

Research Article

A New Hybrid Forecasting Strategy Applied to Mean Hourly Wind Speed Time Series

**Stylianos Sp. Pappas,¹ Dimitrios Ch. Karamousantas,¹
George E. Chatzarakis,² and Christos Sp. Pappas³**

¹ Department of Organic Greenhouse Crops and Floriculture, School of Agricultural Technology, Antikalamos, 24100 Kalamata, Greece

² Department of Electrical and Electronic Engineering Educators, School of Pedagogical and Technological Education (ASPETE),
N. Heraklion, 14121 Athens, Greece

³ School of Engineering, University of Greenwich, Medway Campus, Pembroke Building, Central Avenue, Chatham Maritime,
Kent ME4 4TB, UK

Correspondence should be addressed to Stylianos Sp. Pappas; spappas@aegean.gr

Received 19 March 2014; Accepted 14 May 2014; Published 12 June 2014

Academic Editor: Adrian Ilinca

Copyright © 2014 Stylianos Sp. Pappas et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An alternative electric power source, such as wind power, has to be both reliable and autonomous. An accurate wind speed forecasting method plays the key role in achieving the aforementioned properties and also is a valuable tool in overcoming a variety of economic and technical problems connected to wind power production. The method proposed is based on the reformulation of the problem in the standard state space form and on implementing a bank of Kalman filters (KF), each fitting an ARMA model of different order. The proposed method is to be applied to a greenhouse unit which incorporates an automatized use of renewable energy sources including wind speed power.

1. Introduction

Energy is considered amongst the most significant factors that are closely related to both economic and social developments. It is also a fact that nowadays the majority of the electrical energy production is based on the fossil fuels, which on one hand are, without any doubt, highly efficient but on the other are responsible for the emission of greenhouse gases and their reserves are limited.

Consequently renewable sources of energy, such as wind, biomass, solar power, and wave power, have been already adopted for electric power production. It is well known that the wind power generation raises issues of reliability due to the fact that the wind speed is significantly and directly affected by various factors such as the type of the terrain, the height, season of the year, atmospheric conditions, obstacles present, and many more. This leads to the conclusion that unless the reliability of the wind power generation is at an acceptable level, wind power is not eligible for constant electrical energy supply to the power system [1, 2].

Recent studies have shown that combined forecasting methods can offer robust solutions and can be efficiently implemented to various real-life problems in diverging fields such as chemical processes, economics, load forecasting, tourism demand, environmental issues, medicine, and many more [3–7].

In this study a hybrid model is presented that reveals the advantages of an ARMA and SVM model in wind speed modelling and prediction problem. Initially successful model identification and parameter estimation have to be performed in order to choose the most appropriate ARMA models. For tackling this task the well-established MMPA was used. This approach was introduced by Lainiotis [8, 9] and summarizes the parametric model uncertainty into an unknown, finite dimensional parameter vector whose values are assumed to lie within a known set of finite cardinality. A nonexhaustive list of the reformulation, extension, and application of the MMPA approach as well as its application to a variety of problems can be found in [10–19].

In this research real data were used, provided by Vestas Hellas; the simulation results appear to be very promising.

2. Hybrid Model Presentation

2.1. The ARMA Model. The problem of fitting an ARMA model in a given time series is present for more than half a century and is still appearing in many different fields such as in remote monitoring of civil infrastructure [20], predicting the demand for auto spare parts in China due to the fierce market competition [21], forecasting equipment failures in order to adjust maintenance policies in manufacturing plants [22], estimating retail sales volumes [23], predicting the outbreak and development of animal infectious diseases [24], and many more.

Considering the general case an m -variate (i.e., multi-variate) ARMA model of order (p, q) [ARMA(p, q)] for a stationary time series of vectors y observed at equally spaced instants $k = 1, 2, \dots, n$ is defined as

$$y_k = \sum_{i=1}^p A_i y_{k-i} + \sum_{j=1}^q B_j v_{k-j} + v_k, \quad E[v_k v_k^T] = R, \quad (1)$$

where the m -dimensional vector v_k is uncorrelated random noise, not necessarily Gaussian, with zero mean; covariance matrix R , $\theta = (p, q)$ is the order of the predictor; and $A_1, \dots, A_p, B_1, \dots, B_q$ are the $m \times m$ coefficient matrices of the multivariate (MV) ARMA model.

It is obvious that the problem requires both the predictor's order $\theta = (p, q)$ determination and computation of the predictor's matrix coefficients $\{A_i, B_j\}$.

The major disadvantage of the ARMA models is that their performance can be limited by any significant data non-linearities.

2.2. Multimodel Partition Algorithm (MMPA). Due to the fact that the wind speed does not have a constant or periodic behaviour, it was noted, by trial and error, that not a single ARMA model that was able to describe the whole data set satisfactory. It is actually the combination of various ARMA models, each one used for different time intervals and applied for different time durations that describes in the best manner the existing data. So instead of having various ARMA models of different order θ running in parallel with the SVM it was decided to load all the data to an adaptive filter programmed with the MMPA, and it will be the job of that filter to decide which ARMA model will be used each time.

If we assume that the model order fitting the data is known and is equal to $\theta = (p, q)$, we can rewrite (1) in standard state-space form as

$$\begin{aligned} x(k+1) &= x(k), \\ y(k) &= H(k) x(k) + v(k). \end{aligned} \quad (2)$$

Now assign a new variable λ such as $\lambda = \max(p, q)$. Then $x(k)$ is an $m^2(\lambda+1) \times 1$ vector made up from the coefficients of the matrices $\{A_1, \dots, A_\lambda, B_1, \dots, B_\lambda\}$; $H(k)$ is an $m \times m^2(\lambda+1)$ observation history matrix of the process $\{y(k)\}$ up to time $k - (\lambda + 1)$.

