$ models when the nonlinearity degree, p, is set to 1.

2.2. ANN Models

ANN are statistical learning models which are inspired by central nervous system of animals and composed of elements called nodes, a.k.a. *neurons*. Interconnected neurons exchange messages between each other. These systems are trained via various training algorithms and this learning process updates weights of interconnections. ANN are formed of *layers* which are composed of a number of interconnected neurons and every neuron contains an activation function which is a zero memory nonlinear function. Inputs are fed to the *input layer* which is connected to one or more *hidden layers*. The actual processing is provided by hidden layers via a system of weighted connections. The *output layer* which is linked to the hidden layers provides the final result. Five different ANN structures will be used in this study. These are: *multilayer feed forward* (MLFF), *multilayer perceptron* (MLP), Elman recurrent, *radial basis function* (RBF) and *linear* (for details, see studies [13, 31]). Inputs are past hourly average wind speed/power data and the output is the predicted wind speed/power data. For example, if the number of inputs is r, then the corresponding inputs for the one-step ahead predicted output $\hat{y}(l)$ are $\{y(l-1), y(l-2), \ldots, y(l-r)\}$.

2.3. ANFIS Models

ANFIS models are combinations of two statistical systems and create a successful hybrid system of neural networks and fuzzy inference systems. The working principle of an ANFIS model consists of two steps [13]: i) the system is trained in a similar manner with ANN, ii) trained system will then operate as a fuzzy inference system. ANFIS model is based on Takagi–Sugeno fuzzy inference system [32]. Since ANFIS integrate both ANN and fuzzy logic principles, the advantages of both systems are integrated in a single system. FIS part has the capability to approximate nonlinear functions via a set of *if-then* rules [33]. Basically, ANFIS are constructed with 5 layers. Each layer is formed of nodes which are described by the node functions. The outputs of previous layers are inputs of next layers. ANFIS use either backpropagation or a combination of least squares estimation and backpropagation for the membership function parameter estimation. In addition, ANFIS models have both linear parameters of the consequent *membership function* (MF) and nonlinear parameters of the antecedent MF. The number of consequent parameters depends on the number of rules, which in turn depends on the number of inputs. However, the number of antecedent parameters depends on the number of MFs per input and the MF type in addition to the number of inputs [34].

3. Estimation Methods

3.1. Multistep Ahead Prediction

Prediction study in this paper requires a multistep ahead prediction procedure in order to make predictions up to 24-hours ahead. Predictions have been made for all the models in this study (AR, PAR, ANN and ANFIS) by employing the procedure explained below.

A time series $\mathbf{Y} = [y_1, y_2, \ldots, y_n]$, can be defined as a sequence of recorded observations at each time step. If we take a segment of data with length q up to time l, this segment will be $\mathbf{Y}_{l-q+1}^{l} = [y_l, y_{l-1}, \ldots y_{l-q+1}]$. T-step ahead prediction of $\mathbf{Y}$ is represented as $\mathbf{Y}_{l+1}^{l+t}$ and will be a function of q past values. This can be defined as; [35]

$$\mathbf{Y}_{l+1}^{l+t} = f(y_l, y_{l-1}, \ldots y_{l-q+1}) = f(\mathbf{Y}_{l-q+1}^{l}). \tag{5}$$

The key point in multistep ahead prediction is to use only the measurements up to time l. If measurements at time $l+1, l+2, \ldots$ are needed, predicted values will be used instead of these measurements. An example of 3-step ahead prediction have been shown in Figure 1.

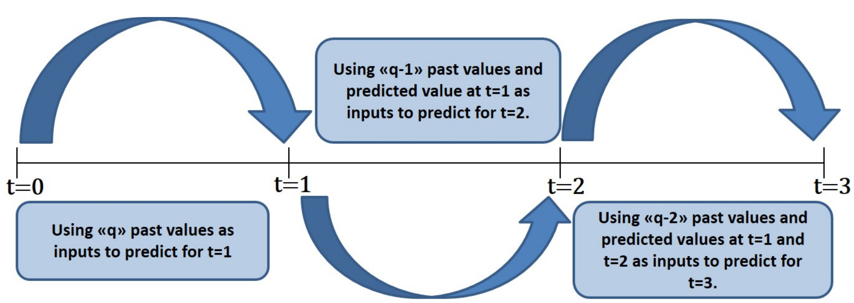


Fig. 1. *3-Step Ahead Prediction Example*

3.2. Coefficient Estimation of PAR models

The *nonlinear least squares* (NLS) estimation generates optimal PAR coefficient estimates under the assumption that the excitation, $\epsilon(l)$, are i.i.d. Gaussian random variables. PAR model coefficients, $\mathbf{a}^{(p,k)}$ which are defined in (3) have been estimated by using the training data for all data sets by employing the NLS method since the data matrix $\mathbf{X}$ includes polynomial products of past data samples. Under the assumption that the model orders p and k are known initially, NLS estimates of PAR model coefficients by [36]:

$$\hat{\mathbf{a}}^{(p,k)}_{\text{NLS}} = (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-\mathbf{1}}\mathbf{X}^{\mathbf{T}}\mathbf{x}. \tag{6}$$

On the other hand, when $\epsilon(l)$ have heavier tails than a Gaussian distribution, computing the variance may not be possible and other approaches can be applied instead of NLS. Using *minimum dispersion* (MD) based estimation methods instead of *minimum mean square error* (MMSE) appears as good alternatives here. *Nonlinear polynomial least L-p norm* (NPLPN) based estimation methods such as *nonlinear iteratively reweighted least squares* (NIRLS) and *normalized polynomial least mean Pth power* (NPLMP), perform optimal coefficient estimations for PAR processes with α-Stable or generalized Gaussian innovations [28].

In this study, we assume a Gaussian excitation for PAR models for simplicity and apply NLS directly since PAR models are linear-in-the parameters models.

3.3. Performance Metrics

There are various error metrics which have been used in the literature to compare prediction performances of the methods. In this study we use NRMSE, NMAPE and bias metrics in order to measure the performances to visualize both L_1 norm (NMAPE) and L_2 norm (NRMSE) based errors and the bias. NRMSE provides a good measure of distortion when the error is Gaussian distributed while NMAPE is successful in dealing with data containing outliers. We have used bias in order to see whether the predictions are under or over estimated. These metrics are defined as:

$$\text{NRMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \widehat{x}_i}{\max(\mathbf{x})}\right)^2}, \quad (7)$$

$$\text{NMAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|x_i - \widehat{x}_i|}{\max(\mathbf{x})} \times 100, \quad (8)$$

$$\text{Bias} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \widehat{x}_i), \quad (9)$$

where $\mathbf{x}$ and $\widehat{\mathbf{x}}$ are the test and the predicted time series of length n, respectively. Best results are shown with lower values of NRMSE and NMAPE.

