J. Wind Eng. Ind. Aerodyn. 112 (2013) 25–38

Contents lists available at SciVerse ScienceDirect

Journal of Wind Engineering and Industrial Aerodynamics

journal homepage: www.elsevier.com/locate/jweia

Wind power prediction based on numerical and statistical models

Christos Stathopoulos [a], Akrivi Kaperoni [a], George Galanis [a,b], George Kallos [a,*]

[a] University of Athens, School of Physics, Division of Applied Physics, Atmospheric Modeling and Weather Forecasting Group, University Campus, Bldg. PHYS-V, 15784 Athens, Greece
[b] Hellenic Naval Academy, Section of Mathematics, Hatzikiriakion, Piraeus 18539, Greece

A R T I C L E   I N F O

Article history:
Received 26 February 2012
Received in revised form
20 August 2012
Accepted 26 September 2012

Keywords:
Wind power
Numerical atmospheric models
Kalman filters

A B S T R A C T

The issue of wind power prediction is approached in this work by means of numerical and statistical prediction models. Two high resolution regional atmospheric systems are employed in order to provide accurate local wind forecasts while a combination of statistical post processes is utilized targeting to the local adaptation of the results and the reduction of possible systematic biases. A variety of power estimation models are employed for the prediction of the wind power potential in real time applications over two areas of Greece: the islands of Crete and Kefalonia. The results obtained prove that accurate power prediction can be reached if the local environmental conditions are credibly estimated while the use of the power output in previous time steps do not contribute significantly to the improvement of the final forecast.

1. Introduction

The exploitation of renewable energy and especially of wind power is receiving increased attention the last years under the influence of novel policies adopted for energy management and the concerns for global warming and climate change. In this framework, the accurate estimation of the available energy potential for short or long forecasting horizons is of primary importance but not always easy to be achieved due to the variable form and the complexity of the environmental conditions that are directly or indirectly involved.

Several approaches to this problem have been proposed. In general, wind power prediction methods are categorized into two groups: physical and statistical. The first ones imply physical considerations such as topography, terrains, local temperature and pressure to estimate more accurately the wind field (Landberg, 1994) and, subsequently, the energy potential. The latter, on the other hand, use statistical models in order to establish the relationship between power and other variables as well as their historical and forecasted values (Giebel et al., 2001).

Initially, wind power forecasting methods (Madsen, 1996) were based on local measurements providing estimations only for short horizons, less than 6–12 h, well known as persistence models that describe the quasi-stationarity of the atmosphere. In Nielsen and Madsen (1997) the use of meteorological forecasts as input to the statistical prediction models for different prediction horizons is introduced. This approach resulted to significant improvements for larger time scales. Nielsen et al. (1999) suggested a new reference forecast model for larger horizons, which is a weighting between the persistence and the mean of the power. Other wind power prediction models have been developed, combining observations from the wind farms with numerical weather predictions (Joensen et al., 1997). The appearance of measured data in the wind turbine sites offered the opportunity for the development of more sophisticated modeling systems (Giebel et al., 2001) taking into consideration weather predictions, local-site characteristics, on-line measurements and accounting data covering several wind farms. Lately, interest is also given in interval forecasts in order to avoid the inherit uncertainty of single point value predictions. Such methods offer a range of predicted values in every time step and become attractive for the management or the trading of wind power generation.

In this work a combination of physical and statistical models is proposed in order to reach accurate local estimation of the wind conditions over the area of interest and to translate them to wind power potential. More precisely, two state-of-the-art atmospheric numerical models are employed: the SKIRON regional atmospheric system (Kallos, 1997; Papadopoulos et al., 2001), an Eta/NCEP (Janjic, 1994) based non-hydrostatic model, developed at the University of Athens by the Atmospheric Modeling and Weather Forecasting Group (AM&WFG) and the RAMS modeling system (Kallos and Lagouvardos, 1997; Lagouvardos et al., 1996; Mavromatidis and Kallos 2003), a merger of a non-hydrostatic cloud model and a hydrostatic mesoscale model. These numerical prediction systems have been proved capable of simulating atmospheric phenomena with resolution ranging from tens of kilometers to a few meters.

* Corresponding author.
E-mail address: kallos@mg.uoa.gr (G. Kallos).

0167-6105/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jweia.2012.09.004

On the other hand, Kalman filtering algorithms (Crochet, 2004; Galanis and Anadranistakis, 2002; Galanis et al., 2006; Kalman, 1960; Kalman and Bucy, 1961; Kalnay, 2002; Pelland et al., 2011) are employed in order to reduce possible systematic biases that numerical weather prediction models usually exhibit when focusing on local applications near the surface. This is a multiparametric problem to which both the shortcoming in the physical parameterization and the inability of such models to simulate successfully sub-grid phenomena contribute. For example, the models resolution in connection with the smoothing of topographic characteristics results to weak representation of local effects on the airflow. The utilization of both linear and non-linear functions in the proposed Kalman algorithms in combination with Kolmogorov–Zurbenko filters (Eskridge et al., 1997; Rao et al., 1997; Zurbenko, 1986), which remove unnecessary variability and reduce the qualitative differences between the modeled and the recorded wind speed values, provides an efficient way of drastically subtracting possible systematic discrepancies.

The obtained, locally adopted, wind information is subsequently used to estimate the wind power potential based on a variety of statistical – regression based – models. In particular, polynomial functions, the power curve of power generators, as well as non-linear hyperbolic functions are developed utilizing the local wind speed and the energy production in previous time steps in order to estimate the local energy potential. The latter is approached by a deterministic and a stochastic point of view, providing point predictions as well as confidence intervals of the forecasted energy yield.

The systems developed are tested in two areas of Greece, both being advantageous for renewable applications due to the increased wind potential: the islands of Crete, in which two different wind power farms have been selected, and Kefalonia.

The paper is organized as follows: Section 2 is devoted to the description of the numerical weather prediction models and the statistical post processes used, in Section 3 the power prediction algorithms are presented, the two cases studied are analyzed in Section 4, while the main results are presented in Section 5. Finally some concluding thoughts are summarized in Section 6.

2. Description of the models

2.1. The Skiron atmospheric system

Skiron (Kallos, 1997; Papadopoulos et al., 2001) consists of a full physics non-hydrostatic atmospheric model and a series of post-processes, developed at the University of Athens by the Atmospheric Modeling and Weather Forecasting Group (AM&WFG), in the basis of the Eta/NCEP model (Janjic, 1994). The combination of convective, turbulence and surface energy budget schemes support successful mesoscale simulations over regions with varying physiographic characteristics. The model implies NCEP/GFS (National Center for Environmental Prediction/Global Forecast System model) initial meteorological data for operational purposes and SST (Sea Surface Temperature) data at a resolution of 0.5°. Vegetation and topography data are applied at a resolution of 30″ and soil texture data at 2'. The model covers a domain that includes Europe, North Africa, Middle East and the entire Mediterranean region with a horizontal increment of $0.05^\circ \times 0.05^\circ$ (Fig. 1a). In the vertical, 45 Eta levels are used from the ground to the model top. The operational use of the model provides 5-day weather forecasts supplying, among other meteorological parameters, outputs for wind speed and direction, air temperature and mean sea level pressure in different locations where wind farms operate.

2.2. The regional atmospheric modeling system RAMS

RAMS is a well-known numerical code developed at Colorado State University and Mission Research Inc/ASTeR Division. The model (version 4.3.0 for the current work) runs operationally by the AM&WFG providing 48-hour forecasts over Greece. It is based on a non-hydrostatic cloud model (Cotton et al., 2003) and a hydrostatic mesoscale model (Mahrer and Pielke, 1977) simulating and forecasting atmospheric processes on scales varying from a few meters to several thousands of kilometers.

The model entails a two-way interactive nesting of any number of grids and uses various levels of complexity turbulence scheme. In the version employed for the present study a sequel of three domains were used with the first covering the major Mediterranean Region at a 36 km horizontal resolution, the second nested grid of 6 km resolution included the major Greece area and the finest one of 3 km horizontal resolution constrained to Crete in the south part of Greece (Fig. 1b). In the vertical, 32 levels are used. The absence of global physical/numerical routines makes RAMS model appropriate for parallelization. In the current paper, RAMS wind speed data output of various vertical levels were used in wind farms sites.

2.3. The Kalman filter method

One of the most successful methods used for the reduction of the inaccuracies in the Numerical Weather Prediction Models (NWPMs) forecasts is the Kalman filtering approach (Galanis and Anadranistakis, 2002; Galanis et al., 2006; Kalman, 1960; Kalman and Bucy, 1961; Kalnay, 2002; Pelland et al., 2011; Persson 1990), an estimation procedure for dynamic systems. The filter combines a set of mathematical equations using weights between observational and recent forecasting data in order to minimize the corresponding bias. The easy adaptation to any alteration of the observations in connection with the need of short series of background information makes the methodology advantageous. A short description of the algorithmic procedure of a classical Kalman filter is given here.