If the general form of the matrices A_λ and B_λ is, respectively,

$$\begin{bmatrix} a_{11}^\lambda & \dots & a_{1m}^\lambda \\ \vdots & \ddots & \vdots \\ a_{m1}^\lambda & \dots & a_{mm}^\lambda \end{bmatrix}, \quad \begin{bmatrix} b_{11}^\lambda & \dots & b_{1m}^\lambda \\ \vdots & \ddots & \vdots \\ b_{m1}^\lambda & \dots & b_{mm}^\lambda \end{bmatrix}, \quad (3)$$

then

$$\begin{aligned} x(k) &\triangleq \left[\alpha_{11}^1 \alpha_{21}^1 \dots \alpha_{m1}^1 : \alpha_{12}^1 \alpha_{22}^1 \dots \alpha_{m2}^1 : \dots : \alpha_{1m}^1 \alpha_{2m}^1 \dots \alpha_{mm}^1 \right. \\ &\quad \left. : b_{11}^1 b_{21}^1 \dots b_{m1}^1 : b_{12}^1 b_{22}^1 \dots b_{m2}^1 : \dots : b_{1m}^1 b_{2m}^1 \dots b_{mm}^1 \right]^T. \end{aligned} \quad (4)$$

If $p > q$ then $\lambda = p$; the last $(m^2(p - q))$ MA coefficients are zero. If $q > p$ then $\lambda = q$; the last $(m^2(q - p))$ AR coefficients are zero. Consider

$$\begin{aligned} H(k) &\triangleq \begin{bmatrix} y_1(k-1) \mathbf{I} \dots y_m(k-1) \mathbf{I} : \dots : y_1(k-\lambda) \mathbf{I} \\ \vdots \\ y_m(k-\lambda) \mathbf{I} : v_1(k-1) \mathbf{I} \dots v_m(k-1) \mathbf{I} \\ \vdots \\ v_1(k-\lambda) \mathbf{I} \dots v_m(k-\lambda) \mathbf{I} \end{bmatrix}, \end{aligned} \quad (5)$$

where \mathbf{I} is the $m \times m$ identity matrix.

Assuming that the system model and its statistics were completely known, the Kalman filter (KF) in its various forms would be the optimal estimator in the minimum variance sense.

However, if the system model is not completely known the MMPA, introduced by Lainiotis [8, 9], is one of the most widely used approaches for similar problems [10–19, 25–28].

In the case under consideration assume that the model uncertainty is the lack of knowledge of the model order θ . Let us further assume that the model order θ lies within a known sample space of finite cardinality; that is, $1 \leq \theta \leq M$, $\theta \in \mathfrak{S}$, where \mathfrak{S} denotes the set of integers. The MMPA operates on the following discrete-time model:

$$\begin{aligned} x(k+1) &= F\left(k+1, \frac{k}{\theta}\right) x(k) + w(k), \\ y(k) &= H\left(\frac{k}{\theta}\right) x(k) + v(k), \end{aligned} \quad (6)$$

where θ is the unknown parameter, the model order in this case; F is the state transition matrix; and $w(k)$ is independent, zero mean, white noise not necessarily Gaussian with covariance Q which is usually set to a small positive nonzero constant. The optimal MMSE (minimum mean square error) estimate of $x(k)$ is given by

$$\hat{x}\left(\frac{k}{k}\right) = \sum_{j=1}^M \hat{x}\left(\frac{k}{k}; \theta_j\right) p\left(\frac{\theta_j}{k}\right). \quad (7)$$

A set of M models is designed, each matching one value of the parameter vector, $\{(1, 1), (2, 2), \dots, (M, M)\}$. The probabilities

$p(\theta_j/k)$ for each model are set to $1/M$, where M is the cardinality of the model set.

A bank of Kalman filters is then applied, one for each model, which can be run in parallel, thus saving enormous computational time. At each iteration, the MMPA selects the model that corresponds to the maximum a posteriori probability as the correct one. This probability tends (asymptotically) to one, while the remaining probabilities tend to zero. The overall optimal estimate can be taken either to be the individual estimate of the elemental filter exhibiting the maximum posterior probability, for example, a value of 0.9 or higher [14], or the weighted average of the estimates produced by the elemental ARMA filters, as described in (7), which is the case used in this paper.

The probabilities are calculated on-line in a recursive manner as it is shown by

$$\begin{aligned} P\left(\frac{\theta_j}{k}\right) &= \frac{L(k/k; \theta_j)}{\sum_{j=1}^M L(k/k; \theta_j) P(\theta_j/k-1)} P\left(\frac{\theta_j}{k-1}\right), \\ L\left(\frac{k}{k}; \theta_j\right) &= \left| \mathbf{P}_{\bar{y}}\left(\frac{k}{k} - 1; \theta_j\right) \right|^{-1/2} \\ &\quad \cdot \exp\left[-\frac{1}{2} \bar{\mathbf{y}}^T \left(\frac{k}{k} - 1; \theta_j\right) \right. \\ &\quad \left. \times \mathbf{P}_{\bar{y}}^{-1}\left(\frac{k}{k} - 1; \theta_j\right) \bar{\mathbf{y}} \left(\frac{k}{k} - 1; \theta_j\right) \right], \end{aligned} \quad (8)$$

where the innovation process

$$\bar{\mathbf{y}}\left(\frac{k}{k} - 1; \theta_j\right) = \mathbf{y}(k) - \mathbf{H}(k; \theta_j) \hat{\mathbf{x}}\left(\frac{k}{k} - 1; \theta_j\right) \quad (9)$$

is a zero mean white process with covariance matrix

$$\mathbf{P}_{\bar{y}}\left(\frac{k}{k} - 1; \theta_j\right) = \mathbf{H}(k; \theta_j) \mathbf{P}\left(\frac{k}{k}; \theta_j\right) \mathbf{H}^T(k; \theta_j) + \mathbf{R}. \quad (10)$$

For equations (8)–(10) $j = 1, 2, \dots, M$.