4. Experimental Analysis

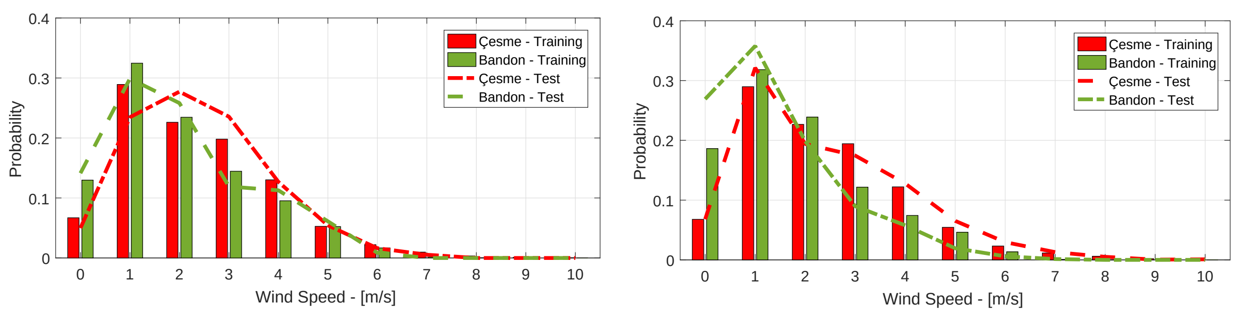
4.1. Data

In this study, wind speed data which are measured at Çeşme (İzmir, Turkey) and Bandon (Oregon, USA), are used to make predictions up to 24-hours ahead. Generally accepted procedure in the literature is to use hourly average values in the studies. In addition, our data sets include only hourly averages. The term ”*hourly average*” corresponds to the average of the samples obtained within an hour with a sampling period of several minutes. Using hourly average values provides computationally cheaper operations in the same time interval. Moreover, short averages generate erratic values than the longer. Thus, using hourly averages is appropriate and the smoothing effect generated by this averaging is assumed to be negligible [37].

In particular, first data set is hourly average wind speed data measured in Çeşme at an altitude of 10 meters by Turkish State Meteorological Service. Bandon data are also hourly average wind speed data and measured by AgriMet (Cooperative Agricultural Weather Network) [38]. The Bandon site is located at 25 meters (80 feets) elevation and at the west coast of Oregon, USA.

For wind speed prediction studies two 3-months-long periods have been selected as test periods in order to measure the performance of the models in different times of the year. In particular, test *period 1* covers January-February-March and *period 2* covers July-August-September and model performances have been measured both for the first months and whole 3 months of the periods. Measurements from 12 and 24 months before these test periods are selected as candidate training data sets. Then, we perform a comparison of the informative content for training sets with different lengths (i.e. 1 year and 2 years) using *Entropy* to quantify the information content of the sequences. As a result, the entropy values for both of the training sets are very close to each other for all the data sets. Hence, using shorter training period will be more appropriate in the sense of computational complexity.

Figure 2 shows hourly average wind speed distributions of test months for 2 locations for both training and test data sets. Also, both of the subfigures depict that the distributions for training and test data have similar characteristics. Specifically, Bandon test data for winter month have quite lower values than training, however, this fact has been compensated by the training data.



(a) Period 1 (b) Period 2

Fig. 2. *Wind Speed Histograms*

Wind power measurements used in this study are from a globally reachable dataset which was published for Global Energy Forecasting Competition 2012 [30] including wind power measurements from several wind farms and electricity load data. Wind power measurements of a single wind farm (*wp1* in the data set excel sheet) have been selected and used in the wind power predictions. Testing data periods are the same dates with wind speed datasets and historical data 12 months before these periods have been selected as training data.

4.2. Modelling Results

Each model explained in previous sections is applied in a training procedure with 1-year length wind speed/power data and 1-month length wind speed/power test data. Both training and test data are normalized in a range between 0 and 1 in order to apply prediction procedures at the same interval for every data set.

4.2.1. Simulated PAR Models: In order to decide the best AR and PAR models in this study, *Akaike information criterion* (AIC) and *Bayesian information criterion* (BIC) methods have been performed to select PAR or AR models for wind speed/power predictions where model space includes nine different PAR models (for p = 1, 2, 3 and k = 1, 2, 3). However, estimated orders by AIC or BIC did not provide the best results in terms of prediction error. Therefore, the "best" PAR and AR models were selected by performing an exhaustive search on models comparing their 24-hours ahead prediction performances. (Please see [39] to estimate best PAR model via AIC and BIC.)

For Çeşme wind speed, $P^{(2)}AR(3)$ and AR(3) models have been found to be the best PAR and AR models according to their NRMSE values for both 24-hours ahead prediction, where for Bandon, $P^{(2)}AR(2)$ and AR(2) models have been selected. For power prediction studies $P^{(2)}AR(1)$ and AR(1) models achieve the lowest NRMSE values.

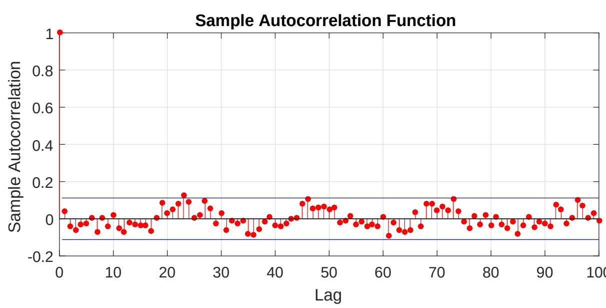
PAR/AR model coefficients for each model in question, have been estimated via NLS method by using the exhaustively estimated model orders. Moreover, in training period, PAR models have been tested for zero and non-zero intercepts and PAR models with zero intercept achieve better prediction results than that of non-zero intercept. That's why the intercept, μ, has been selected as zero for all the datasets in this study. PAR model excitation sequence ϵ_n, is a Gaussian process with mean 0 and variance σ_ϵ^2. Excitation variance values for each data set have been estimated in training process.

By using all these estimated and initially defined parameters, PAR model equation in (1) has been used to make predictions by taking k past wind speed samples as its input. 1000 Monte Carlo simulations have been performed and averages have been used as the predictions in order to obtain stable results.

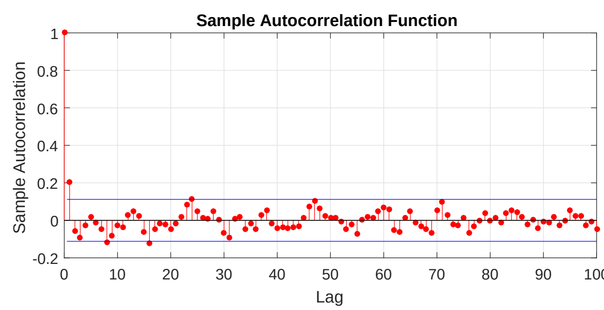
4.3. Validation of PAR models

In order to validate the PAR models estimated for prediction, firstly a statistical significance study has been run based on the residual autocorrelations. Each training data set has a length of 8760 samples and to perform residual autocorrelations a 1-month-long (720 samples) period inside of the training set has been randomly selected for each data set.

