Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering

P. Louka[a,b], G. Galanis[a,c], N. Siebert[d], G. Kariniotakis[d,*],

P. Katsafados[a], G. Kallos[a], I. Pytharoulis[a,b]

[a] *University of Athens, School of Physics, Division of Applied Physics, Atmospheric Modelling and Weather Forecasting Group, University Campus, Bldg. PHYS-V, 15784 Athens, Greece*

[b] *Hellenic National Meteorological Service, El. Venizelou 14, Hellinikon 167 77, Athens, Greece*

[c] *Naval Academy of Greece, Section of Mathematics, Xatzikyriakion, Piraeus 185 39, Greece*

[d] *Ecole Nationale Superieure des Mines de Paris, Rue Claude Daunesse. Sophia Antipolis - Les Lucioles, France*

Abstract

This paper studies the application of Kalman filtering as a post-processing method in numerical predictions of wind speed. Two limited-area atmospheric models have been employed, with different options/capabilities of horizontal resolution, to provide wind speed forecasts. The application of Kalman filter to these data leads to the elimination of any possible systematic errors, even in the lower resolution cases, contributing further to the significant reduction of the required CPU time. The potential of this method in wind power applications is also exploited. In particular, in the case of wind power prediction, the results obtained showed a remarkable improvement in the model forecasting skill.

Keywords: wind speed forecasting; Kalman filtering; wind power forecasting

*Corresponding author. Tel.: +33(0)4.93.95.75.01; fax:+33(0)4.93.95.75.35.
E-mail address: georges.kariniotakis@ensmp.fr (G. Kariniotakis)

1. Introduction

Limited Area Models (LAMs) are widely applied for providing weather forecasts up to three days with their forecast skill ranging between 80-90%. The current requirements for long-term as well as accurate predictions have forced meteorologists to broaden the ability of Numerical Weather Prediction (NWP) models supplying reliable forecasts maintaining their skill for periods longer than three days.

It is well known that NWP models usually exhibit systematic errors in the forecasts of certain meteorological parameters, such as wind speed, especially near the surface. This drawback is a result not only of the shortcoming in the physical parameterization, but also of the inability of these models to successfully handle sub-grid phenomena. The model horizontal resolution associated with smoothing/averaging the orographic and landscape characteristics leads to weak representation of local effects on the airflow. For example, winds induced by the orography of a region are usually underestimated systematically.

A way of counteracting such a drawback is to increase the model resolution that may provide considerable improvement in the representation of smaller scale flow characteristics. Nevertheless, an open question remains as to whether the use of higher resolution LAMs improves the forecast skill considerably. Even in the case that this is true, it is still uncertain whether such improvement compensates the usage of computationally costly resources that are required for these applications (Mass et al., 2002).

An alternative way to reduce the limitation of the NWP models to accurately predict sub-grid phenomena is the use of post-processing approaches based on statistical methods. One of the most successful methods in this issue is the use of Kalman filters

(Kalman, 1960; Kalman and Bucy, 1961; Persson, 1991; Dragulanescu, 1993; Galanis and Anadranistakis 2002; Crochet, 2004; Kalnay, 2002). They consist of a set of mathematical equations that provides an efficient computational solution of the least square method with minor computational cost and easy adaptation to any alteration of the observations.

The aim of this paper is, on the one hand, to investigate the rate of improvement in wind speed predictions with the application of Kalman filtering and to study the effect of such a post processing method on different horizontal resolution outputs. For this reason, an optimal polynomial Kalman filter is employed to two NWP models with different characteristics and horizontal resolution, and a detailed statistical analysis is performed. The discussion is focused on the capability of the filter to improve the direct model outputs even in cases of lower resolution taking also into account the requirements in CPU time.

On the other hand, the filtered wind speed predictions are used as input in wind power prediction models and the improvement in the final wind power forecasts is examined. Such forecasts are recognized today as useful tools for the management of power systems where wind penetration is important.

2. The modelling systems

In this Section a general description of the models used is provided.

2.1. The SKIRON modelling system

The SKIRON modelling system (e.g. Kallos, 1997; Papadopoulos et al., 2001) runs operationally at the University of Athens providing 5-day weather forecasts

(http://forecast.mg.uoa.gr). It has been developed at the University of Athens by the Atmospheric Modelling and Weather Forecasting Group (AM&WFG), based on the Eta/NCEP model (Janjic, 1994) and consists of various modules for pre- and post-processing together with a version of the Eta model appropriately coded in order to run on any parallel computer platform. SKIRON is a full physics non-hydrostatic model with sophisticated convective, turbulence and surface energy budget schemes. It is appropriate for regional/mesoscale simulations in regions with varying physiographic characteristics. The fact that it is a non-hydrostatic modelling system makes it computationally robust at all resolutions and efficient in NWP applications.

SKIRON uses NCEP/GFS initial meteorological data for operational purposes at a resolution of 1º, and SST data at a resolution of 0.5º. Vegetation and topography data are applied at a resolution of 30" and soil texture data at a resolution of 2'. The domain of the model covers the entire Mediterranean region with a horizontal increment of 0.1ºx0.1º (Figure 1).

SKIRON has been successfully applied to a large number of different regions and for long forecasting periods (e.g. Papadopoulos et al., 2001; Papadopoulos et al., 2002; Katsafados, 2003). However, local adaptation problems emerge, especially for near surface parameters, leading to systematic errors that cannot be confronted by the model itself. In particular, the wind speed prediction, being strongly dependent on local airflow characteristics, is one of the most commonly biased parameters closely related to the resolution of the model. In this paper, a way out of these problems is proposed based on the use of Kalman filter as a post processing method with low computational cost.