2.3. Support Vector Machines (SVM). In support vector machines, as they were proposed in [29], the training data set $\mathbf{x}_i \in R^d$, $i = 1, \dots, N$, is mapped into a higher dimensional feature space via an operator Φ .

A mathematical representation of the SVM function is

$$y = \omega \cdot \Phi(x) + b, \quad (11)$$

where ω and b can be found by the minimization of the following equations:

$$S(C) = C \frac{1}{M} \sum_{j=1}^M Lf_e(h_j - y_j) + \frac{1}{2} \|\omega\|^2, \quad (12)$$

$$Lf(h, y) = \begin{cases} |h - y| - e & |h - y| \geq e, \\ 0 & \text{others,} \end{cases} \quad (13)$$

where parameters C and e are user-defined. The term h_i is the actual wind speed at the time instant j and term $Lf_e(h_i - y_i)$

is the *loss function*. By looking at (13) it is obvious that there is any penalty for errors below e . The width of the function is given by the term $(1/2)\|\omega\|^2$ and finally the training error term is given by $C(1/M) \sum_{j=1}^M Lf_e(h_i - y_i)$, where C is the tradeoff between the width of the function and the minimum training error. For dealing with nonlinear cases, like wind speed data, one may introduce slack variables ξ and ξ^* into (11) such that

$$\begin{aligned} \omega \cdot \Phi(j) + b_j - h_j &\leq e + \xi_j^*, \\ -(\omega \cdot \Phi(j) + b_j) + h_j &\leq e + \xi_j, \end{aligned} \quad (14)$$

where $\xi_j, \xi_j^* \geq 0$, and $j = 1, 2, \dots, M$.

By considering the above slack variables and in order to include any extra cost of the training errors, (12), which represents the objective function to be minimized, is rearranged to

$$S(\omega, \xi, \xi^*) = C^* \left(\sum_{j=1}^M (\xi_j + \xi_j^*) \right) + \frac{1}{2} \omega \omega^T, \quad (15)$$

where again C^* is user-defined and is the tradeoff between the maximum margin defined by $\|\omega\|$ and the minimum training error as defined by $\sum_{j=1}^M (\xi_j + \xi_j^*)$.

Finally by introducing positive Lagrangian multipliers and maximizing (15) the latter equation is reformed to

$$\begin{aligned} S(a_j - a_j^*) &= \sum_{j=1}^M h_i (a_j - a_j^*) - e \sum_{j=1}^M (a_j - a_j^*) \\ &\quad - \frac{1}{2} \sum_{j=1}^M \sum_{i=1}^M (a_j - a_j^*) \times (a_i - a_i^*) K(x_j - x_i) \end{aligned} \quad (16)$$

subject to

$$\sum_{j=1}^N (a_j - a_j^*) = 0, \quad 0 \leq a_j, a_j^* \leq C, \quad (17)$$

where $j = 1, 2, \dots, M$.

The Lagrangian multipliers a_j, a_j^* satisfy $a_j * a_j^* = 0$ and also

$$f(x, a, a^*) = \sum_{j=1}^l (a_j - a_j^*) K(x - x_j) + b. \quad (18)$$

The *Kernel* function $K(x - x_i)$ introduced in (18) is defined such that $K(x_j - x_i) = \Phi(x_i) \cdot \Phi(x_j)$, meaning that its value is equal to the inner product of the vectors x_i and x_j , included in the featured space $\Phi(x_i)$ and $\Phi(x_j)$.

In this study the Gaussian kernel function (19), also known as radial basis function (RBF), is used. Consider

$$K(x_j - x_i) = e^{(-\|x_j - x_i\|^2 / 2\sigma^2)}. \quad (19)$$

The most significant feature of the SVM compared to other similar algorithms is that they manage to achieve optimum performance by restricting the decision's function complexity so that is the most suitable according to the quantity of the data present.

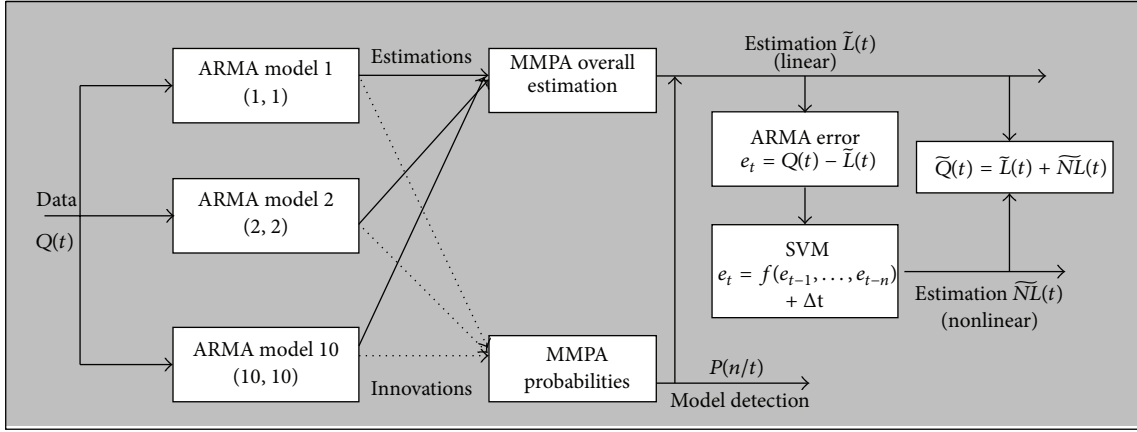


FIGURE 1: Proposed method.