In Figure 3, residual autocorrelations have been depicted for 2 example data sets (Çeşme Period 1 and Power Period 2). Examining both of the subfigures, it can be clearly stated that the residual autocorrelations are generally below the limits and are contained. However, they both show a cyclic character with a period of 24 hours (one day). Notwithstanding, correlation values of the residuals for each lag are low. In order not to repeat the same results, we only show 2 out of 6 figures. All other 4 data sets show same characteristics like the examples. Thus, we can extend the comments above to all data sets that we utilized in the study.



(a) Çeşme (period 1)



(b) Power (period 2)

Fig. 3. *Residual autocorrelations for randomly selected periods with length of 720 samples. Limits in the figures refer to* $\pm 3\sigma$.

Apart from residual autocorrelations, PAR models obtain reasonable fitting results which are measured using Adjusted-R^2 and F-Test. Both tests measure the fitting performance of an estimated model. Estimated PAR models ($P^{(2)}AR(3)$ for Çeşme, $P^{(2)}AR(2)$ for Bandon and $P^{(2)}AR(1)$ for Power) for each data set have achieved at least 65% Adjusted-R^2 score while rival methods also achieve nearly the same values. Although all the models have p-values lower than 0.0001 after F-Test, F-score of PAR model is higher and it can be stated that its fitting performance is slightly better than the others.

Table 1. *ANN simulation parameters*

	MLFF	RBF(1)	MLP	Elman(2)	L-ANN
Training function	TRAINLM(5)	—	TRAINLM	TRAINLM	TRAINGDX(5)
Learning function	LEARNGD(5)	LEARNGD	LEARNP(5)	LEARNGD	LEARNGD
Number of layers	4	3	3	3	2
Neurons – inputs(3) (I)	3/1	3/1	3/1	3/1	3/1
Neurons - Hidden 1(4) (H_1)	10/20	Variable (max. 30)	10/20	10/20	—
Neurons - Hidden 2(4) (H_2)	10/20	—	—	—	—
Notation	I-H_1-H_2-1	I-H_1-1	I-H_1-1	I-H_1-1	I-1
Activation func. (hidden)	TANSIG(5)	RADBAS(5)	TANSIG	TANSIG	—
Activation func. (output)	PURELIN	PURELIN	PURELIN	PURELIN	—
Epochs	200	200	200	200	200

(1) Neurons are iteratively added to the RBF network one by one until the sum-squared error falls beneath 0.005 or a maximum number of neurons has been reached.
(2) Elman R-ANN model is selected with 1 feedback input.
(3) 1 input neuron has been used for wind power predictions.
(4) 20 hidden neurons have been used for wind power predictions.
(5) LM: Levenberg–Marquardt algorithm, GDX: gradient descent with momentum and adaptive learning rate backpropagation, GD: gradient descent, P: perceptron, TANSIG: hyperbolic tangent sigmoid, RADBAS: radial basis function.

4.3.1. Simulated ANN and ANFIS models: ANN models are implemented by using the parameters given in Table 1. ANN and ANFIS model structures are created in Matlab®'s ANN and FIS toolboxes and models are oriented to minimize the mean square error (MSE). Some of the models and their parameters for ANN and ANFIS models have been taken from a previous study [13] on wind prediction using several ANN and ANFIS models. In addition to the ANN models in Table 1, two ANFIS structures with 3 inputs have been implemented where both use *linear* output membership function and *hybrid* optimization method. The only difference is the MF types in question where ANFIS-1 uses *generalized bell* (GBELL) and ANFIS-2 employs *triangular* shaped MFs.

Number of inputs have been decided on after an exhaustive search for each model (ANN and ANFIS) between 1 and 6 by choosing the one giving the best performance. Consequently, number of inputs have been decided as 3 in wind speed prediction for both ANN and ANFIS for both locations. For wind power prediction, number of inputs are selected for ANN models as 1 for both of the test months.

Table 2 depicts the number of adjustable parameters for each simulated model. Linear models AR and L-ANN have a small number of parameters. PAR models are nonlinear models but their complexity is low when compared to other nonlinear models in terms of number of adjustable

Table 2 Number of adjustable parameters for the simulated models

Model	Number of parameters	Model	Number of parameters	Model	Number of parameters	Model	Number of parameters
AR(1)	1	$P^{(2)}$AR(3)	9	L-ANN (1-1)	2	MLFF (3-10-10-1)	161
AR(2)	2	Elman (3-10-1)	151	L-ANN (3-1)	4	MLFF (1-20-20-1)	481
AR(3)	3	Elman (1-20-1)	461	MLP (3-10-1)	51	ANFIS-1	135
$P^{(2)}$AR(1)	2	RBF (1-Var-1)	31-91	MLP (1-20-1)	61	ANFIS-2	135
$P^{(2)}$AR(2)	5	RBF (3-Var-1)	51-151				

parameters. In addition, complexity of the NLS used for estimating the parameters of PAR and AR models is lower than those of the training algorithms which ANN and ANFIS models have employed.

4.4. Wind Speed Prediction Results

Tables between 3 and 6 show the performance of all the implemented models for 4 different cases which are Çeşme-period1, Bandon-period1, Çeşme-period2 and Bandon-period2, respectively. In these tables, one PAR model is indicated for each case. These models are the best PAR models which achieve the lowest NRMSE value for 24-hour ahead prediction as stated in section 4.2.1.

Examining the values in Table 3, the performance of $P^{(2)}$AR(3) model is clearly much better than all other models when prediction horizon is longer than 6 hours. RBF, MLP and ANFIS-2 have achieved the second best performance after PAR models and they are preferable to all other models for a longer prediction horizon. For all prediction horizons, all the models except $P^{(2)}$AR(3) have underestimated the wind speed (bias values are over zero). Below 12 hours linear model AR(3) model is preferable which is a fact that supports the previous studies [8–10] about wind speed prediction that linear models are more appropriate models for short term prediction whereas nonlinear models work better for long term wind speed prediction.

Prediction performance for Bandon in test period 1 shown in Table 4 provides nearly the same results as that for Çeşme. $P^{(2)}$AR(2) model appears to be preferable to all other models. MLFF is the best performing ANN model and ANFIS-1 is the best performing ANFIS model for Bandon data at 24 hours ahead prediction. For all prediction horizons, all the models' predictions have been underestimated for Bandon.