2.2. The RAMS model.

The Regional Atmospheric Modelling System (RAMS v4.3.0) is a highly versatile numerical code, developed at Colorado State University and Mission Research Inc/ASTeR Division. RAMS runs operationally by the Atmospheric Modelling and Weather Forecasting Group providing 48-hour forecasts over Greece and has been employed in a number of different applications (Kallos and Lagouvardos, 1997; Lagouvardos et al., 1996; Lagouvardos et al., 1997; Mavromatidis and Kallos, 2002). It is considered as an advanced modelling system being the merger of a non-hydrostatic cloud model and a hydrostatic mesoscale model. It has been developed in order to simulate atmospheric phenomena with resolution ranging from tens of kilometres to a few meters with the capability of using two-way interactive nesting of any number of grids. Furthermore, it uses various levels of complexity turbulence scheme.

RAMS is well suited for parallelization since it does not use physical/numerical routines that are global. A general description of the model and its capacities is given in Cotton *et al*. (2003) focussing on the new developments in the RAMS physics and computational algorithms since 1992.

In this paper RAMS wind speed data of various grid resolutions have been filtered, using the developed Kalman filtering technique, in order to investigate the improvement achieved to the different resolution grid data in association with the CPU time required.

2.3. The wind power prediction model

Adaptive Fuzzy-Neural Networks (F-NN) are applied here for wind power prediction (Kariniotakis and Mayer 2002, Kariniotakis and Pinson 2003). The core F-NN

model is generic and can be trained on appropriate input depending on the final use, which can be either very short-term (few seconds to a few minutes) or short-term prediction. For short-term horizons up to 120 hours ahead, it is necessary to include numerical weather forecasts as explanatory input to the model in order to have an acceptable performance.

The fuzzy model can be expressed in the form of rules of the type:

"IF $\underline{x}$ is A THEN y is B" (1)

where $\underline{x}$, y are linguistic variables and A, B are fuzzy sets. In the case of time-series prediction, rules may have the form:

$$R: \quad \text{IF} \quad x_1 \text{ is } A_1, \text{ and } \ldots, \text{ and } x_n \text{ is } A_n \quad \text{THEN} \quad y = g(x_1, \ldots, x_n) \tag{2}$$

where:

$x_1,\ldots,x_n$ are real-valued variables representing input variables of the system defined in the universes of discourse $X_1,\ldots,X_n$ respectively,

$A_1,\ldots,A_n$ are fuzzy sets represented by membership functions $\mu_{A_j}(x_j)$,

y is the variable of the consequent part of the rule whose value is inferred. In the specific problem it represents future wind power $(\hat{P}(t+1), \hat{P}(t+2), \ldots)$,

$g(.)$ is a function that implies the value of y when $x_1,\ldots,x_n$ satisfy the premise.

The function $g(.)$ in the consequent part of the rules may be linear, non-linear or constant. In the case of a linear function, the fuzzy rule-base takes the form:

$$\begin{aligned} R^1: &\quad \text{IF} \quad x_1 \text{ is } A_1^1, \ldots, \text{ and } x_n \text{ is } A_n^1 \quad \text{THEN} \quad y^1 = p_0^1 + p_1^1 x_1 + \ldots + p_n^1 x_n \\ &\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ R^m: &\quad \text{IF} \quad x_1 \text{ is } A_1^m, \ldots, \text{ and } x_n \text{ is } A_n^m \quad \text{THEN} \quad y^m = p_0^m + p_1^m x_1 + \ldots + p_n^m x_n \end{aligned}$$

where p_j^i are the weights of the linear function $g^i(.)$.

Each rule gives an estimation of the output y^i according to the conditions defined by the fuzzy sets in the premises. In the context of time-series prediction, each variable x_j in the premise corresponds to a past value of the process (i.e. power: $P(t)$, $P(t\text{-}1)$…), or past values of explanatory input (i.e. wind speed: $WS(t)$, $WS(t\text{-}1)$…) or meteorological forecasts (i.e. wind speed $WS(t+1)$, $WS(t+2)$, …).

A linear function in the consequence is indeed an ARX (autoregressive with exogenous variables) model. It is clear that with the above definitions, the rule-base consists of an ensemble of "local" models. Local modelling is a desired property of the model, especially in the case of a non-stationary process such as wind generation.

Fuzzy sets in the premises are modelled here using Gaussian functions. In the case of a linear function in the consequence, the F-NN model may be written analytically as:

$$\hat{y} = \frac{\sum_{i=1}^{m}\left(p_0^i + \sum_{j=1}^{n} p_j^i x_j\right)\cdot\prod_{j=1}^{n}\mu_{A_j^i}(x_j)}{\sum_{i=1}^{m}\prod_{j=1}^{n}\mu_{A_j^i}(x_j)} \tag{3}$$

where: $\mu_{A_j^i}$ is the membership function of x_j to the fuzzy set A_j^i.

3. The Kalman filtering methodology

Kalman filters (Kalman, 1960; Kalman and Bucy, 1961; Kalnay, 2002; Persson, 1991; Dragulanescu, 1993; Galanis and Anadranistakis 2002; Crochet, 2004) are the statistically optimal sequential estimation procedure for dynamic systems. Observations are recursively combined with recent forecasts with weights that minimize the corresponding biases.

The main advantage of this methodology is the easy adaptation to any alteration of the observations as well as the fact that it needs short series of background

information. A short description of the algorithmic procedure of a classical Kalman filter is given here.

The main goal is the simulation of an unknown process $\mathbf{x_t}$, t denoting the discrete time step. The change of **x** in time is described by the *System Equation*

$$\mathbf{x_t} = \mathbf{F_t}\cdot\mathbf{x_{t-1}} + \mathbf{w_t} \tag{4}$$

On the other hand, a known array $\mathbf{y_t}$ is also used. It is connected with the unknown process by the *Observation Equation* :

$$\mathbf{y_t} = \mathbf{H_t}\cdot\mathbf{x_t} + \mathbf{v_t} \tag{5}$$

The coefficient matrices $\mathbf{F_t}$ and $\mathbf{H_t}$ have to be determined before the application of the filter. The same holds also for the covariance matrices $\mathbf{W_t}$, $\mathbf{V_t}$ of the Gaussian and independent random vectors $\mathbf{w_t}$ and $\mathbf{v_t}$.