The main aim of the method is the simulation of an unknown process $\boldsymbol{x}_t$, t denoting the discrete time step. The change of $\boldsymbol{x}$ in

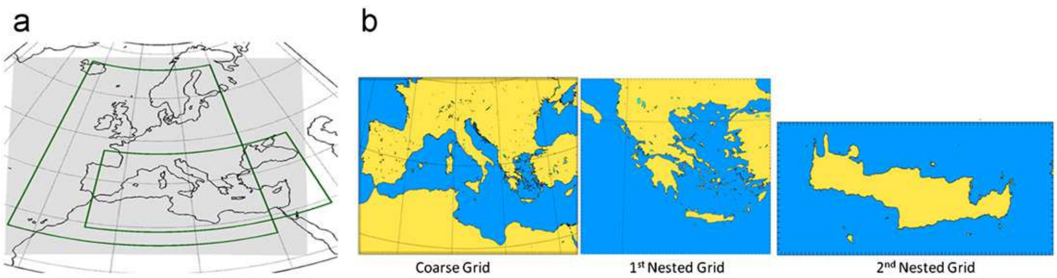

Fig. 1. Numerical weather prediction models domain: (a) Skiron and (b) RAMS models.

time is described by the *System equation*:

$$\mathbf{x_t} = \mathbf{F_t} \cdot \mathbf{x_{t-1}} + \mathbf{w_t} \tag{1}$$

Additionally, a known array $\mathbf{y_t}$ is used. The connection between $\mathbf{y_t}$ and the unknown process is described through the *Observation Equation*:

$$\mathbf{y_t} = \mathbf{H_t} \cdot \mathbf{x_t} + \mathbf{v_t} \tag{2}$$

The covariance matrices $\mathbf{W_t}$ and $\mathbf{V_t}$ of the Gaussian and independent random vectors $\boldsymbol{w_t}$ and $\boldsymbol{v_t}$ along with the coefficient matrices $\mathbf{F_t}$ and $\mathbf{H_t}$ have to be determined before the application of the filter. Kalman filter gives a method for the recursive estimation of the unknown state $\boldsymbol{x_t}$ based on observation values $\boldsymbol{y}$ up to time t. A first estimate of $\boldsymbol{x_t}$ and its covariance matrix $\boldsymbol{P_t}$, based on the previous time step values, is given by

$$\mathbf{x_{t/t-1}} = \mathbf{F_t} \cdot \mathbf{x_{t-1}}, \quad \mathbf{P_{t/t-1}} = \mathbf{F_t} \cdot \mathbf{P_{t-1}} \cdot \mathbf{F_t^T} + \mathbf{W_t} \tag{3}$$

Since the new observation value $\mathbf{y_t}$ is obtained, the estimate of $\boldsymbol{x}$ at time t becomes:

$$\mathbf{x_t} = \mathbf{x_{t/t-1}} + \mathbf{K_t} \cdot (\mathbf{y_t} - \mathbf{H_t} \cdot \mathbf{x_{t/t-1}}) \tag{4}$$

where

$$\mathbf{K_t} = \mathbf{P_{t/t-1}} \cdot \mathbf{H_t^T} \cdot \left(\mathbf{H_t} \cdot \mathbf{P_{t/t-1}} \cdot \mathbf{H_t^T} + \mathbf{V_t}\right)^{-1} \tag{5}$$

is the Kalman gain which arranges how easily the filter adjusts to possible new conditions. The final estimate of $\boldsymbol{P_t}$ is

$$\mathbf{P_t} = (\mathbf{I} - \mathbf{K_t} \cdot \mathbf{H_t}) \cdot \mathbf{P_{t/t-1}} \tag{6}$$

Eqs. (4)–(6) update the Kalman algorithm between two successive time steps.

The above general algorithm can be utilized in various ways, linear or non-linear, depending on the parameters studied. In the current approach a non-linear Kalman filter is applied targeting to the estimation of the bias in time for a single meteorological parameter as a function of the forecasting model direct output. Namely, considering as m_t the direct output of the NWP model at time t, the bias $\mathbf{y_t}$ of this forecast is realized as a polynomial:

$$y_t = x_{0,t} + x_{1,t} \cdot m_t + x_{2,t} \cdot m_t^2 + \cdots + x_{n,t} \cdot m_t^n + v_t \tag{7}$$

where the coefficients $(x_{i,t})$ are the parameters that should be estimated by the filter and v_t the Gaussian non-systematic error. As a result, the state vector is the one formed by the coefficients $(x_{i,t})$:

$$\mathbf{x_t} = \begin{bmatrix} x_{0,t} & x_{1,t} & x_{2,t} & \ldots & x_{n,t} \end{bmatrix}^T \tag{8}$$

and the observation procedure is the scalar bias $\mathbf{y_t}$. The observation matrix takes the form $\mathbf{H_t} = \begin{bmatrix} 1 & m_t & m_t^2 & \cdots & m_t^n \end{bmatrix}$ while as system matrix we use the identity. The system and observation equations (1) and (2) correspondingly become:

$$\mathbf{x_t} = \mathbf{x_{t-1}} + \mathbf{w_t}, \quad \mathbf{y_t} = \mathbf{H_t} \cdot \mathbf{x_t} + \mathbf{v_t} \tag{9}$$

It is worth to be mentioned that the order of the polynomial in Eq. (7) has alternatives: one may employ a function of arbitrary order. The optimum choice depends heavily on the specific case study, on the meteorological parameter and the location of the area. Although, 3rd order polynomials proved to have good contribution in the elimination of the systematic part of the bias with less need in CPU time in previous studies (Galanis et al., 2011), in the present work 1st and 2nd order polynomials turned out to be the optimal choice for the simulation and correction of wind speed evolution.

The variance matrices $\mathbf{W_t}$, of the system equation, and $\mathbf{V_t}$, of the observation equation, are estimated based on the sample of the last three values of $\mathbf{w_t} = \mathbf{x_t} - \mathbf{x_{t-1}}$ and $\mathbf{v_t} = \mathbf{y_t} - \mathbf{x_t}$, respectively:

$$W_t \equiv \frac{1}{3} \cdot \sum_{i=0}^{2} \left((x_{t-i} - x_{t-i-1}) - \left(\frac{\sum_{i=0}^{2} (x_{t-i} - x_{t-i-1})}{3} \right) \right)^2 \tag{10a}$$

$$V_t \equiv \frac{1}{3} \cdot \sum_{i=0}^{2} \left((y_{t-i} - H_{t-i} \cdot x_{t-i}) - \left(\frac{\sum_{i=0}^{2} (y_{t-i} - H_{t-i} \cdot x_{t-i})}{3} \right) \right)^2 \tag{10b}$$

The latter are objective estimators of $\mathbf{W_t}$, $\mathbf{V_t}$ respectively, since the variables $\mathbf{w_t}$ and $\mathbf{v_t}$ denoting the non-systematic part of errors in Eqs. (1) and (2), follow the normal distribution by assumption. The choice of time period of 3 days was proven to be the optimal in the present study as leaded to the achievement of successful corrections and fast adaptability simultaneously. Nevertheless, this choice might vary under different geographic or climatological environments (Galanis et al., 2006).

2.4. *The Kolmogorov–Zurbenko filters*

It is well known that wind speed observations are point measurements recorded at discrete times without smoothing and are therefore discontinuous and highly variable. On the other hand, model outputs are always smoothed in time and space having, therefore, a continuous and mild evolution. In order to reduce the influence of these qualitative differences between the time series under consideration, the use of low pass filters is proposed. Kolmogorov–Zurbenko (KZ) filters (Eskridge et al., 1997; Galanis et al., 2011; Rao et al., 1997) are based on iterative moving averages and are able to remove high frequency variations. Specifically, if $(x_i^0)_i$ stands for the initial values of a series, then the first iteration of the filter results in

$$x_i^1 = \frac{1}{2q+1} \sum_{j=-q}^{q} x_{i+j}^0 \tag{11}$$

where parameter q shapes the length of the filter window: $m=2q+1$. In a next step, these values $(x_i^1)_i$ become the input for the second iteration: $x_i^2 = (1/2q+1) \sum_{j=-q}^{q} x_{i+j}^1$, and so on. The parameter m and the number n of iterations used depend on the portion of the variability that needs to be excluded. The desired separating frequency is

$$\omega_o = \frac{\sqrt{6}}{\pi} \sqrt{\frac{1-(1/2)^{1/2n}}{m^2-(1/2)^{1/2n}}} \tag{12}$$

3. Wind power forecasting approaches

The models applied, targeting to the prediction of the power output, are essentially based on the wind speed w_{obs} and wind power p_{obs} observation data as recorded by the wind turbines at the sites of interest. Long term series of these data allow the decoding of the existing relationship between the inserted wind speed and the energy yield of a turbine. By this way, the actual generation is counted on and several factors that interfere in the production efficiency, like local characteristics, site effects, interaction with neighboring turbines and other features, which are not taken into account in characteristic curves calculated by manufactures in ideal laboratory conditions, are incorporated.

With the use of statistical regression methods, a variety of models are fitted to the data initially. At a second stage, wind speed forecasts w_{for} from NWPMs, optimized by Kolmogorov–Zurbenko and Kalman filters, are utilized in order to estimate the power output of single turbines. Within this framework, the following categories of power estimation models have been utilized:

3.1. Estimation using polynomial regression

One of the basic statistical methods for power prediction is the description of the energy output of a single turbine as a cubic function of the wind speed. The current approach is equivalent to the one initially proposed by Joensen et al. (1999), estimating the derived energy yield by applying multiple regression methods on the predicted wind speed and the measured power of the turbine:

$$p_{t+k}=a_k w_{t+k}+b_k w_{t+k}^2+c_k w_{t+k}^3+d_k p_t+l_k+e_{t+k} \tag{13}$$

Here p_{t+k} is the predicted wind power, w_{t+k} the forecasted wind speed, p_t the power output in a previous time step and a, b, c, d, l, e the corresponding weights.

A simpler approach comes by the application of the previous relation in the absence of the term containing the previous energy production:

$$p_{t+k}=a_k w_{t+k}+b_k w_{t+k}^2+c_k w_{t+k}^3+l_k+e_{t+k} \tag{14}$$

This scheme can obtain different statistical weights depending on wind speed values, as this non-linear type of equation presents variable behavior over different scales of input data.

3.2. Estimation based on the power curve

The implementation of the characteristic curve described below is widely used in power prediction applications (Joensen et al., 1997; Landberg, 1998; Madsen et al., 1995) and provides a safe tool for converting wind speed into power.

Every single wind turbine is associated with a characteristic curve which describes the ideal process of the conversion of wind speed kinetic energy to electric power. The shape of the characteristic curve is similar for different types of wind turbines and is provided by the manufacturer through calibration in identical laboratory conditions.

In order to formulate the correspondence of wind speed and wind power output, two aspects are taken into consideration: The correlation between wind speed and wind power as was specified by the manufacturer and the theoretical expression of the available power described by the equation $P=½^*\rho_{\mathrm{air}}A_{\mathrm{r}}u^3$, noting that the available power P in the free flowing stream of wind is a cubic function of wind speed u, where A_{r} is the rotor swept area exposed to the wind and ρ_{air} the air density. It should be noted that in the present study the – normally piece wised – values of the characteristic curve were fitted to a cubic polynomial in order to avoid discontinuities.