Table 3. *Performance comparison of Çeşme wind speed time series for the test period 1*

	1-Month Test									3-months Test		
	6 Hours			12 hours			24 Hours			24 hours		
	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias
P(2)AR(3)	0,0554	5,6688	**0,0001**	**0,0762**	**8,2396**	-0,0056	**0,1026**	**11,7231**	-0,0187	**0,1015**	**11,7429**	-0,0188
AR(3)	**0,0511**	**5,5591**	0,0361	0,0801	9,2375	0,0649	0,1297	15,5834	0,1108	0,1294	15,4769	0,1100
MLFF	0,0638	6,6579	0,0084	0,0882	9,4067	0,0137	0,1148	12,4311	0,0183	0,1155	12,8892	0,0165
MLP	0,0653	6,8056	0,0047	0,0894	9,4993	0,0072	0,1145	12,3899	0,0083	0,1152	12,8364	0,0064
Elman	0,0666	7,0689	0,0034	0,0920	9,8633	**0,0047**	0,1160	12,6288	**0,0046**	0,1169	13,0848	**0,0026**
L-ANN	0,0800	8,7943	0,0035	0,1045	11,4765	0,0052	0,1245	13,6713	0,0067	0,1257	14,1844	0,0046
RBF	0,0662	6,7510	0,0119	0,0908	9,3135	0,0203	0,1147	12,0531	0,0285	0,1151	12,5064	0,0264
ANFIS-1	0,0677	6,8119	0,0156	0,0926	9,4747	0,0279	0,1178	12,1115	0,0398	0,1177	12,5584	0,0378
ANFIS-2	0,0612	6,5445	0,0086	0,0863	9,4519	0,0158	0,1148	12,5259	0,0234	0,1139	12,9248	0,0209

Table 4. *Performance comparison of Bandon wind speed time series for the test period 1*

	1-Month Test									3-months Test		
	6 Hours			12 hours			24 Hours			24 hours		
	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias
P(2)AR(2)	0,0683	**6,1242**	**0,0235**	**0,0984**	**9,4282**	**0,0293**	**0,1291**	13,4536	**0,0272**	**0,1122**	**11,8619**	**0,0026**
AR(2)	**0,0594**	6,7273	0,0498	0,1069	12,2032	0,0907	0,1798	20,5896	0,1532	0,1574	17,1953	0,1278
MLFF	0,0864	8,8088	0,0311	0,1332	13,4365	0,0489	0,1660	16,9221	0,0566	0,1505	15,7561	0,0153
MLP	0,0865	9,1269	0,0349	0,1377	14,3080	0,0579	0,1799	18,5120	0,0751	0,1618	16,9407	0,0309
Elman	0,0834	8,8819	0,0296	0,1328	13,9725	0,0501	0,1765	18,3858	0,0687	0,1605	17,0611	0,0244
L-ANN	0,0927	9,8636	0,0270	0,1406	14,9318	0,0408	0,1764	18,7244	0,0511	0,1643	17,9314	0,0054
RBF	0,0884	9,3551	0,0365	0,1412	14,7057	0,0623	0,1858	19,1139	0,0848	0,1661	17,2954	0,0397
ANFIS-1	0,0829	7,9212	0,0284	0,1192	11,0680	0,0410	0,1343	**12,7252**	0,0438	0,1215	11,9520	0,0114
ANFIS-2	0,0895	9,3282	0,0348	0,1367	14,4478	0,0567	0,1860	19,2294	0,0793	0,1672	17,5688	0,0336

In Table 5, it has been shown that the performance of $P^{(2)}AR(3)$ outperforms all other models when prediction horizon is over 12 hours. RBF and linear AR(3) perform well for all horizons in test period 2. Generally, most of the models have overestimated the wind speed (bias values are above zero). Moreover, in Table 6, the proposed model $P^{(2)}AR(2)$ shows remarkable performance in all the cases. In addition, linear AR(2) model performs as the second best model when all other models achieve very low prediction performance in test period 2 in Bandon site. Bias results appear similar with Çeşme test period 2 results. Specifically, most of the models overestimated wind speed in all the prediction horizons.

Figures 4-(a) and (b) show the 24-hours ahead prediction results vs. observed test data in a randomly selected week in July. Simulated models are divided into 4 groups, namely PAR, ANN (MLFF, MLP, Elman, RBF), ANFIS and linear (AR and L-ANN). The best model for prediction horizon of 24 hours from each group is selected and plotted. PAR models appear remarkable

Table 5. *Performance comparison of Çeşme wind speed time series for the test period 2*

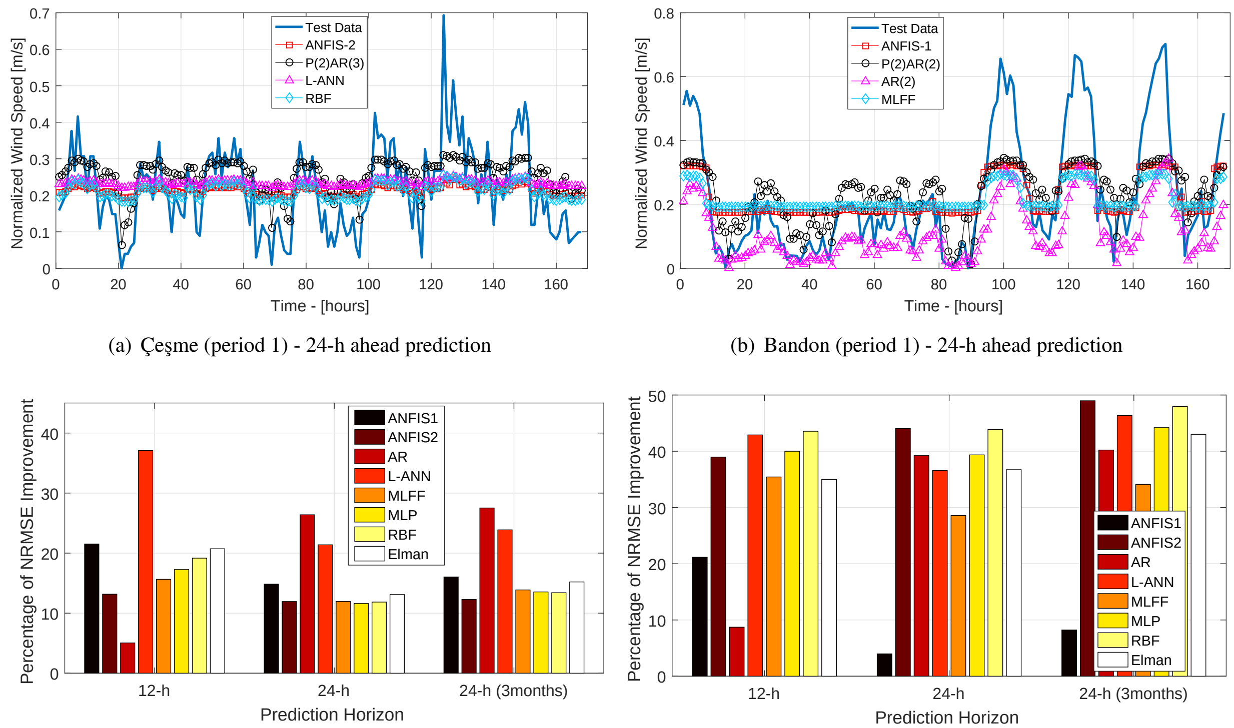
	1-Month Test									3-months Test		
	6 Hours			12 hours			24 Hours			24 hours		
	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias
P(2)AR(3)	0,0436	3,5922	**0,0033**	0,0591	**5,1160**	**0,0015**	**0,0789**	**7,5981**	-0,0061	**0,0782**	**7,8179**	0,0045
AR(3)	**0,0371**	**3,5343**	0,0217	**0,0558**	5,4451	0,0391	0,0894	9,0624	0,0678	0,0972	10,2940	0,0773
MLFF	0,0545	5,8218	-0,0058	0,0785	8,3700	-0,0070	0,1034	11,0578	-0,0094	0,1059	11,3404	0,0102
MLP	0,0580	6,2713	-0,0100	0,0834	8,9522	-0,0133	0,1086	11,6392	-0,0181	0,1077	11,6112	0,0011
Elman	0,0580	6,3409	-0,0117	0,0842	9,1503	-0,0165	0,1101	11,8817	-0,0207	0,1082	11,7624	-0,0013
L-ANN	0,0588	6,3336	-0,0089	0,0848	9,1286	-0,0138	0,1103	11,8562	-0,0186	0,1086	11,7773	**0,0007**
RBF	0,0580	6,6109	-0,0080	0,0766	8,7526	-0,0040	0,0963	10,6481	**0,0032**	0,1072	11,7528	0,0282
ANFIS-1	0,0653	6,7592	-0,0040	0,0927	9,7987	-0,0060	0,1167	12,2866	-0,0068	0,1171	12,5301	0,0141
ANFIS-2	0,0665	6,4903	-0,0092	0,0889	9,0766	-0,0107	0,1107	11,4970	-0,0093	0,1097	11,6950	0,0125