The Kalman filter gives a method for the recursive estimation of the unknown state $\mathbf{x_t}$ based on observation values **y** up to time *t*. A first estimate of $\mathbf{x_t}$ and $\mathbf{P_t}$ based on the previous time step values is given by:

$$\mathbf{x_{t/t-1}} = \mathbf{F_t}\cdot\mathbf{x_{t-1}}\,, \qquad \mathbf{P_{t/t-1}} = \mathbf{F_t}\cdot\mathbf{P_{t-1}}\cdot\mathbf{F_t}^{\mathbf{T}} + \mathbf{W_t}\,. \tag{6}$$

As soon as the new observation value $\mathbf{y_t}$ is known, the estimate of **x** at time *t* becomes:

$$\mathbf{x_t} = \mathbf{x_{t/t-1}} + \mathbf{K_t}\cdot(\mathbf{y_t}-\mathbf{H_t}\cdot\mathbf{x_{t/t-1}})\,, \tag{7}$$

where

$$\mathbf{K_t} = \mathbf{P_{t/t-1}}\cdot\mathbf{H_t}^{\mathbf{T}}\cdot(\mathbf{H_t}\cdot\mathbf{P_{t/t-1}}\cdot\mathbf{H_t}^{\mathbf{T}}+\mathbf{V_t})^{-1} \tag{8}$$

is the Kalman gain. $\mathbf{K_t}$ arranges how easily the filter adjusts to possible new conditions.

The final estimate of $\mathbf{P_t}$ is

$$\mathbf{P_t} = (\mathbf{I}-\mathbf{K_t}\cdot\mathbf{H_t})\cdot\mathbf{P_{t/t-1}}. \tag{9}$$

Equations (6)-(9) update the Kalman algorithm from time *t*-1 to *t*.

Our approach is based on the application of non linear Kalman filters. In particular, it focuses on the study of a single meteorological parameter in time, based on the estimation of the bias of this parameter as a function of the forecasting model direct output. Specifically, if we denote by m_t the direct output of the NWP model at time t and by y_t the bias of this forecast, we realize y_t by means of m_t as a *3*-rd order polynomial:

$$y_t = x_{0,t} + x_{1,t} \cdot m_t + x_{2,t} \cdot m_t^2 + x_{3,t} \cdot m_t^3 + v_t \tag{10}$$

where the coefficients ($x_{i,t}$) are the parameters that have to be estimated by the filter and $\mathbf{v_t}$ the Gaussian non systematic error in the previous procedure.

In this way, the state vector is the one formed by the coefficients ($x_{i,t}$): $\mathbf{x_t} = \begin{bmatrix} x_{0,t} & x_{1,t} & x_{2,t} & x_{3,t} \end{bmatrix}^T$ and the observation procedure is the scalar bias y_t. On the other hand, the observation matrix takes the form $\mathbf{H_t} = \begin{bmatrix} 1 & m_t & m_t^2 & m_t^3 \end{bmatrix}$ while as system matrix we use the identity. As a result, the system and observation equations, (4) and (5) correspondingly become, respectively:

$$\mathbf{x_t} = \mathbf{x_{t-1}} + \mathbf{w_t}, \quad \mathbf{y_t} = \mathbf{H_t} \cdot \mathbf{x_t} + \mathbf{v_t} \tag{11}$$

It is worth noting that instead of the 3-rd order polynomial of relation (10), one may employ a function of arbitrary order. However, a plain – 1-st order form – leads, in most of the cases, to difficulties in the simulation of non linear procedures, as the wind speed evolution, while higher order polynomials result to increased "noise" due to instabilities with no essential contribution in filters accuracy, and hence model output improvement. A detailed study for the determination of the optimum order for a filter polynomial has been investigated in Galanis et al. (2006) where the 3-rd order proved to have the best contribution in the elimination of the systematic part of the bias with less need in CPU time.

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 7 values of $\mathbf{w_t=x_t-x_{t-1}}$ and $\mathbf{v_t=y_t-x_t}$, respectively:

$$\boldsymbol{W_t} \equiv \frac{1}{6} \cdot \sum_{i=0}^{6} (((\boldsymbol{x_{t-i}} - \boldsymbol{x_{t-i-1}}) - (\frac{\sum_{i=0}^{6} (\boldsymbol{x_{t-i}} - \boldsymbol{x_{t-i-1}})}{7}))^2, \tag{12}$$

$$\boldsymbol{V_t} \equiv \frac{1}{6} \cdot \sum_{i=0}^{6} (((\boldsymbol{y_{t-i}} - \boldsymbol{H_{t-i}} \cdot \boldsymbol{x_{t-i}}) - (\frac{\sum_{i=0}^{6} (\boldsymbol{y_{t-i}} - \boldsymbol{H_{t-i}} \cdot \boldsymbol{x_{t-i}})}{7}))^2. \tag{13}$$

The later 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 equations (11), follow the normal distribution by assumption. This time period of 7 days was proven to be the optimal choice in our study in order to achieve successful corrections and fast adaptability simultaneously (Galanis et al., 2006). However it is possible to vary under different geographic or climatological environments.