Taking into account that the energy production is depended on the performance of the wind turbine, two mathematical formulations for the energy yield were adopted.

The first expresses the provided power p_{t+k} as a linear analogue of the power stemming from the characteristic curve $po(w_{t+k})$:

$$p_{t+k}=a\cdot po(w_{t+k})+e \tag{15}$$

A relative form of the above equation can be obtained by adding an extra term denoting that the provided power p_{t+k}, is related with the power output in a previous time step:

$$p_{t+k}=a\cdot po(w_{t+k})+b\cdot p_t+e \tag{16}$$

3.3. Estimation based on non-polynomial equations

A typical problem of the previous models for the estimation of wind power concerns the control of the non-linear shape of the obtained power curves which is highly sensitive to wind variations. This sensitivity cannot be easily captured by polynomial models. A way out is proposed here by the adaptation of non-linear hyperbolic functions containing trigonometric terms that describe successfully non-linear behaviors, like the dependency of the power output on forcing wind speed. Analogous methods implementing trigonometric terms have been tested in previous works either to include the diurnal cycle of wind direction (Nielsen and Madsen, 1997) or to entail a correction term of the diurnal fluctuation of the power output (Madsen et al., 1995). In the present approach, we focus on the non-linear relationship between wind speed and wind power output. More precisely, the function of hyperbolic tangent is utilized in a model of the form:

$$p_{t+k}=a_0\cdot \tanh\left(\frac{w_{t+k}-a_1}{a_2}\right)+a_3 \tag{17}$$

where, p_{t+k} is the power output, w_{t+k} the wind speed forcing and a_i the regression coefficients that obtain specific physical meaning: a_0 is a scale factor of the curve, a_1 describes the range of the horizontal axis where the fluctuation of the curve appears, a_2 shapes the curvature and a_3 is a correction term.

An alternative form of the above model can be derived by adding an extra term p_t denoting the measured energy production in a previous stage:

$$p_{t+k}=a_0\cdot \tanh\left(\frac{w_{t+k}-a_1}{a_2}\right)+b\cdot p_t+a_3 \tag{18}$$

3.4. Methods for point prediction

A critical issue for statistically based prediction methods is the estimation of the optimum weights. The determination of these coefficients stems from observations in the site of interest, either from scada systems placed on the wind turbine, or a meteorological mast. As mentioned earlier, observational wind speed data presents high variability that can be smoothed with the application of KZ filters. The regression coefficients express site characteristics and the operational process of the wind turbine during the study period. Following the assumption that this behavior will remain similar in the future, power prediction is feasible based on wind speed forecasts for the area under study. The choice of the proper forecasted wind speed values results from statistical analysis between forecasted and measured data during a training period. Fig. 2 illustrates the steps followed during the point prediction methodology.

3.5. Methods for confidence interval prediction

This methodology, widely applied in recent studies for wind power prediction (Bremnes, 2004; Nielsen et al., 2004; Silva and

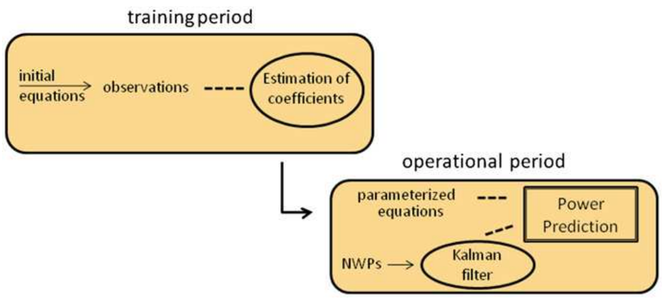

Fig. 2. Flow chart of point prediction method.

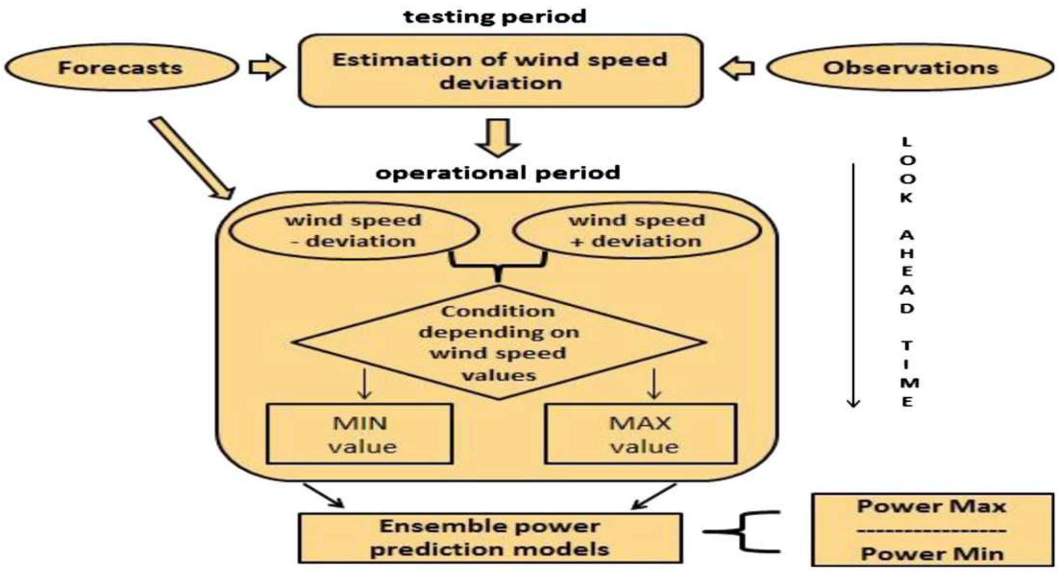

Fig. 3. Flow chart of interval prediction method.

Moulin, 2000), aims to estimate wind power confidence intervals instead of single point forecasts. In non-linear applications, like the conversion of wind speed into energy potential, it is expected that the resulted intervals may not be symmetric.

In the current study the estimation of forecasting intervals depends on the upper and lower value of wind speed. Namely, a testing period for the estimation of the deviation between forecasted and observed values is followed by the application of filters for limiting the interval range. Finally, simple ensembles of these power prediction models provide the desired confidence prediction interval. The methodology adopted is briefly outlined in Fig. 3.

4. The cases studied

In the present work, wind power prediction models were implemented in three operational wind parks, located in Kefalonia (Fig. 4) and Crete (Fig. 5) islands, Greece. Kefalonia is situated in the Ionian Sea, in western Greece and is characterized by strong winds. On the other hand, Crete in the southern part of the Aegean Sea is characterized by high mountains crossing the island and increased wind potential.

The evaluation of the prediction methods performance is based on thorough statistical analysis of SKIRON and RAMS outputs, the Kalman filtered outcomes and the wind power prediction products derived from the trained models.

In particular, the following statistical indexes of accuracy were utilized:

- Mean Bias Error of forecasted values:

$$MBE = \frac{1}{k}\sum_{i=1}^{k}(for(i) - obs(i)) \tag{19}$$

where $obs(i)$ denotes the observed value at time i, $for(i)$ the respective forecasted value (direct model output or filtered forecast), and k the size of the sample. Mean bias error is simply the difference between average forecast and average observation (Wilks, 1995). It provides information for possible prediction overestimations ($MBE > 0$) or underestimations ($MBE < 0$) of the observed values. However, only by itself this statistical index does not offer enough information for the accuracy of individual forecast errors. Additionally, MBE divided by the average of observed values forms the non-dimensional normalized bias $Nbias$.

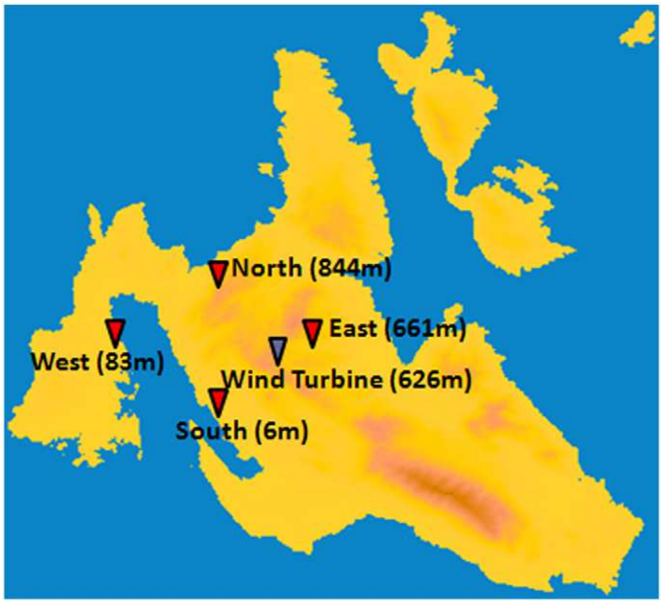

Fig. 4. The map of Kefalonia Island where the four grid points of SKIRON and the location of the wind turbine are depicted. In the brackets, the elevation of each location is given. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

Fig. 5. The two wind farms studied in the east and west part of Crete.

- Root mean square error:

$$RMSE = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(for(i)-obs(i))^2} \quad (20)$$

a classical and widely used divergence measure providing information for the forecast accuracy and variability (Wilks, 1995). Successful forecasts correspond to low values of RMSE, while higher indicate deviations.

- Coefficient of determination-R^2:

$$R^2 = 1-\frac{\sum_{i=1}^{k}(obs(i)-for(i))^2}{\sum_{i=1}^{k}(obs(i)-\overline{for})^2} \quad (21)$$

where $\overline{for}$ is the average of the forecasted values. Coefficient of determination or R^2 is a typical tool for determining the degree of linear-correlation of forecasted and measured values. R^2 varies between 0 and 1 and values close to unity indicate a linear agreement of the tested data sets (Wilks, 1995).