Table 6. *Performance comparison of Bandon wind speed time series for the test period 2*

	1-Month Test									3-months Test		
	6 Hours			12 hours			24 Hours			24 hours		
	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias
P(2)AR(2)	**0,0337**	**2,3707**	**-0,0004**	**0,0496**	**3,9644**	**-0,0043**	**0,0693**	**6,5320**	**-0,0148**	**0,0878**	**7,1058**	**0,0015**
AR(2)	0,0341	3,1922	0,0233	0,0608	5,7136	0,0422	0,1010	9,4877	0,0704	0,1281	11,5965	0,0960
MLFF	0,0728	7,9022	-0,0252	0,1098	12,0676	-0,0396	0,1355	15,7152	-0,0635	0,1441	14,0129	-0,0213
MLP	0,0725	8,0474	-0,0288	0,1108	12,4097	-0,0433	0,1393	16,1842	-0,0652	0,1472	14,3800	-0,0214
Elman	0,0715	7,9189	-0,0270	0,1080	11,9932	-0,0368	0,1327	14,9218	-0,0465	0,1460	13,8345	-0,0020
L-ANN	0,0853	10,2751	-0,0515	0,1293	15,5128	-0,0769	0,1568	18,7881	-0,0928	0,1556	15,8369	-0,0475
RBF	0,0711	7,7178	-0,0253	0,1073	11,6357	-0,0343	0,1307	14,6585	-0,0471	0,1439	13,5601	-0,0038
ANFIS-1	0,0757	8,2571	-0,0306	0,1105	12,2874	-0,0499	0,1297	15,5006	-0,0782	0,1311	13,2829	-0,0439
ANFIS-2	0,0784	8,7240	-0,0315	0,1196	13,6140	-0,0508	0,1468	17,2387	-0,0738	0,1527	15,1032	-0,0285

for this prediction horizon when ANFIS and ANN models did not noticeably capture their performance. Figures 4-(c) and (d) show percentage of NRMSE improvement obtained by the PAR model with respect to other models. Observing the values in Figure 4-(c) for all the prediction horizon values over 12 hours, PAR model improves the prediction performance in NRMSE over ANN and ANFIS models by about 20%. In Figure 4-(d), again PAR model has an improvement over ANN and ANFIS models by about 35%.

In Figures 5-(a) and (b) the 24-hours ahead prediction results for January have been shown for both locations. PAR model outperforms all other models again for winter. Figures 5-(c) and (d) show NMAPE improvement obtained by the PAR model with respect to other models. NMAPE improvement means the reduction in NMAPE obtained by PAR with respect to those of the other simulated models. Observing these values, for all the prediction horizon values, PAR model improves the prediction performance in NMAPE by about 4% for Çeşme and 10% for Bandon.



(a) Çeşme (period 1) - 24-h ahead prediction

(b) Bandon (period 1) - 24-h ahead prediction

(c) Çeşme (period 1) - Percentage of NRMSE Improvement by PAR wrt other models

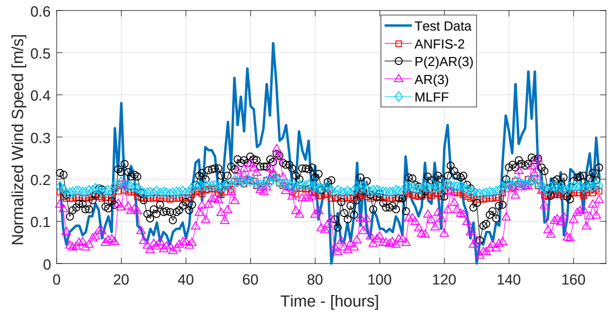
(d) Bandon (period 1) - Percentage of NRMSE Improvement by PAR wrt other models

Fig. 4. *Wind speed prediction comparison figures for the test period 1*

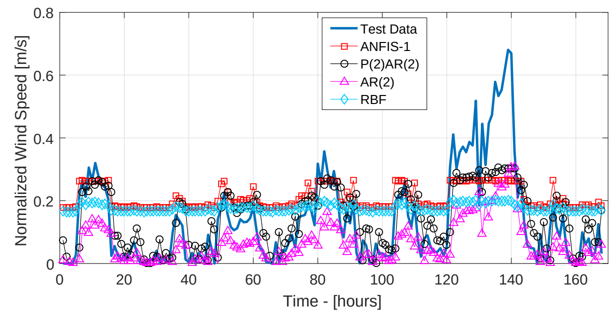
4.5. Wind Power Prediction Results

Examining the Table in 7, the performance of $P^{(2)}AR(1)$ in test period 1 is dominantly better than all other models for all the prediction horizons. Linear AR(1) and MLFF have achieved the second and third best performances, respectively and they are preferable to all other models for a longer prediction horizon. For all the prediction horizons most of the models have overestimated the wind power. Wind power prediction performances in test period 2 are shown in Table 8. As seen clearly, $P^{(2)}AR(1)$ performs better than all other models for all the prediction horizons generally. MLFF and MLP perform considerable for longer horizons.