4. The case study

SKIRON NWP data of a time period statistically long for wind power purposes, i.e. one year, have been provided in the framework of E.U. ANEMOS project (http://anemos.cma.fr). In particular, SKIRON forecasts for wind speed and direction, air temperature and mean sea level pressure have been supplied for different locations in the Mediterranean Region where wind farms are operated. For the specific case study, SKIRON wind speed at 10 m above the ground for the area of Rokas in Crete, Greece (Lon: 26.2º, Lat: 35.2º, 480m above ground) for the whole year 2003 were available. At

the same time observations of wind speed at 40 m and power from the wind farm installed at the site were provided within the framework of the project.

On the other hand, RAMS run in a hindcasting mode for a selected period of 2 days, namely 4 to 6 September 2003. Although, this period is short due to CPU time limitations, it is, nevertheless, indicative of the potential benefits that may be gained by the application of Kalman filtering in various resolution domains/cases. Five two-way nested grids were used with the outer one to cover the whole Mediterranean Region with 48km horizontal resolution; the next grid of 12 km resolution included the whole Greece; the third grid of 6 km resolution contained Crete; the fourth grid of 1.5 km resolution included the east part of Crete, while the smallest grid included the edge of the east part of Crete with 0.5 km resolution where the Rokas wind farm is located (c.f. Figure 2). The model was initiated at 12UTC using ECMWF gridded analysis meteorological fields. The modelled horizontal wind was extracted at the location of Rokas for the nested grids of 12 km, 6 km, 1.5 and 0.5 km. In this way the effect of the grid resolution on the predicted wind was investigated.

5. Results

The results presented here are based on detailed statistical analysis of SKIRON and RAMS outputs, the Kalman filtered outcomes and the wind power prediction product presented as an application.

The statistical analysis was based on the calculation of the:

- **Bias** of direct model output in comparison with the corresponding bias of the improved, by the Kalman filter, forecasts:

$$Bias = \frac{1}{k} \cdot \sum_{i=1}^{k} \left(for(i) - obs(i) \right)$$

where *obs(i)* denotes the observed value at time *i*, *for(i)* the corresponding forecasted value (direct model output or filtered forecast), and *k* is the size of our sample.

Bias is a crucial parameter for Kalman filtering since any type of Kalman method aims at considerably reducing the corresponding biases. The success on this issue is the main criterion ensuring the credibility of the filter.

- **Absolute bias**, assessing the "real" improvement in the model predictions despite possible changes in the type of errors:

$$Absolute\ Bias = \frac{1}{k} \cdot \sum_{i=1}^{k} \left| for(i) - obs(i) \right|$$

where | | denotes the absolute value.

- **Standard deviation of the bias and of the absolute bias** that estimates the variability of the results:

$$St.\ Dev.\ of\ \ Bias = \sqrt{\frac{1}{k} \cdot \sum_{i=1}^{k} \left(\left(for(i) - obs(i) \right) - Bias \right)^2} .$$

- **Root Mean Square Error** assessing the error in the predicted values:

$$RMSE = \sqrt{\frac{1}{k} \sum_{i=1}^{k} \left(for(i) - obs(i) \right)^2} .$$

5.1. Kalman filtering application to SKIRON wind speed

The Kalman filter with polynomial of 3-rd order was applied to all wind speed SKIRON data of Rokas wind farm for different forecasting periods, i.e. 24, 48, 72, 96 and 120 hours, for the whole year 2003.

Table 1 includes the bias, absolute bias and standard deviation of bias of the model direct output and the Kalman filtered outcome for each forecasting period. It is clear that SKIRON underestimates the wind speed in all forecasting periods. In all cases the application of Kalman filtering improves the model output reducing the bias considerably to approach zero suggesting that the main goal of the filter is fulfilled. Moreover, the significant reduction of all three statistical quantities for the Kalman filtered results ensures that the discrepancy between the two time series (observed and forecasted) has been decreased despite any possible changes in the type of error (underestimation or overestimation). Hence, the filtered model output approximates the observations considerably.

Figures 3 to 5 illustrate the variation of bias, absolute bias and standard deviation of bias of the direct model output and the Kalman filtered wind speed with the forecasting period. Apart from the obvious improvement in the initial forecast, it is worth noticing that this positive influence remains invariant with the forecast time for the bias.

The time series illustrated in Figure 6 further support the significant improvement in the model output after the application of Kalman filtering to wind speed predictions. It is clear that the systematic error emerged is eliminated regardless of its type.

In the framework of ANEMOS project, the developed Kalman filtering technique was successfully applied also to SKIRON wind speed numerical predictions for other wind farms in Spain and Corsica with similar improvements in the direct model outputs.

5.2. Kalman filtering application to RAMS wind speed

RAMS wind speed data of various grid resolutions have been filtered using the developed Kalman technique in order to investigate the improvement achieved to the

different resolution grid data in association with the CPU time required. Figure 7 shows time series of the predicted horizontal wind and the corresponding Kalman filtered for four of the domains, i.e., those with horizontal resolution of 12 km, 6 km, 1.5km and 0.5 km. It is shown that RAMS follows the evolution in time of the measurements fairly well with the wind resulting from the 0.5 km grid resolution being closest to the observations than the results of all other grids. An underestimation of the predictions is observed for winds stronger than approximately 20 to 25 m/s dependent of the grid resolution, with the highest resolution grid being able to capture some of the local peaks observed. The main underestimation is obtained during 00UTC and 15UTC on 5/9/2003, when the highest winds were measured. Figure 7 suggests that regional and mesoscale features of the airflow are captured satisfactorily by the coarser grids (12km and 6km), while the small scale features (e.g. topography induced) are better described by the finest grids (1.5km and 0.5km). However, the benefits gained beyond 6km resolution are not always worth the computational expenses.

The use of the developed Kalman filtering technique investigated the improvement that may be achieved when applied to the modelled wind extracted by each domain. It is shown (Figure 7) that the proposed technique improves the predicted wind speed in all cases. This result is further supported by the values of Bias and RMSE (c.f. Table 2) that were considerably diminished after Kalman filtering implementation for all domains, while the mean wind speed was increased approaching the observed value.