- Index of Agreement-IOA:

$$IOA = 1-\frac{\left(\overline{for}-\overline{obs}\right)}{\sum_{i=1}^{k}\left(|for(i)-\overline{obs}|+|obs(i)-\overline{obs}|\right)^2} \quad (22)$$

with $\overline{obs}$ denoting the average of observation data. This index evaluates the skill of the model in predicting variations about the observed mean (Willmott et al., 2011). IOA ranges between zero and one and values above 0.5 imply efficiency in the prediction.

For wind speed data, three more measures have been mobilized:

- Kurtosis:

$$kurtosis = \left\{\frac{k(k+1)}{(k-1)(k-2)(k-3)}\sum_{i=1}^{k}\left(\frac{x_i-\overline{x}}{s}\right)^4\right\}-\frac{3(k-1)^2}{(k-2)(k-3)} \quad (23)$$

where k denotes the size of the sample, $\overline{x}$ is the sample mean and s is the sample standard deviation. Kurtosis is the degree of the relative peakness of a distribution, compared with the standard distribution. Positive/negative kurtosis indicates a relatively peaked/flat distribution.

- Skewness:

$$skewness = \frac{k}{(k-1)(k-2)}\sum_{i=1}^{k}\left(\frac{x_i-\overline{x}}{s}\right)^3 \quad (24)$$

a parameter that characterizes the degree of asymmetry of a distribution around its mean. Positive skewness indicates a distribution with the long tail extending towards more positive values, whereas if the reverse is true the distribution has negative skewness.

- Scatter index:

$$sc\_Ind = \frac{standard\ deviation\ of\ error}{mean\ observed\ value} \quad (25)$$

which measures the deviation between observations and predictions as a portion of the average observed value (Clancy et al., 1986). Lower values indicate better forecasting skill.

4.1. Case 1: Kefalonia wind park

In the case of Kefalonia, an offline prediction method was applied.

SKIRON NWP data of a statistically long time period (from May 2009 to April 2010), have been provided for the area of Tetrapolis wind farm (blue pointed in Fig. 4). In particular, SKIRON forecasts for wind speed at 10 m and 116 m above the ground and for four grid points around the wind park, i.e. East (Lon: 20.58°, Lat: 38.22°), West (Lon: 20.43°, Lat: 38.23°), North (Lon: 20.50°, Lat: 38.29°), South (Lon: 20.50°, Lat: 38.17°), and for an additional point coinciding with park's site, obtained by interpolation (noted as Int in Table 1), were available. At the same time, observations of wind speed, energy production, as well as, their maximum and minimum values, from a specific wind turbine of the park, were provided. In Fig. 4 the four grid points of the NWP model along with the location of the wind turbine (Lon: 20.54°, Lat: 38.22°) are presented.

A study for the optimum choice of the model level and grid point was performed based on statistical analysis of SKIRON outputs as presented on Table 1. The strongest correlation between wind speed forecasts and the corresponding observations was succeeded in the case of the South grid point and the first model level (116 m). In particular, *RMSE* and *MBE* imply that this grid point presents more accurate behavior in the mesoscale range, while the rest of the statistical indexes do not seem to present any significant differences.

4.2. Case 2: Crete wind parks

For the area of Crete, RAMS wind speed forecasts were provided for two different locations: Xirolimni (Lon: 23.82°, Lat: 35.36°) and Vardia (Lon: 26.18°, Lat: 35.16°) pointed out in Fig. 5, where the wind farms of interest are operated. Wind data were available on a daily basis for a period of 9 months, namely from June 2010 to February 2011. The power prediction system designed was updated in daily basis with wind speed and total power measurements from each wind farm.

For both the sites under study, a Kolmogorov–Zurbenko filter with coefficients ($m=3$, $n=26$) was employed, in order to smooth observations and reduce the qualitative differences between them and the corresponding forecasts. This particular filter is equivalent

Table 1
Statistical analysis for the NWP model output.

	Wind speed at 10 m above ground					Wind speed at 116 m above ground				
	Int	S	N	E	W	Int	S	N	E	W
MBE (m/s)	−4.31	−4.23	−4.41	−3.64	−3.59	−3.00	−2.14	−3.01	−3.62	−2.03
RMSE (m/s)	5.59	5.70	5.71	4.89	5.19	4.33	3.87	4.35	4.97	4.27
R^2	0.52	0.39	0.50	0.58	0.40	0.60	0.55	0.59	0.53	0.44
IOA	0.40	0.40	0.39	0.48	0.48	0.55	0.60	0.54	0.48	0.56
St_Dv (m/s)	3.55	3.82	3.63	3.26	3.75	3.12	3.22	3.15	3.41	3.75
Sc_Ind	0.44	0.47	0.45	0.40	0.46	0.39	0.40	0.39	0.42	0.46

to a cut-off frequency of 0.031711 (see the corresponding criterion in Eq. (12)).

Moreover, a Kalman filter based on a polynomial of second order was applied to all wind speed SKIRON data of Tetrapolis wind farm in Kefalonia targeting to the reduction of possible systematic biases. On the other hand, the Kalman filter implemented to RAMS wind speed forecasts for the Crete area was characterized by a polynomial of first order as it was proved to have a better performance in this case.

The wind power prediction models discussed in Section 3 after trained over a 6 month learning period (June – November 2009) resulted in the following analytical form for the Kefalonia case:

$$M1: p_{t+k} = 10.47 \cdot w_{t+k} + 7.09 \cdot w_{t+k}^2 - 0.25 \cdot w_{t+k}^3 + 0.0023 \cdot p_t - 108.28$$

$$M2: p_{t+k} = 10.45 \cdot w_{t+k} + 7.08 \cdot w_{t+k}^2 - 0.246 \cdot w_{t+k}^3 - 108.5$$

M3 :

- If the forecasted wind speed is $0 < w_{t+k} \leq 4$ m/s, the predicted energy production is $p_{t+k} = 0$.
- If the forecasted wind speed is $4 < w_{t+k} \leq 10$ m/s, energy production is
$$p_{t+k} = -22.5523 \cdot w_{t+k} + 4.046 \cdot w_{t+k}^2 + 0.313 \cdot w_{t+k}^3 + 0.0104 \cdot p_t + 21.39$$
- If $10 < w_{t+k} \leq 14$ m/s, then
$$p_{t+k} = 1847.8 \cdot w_{t+k} - 133.5 \cdot w_{t+k}^2 + 3.3 \cdot w_{t+k}^3 - 7917.6$$
- If $w_{t+k} \geq 14$ m/s, then $p_{t+k} = 810$

$$M4: p_{t+k} = 1.019 \cdot po(w_{t+k}) - 9.741$$

$$M5: p_{t+k} = 1.0119 \cdot po(w_{t+k}) - 0.0028 \cdot p_t - 9.1904$$

$$M6; p_{t+k} = 404 \cdot \tanh\left(\frac{w_{t+k} - 9.1}{2.864}\right) + 409.4$$

$$M7: p_{t+k} = 403.51 \cdot \tanh\left(\frac{w_{t+k} - 9.1}{2.864}\right) + 0.025 \cdot p_t + 407.6$$

where p_{t+k} denotes the predicted wind power, w_{t+k} the forecasted wind speed, p_t the measured energy yield in a previous time step and $po(w_{t+k})$ the power derived from the characteristic power curve.

In the Crete case, wind speed and power output observations for the time period from June 2010 to September 2011, were used to train the power prediction models. The obtained formulas were as follows.

4.2.1. Xirolimni park

$$M1: p_{t+k} = -41.53 \cdot w_{t+k} + 10.05 \cdot w_{t+k}^2 - 0.328 \cdot w_{t+k}^3 + 0.0341 \cdot p_t + 26.08$$

$$M2: p_{t+k} = -42.33 \cdot w_{t+k} + 10.29 \cdot w_{t+k}^2 - 0.336 \cdot w_{t+k}^3 + 28.14$$

M3 :

- If the forecasted wind speed is $0 < w_{t+k} \leq 4$ m/s, the predicted energy production is $p_{t+k} = 0$.
- If the forecasted wind speed is $4 < w_{t+k} \leq 10$ m/s, energy production is
$$p_{t+k} = -88.24 \cdot w_{t+k} + 14.45 \cdot w_{t+k}^2 - 0.43 \cdot w_{t+k}^3 + 0.0141 \cdot p_t + 164.43$$
- If $10 < w_{t+k} \leq 14$ m/s, then
$$p_{t+k} = -722.27 \cdot w_{t+k} + 70.8 \cdot w_{t+k}^2 - 2.1 \cdot w_{t+k}^3 + 2537.7$$
- If $w_{t+k} \geq 14$ m/s, then $p_{t+k} = 600$

$$M4: p_{t+k} = 0.9245 \cdot po(w_{t+k}) - 22.5717$$

$$M6: p_{t+k} = 305.9 \cdot \tanh\left(\frac{w_{t+k} - 9.933}{3.68}\right) + 298.8$$

4.2.2. Vardia park

$$M1: p_{t+k} = -58.54 \cdot w_{t+k} + 11.2 \cdot w_{t+k}^2 - 0.346 \cdot w_{t+k}^3 + 0.008 \cdot p_t + 78.067$$

$$M2: p_{t+k} = -58.81 \cdot w_{t+k} + 101.25 \cdot w_{t+k}^2 - 0.347 \cdot w_{t+k}^3 + 78.89$$

M3 :

- If the forecasted wind speed is $0 < w_{t+k} \leq 4\ m/s$, the predicted energy production is $p_{t+k} = 0$.
- If the forecasted wind speed is $4 < w_{t+k} \leq 10$ m/s, energy production is
$$p_{t+k} = -41.29 \cdot w_{t+k} + 7.08 \cdot w_{t+k}^2 - 0.107 \cdot w_{t+k}^3 + 0.019 \cdot p_t + 69.67$$
- If $10 < w_{t+k} \leq 14$ m/s, then
$$p_{t+k} = -432.5 \cdot w_{t+k} + 46.2 \cdot w_{t+k}^2 - 1.4 \cdot w_{t+k}^3 + 1340.9$$
- If $w_{t+k} \geq 14$ m/s, then $p_{t+k} = 600$

$$M4: p_{t+k} = 0.8442 \cdot po(w_{t+k}) - 14.9197$$

$$M5: p_{t+k} = 0.797 \cdot po(w_{t+k}) + 0.083 \cdot p_t - 18.8637$$

$$M6: p_{t+k} = 315.3 \cdot \tanh\left(\frac{w_{t+k} - 10.47}{4.106}\right) + 304.5$$

The evaluation period of the power prediction models described above lasted from October to December 2010 for the wind farm operating in Xirolimni while its duration was about 2 months longer in the case of Vardia.