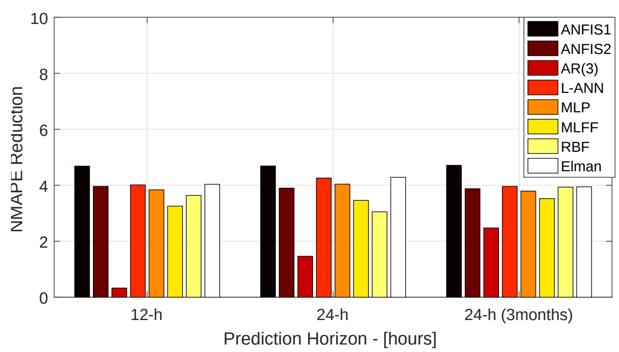
Figure 6-(a) and (b) show the 24-h ahead wind power prediction results vs. observed test data in a week. Examining the figures, PAR and MLFF capture the test data trend better than any other model for 24-hours of prediction horizon. PAR captures lower power values better than MLFF, when MLFF is better for higher power values. Performance of ANFIS and linear models fall behind PAR for both test periods. Figure 6-(c) and (d) show the NMAPE improvement of PAR



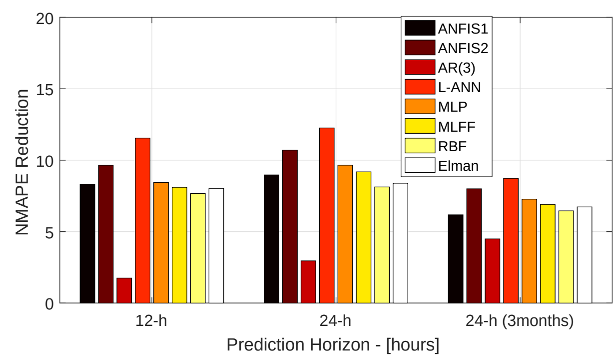
(a) Çeşme (period 2) - 24-h ahead prediction



(b) Bandon (period 2) - 24-h ahead prediction



(c) Çeşme (period 2) - NMAPE Reduction obtained using PAR wrt other models



(d) Bandon (period 2) - NMAPE Reduction obtained using PAR wrt other models

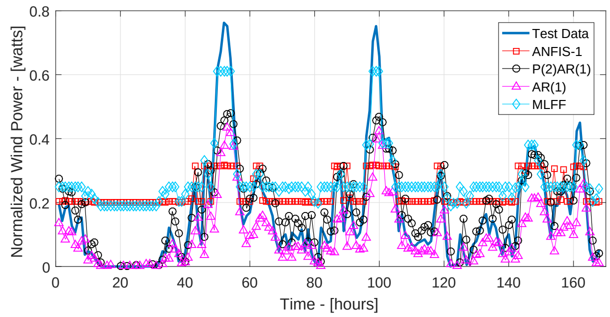
Fig. 5. *Wind speed prediction comparison figures for the test period 2*

over all other models. For 24-hours ahead prediction, NMAPE values of PAR are around 5%-10%, however the rest are around 10%-20%. Thus, improvement of PAR is at least 5% for both seasons.

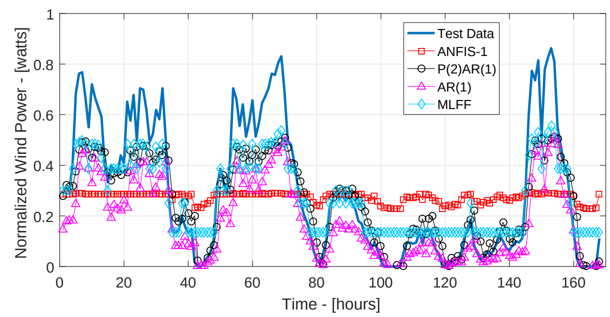
Examining the error metrics for 1 month and 3 months, the results generally follow the same characteristics. Increasing the test length did not improve the performance of the proposed method and the others. However, using two different test length proves the applicability of the proposed method to whole year.

5. Discussion & Conclusion

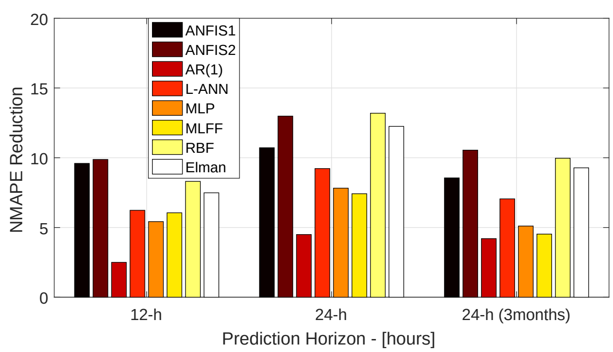
A detailed comparative study has been carried out for PAR model performance on wind speed/power predictions up to 24 hours ahead of time over several ANN, ANFIS and linear AR models. Two data sets for two different locations in Turkey and USA are used to train and test the models. Results are presented in both figures and performance metrics like NRMSE, NMAPE and bias to enrich the conclusion of the study.



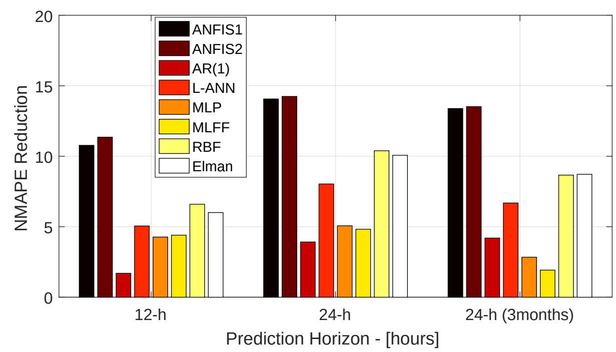
(a) Test period 1 - 24-h ahead prediction



(b) Test period 2 - 24-h ahead prediction



(c) Test period 1 - NMAPE Reduction obtained using PAR wrt other models



(d) Test period 2 - NMAPE Reduction obtained using PAR wrt other models

Fig. 6. *Wind power prediction comparison figures*

In particular, wind speed prediction results are similar to many other studies in the literature which claim that the prediction performance decreases when prediction horizon increases. In addition, linear model performance for longer prediction horizon gets worse than nonlinear models. This fact reveals that nonlinear characteristics of wind speed is more visible when prediction horizon is longer. Examining the performance metric values, linear AR achieves better wind speed prediction results in the first 6-10 hours of prediction horizon while PAR appears to be more successful for prediction horizons over 6-10 hours. RBF, MLFF and MLP show close but inferior performance to PAR for longer horizons. Both ANFIS models achieve similar performance with most of the ANN. Performances of ANN and ANFIS are relatively worse when test period 2 is used for predictions and did not even reach the performance of linear AR.