2.4. The Hybrid Model. The wind speed behaviour is unpredictable and it is difficult to be represented. This is the reason for combining two different techniques for modelling the linear and the nonlinear parts of the series. The hybrid model proposed is based on a linear pattern, $L(t)$ produced by the MMPA, and a nonlinear one, $NL(t)$ produced by the SVM. It can be represented as

$$Q(t) = L(t) + NL(t). \quad (20)$$

Both parts are directly calculated from the wind speed time series.

If $e(t)$ is the MMPA estimation error at any time instant t , then

$$e(t) = Z(t) - \tilde{L}(t). \quad (21)$$

It is now the SVM that models these residuals as

$$e(t) = f(e_{(t-1)}, e_{(t-2)}, \dots, e_{(t-n)}) + \Delta t, \quad (22)$$

where f is nonlinear and Δt is random error. It is obvious that $\tilde{NL}(t)$ is the forecast of (22).

Consequently the forecast of the hybrid model is

$$\tilde{Q}(t) = \tilde{L}(t) + \tilde{NL}(t). \quad (23)$$

The proposed method is schematically represented in Figure 1.

3. Results and Discussion

3.1. Results. In this method the weighted average of the estimates produced by the elemental ARMA filters was used as a data preprocessor in order to detect the data's linearities. This was succeeded using a bank of 10 Kalman filters of order $(1, 1), (2, 2), (3, 3), \dots, (10, 10)$ programmed with the MMPA. Then the MMPA's estimation error was applied as input to the SVM that were able to achieve a further error reduction and come up with a better forecasting outcome. As far as the SVM are concerned the three parameters (C , σ , and e) had to be carefully adjusted. Unsuitable values for these parameters

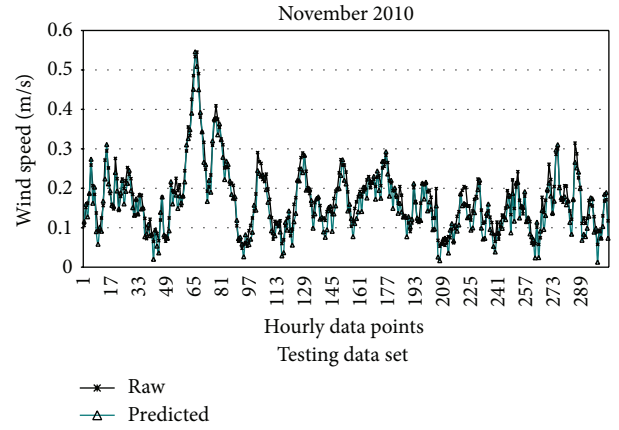


FIGURE 2: November 2010: predicted and observed (raw) time series.

may lead to either overfitting or underfitting of the training data. The values used in this research were $C = 35$, $\sigma = 2.5$, and $e = 0.6$.

This research was conducted based on the hour average of daily wind speed recorded by the Vestas Hellas from November 2010 up to February 2011. The obtained time series did not follow any periodic pattern and it was also presenting irregular amplitudes, making it hard to both model and predict (Figures 2, 3, 4, and 5; “raw” data).

The aim of this work is to generate a single-step prediction based on past observations. The data were normalized to take values from zero to one, before using them as input data to the hybrid model.

From the 2725 available data points, 720 were for November, 744 for December, 744 for January, and 517 for February. For each month 20% of the available data was used for training, 20% for validation, and 60% for testing.

The performance of the hybrid method is judged by (a) comparing the predicted and the observed (raw) wind time series (Figures 2, 3, 4, and 5), (b) drawing scatter diagrams of the predicted and the observed sequences (Figures 6, 7, 8, and 9), and (c) computing the mean percentage absolute error

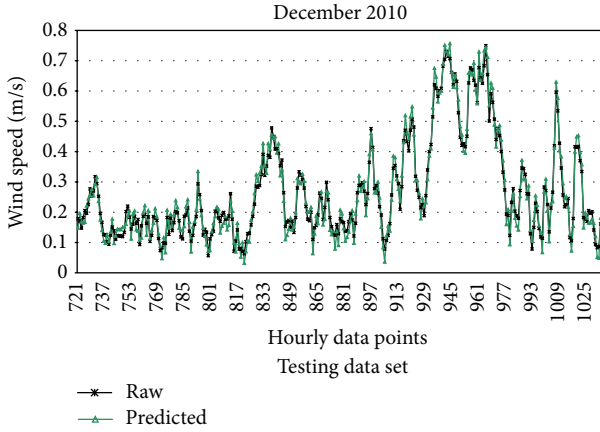


FIGURE 3: December 2010: predicted and observed (raw) time series.