The importance of this application lies on the fact that the filtered wind extracted from the 12 km resolution grid obtains better statistics than the direct modelled wind of 0.5 km grid. Moreover, the proposed method offers improved wind predictions in significantly reduced CPU time. Specifically, while the requirements in CPU time for a

48-hour run with RAMS using 5 two-way nested grids was approximately 48 hours, the required corresponding time for the same run with just 2 two-way nested grids (48 km and 12 km) was 1.7 hours plus a few minutes for the Kalman filter application.

5.3. Results for wind-power prediction

This section provides an example of the use of Kalman filtered NWP data to wind power applications. The F-NN model was used to predict the power production of Rokas wind farm using as input the SKIRON and alternatively Kalman filtered SKIRON NWPs for different forecasting periods, i.e. 24, 48, 72, 96 and 120 hours for the year 2003. The F-NN model needs a training set to "learn" how to predict. In the following study, the first 3850 observations of the year 2003 were used for training; the remaining observations were used to evaluate the performance of the model. Statistical analysis of the errors was performed based on the calculation of the:

- **Normalized Mean Absolute Error** (NMAE). This criterion gives equal weight to all errors and is easy to interpret in practice since it is directly related to the quantity of power not predicted.

$$NMAE(h) = \frac{1}{P_{inst} N} \sum_{t=1}^{N} \left| P(t+h) - \hat{P}(t+h \mid t) \right|$$

where $P(t+h)$ is the power measured at time $t+h$, $\hat{P}(t+h \mid t)$ is the forecast of $P(t+h)$ computed at time t, N the number of computed forecasts, and $|\ |$ denotes the absolute value.

- **Normalized Root Mean Square Error** (NRMSE), gives more weight to larger errors thus providing information on the relative size of the large errors.

$$NRMSE(h) = \frac{1}{P_{inst}} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(P(t+h) - \hat{P}(t+h \mid t) \right)^2}$$

where P_{inst} denotes the installed capacity for which the forecasts are computed.

- **Bias** of fuzzy model output using raw and filtered SKIRON wind speed forecasts.

In Figure 8 the NMAE and NRMS are presented for all horizons and both sets of data. For the 24 and 48 hour-ahead horizons the performance obtained with the SKIRON data is comparable to that obtained with other NWP models on other wind farms (Kariniotakis, G., et al. 2004). The performance decrease and the error values for the longer-term forecasts remain acceptable. When using the Kalman data, the performance increase is evident; the improvement with respect to the SKIRON data is in the order of 22% of the NMAE for the 48 to 120 hours-ahead horizons.

The bias correction witnessed with the NWP data directly translates to a correction of the bias of the wind power forecasts. In Figure 9 the biases of the wind power forecasts using the raw and filtered SKIRON NWP are shown. The wind power forecast biases are similar to that of the NWP biases.

As the Kalman filtering contributes to the reduction of the bias in the wind speed prediction input, investigating the distribution of the power prediction errors is necessary. In figures 10 and 11 the error distributions for the 48 and 96 hours-ahead horizons are presented. The histograms clearly show the reduction of the bias. They are more symmetrical for the corrected NWP than for the raw NWP. Further, the filtered data lead to a sharper distribution of the errors. As a result, the forecasts obtained with the filtered data present a lower uncertainty. This can be illustrated by analysing the number of forecast errors inferior to a certain error margin. In Figure 12 the number of forecasts between +/–5 %, +-15% and +/-30% error margins are given for the power forecasts

obtained from the filtered and raw NWP for each forecasting period. For example, for 24 hours ahead, 32% of the errors are between +/-5% of the nominal power of the farm for the raw SKIRON data, whereas for the filtered data, 38% of the errors are in the same error margin for the same horizon. From this figure it is clear that the filtering reduces the error for all error bins. However, the most important benefit seems to be for the smaller errors (+/- 5% and +/- 15%). The reduction of larger error (+/-30% and above) is quite limited.

The overall improvement in wind power forecast accuracy is clearly illustrated in Figure 13, where the measured power is plotted along with the forecasts computed using the raw and filtered SKIRON forecasts.

6. Conclusions

A new technique based on the implementation of non-linear polynomial functions (3-rd order) in Kalman filter algorithms was applied to wind speed numerical predictions obtained at a particular wind farm in Crete, Greece. This method was applied to the outputs of two atmospheric numerical models with different capabilities in horizontal resolution. The methodology showed high performance for all cases leading to the elimination of any type of systematic errors. In particular, all error parameters used were significantly reduced, the forecast skill was maintained for all forecasting periods and most importantly it reduced the requirements in CPU time since its application to lower resolution data led to similar or even more accurate results compared to the costly high resolution direct model outputs.

Beyond the traditional meteorological use this technique finds a wide range of applications in the engineering sector. In this case, it was used for providing improved

input to a wind power prediction model. It was shown that the Kalman filtered wind data ameliorated the ability of power forecasting models to provide useful predictions for very long horizons. All the performance criteria were improved significantly for all predictions horizons.

Therefore, this work suggests that the use of very expensive computational facilities to perform high-resolution (< 6 km) applications for wind energy predictions may be avoided by the combined use of moderate NWP model resolution and an adaptive statistical technique such as Kalman filtering; providing similar or even more accurate predictions at wind farm scale.

In an operational setting, this can greatly benefit unit commitment, economic dispatch and electricity market biding strategies as well as allowing better maintenance scheduling. Such improvements can greatly ease wind power integration into conventional power systems thus favouring an increase in the use of wind as a renewable energy source.

The credibility of the proposed method, gives the opportunity for further use and applications. In this way, it can be implemented in the main forecast model activated in each time step of integration, smoothing, in this way, any possible temporary discontinuity that a rapid change in the time series could produce. It can be also used as a pre-processing assimilation method for the correction of the initial conditions in a numerical prediction model.