The proposed method has been used also in wind power predictions. Examining the wind power prediction results shows us that performance of PAR has been distinctive and it outperforms other models for all the cases. Although MLFF and MLP perform worse than PAR, they appear better

Table 7. *Performance comparison of Wind power prediction for Summer month*

	1-Month Test									3-months Test		
	6 Hours			12 hours			24 Hours			24 hours		
	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias
P(2)AR(1)	**0,0210**	**1,5283**	**0,0008**	**0,0359**	**2,7777**	**-0,0002**	**0,0567**	**4,7827**	**-0,0040**	**0,0913**	**5,5466**	**0,0129**
AR(1)	0,0311	2,8262	0,0215	0,0583	5,2799	0,0403	0,1023	9,2787	0,0710	0,1284	9,7502	0,0904
MLFF	0,0462	5,2700	-0,0275	0,0792	8,8320	-0,0512	0,1124	12,2036	-0,0749	0,1163	10,0737	-0,0520
MLP	0,0417	4,6074	-0,0207	0,0713	8,2005	-0,0427	0,1097	12,6027	-0,0714	0,1169	10,6537	-0,0504
Elman	0,0479	5,5514	-0,0215	0,0878	10,2639	-0,0437	0,1455	17,0304	-0,0807	0,1628	14,8224	-0,0477
L-ANN	0,0461	5,1694	-0,0173	0,0804	9,0109	-0,0301	0,1249	14,0076	-0,0469	0,1450	12,5989	-0,0159
RBF	0,0503	5,8857	-0,0241	0,0943	11,0839	-0,0498	0,1537	17,9750	-0,0874	0,1693	15,5102	-0,0533
ANFIS-1	0,0672	7,2553	-0,0228	0,1131	12,3731	-0,0429	0,1435	15,4991	-0,0575	0,1723	14,1025	-0,0208
ANFIS-2	0,0689	7,4776	-0,0214	0,1155	12,6548	-0,0360	0,1584	17,7685	-0,0599	0,1870	16,0925	-0,0199

Table 8. *Performance comparison of Wind power prediction for Winter month*

	1-Month Test									3-months Test		
	6 Hours			12 hours			24 Hours			24 hours		
	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias	NRMSE	NMAPE	Bias
P(2)AR(1)	0,0519	**2,6401**	0,0165	**0,0833**	**4,4254**	0,0259	**0,1199**	**6,8146**	0,0345	0,1467	**9,1705**	0,0548
AR(1)	**0,0451**	3,2898	0,0306	0,0843	6,1249	0,0572	0,1483	10,7362	0,1006	0,1726	13,3679	0,1263
MLFF	0,0633	5,5255	-0,0132	0,1019	8,8270	-0,0320	0,1332	11,6392	-0,0535	**0,1316**	11,0963	-0,0245
MLP	0,0596	5,2008	-0,0085	0,1002	8,6963	-0,0203	0,1336	11,8839	-0,0373	0,1405	12,0112	**-0,0078**
Elman	0,0662	5,7053	-0,0026	0,1198	10,4319	-0,0099	0,1898	16,8847	-0,0289	0,2070	17,8914	0,0136
L-ANN	0,0625	5,4174	**-0,0006**	0,1095	9,4819	**-0,0009**	0,1715	14,8508	**-0,0014**	0,1896	15,8575	0,0389
RBF	0,0687	5,9576	-0,0054	0,1244	11,0212	-0,0170	0,1916	17,2038	-0,0333	0,2050	17,8344	0,0105
ANFIS-1	0,1084	9,2020	0,0064	0,1751	15,2000	0,0017	0,2408	20,8790	-0,0106	0,2682	22,5564	0,0441
ANFIS-2	0,1094	9,2455	0,0112	0,1828	15,7807	0,0133	0,2456	21,0580	0,0088	0,2739	22,6919	0,0656

when compared to the other ANN and ANFIS models. Examining bias values of all the models for power prediction, we can conclude that over or under estimation of the wind power changes according to the data set in use. Specifically, PAR models generally underestimate the power in test period 1 and overestimate in test period 2. Moreover, for increasing order of prediction horizon values, bias values for PAR models generally follow a slightly increasing characteristic and for all the data sets bias values are between -0.05 and 0.05. Apart from its performance with respect to NRMSE and NMAPE for all the data sets, PAR performs predictions with very low bias values. In particular, PAR has the lowest bias for 24-hours ahead predictions in some of the cases, and in the other cases, it is among the best five methods in terms of bias values.

Moreover, PAR model is computationally more efficient than other nonlinear models which are used in this study. PAR models have very lower number of adjustable parameters (close to the linear models) and however, achieve the best results for a prediction horizon values higher than 6 hours. This superior performance makes PAR more preferable compared to the models which have widespread usage in wind speed/power prediction studies.

We should state that the nonlinear functions that have been employed by ANN and ANFIS, can be represented by polynomials with infinite order. However, PAR prediction performance shows that lower degree polynomials (order 2 for this study) is enough for 24-hours ahead wind speed/power predictions. Thus, we are of the opinion that the effects of these higher ordered polynomial components in nonlinear functions of ANN and ANFIS cause failure on prediction results.

6. References

[1] O. Maatallah, A. Achuthan, K. Janoyan, P. Marzocca. "Recursive wind speed forecasting based on Hammerstein Auto-Regressive model", *Applied Energy*, **145**, pp. 191–197, (2015).

[2] S. S. Soman, H. Zareipour, O. Malik, P. Mandal. "A review of wind power and wind speed forecasting methods with different time horizons", *North American Power Symposium (NAPS), 2010*, pp. 1–8, (IEEE, 2010).

[3] G. Kariniotakis, P. Pinson, N. Siebert, G. Giebel, R. Barthelmie, et al. "The state of the art in short-term prediction of wind power-from an offshore perspective", *Proceedings of*, pp. 20–21, (2004).

[4] I. Colak, S. Sagiroglu, M. Yesilbudak. "Data mining and wind power prediction: A literature review", *Renewable Energy*, **46**, pp. 241–247, (2012).

[5] S. Fang, H.-D. Chiang. "Improving supervised wind power forecasting models using extended numerical weather variables and unlabelled data", *IET Renewable Power Generation*, (2016).

[6] A. U. Haque, P. Mandal, J. Meng, M. Negnevitsky. "Wind speed forecast model for wind farm based on a hybrid machine learning algorithm", *International Journal of Sustainable Energy*, **34(1)**, pp. 38–51, (2015).