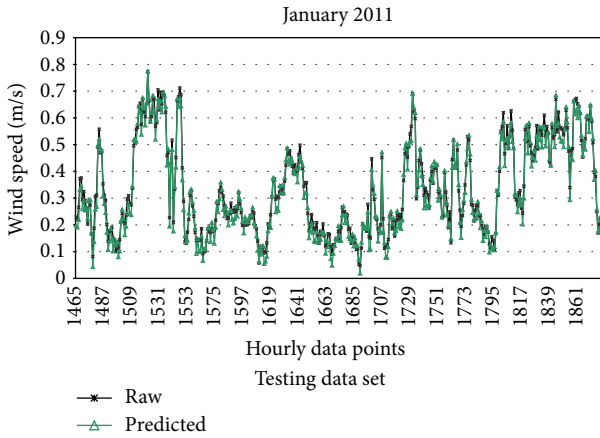


FIGURE 4: January 2011: predicted and observed (raw) time series.

TABLE 1: Performance of the hybrid model.

Month	MAPE %	R^2
November	2.78	0.8815
December	3.27	0.8533
January	3.19	0.8635
February	3.02	0.8758
Average	3.01	0.8685

(MAPE) for the testing data set, using the mathematical formula given in (22). Table 1 summarizes the results. Consider

$$\text{MAPE} = \frac{1}{M} \sum_{j=1}^M \left(\frac{|P_{\text{real},j} - P_{\text{predicted},j}|}{P_{\text{real},j}} \right) \times 100. \quad (24)$$

3.2. Discussion. Figures 2–5 indicate that the predictions are very close to the real values of the wind speed time series. Additionally the hybrid model is capable of following satisfactorily the irregularities of the observed series. The value of MAPE between the real and the predicted series can be considered quite low since it has an average value below 3.01% and the individual value for each month is not above 3.3%.

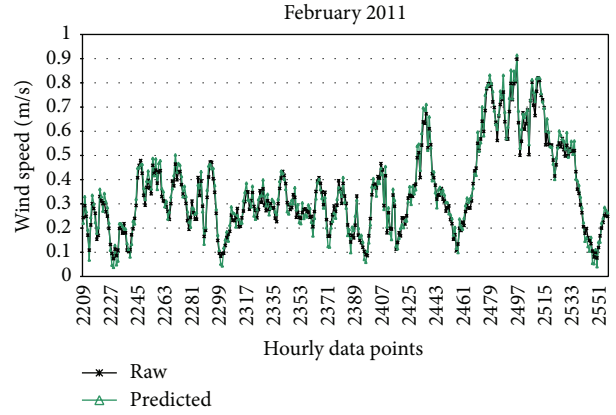


FIGURE 5: February 2011: forecasted predicted and observed (raw) time series.

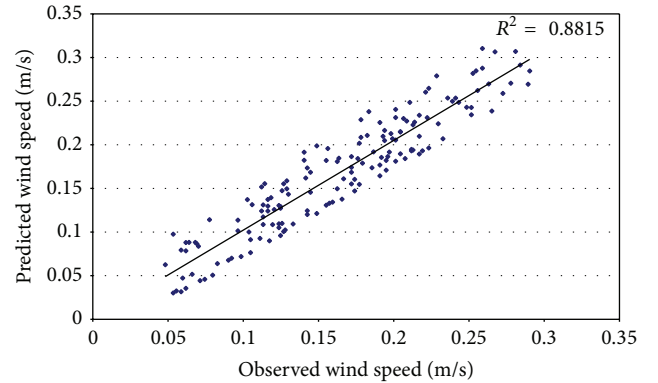


FIGURE 6: November 2010: scatter of predicted and observed wind speed time series.

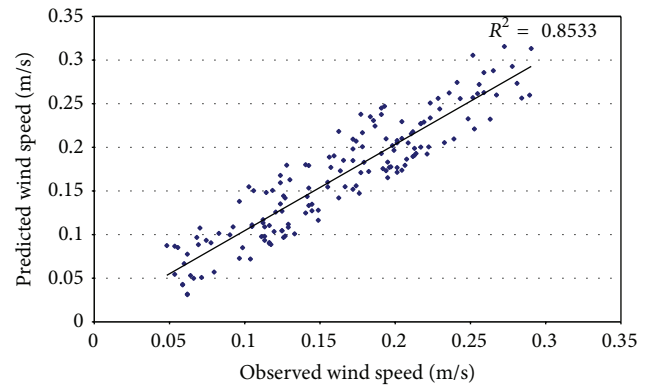


FIGURE 7: December 2010: scatter of predicted and observed wind speed time series.

Furthermore the value of R^2 is another indication of close is the predicted to real series. A value of $R^2 = 1$ indicates a straight line of the form $y = ax$ meaning that the predicted values are exactly the same as the real ones. In this study R^2 is at an average of 0.8685 which means that the predicted series is indeed very close to the real one.

TABLE 2: Summary of results.

Month	MAPE %			R^2		
	Hybrid	MMPA-ARMA	SVM	Hybrid	MMPA-ARMA	SVM
November	2.78	4.93	4.73	0.8815	0.7462	0.7861
December	3.27	4.75	4.32	0.8533	0.7225	0.7689
January	3.19	5.23	4.12	0.8635	0.7349	0.7921
February	3.02	5.64	4.43	0.8758	0.7182	0.7783
Average	3.01	5.14	4.25	0.8685	0.73045	0.78135

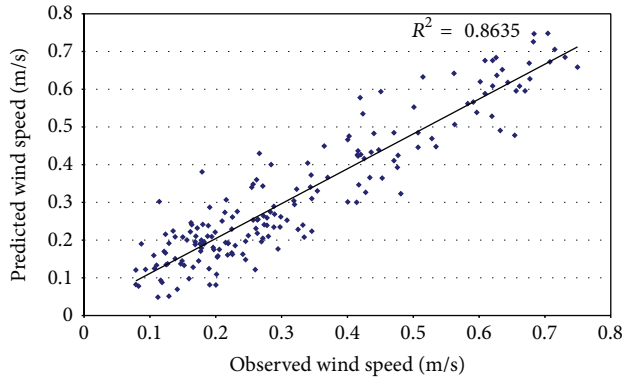


FIGURE 8: January 2011: scatter of predicted and observed wind speed time series.