5. Results

5.1. Evaluation of the wind speed forecasts

A significant part of the wind power forecasting error is usually due to the discrepancies introduced by the NWP wind forecasts. This is a multiparametric problem to which different factors contribute with the local area characteristics keeping a primary role. Kalman filters provide an efficient way to deal with these problems.

In the present work, Kalman filters with a characteristic polynomial of different degree were applied. More specifically, both 1st and 2nd order filters were used in the case of Kefalonia, whereas, Crete island wind speed forecasts were improved by applying a filter with polynomial of 1st order alone, as it was explicitly proven to lead to forecasts of better quality.

In Table 2 the main statistical indexes of accuracy used for the evaluation of wind speed forecasts are given: More precisely, the bias, root mean square error, coefficient of determination, index of agreement, kurtosis and skewness of the model direct output, the Kalman filtered outcome and the observations for the forecasting period designated for each wind farm are provided. It is clear that both NWP models underestimate the wind speed. The application of Kalman filtering improves the model output by reducing the bias significantly to approach zero. However, no significant improvement, concerning the RMSE is observed. On the other hand, R^2 and IOA values indicate that wind speed predictions both before and after the application of the filter do not present linear relationship with the observational data set. As far as kurtosis and skewness concerned, the distribution of the filtered forecasts is more symmetric and better fits to that of the observed wind speed data.

Comparing the two different Kalman filters applied in the case of Kefalonia island, the 2nd order filter seems to prevail, yet the particular filter leads to a slight increased RMSE. Another critical remark forthcoming is the better performance of the RAMS model at its second level (154 m) where the statistical indexes reveal more accurate direct and filtered results.

The fitting of the observed wind speed data, model direct output and Kalman filtered outcomes to the Weibull distribution,

Table 2
Bias, RMSE, Kurtosis and Skewness for the direct model output and the Kalman filtered results.

	Wind speed ("Tetrapolis", Kefalonia)			
	SKIRON (116m)	Kalman (1st order)	Kalman (2nd order)	Observations
MBE (m/s)	−2.13	−0.60	−0.27	–
RMSE (m/s)	3.87	3.89	4.29	–
R^2	0.55	0.43	0.35	–
IOA	0.60	0.61	0.56	–
Kurt	0.12	0.21	0.21	1.29
Skew	0.71	0.65	0.65	1.08

	Wind speed ("Xirolimni", E Crete)				
	RAMS (47 m)	Kalman	RAMS (154 m)	Kalman	Observations
MBE (m/s)	−2.73	−0.82	−2.60	−0.26	–
RMSE (m/s)	4.20	3.87	3.89	3.63	–
R^2	0.28	0.19	0.52	0.48	–
IOA	0.57	0.51	0.69	0.63	–
Kurt	−0.86	−0.74	−0.80	−0.58	−0.36
Skew	0.09	0.12	0.07	−0.08	0.61

	Wind speed ("Vardia", W Crete)				
	RAMS (47 m)	Kalman	RAMS (154 m)	Kalman	Observations
MBE (m/s)	−2.76	−0.28	−1.11	−0.09	–
RMSE (m/s)	5.67	4	3.53	3.85	–
R^2	0.29	0.49	0.61	0.56	–
IOA	0.58	0.68	0.72	0.68	–
Kurt	−0.67	-0.12	0.78	0.66	0.28
Skew	0.62	0.74	0.79	0.90	0.98

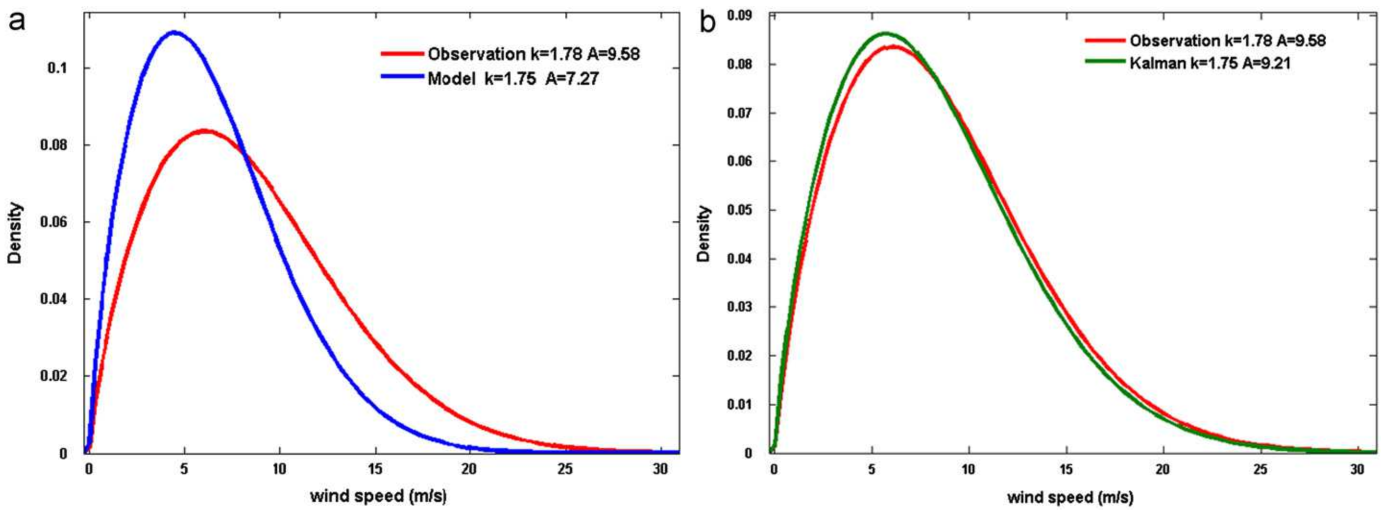

Fig. 6. Weibull distribution curves for (a) Direct Skiron model output prediction (blue) versus corresponding observations (red), (b) Kalman filtered outcome (green) versus corresponding observations (red). Both for wind farm "Tetrapolis" at Kefalonia. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

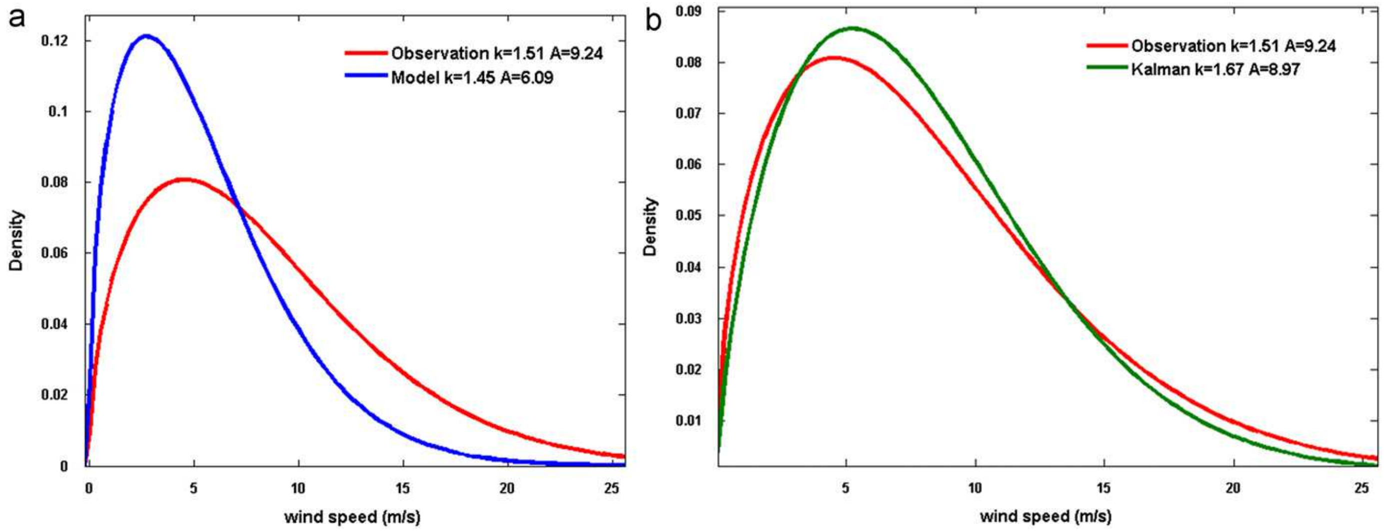

Fig. 7. Weibull distribution curves for (a) Direct RAMS – 1st level model output prediction (blue) versus corresponding observations (red), (b) Kalman filtered outcome (green) versus corresponding observations (red). Both for Vardia wind farm, W. Crete. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