7. Acknowledgements

The present work was carried out in the framework of the European project ANEMOS (Development of a next generation wind resource forecasting system for the

large-scale integration of onshore and offshore wind farms – Contract No ENK5-CT-2002-00665).

8. References

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

Crochet, P., 2004: Adaptive Kalman filtering of 2-metre temperature and 10-metre wind-speed forecasts in Iceland. *Meteor. Appl*., **11**, 173-187.

Dragulanescu, L., 1993: Application des filtres Kalman pour ajuster les temperatures prognosees avec un modele numerigue, *Meteor. Hydrol.*, **23**, 11-14.

Galanis, G., and M. Anadranistakis, 2002: A one dimensional Kalman filter for the correction of near surface temperature forecasts, *Meteor. Appl*., **9**, 437-441.

Galanis, G., P. Louka, P. Katsafados, G. Kallos and I. Pytharoulis, 2006: Applications of non-linear Kalman filters to numerical weather predictions (submitted).

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

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

Kallos, G., and K. Lagouvardos, 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, *Trans. ASME*, Ser. D, **82**, 35-45.

Kalman, R.E., and R.S. Bucy, 1961: New results in linear filtering and prediction problems, *Trans. ASME*, Ser. D, **83**, 95-108.

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

Katsafados, P., 2003: Factors and parameterizations that determine the performance of limited area models in long-range forecasts. PhD thesis, School of Physics, Univ. of Athens, Greece, pp. 257 (in Greek).

Kariniotakis, G., and D. Mayer, 2002: An advanced on-line wind resource prediction system for the optimal management of wind parks. *Proceedings of the 2002 MedPower Conference*, 4-6 November 2002, Athens, Greece.

Kariniotakis, G. and P. Pinson 2003: Evaluation of the MORE-CARE wind power prediction platform. Performance of the fuzzy logic based models. *CD-Proceedings of the EuropeanWind Energy Conference 2003*, June 20003, Madrid, Spain.

Kariniotakis, G., I. Marti, et al. 2004: What Performance Can Be Expected by Short-term Wind Power Prediction Models Depending on Site Characteristics? *CD-Proceedings of the European Wind Energy Conference EWEC 2004*, 22-25 November 2004, London, UK.

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

Lagouvardos K., V. Kotroni and G. Kallos, 1997: Application of CSU-RAMS for the study of severe weather phenomena. *Proceedings of Symposium on Regional Weather Prediction on Parallel Computer Environments*, 15-17 October, Athens, Greece, pp 6.

Mass, C.F., D. Ovens, K. Westrick, and B.A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc*., **83**, No 3, 407-430.

Mavromatidis, E. and G. Kallos, 2002: A study of Cold and Warm Cloud Formation with RAMS. *Atmospheric Modelling From Microscale to Global, 5th RAMS Workshop and Related Applications*, 30 September- 3 October 2002, Santorini, Greece.

Papadopoulos, A., P. Katsafados, and G. Kallos, 2001: Regional weather forecasting for marine application. Global Atmos. Ocean Syst., 8, No 2-3, 219-237.

Papadopoulos, A., L. Perivoliotis, K. Nittis and P. Katsafados, 2002: Evaluation of "POSEIDON" forecasts in the Aegean Sea for a three-year period. *Eurogoos Annual Meeting*, 7-10 December 2002, Athens Greece.

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

Pielke, Sr. R.A., 2001: Further Comments on "The Differentiation between Grid Spacing and Resolution and Their Application to Numerical Modeling". *Bull. Amer. Meteor. Soc.*, **82**, No. 4, 699–700.

Figure Captions

Figure 1: The orography of the domain that the non-hydrostatic SKIRON modelling system currently uses

Figure 2: The five two-way nested grids used for the RAMS application for the case of Crete wind farms

Figure 3: Variation of wind speed bias with forecast time for the direct model output and after the application of Kalman filter, for the year 2003

Figure 4: Variation of wind speed Absolute Bias with forecast time for the direct model output and after the application of Kalman filter, for the year 2003

Figure 5: Variation of wind speed Standard Deviation of Bias with forecast time for the direct model output and after the application of Kalman filter, for the year 2003

Figure 6: Time series of the wind speed observed, forecasted (24-hour forecasts) and Kalman filtered

Figure 7: Time series of the wind speed observed, modelled with RAMS and Kalman filtered for the grids with 12km, 6km, 1.5km and 0.5km horizontal resolution

Figure 8: Comparison of NMAE and RMS values of the wind power forecasts using SKIRON and filtered SKIRON data.

Figure 9: Variation of wind power bias with forecast time for the wind power forecasts using SKIRON and filtered SKIRON data.

Figure 10: Error distribution of wind power forecasts for the 48 hour-ahead horizon using SKIRON and filtered SKIRON data.

Figure 11: Error distribution of wind power forecasts for the 96 hour-ahead horizon using SKIRON and filtered SKIRON data.

Figure 12: Number of power forecast errors between +/-5 %, +/- 15 %, +/- 30 % of nominal power obtained using raw SKIRON and filtered SKIRON data.

Figure 13: Time series of the observed wind power and wind power forecasts using raw SKIRON and filtered SKIRON NWP data (in percentage of the wind farm nominal capacity).