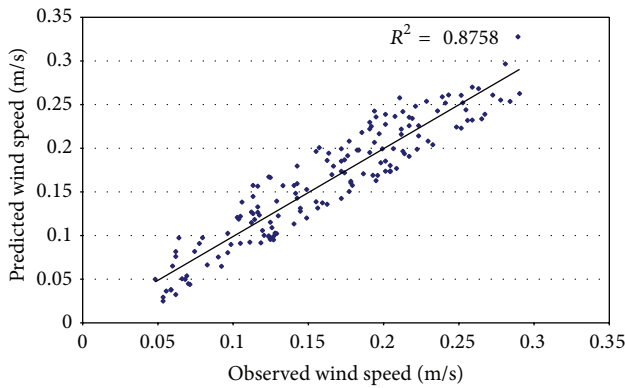


FIGURE 9: February 2011: scatter of predicted and observed wind speed time series.

For the sake of completeness of this research the results of forecasting using only the adaptive combination of MMPA with the ARMA models as well as using only the ANN architecture with the SVM are also presented. The first method was applied successfully in [22] for electric load modeling and forecasting. The difference is that electric load data is seasonal and after careful manipulation it can be converted to Gaussian, something that cannot be done with a wind speed time series.

Using the data for November, Figures 10 and 11 compare the predicted and the observed (raw) wind time series and Figures 12 and 13 show the scatter diagram of the predicted and observed sequences for MMPA-ARMA and SVM,

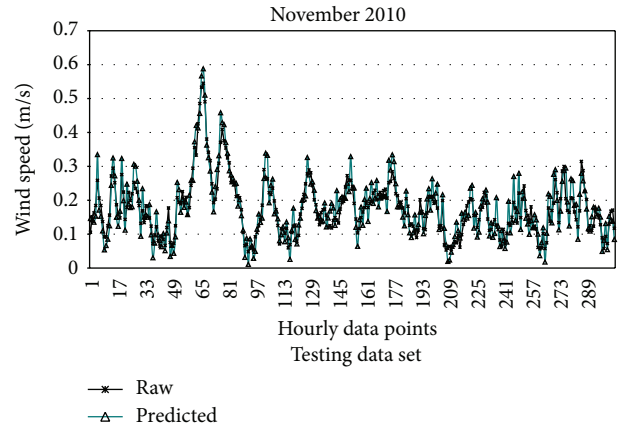


FIGURE 10: November 2010: predicted and observed (raw) time series, MMPA-ARMA.

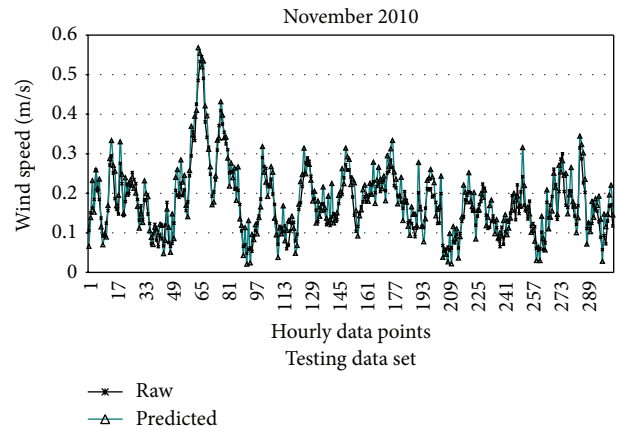


FIGURE 11: November 2010: predicted and observed (raw) time series, SVM.

respectively. Table 2 summarizes the performance of the aforementioned methods and the hybrid method proposed for the whole data set.

At this point it should be mentioned that when dealing with the wind speed forecasting problem it is quite difficult to attain very high prediction accuracy. The hybrid method proposed gave some very satisfactory results and the accuracy level reached can be considered sufficient for decision making as far as electric power production is concerned.

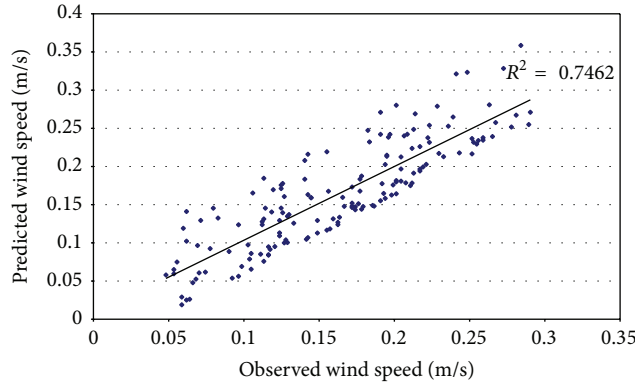


FIGURE 12: November 2010: scatter of predicted and observed wind speed time series, MMPA-ARMA.

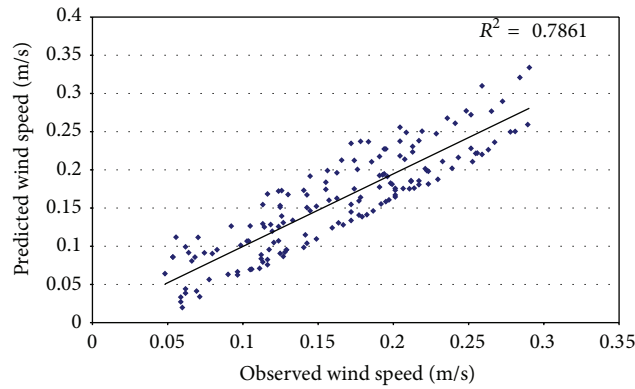


FIGURE 13: November 2010: scatter of predicted and observed wind speed time series, SVM.