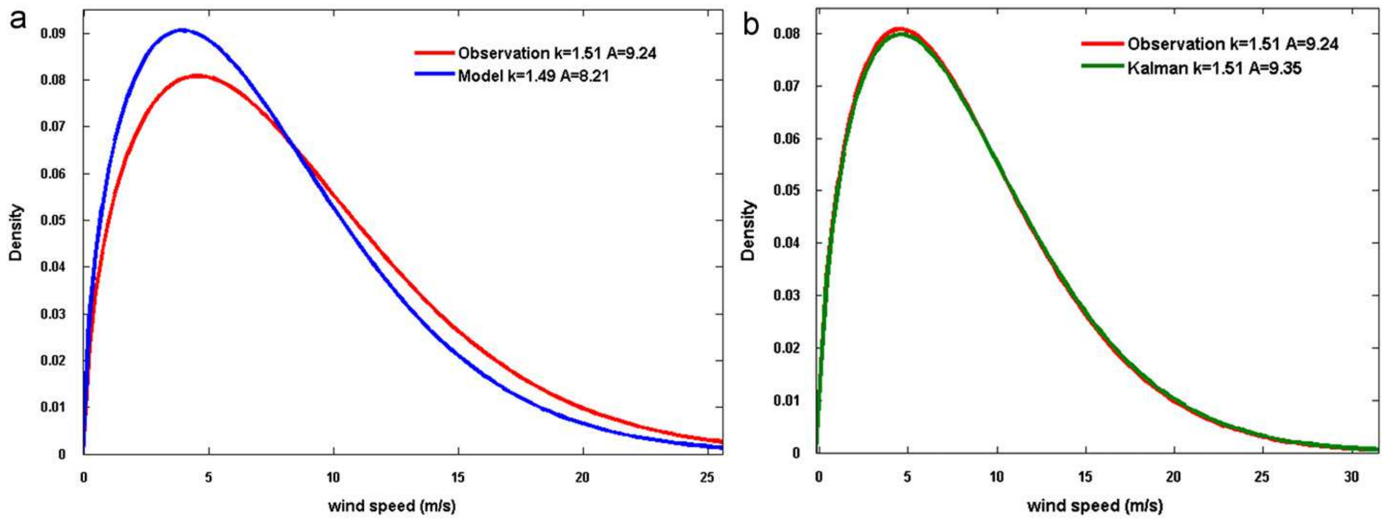

Fig. 8. Weibull distribution curves for (a) Direct RAMS – 2nd level model output prediction (blue) versus corresponding observations (red), (b) Kalman filtered outcome (green) versus corresponding observations (red). Both for wind farm (Vardia W. Crete). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

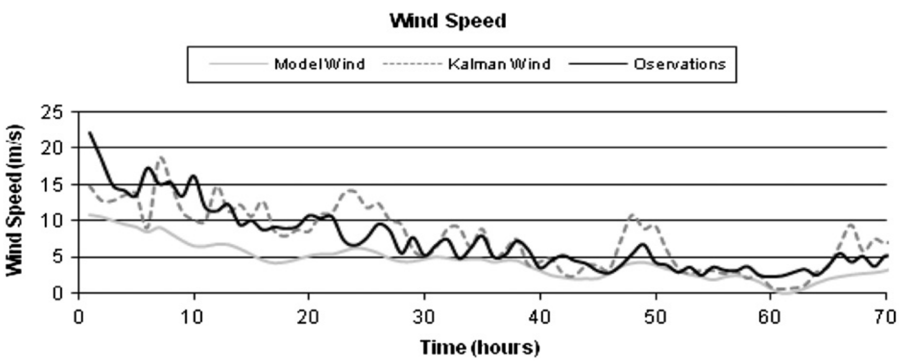

Fig. 9. Time series of the wind speed forecasted, Kalman filtered and observed ("Tetrapolis" Kefalonia).

illustrated in Figs. 6–8, further support not only the significant improvement in wind speed forecasts after the implementation of the Kalman filter but also the better performance of the RAMS model at its 2nd level. In Fig. 9, the elimination of the systematic error emerged, regardless its type, becomes clearer through a characteristic set of wind speed time series.

5.2. Evaluation of the power prediction models

In order to compare the different types of wind power prediction models implied in the present study, a variety of statistical indexes of accuracy is given.

Table 3 indicates slight differences in the performance of all the seven models with an exception in the case of Kefalonia, where a particularly small bias is resultant for those models that use a 3rd order polynomial of the wind (M1, M2). Thus, in the examination of the rest wind park, models M5, M7 for Xirolimni and M7 for Vardia did not offer any extra information and for this reason are not presented. The term of the power output in a previous time step added in some models (M1, M5 and M7) did not lead to any improvement of the predictions.

In the present study 12 h was the previous time step used as observational power output, i.e. 12 h is the *k*-step forward in the case of models using past measured energy production. Employment of other time steps such as 2, 6 and 24 h was tested (not presented) and led neither to finer results nor to any remarkable difference.

It is concluded, in this way, that the variational nature of wind plays the most important role in wind power prediction and restricts the influence of past power observational values. This outcome might be restricted for similar to the examined test cases and is associated with the hourly time step that information was available.

Furthermore, attempting a deeper analysis, evaluation indexes *MBE* and *Nbias* for Kefalonia park site, evident better results compared to the two sites of Crete. This might be explained from the wider nominal power (800 kw) in the examined wind turbine instead of 600 kw that the two other turbines in Crete encompass. Results for R^2 do not indicate remarkable linearity between forecasted and measured wind power values, though the goal of this study is the elimination of errors in long term horizon. Index of agreement shows sufficient prediction around the central tendency of the measured sets both for Kefalonia and Vardia wind parks.

In Table 4, an alternative statistical analysis is presented. Forecasted wind speed is now replaced by observations in wind power prediction models. The results provide a more accurate view of the models performance, as the uncertainty of wind forecasts is excluded. The *MBE*, *RMSE* reduction and R^2 results witness an overall good performance of the models regardless of whether they are based on wind speed or the power curve for Kefalonia case. For Vardia and Xirolimni cases, statistics reflect the potential ability of the models to simulate wind speed into the energy yield, however with less success compared to Kefalonia case.

Relevant behavior appears in Fig. 10 where the correspondence between the measured power output and the one estimated from the models is presented.

The strong relationship between the errors emerged in wind speed forecasts and the power generation predictions is further supported in Fig. 11, where the wind speed bias and the power bias are illustrated indicatively for the case of Kefalonia and the prediction models M2 and M4. It is worth noticing here the linear association of the two biases. As a result, temporal periods with minor divergence between wind speed forecasts and observations, equivalently leads to reduced divergence in the predicted power (Fig. 12).

5.3. Evaluation of the prediction interval method

An alternative approach to the wind power prediction problem is the development of a probabilistic model stated hereafter as prediction interval (see also Chatfied, 2000). In contrast to the usual deterministic prediction methods, an interval forecast is a range of values within which the real observational value is expected to be found. These forecast intervals are not necessarily symmetric and, therefore, they cannot be evaluated by the dominant statistic rules of confidence intervals (Christoffersen, 1998). If $B_{p_{t+k}}$ denotes the bottom limit and $H_{p_{t+k}}$ the upper limit, respectively of the power prediction interval, then

$$I_{p_{t+k}} = H_{p_{t+k}} - B_{p_{t+k}} \tag{26}$$

depicts the forecast interval at the specific time.

Table 3
Statistical indicators for wind power predictions.

	Wind power forecasting results analysis ("Tetrapolis", Kefalonia)						
	M1	M2	M3	M4	M5	M6	M7
MBE (kW)	3.94	2.74	16.30	27.53	24.35	12.75	13.16
RMSE (kW)	267.39	267.45	285.87	284	284	280	281.3
Nbias	0.08	0.02	0.04	0.07	0.06	0.03	0.03
R^2	0.36	0.37	0.35	0.36	0.36	0.37	0.36
IOA	0.68	0.67	0.66	0.66	0.67	0.66	0.66

	Wind power forecasting results analysis (Xirolimni, E. Crete)				
	M1	M2	M3	M4	M6
MBE (kW)	48.25	53.14	54.70	56.28	41.85
RMSE(kW)	187.91	190.78	195.91	208.30	200.88
Nbias	0.17	0.23	0.29	0.32	0.34
R^2	0.13	0.12	0.12	0.11	0.12
IOA	0.46	0.39	0.46	0.40	0.46

	Wind power forecasting results analysis (Vardia, W. Crete)					
	M1	M2	M3	M4	M5	M6
MBE (kW)	30.88	31.52	37.85	29.91	20.63	36.71
RMSE(kW)	180.46	180.88	187.68	159.38	153.65	169.89
Nbias	0.18	0.17	0.19	0.12	0.17	0.21
R^2	0.44	0.43	0.45	0.46	0.46	0.47
IOA	0.63	0.63	0.63	0.64	0.64	0.64

Table 4
Wind power prediction statistics when wind speed observations are used as an input to the prediction models.

	Power prediction model performance – input wobs ("Tetrapolis", Kefalonia)						
	M1	M2	M3	M4	M5	M6	M7
MBE (kW)	−3.03	−4.26	3.33	16.76	13.06	2.37	2.37
RMSE(kW)	76.85	77.13	59.38	59.69	52.16	46.76	46.76
Nbias	0.02	0.01	0.01	0.03	0.03	0.01	0.01
R^2	0.88	0.92	0.98	0.97	0.97	0.98	0.98
IOA	0.82	0.88	0.96	0.95	0.95	0.96	0.96

	Power prediction model performance – input wobs (Xirolimni, E. Crete)				
	M1	M2	M3	M4	M6
MBE (kW)	21.69	22.37	35.40	30.18	29.04
RMSE(kW)	115.56	116.02	115.83	104.25	108.51
Nbias	0.16	0.21	0.28	0.22	0.21
R^2	0.79	0.79	0.74	0.78	0.79
IOA	0.81	0.81	0.80	0.84	0.85

	Power prediction model performance – input wobs (Vardia, W. Crete)					
	M1	M2	M3	M4	M5	M6
MBE (kW)	24.07	19.49	23.82	22.15	22.15	29.77
RMSE(kW)	112.77	121.53	112.60	97.54	97.54	107.65
Nbias	0.14	0.12	0.09	0.10	0.11	0.08
R^2	0.78	0.77	0.79	0.72	0.70	0.77
IOA	0.82	0.81	0.83	0.81	0.80	0.82