Table 1: Bias, Average Bias and Standard deviation of Bias for the direct model output and the Kalman filtered

	Wind Speed					
Forecast period	**Model Bias**	**Kalman Bias**	**Model Abs. Bias**	**Kalman Abs. Bias**	**Model Std. Dev. Bias**	**Kalman Std. Dev. Bias**
T+24 h	-2.11	-0.13	2.79	1.75	2.78	2.38
T+48 h	-2.23	-0.08	3.12	1.86	3.17	2.58
T+72 h	-2.34	-0.14	3.45	2.00	3.58	2.81
T+96 h	-2.72	-0.03	3.99	2.23	4.12	3.20
T+120 h	-2.99	-0.07	4.32	2.35	4.51	3.43
Average	**-2.48**	**-0.09**	**3.53**	**2.04**	**3.63**	**2.88**

Table 2: Bias, RMSE and the corresponding mean value for the direct RAMS output and the Kalman filtered for the grids with 12km, 6km, 1.5km and 0.5km horizontal resolution

	Model				**Kalman**			
Resolution	***12 km***	***6 km***	***1.5 km***	***0.5 km***	***12 km***	***6 km***	***1.5 km***	***0.5 km***
Bias	-2.67	-0.65	-1.89	-1.67	-0.02	0.40	0.35	0.03
RMSE	4.07	3.54	3.50	3.36	2.37	2.65	2.22	2.25
Mean	17.43	19.45	18.21	18.43	20.08	20.50	20.45	20.13

Figure 1

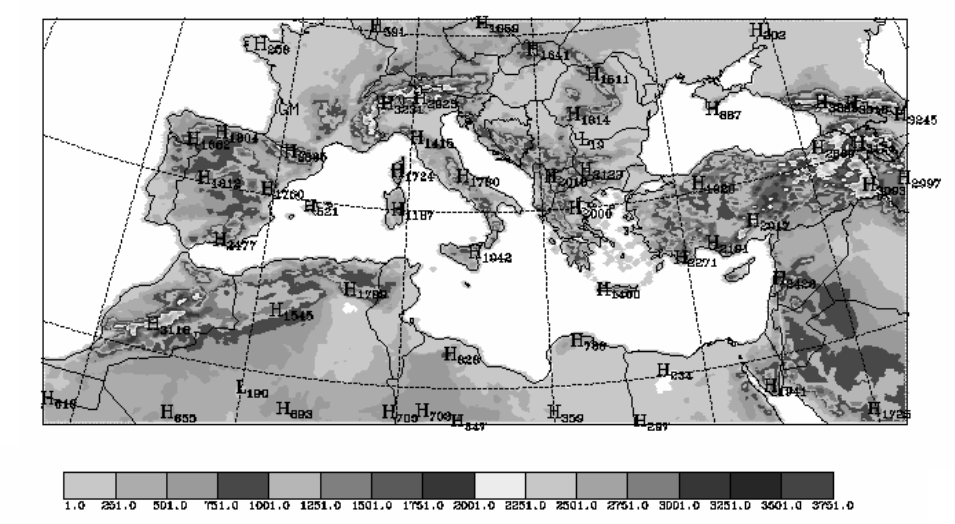

1.0
251.0
501.0
751.0
1001.0
1251.0
1501.0
1751.0
2001.0
2251.0
2501.0
2751.0
3001.0
3251.0
3501.0
3751.0

Figure 2

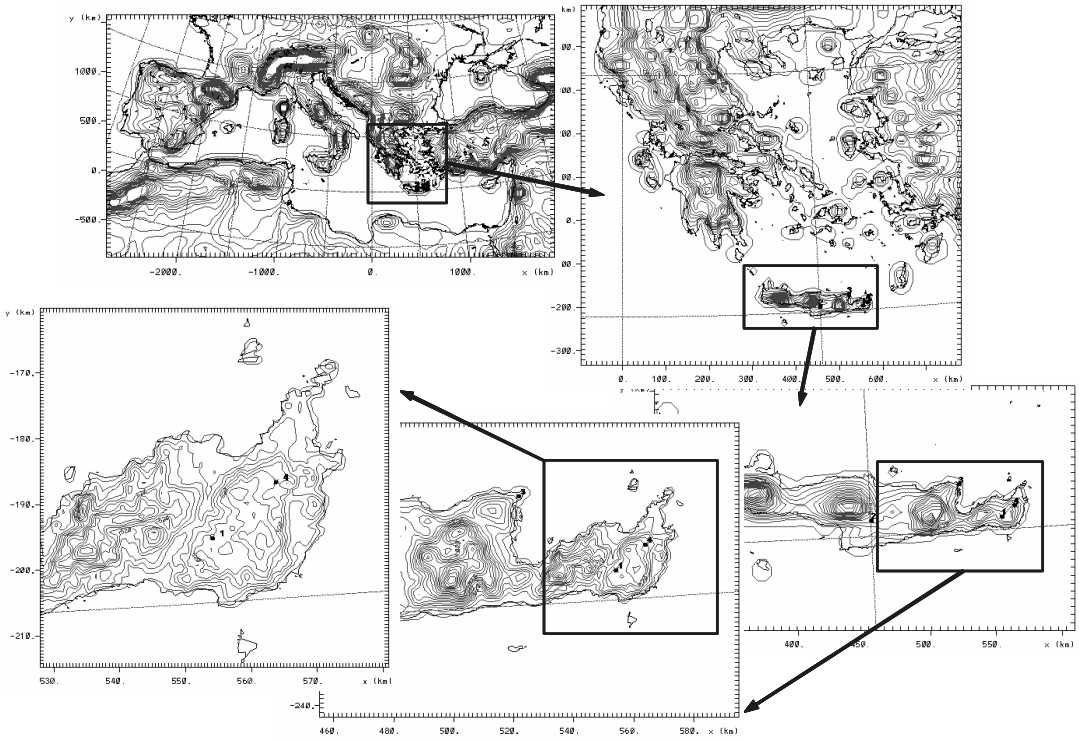

y (km)
1000.
500.
0.
-500.
-2000.
-1000.
0.
1000.
x (km)
-200
-300.
0.
100.
200.
300.
400.
500.
600.
x (km)
-170.
-180.
-190.
-200.
-210.
530.
540.
550.
560.
570.
x (km)
-240.
460.
480.
500.
520.
540.
560.
580.
x (km)
400.
450.
500.
550.
x (km)