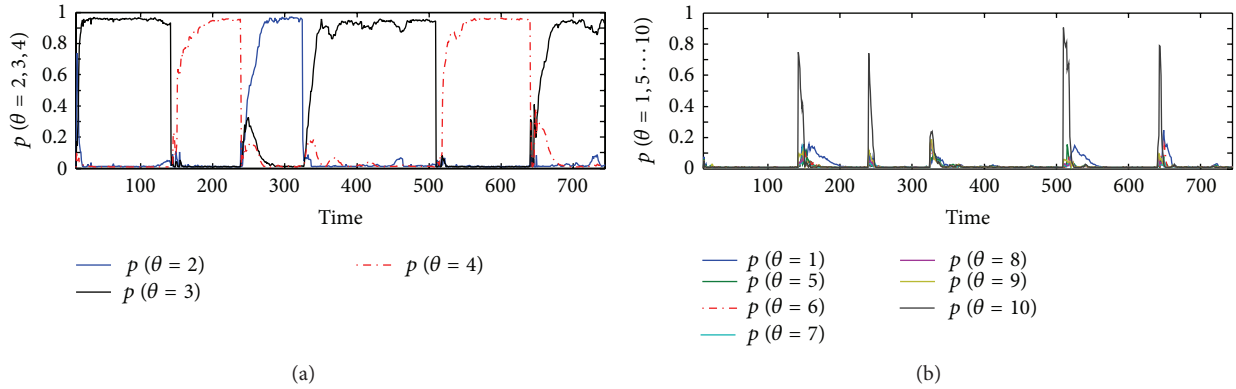


FIGURE 14: December 2010: MMPA probability sequence (ARMA model order selection) for the whole data set provided.

Last but not least it should be also mentioned that a designer's crucial task when using MMPA is to assign a proper value for the cardinality M . Trying to find an answer to that problem one has to take in mind the following considerations:

- (i) the quality of the overall MMPA estimate (in terms of MAPE) increases with the number of the elemental filters applied, so a small value of M leads to estimates of a poor quality;
- (ii) the computational load of the MMPA is proportional to the number of filters implemented; so a large value

M increases the computational burden, a factor that makes the implementation of the MMPA for real time problems difficult.

A way of assigning the proper value of M is to record the behaviour of MMPA during its operation. It is already mentioned that, at each iteration, the MMPA selects the model that corresponds to the maximum a posteriori probability as the correct one. This probability tends (asymptotically) to one, while the remaining probabilities tend to zero. Figure 14

shows the behaviour of the MMPA for the whole data set concerning December.

It can be seen that the highest order θ of the appropriate ARMA model for the time series under investigation is (4, 4) which makes the choice of $M = 10$ acceptable. Another also significant conclusion is that if a model change is required, the algorithm senses the variation and increases its corresponding a posteriori probability while decreasing the remaining ones. Thus, the algorithm is adaptive in the sense of being able to track model changes in real time. This procedure incorporates the algorithm's intelligence.

4. Conclusions

The area of forecasting is very demanding and it is over 50 years that ARMA models were exclusively used for tackling real-life problems. Recently ANN were applied in difficult prediction problems showing very satisfactory results especially due to their ability of manipulating the nonlinearities of the dataset. The aim of this work was not to just add yet another technique of wind speed prediction but to actually validate the fact that different forecasting methods fulfil each other and lead to accurate results. As it was shown the two individual forecasting methods, adaptive MMPA-ARMA and SVM, cannot match the performance of the hybrid method proposed. The results of the proposed method are to be applied to a greenhouse unit which incorporates an automatized use of renewable energy sources including wind speed power. Future work can include adjustments for wind speed prediction for time intervals smaller than 1 hour, say ten-minute intervals, and also for on-line wind speed prediction.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported in part by the European Union (European Social Fund-ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF), Research Funding Program: ARCHIMEDES III. Investing in knowledge society through the European Social Fund.