[7] T. Nielsen, A. Joensen, H. Madsen, L. Landberg, G. Giebel. "A new reference for wind power forecasting", *Wind energy*, **1(1)**, pp. 29–34, (1998).

[8] J. Palomares-Salas, J. De la Rosa, J. Ramiro, J. Melgar, A. Agüera, A. Moreno. "Comparison of models for wind speed forecasting", *Proc. International Conference on Computational Science*, (2009).

[9] P. Gomes, R. Castro. "Wind speed and wind power forecasting using statistical models: AutoRegressive Moving Average (ARMA) and Artificial Neural Networks (ANN)", *International Journal of Sustainable Energy Development*, **1(1/2)**, pp. 275–293, (2012).

[10] J. Torres, A. Garcia, M. De Blas, A. De Francisco. "Forecast of hourly average wind speed with ARMA models in Navarre (Spain)", *Solar Energy*, **79(1)**, pp. 65–77, (2005).

[11] A. Haque, P. Mandal, M. Kaye, J. Meng, L. Chang, T. Senjyu. "A new strategy for predicting short-term wind speed using soft computing models", *Renewable and Sustainable Energy Reviews*, **16(7)**, pp. 4563–4573, (2012).

[12] B. Doucoure, K. Agbossou, A. Cardenas. "Time series prediction using artificial wavelet neural network and multi-resolution analysis: Application to wind speed data", *Renewable Energy*, **92**, pp. 202–211, (2016).

[13] M. De Giorgi, A. Ficarella, M. Tarantino. "Error analysis of short term wind power prediction models", *Applied Energy*, **88(4)**, pp. 1298–1311, (2011).

[14] E. T. Renani, M. F. M. Elias, N. A. Rahim. "Using data-driven approach for wind power prediction: A comparative study", *Energy Conversion and Management*, **118**, pp. 193–203, (2016).

[15] Y. Katsigiannis, A. Tsikalakis, P. Georgilakis, N. Hatziargyriou. "Improved wind power forecasting using a combined neuro-fuzzy and artificial neural network model", *Advances in Artificial Intelligence*, pp. 105–115, (Springer, 2006).

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

[17] J. Catalao, H. Pousinho, V. Mendes. "Hybrid intelligent approach for short-term wind power forecasting in portugal", *IET Renewable Power Generation*, **5(3)**, pp. 251–257, (2011).

[18] S. Fang, H.-D. Chiang. "A high-accuracy wind power forecasting model", *IEEE Transactions on Power Systems*, (2016).

[19] L. Fugon, J. Juban, G. Kariniotakis. "Data mining for wind power forecasting", *European Wind Energy Conference & Exhibition EWEC 2008*, pp. 6–pages, (EWEC, 2008).

[20] A. Kusiak, Z. Zhang. "Short-horizon prediction of wind power: A data-driven approach", *IEEE Transactions on Energy Conversion*, **25(4)**, pp. 1112–1122, (2010).

[21] D. Lee. "Short-term prediction of wind farm output using the recurrent quadratic Volterra model", *Power and Energy Society General Meeting, 2011 IEEE*, pp. 1–8, (IEEE, 2011).

[22] M. Lydia, S. S. Kumar, A. I. Selvakumar, G. E. P. Kumar. "Linear and non-linear autoregressive models for short-term wind speed forecasting", *Energy Conversion and Management*, **112**, pp. 115–124, (2016).

[23] E. Cadenas, W. Rivera, R. Campos-Amezcua, C. Heard. "Wind speed prediction using a univariate arima model and a multivariate narx model", *Energies*, **9(2)**, p. 109, (2016).

[24] P.-J. Lahaye, J.-B. Poline, G. Flandin, S. Dodel, L. Garnero. "Functional connectivity: studying nonlinear, delayed interactions between BOLD signals", *Neuroimage*, **20(2)**, pp. 962–974, (2003).

[25] J. Gruber, C. Bordons, R. Bars, R. Haber. "Nonlinear predictive control of smooth nonlinear systems based on Volterra models. Application to a pilot plant", *International Journal of Robust and Nonlinear Control*, **20(16)**, pp. 1817–1835, (2010).

[26] C. Fernandes, J. Mota, G. Favier, et al. "MIMO Volterra modeling for nonlinear communication channels", *Learning and Nonlinear Models*, **2(8)**, pp. 71–92, (2010).

[27] S. Bekleric. "Nonlinear prediction via Volterra series and applications to geophysical data", *Masters Abstracts International*, volume 47, (2008).

[28] E. Kuruoğlu. “Nonlinear least l_p-norm filters for nonlinear autoregressive α-stable processes”, *Digital Signal Processing*, **12(1)**, pp. 119–142, (2002).

[29] O. Karakus, E. Kuruoglu, M. Altinkaya. “Long term wind speed prediction with polynomial autoregressive model”, *Signal Processing and Communications Applications Conference (SIU), 2015 23th*, pp. 645–648, (IEEE, 2015).

[30] T. Hong, P. Pinson, S. Fan. “Global energy forecasting competition 2012”, *International Journal of Forecasting*, **30(2)**, pp. 357–363, (2014).

[31] A. Sfetsos. “A comparison of various forecasting techniques applied to mean hourly wind speed time series”, *Renewable energy*, **21(1)**, pp. 23–35, (2000).

[32] T. Takagi, M. Sugeno. “Fuzzy identification of systems and its applications to modeling and control”, *Systems, Man and Cybernetics, IEEE Transactions on*, **(1)**, pp. 116–132, (1985).

[33] A. Abraham. “Adaptation of fuzzy inference system using neural learning”, *Fuzzy systems engineering*, pp. 53–83, (Springer, 2005).

[34] M. Bernardo-Faura, S. Massen, C. Falk, N. Brady, R. Eils. “Data-derived modeling characterizes plasticity of mapk signaling in melanoma”, *PLoS Computational Biology, 2014,*, **10(9)**, (2014).

[35] H. Cheng, P.-N. Tan, J. Gao, J. Scripps. “Multistep-ahead time series prediction”, *Advances in Knowledge Discovery and Data Mining*, pp. 765–774, (Springer, 2006).

[36] Y. Bard. *Nonlinear parameter estimation*, (Academic Press, 1974).

[37] B. Harper, J. Kepert, J. Ginger. “Guidelines for converting between various wind averaging periods in tropical cyclone conditions”, *World Meteorological Organization, WMO/TD*, **1555**, (2010).

[38] “Agrimet-Cooperative Agricultural Weather Network”, , (Accessed: Aug 10, 2016), [Online]. Available: http://www.usbr.gov/pn/agrimet/webaghrread.html.

[39] O. Karakus, E. E. Kuruoglu, M. A. Altinkaya. “Estimation of the nonlinearity degree for polynomial autoregressive processes with RJMCMC”, *23rd European Signal Processing Conference (EUSIPCO)*, pp. 953–957, (IEEE, 2015).