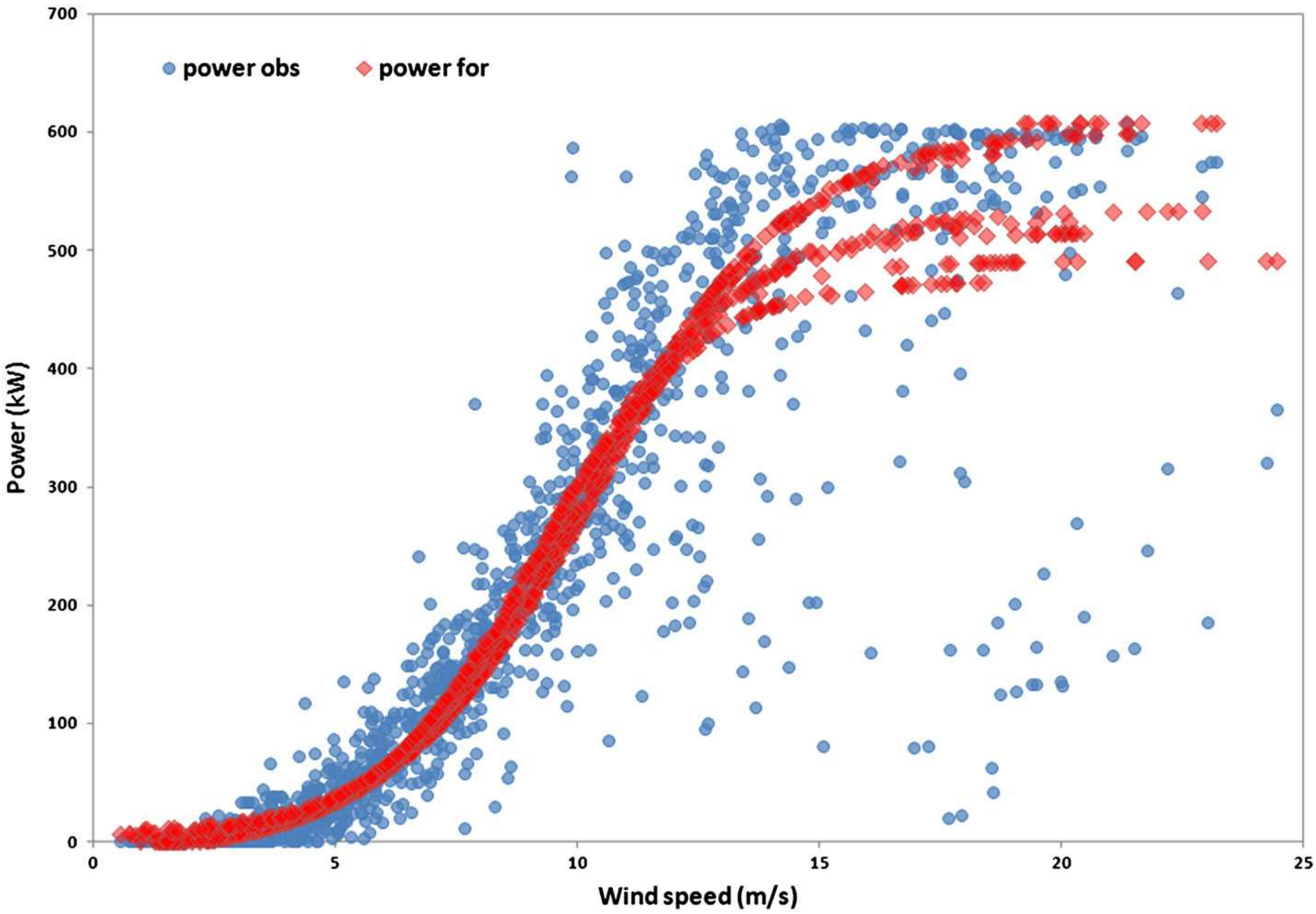

Fig. 10. Measured and estimated power derived from observed wind speed values (case of Vardia, W. Crete).

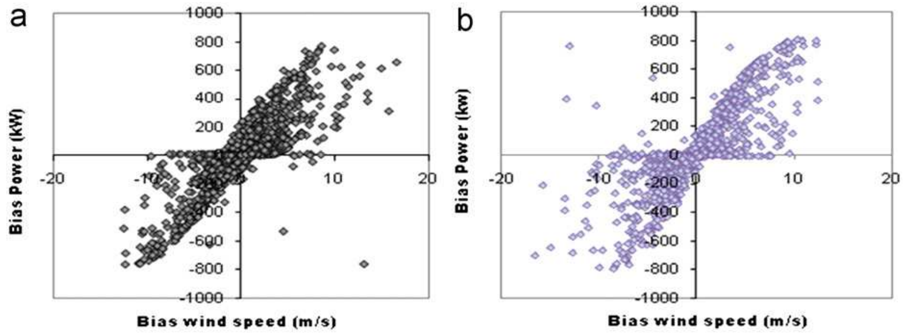

Fig. 11. Wind speed Bias versus Power Bias in the case of Kefalonia Island: (a) for the model M2 and (b) for the model M4.

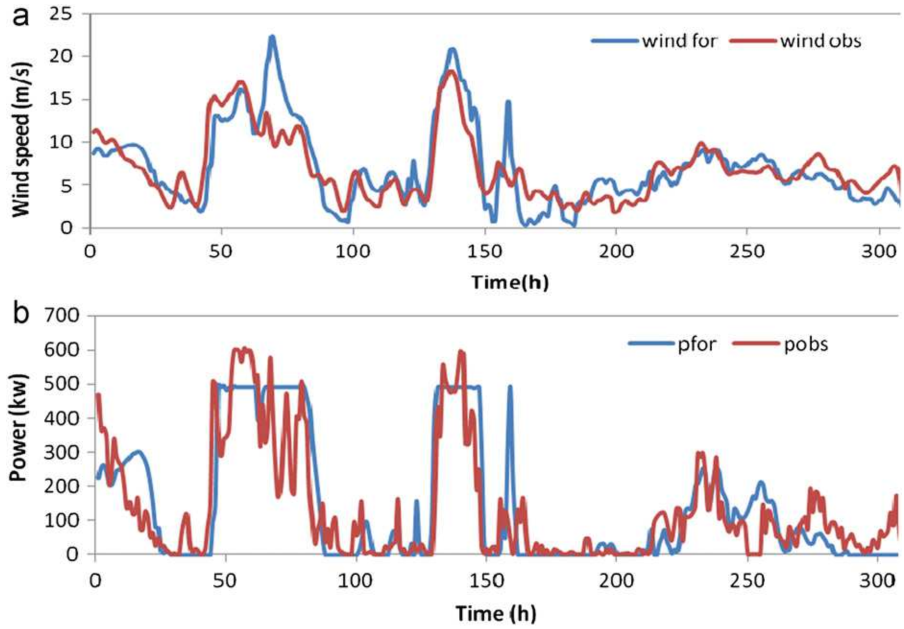

Fig. 12. Time series of observations and forecasts for the (a) wind speed and (b) wind power resulting from the model M4 (case of Vardia, W. Crete).

Table 5
Standard deviation and score skill for wind speed and power prediction interval.

	Wind speed interval (Kefallonia)	Power interval (Kefallonia)	Wind speed interval (Vardia)	Power interval (Vardia)
Standard deviation	4.37	281	5.44	217
Score skill	54	67	66	72

An optimization procedure is also introduced in order to evaluate the statistical performance of the interval method. The score skill

$$S_c = \frac{\sum_{i=1}^{n} p_{obs_{t+k}} \in I_{p_{t+k}}}{n} \tag{27}$$

indicates the sum of the cases were the recorded prediction value $p_{obs_{t+k}}$ was found inside the interval bounds compared to the entire test horizon. This statistical test is simply the number of successful cases (percentage), where the observed value was situated inside the predicted margin (an alternative definition of indicator variable stated in Christoffersen (1998)). Since the margin $I_{p_{t+k}}$ is not stable, the direct comparison with the nominal power is not possible. It is obvious that an interval of 600 kw in a nominal power of 800 kw, for example, could be considered successful but it would be of no prediction value.

An empirical tool for evaluation is adopted here based on the calculation of the main distance between the two forecasting bounds stemming from the standard deviation $\sigma(p_{\min}, p_{\max})$. A comparison between the value of σ and the nominal power provides the opportunity to characterize the uncertainty of the method followed, especially for the computation of the two bounds. Table 5 presents the standard deviation between the two bounds of wind speed ($w_{\min}, w_{\max}$) and power output ($p_{\min}, p_{\max}$) as resulted for the prediction method. Moreover the score skill of these variables is given:

As expected, the success of the power forecast interval follows the one of the wind speed. The standard deviation in the power interval, which is the key parameter in terms of the method's reliability, shows that the forecasted interval covers an amount of almost 35 percentage of the nominal power in

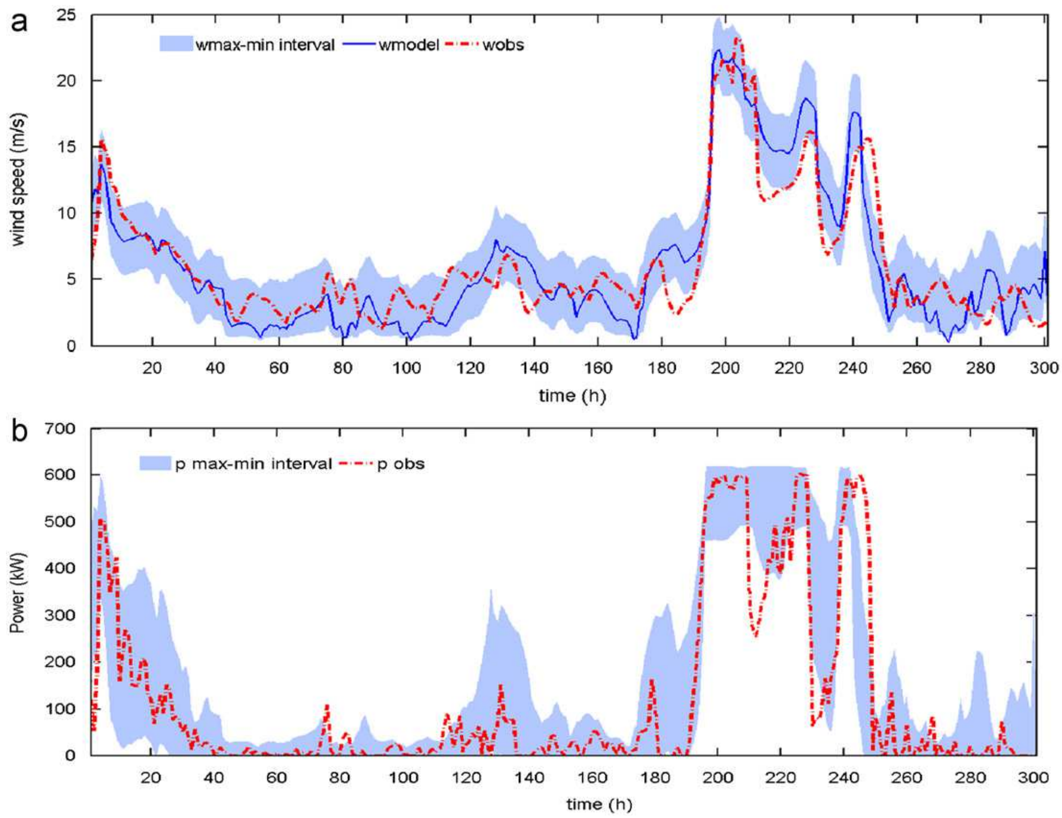

Fig. 13. Time series of the observations versus forecast intervals (a) wind speed and (b) wind power (case of Vardia, W. Crete).

both test cases. On the other hand, the standard deviations of the wind speed and power forecast interval do not have the same correspondence in the two test cases. This can be attributed to the ensemble of different equations for the calculation of wind power.