References

- [1] M. Lei, L. Shiyang, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 4, pp. 915–920, 2009.
- [2] M. Marciukaitis, V. Katinas, and A. Kavaliauskas, "Wind power usage and prediction prospects in Lithuania," *Renewable and Sustainable Energy Reviews*, vol. 12, no. 1, pp. 265–277, 2008.
- [3] C.-H. Wen and C. A. Vassiliadis, "Applying hybrid artificial intelligence techniques in wastewater treatment," *Engineering Applications of Artificial Intelligence*, vol. 11, no. 6, pp. 685–705, 1998.
- [4] M. Shafie-Khah, M. P. Moghaddam, and M. K. Sheikh-El-Eslami, "Price forecasting of day-ahead electricity markets using a hybrid forecast method," *Energy Conversion and Management*, vol. 52, no. 5, pp. 2165–2169, 2011.
- [5] K.-Y. Chen, "Combining linear and nonlinear model in forecasting tourism demand," *Expert Systems with Applications*, vol. 38, no. 8, pp. 10368–10376, 2011.
- [6] B. Baruque, E. Corchado, A. Mata, and J. M. Corchado, "A forecasting solution to the oil spill problem based on a hybrid intelligent system," *Information Sciences*, vol. 180, no. 10, pp. 2029–2043, 2010.
- [7] W.-C. Yeh, W.-W. Chang, and Y. Y. Chung, "A new hybrid approach for mining breast cancer pattern using discrete particle swarm optimization and statistical method," *Expert Systems with Applications*, vol. 36, no. 4, pp. 8204–8211, 2009.
- [8] D. G. Lainiotis, "Partitioning: a unifying framework for adaptive systems—I: estimation," *Proceedings of the IEEE*, vol. 64, no. 8, pp. 1126–1143, 1976.
- [9] D. G. Lainiotis, "A unifying framework for adaptive systems—II: control," *Proceedings of the IEEE*, vol. 64, no. 8, pp. 1182–1198, 1976.
- [10] S. S. Pappas, A. K. Leros, and S. K. Katsikas, "Joint order and parameter estimation of multivariate autoregressive models using multi-model partitioning theory," *Digital Signal Processing*, vol. 16, no. 6, pp. 782–795, 2006.
- [11] S. K. Katsikas, S. D. Likothanassis, G. N. Beligiannis, K. G. Berketis, and D. A. Fotakis, "Genetically determined variable structure multiple model estimation," *IEEE Transactions on Signal Processing*, vol. 49, no. 10, pp. 2253–2261, 2001.
- [12] P. Papaparaskeva, "A partitioned adaptive approach to nonlinear channel equalization demetrios G. lainiotis and," *IEEE Transactions on Communications*, vol. 46, no. 10, pp. 1325–1336, 1998.
- [13] V. C. Moussas, S. D. Likothanassis, S. K. Katsikas, and A. K. Leros, "Adaptive on-line multiple source detection," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, pp. 1029–1032, March 2005.
- [14] D. G. Lainiotis, K. S. Katsikas, and S. D. Likothanassis, "Adaptive deconvolution of seismic signals—performance, computational analysis, parallelism," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 11, pp. 1715–1734, 1988.
- [15] N. V. Nikitakos, A. K. Leros, and S. K. Katsikas, "Towed array shape estimation using multimodel partitioning filters," *IEEE Journal of Oceanic Engineering*, vol. 23, no. 4, pp. 380–384, 1998.
- [16] V. C. Moussas and S. S. Pappas, "Adaptive network anomaly detection using bandwidth utilization data," in *Proceedings of the 1st International Conference on Experiments/Processes/System Modelling/Simulation/Optimization*, Athens, Greece, 2005.
- [17] S. S. Pappas, L. Ekonomou, P. Karampelas, S. K. Katsikas, and P. Liatsis, "Modeling of the grounding resistance variation using ARMA models," *Simulation Modelling Practice and Theory*, vol. 16, no. 5, pp. 560–570, 2008.
- [18] S. S. Pappas, L. Ekonomou, D. C. Karamousantas, G. E. Chatzarakis, S. K. Katsikas, and P. Liatsis, "Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models," *Energy*, vol. 33, no. 9, pp. 1353–1360, 2008.
- [19] S. S. Pappas, L. Ekonomou, P. Karampelas et al., "Electricity demand load forecasting of the Hellenic power system using an ARMA model," *Electric Power Systems Research*, vol. 80, no. 3, pp. 256–264, 2010.

- [20] E. Peter Carden and J. M. W. Brownjohn, "ARMA modelled time-series classification for structural health monitoring of civil infrastructure," *Mechanical Systems and Signal Processing*, vol. 22, no. 2, pp. 295–314, 2008.
- [21] Y. Chen, H. Zhao, and L. Yu, "Demand forecasting in automotive aftermarket based on ARMA model," in *Proceedings of the International Conference on Management and Service Science (MASS '10)*, pp. 1–4, 2010.
- [22] Z. Jie, X. Limei, and L. Lin, "Equipment fault forecasting based on ARMA model," in *Proceedings of the IEEE International Conference on Mechatronics and Automation (ICMA '07)*, pp. 3514–3518, August 2007.
- [23] R. Majhi, S. Mishra, B. Majhi, G. Panda, and M. Rout, "Efficient sales forecasting using PSO based adaptive ARMA model," in *Proceedings of the World Congress on Nature and Biologically Inspired Computing (NABIC '09)*, pp. 1333–1337, December 2009.
- [24] F. Li and P. Luan, "ARMA model for predicting the number of new outbreaks of newcastle disease during the month," in *Proceedings of the IEEE International Conference on Computer Science and Automation Engineering (CSAE '11)*, vol. 4, pp. 660–663, June 2011.
- [25] S. S. Pappas, L. Ekonomou, V. C. Moussas, P. Karampelas, and S. K. Katsikas, "Adaptive load forecasting of the Hellenic electric grid," *Journal of Zhejiang University: Science A*, vol. 9, no. 12, pp. 1724–1730, 2008.
- [26] G. Beligiannis, L. Skarlas, and S. Likothanassis, "A generic applied evolutionary hybrid technique," *IEEE Signal Processing Magazine*, vol. 21, no. 3, pp. 28–38, 2004.
- [27] S. S. Pappas, N. Harkiolakis, P. Karampelas, L. Ekonomou, and S. K. Katsikas, "A new algorithm for on-line Multivariate ARMA identification using Multimodel Partitioning Theory," in *Proceedings of the 12th Pan-Hellenic Conference on Informatics (PCI '08)*, pp. 222–226, August 2008.
- [28] V. C. Moussas, S. K. Katsikas, and D. G. Lainiotis, "Adaptive estimation of FCG using nonlinear state-space models," *Stochastic Analysis and Applications*, vol. 23, no. 4, pp. 705–722, 2005.
- [29] V. Vapnik, *The Nature of Statistic Learning Theory*, Springer, New York, NY, 1995.