As discussed earlier, the application of filters in order to correct the initial wind speed forecasted bounds leads to asymmetric wind speed forecasts. A similar behavior appears in the evolution of wind power values in time (Fig. 13).

6. Conclusions

The issue of wind power prediction has been studied in this work by means of physical and statistical modeling. High resolution numerical weather prediction systems (the Skiron/Eta and the RAMS model) are employed for simulating the local air flow over the sites of interest while mathematical post processes, based on Kalman and Kolmogorov–Zurbenko filters, are used in order to adopt local area characteristics and to eliminate possible systematic errors. The obtained wind simulations are used for the estimation of wind power potential by statistical regression models that engage a variety of linear and non-linear approaches.

The above systems have been applied in three wind farms located at the islands of Crete and Kefalonia, Greece. The main results/conclusions made can be summarized as follows:

- Kalman filters improved significantly the forecasting performance of the numerical weather prediction models by almost eliminating the corresponding biases. However, no significant improvement of the variability of the error was succeeded.
- The distribution of the modeled wind speed values was corrected by the use of Kalman filters being translated very close to that of the corresponding observations.
- All the power prediction models used provide comparable, successful in general, results. The use of non-linear hyperbolic functions is advantageous since it provides more sensitive ways of controlling the steepness of the power prediction curve.
- The use of previous power prediction values did not seem to support the improvement of power forecasts. The variational nature of wind speed prevails and the power predictions are strongly sensitive on any biases induced by the wind forecasting systems.
- Forecast confidence intervals obtained by stochastic power predictions models are able to cover a significant portion of the nominal power.

The issue of wind power forecasting is considered today one of the hottest topics of applied research and provides new challenges to the research community. The present paper is aiming to contribute to this discussion and to trigger new activities and studies towards accurate wind power predictions.

Acknowledgments

This work is supported by the FP7 projects WAUDIT (Project reference number: 238576) and MARINA PLATFORM (Grant agreement number 241402). Wind data for Kefallonia were provided by ELLINIKI TECHNODOMIKI ANEMOS and wind data for Crete were provided by the administrator of PPC in Crete.

References

Bremnes, J.B., 2004. Probabilistic wind power forecasts using local quantile regression. Wind Energy 7 (1), 47–54.
Chatfied, C., 2000. Time-series Forecasting. Chapman & Hall/ CRC.

Christoffersen, P.F., 1998. Evaluating interval forecasts. International Economic Review, 39, 841–862.

Clancy, R., Kaitala, J., Zambreski, L., 1986. The Fleet Numerical Oceanography Center Global Spectral Ocean Wave Model. Bulletin of the American Meteorological Society 67, 498–512.

Cotton, W.R., Pielke Sr., R.A., Walko, R.L., Liston, G.E., Tremback, C.J., Jiang, H., McAnelly, R.L., Harrington, J.Y., Nicholls, M.E., Carrio, G.G., McFadden, J.P., 2003. RAMS 2001: current status and future directions. Meteor. Atmospheric Physics 82, 5–29.

Crochet, P., 2004. Adaptive Kalman filtering of 2-metre temperature and 10-metre wind-speed forecasts in Iceland. Meteorological Applications 11, 173–187.

Eskridge, R.E., Ku, J.Y., Rao, S.T., Porter, P.S., Zurbenko, I.G., 1997. Separating different scales of motion in time series of meteorological variables. Bulletin of the American Meteorological Society 78 (7), 1473–1483.

Galanis, G., Anadranistakis, M., 2002. A one dimensional Kalman filter for the correction of near surface temperature forecasts. Meteorological Applications 9, 437–441.

Galanis, G., Louka, P., Katsafados, P., Kallos, G., Pytharoulis, I., 2006. Applications of Kalman filters based on non-linear functions to numerical weather predictions. Annales Geophysicae 24, 2451–2460.

Galanis, G., Chu, P.C., Kallos, G., 2011. Statistical post processes for the improvement of the results of numerical wave models. A combination of Kolmogorov–Zurbenko and Kalman filters. Journal of Operational Oceanography 4 (1), 23–31.

Giebel, G., Landberg, L., Nielsen, T.S., 2001, The ZEPHYR project: The next generation prediction system. In: Proceedings of the 2001 European Wind Energy Conference, EWEC'01, Copenhagen, Denmark, June 2001, pp. 777–780.

Janjic, Z.I., 1994. The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. Monthly Weather Review 122, 927–945.

Joensen, A.K., Madsen, H., Nielsen, T.S., 1997. Non-Parametric Statistical Methods for Wind Power Prediction. EWEC'97, Dublin, Denmark.

Joensen, A.K., Giebel, G., Landberg, L., Madsen, H., Nielsen, H.A., 1999, Model output statistics applied to wind power prediction. In: Wind Energy for the Next Millennium, European Wind Energy Conference, Nice, France, 1–5 March 1999, pp. 1177–1180.

Kallos G., 1997, The regional weather forecasting system SKIRON. In: Proceedings, Symposium on Regional Weather Prediction on Parallel Computer Environments, 15–17 October 1997, Athens, Greece, 9 pp.

Kallos, G., Lagouvardos, K., 1997: Atmospheric Modeling Simulations over the Eastern USA with the RAMS3b Model for the Summer 1995. Final Report to Electric Power Research Institute, Palo Alto, CA, 69 pp.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Transactions on ASME Series D 82, 35–45.

Kalman, R.E., Bucy, R.S., 1961. New results in linear filtering and prediction problems. Transactions on ASME Series D 83, 95–108.

Kalnay, E., 2002. Atmospheric Modeling, Data Assimilation and Predictability, 341. Cambridge University Press.

Lagouvardos, K., Kotroni, V., Nickovic, S., Kallos, G., 1996. Evidence of a winter tropical storm over eastern Mediterranean: simulations with the Regional Atmospheric Modelling System (RAMS) and the ETA/NMC model. In: Proceedings of the 7th International Conference on Mesoscale Processes, 9–13 September, Reading, UK.

Landberg, L., 1994. Short-term Predictions of Local Wind Conditions. Ph.D. thesis, Riso National Laboratory, Denmark.

Landberg, L.A., 1998. Mathematical look at a physical power prediction model. Wind Energy 1, 23–28.

Madsen, H., 1996. Models and Methods for Predicting Wind Power, Department of Mathematical Modelling. Technical University of Denmark, Denmark.

Madsen, H.K., Sejling, T.S., Nielsen, H.Aa., U.S., 1995,Wind power prediction tool in control dispatch centres, ELSAM-report (ISBN 87-87090-25-2), 190 pp.

Mahrer, Y., Pielke, R.A., 1977. The effects of topography on sea and land breezes in a two-dimensional numerical model. Monthly Weather Review 105, 1151–1162.

Mavromatidis, E., Kallos, G., 2003. An investigation of cold cloud formation with a 3-D model with explicit microphysics. Journal of Geophysical Research 108, 4420–4441.

Nielsen, H. Aa., Nielsen, T.S., Madsen, H., 2004. Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts. In Proceedings of the 2004 European Wind Energy Conference, EWEC'04, Scientific Track, London, United Kingdom, pp. 34–38.

Nielsen, T.S., Madsen, H., 1997, Using Meteorological Forecasts in On-line Predictions of Wind Power, Technical Report IMM-REP-1997-31, Informatics and Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark.

Nielsen, T.S., Joensen, A., Madsen, H., Landberg, L., Giebel, G., 1999. A new reference for predicting wind power. Wind Energy 1, 29–34.

Papadopoulos, A., Katsafados, P., Kallos., G., 2001. Regional weather forecasting for marine application. Global Atmosphere and Ocean System 8 (2–3), 219–237.

Pelland, S., Galanis, G., Kallos, G., 2011. Solar and photovoltaic forecasting through post-processing of the global environmental multiscale numerical weather prediction model. **Progress in Photovoltaics**: **Research** and **Applications** , htt p://dx.doi.org/10.1002/pip.1180.

Persson, A., 1990. Kalman filtering a new approach to adaptive statistical interpretation of numerical meteorological forecasts. ECMWF Newsletter.

Rao, S.T., Zurbenko, I.G., Neagu, R., Porter, P.S., Ku, J.Y., Henry, R.F., 1997. Space and Time Scales in Ambient Ozone Data. Bulletin of the American Meteorological Society 78 (10), 2153–2166.

Silva, A.P., Moulin, L.S., 2000. Confidence intervals for neural network based short-term load forecasting. IEEE Transactions on Power Systems 15 (4).

Wilks, D., 1995. Statistical Methods in the Atmospheric Sciences. Academic Press.

Willmott, C., Robesonb Scott, M., Matsuuraa, Kenji, 2011. Short communication. A refined index of model performance. International Journal of Climatology.

Zurbenko, I.G.: The spectral Analysis of Time Series. North-Holland, 248 pp,1986.