Figure 3

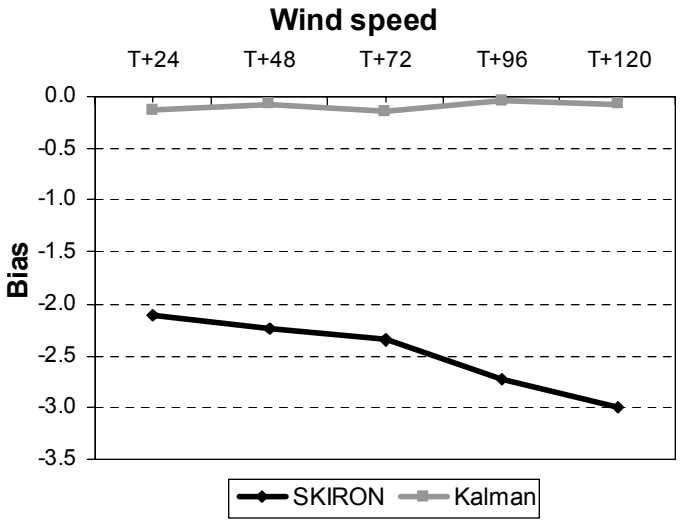

Wind speed
T+24
T+48
T+72
T+96
T+120
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
Bias
SKIRON
Kalman

Figure 4

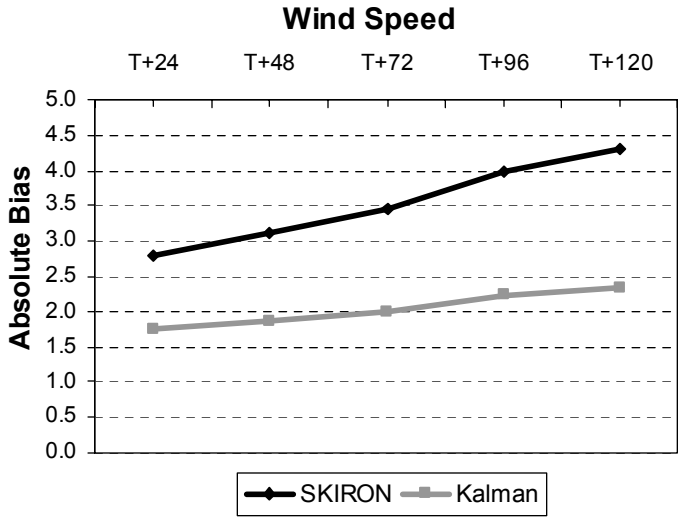

Wind Speed
T+24
T+48
T+72
T+96
T+120
Absolute Bias
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
SKIRON
Kalman

Figure 5

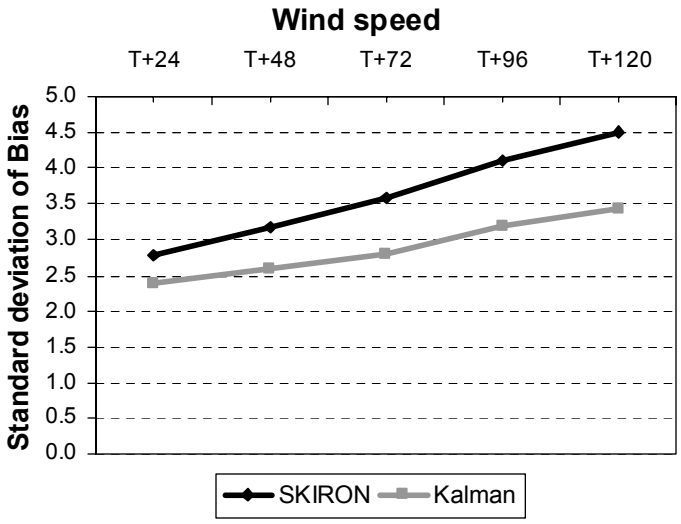

Wind speed
T+24
T+48
T+72
T+96
T+120
Standard deviation of Bias
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
SKIRON
Kalman

Figure 6

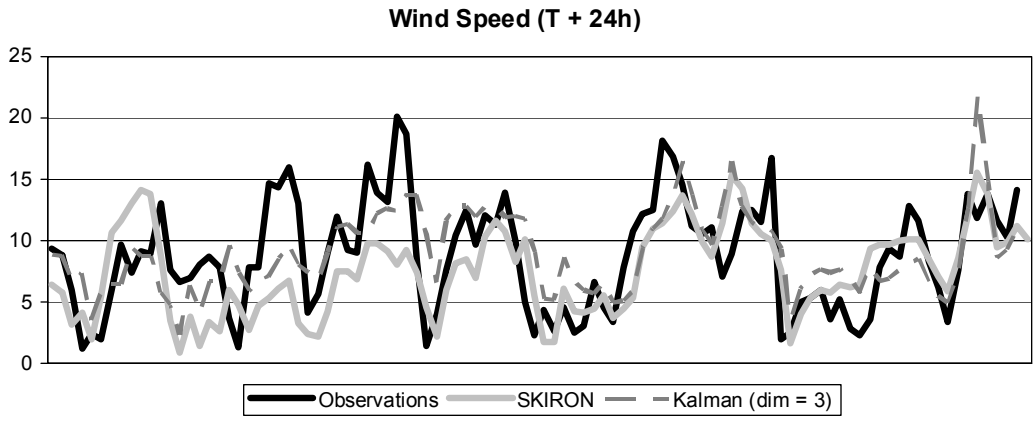

Wind Speed (T + 24h)
25
20
15
10
5
0
Observations
SKIRON
Kalman (dim = 3)

Figure 7

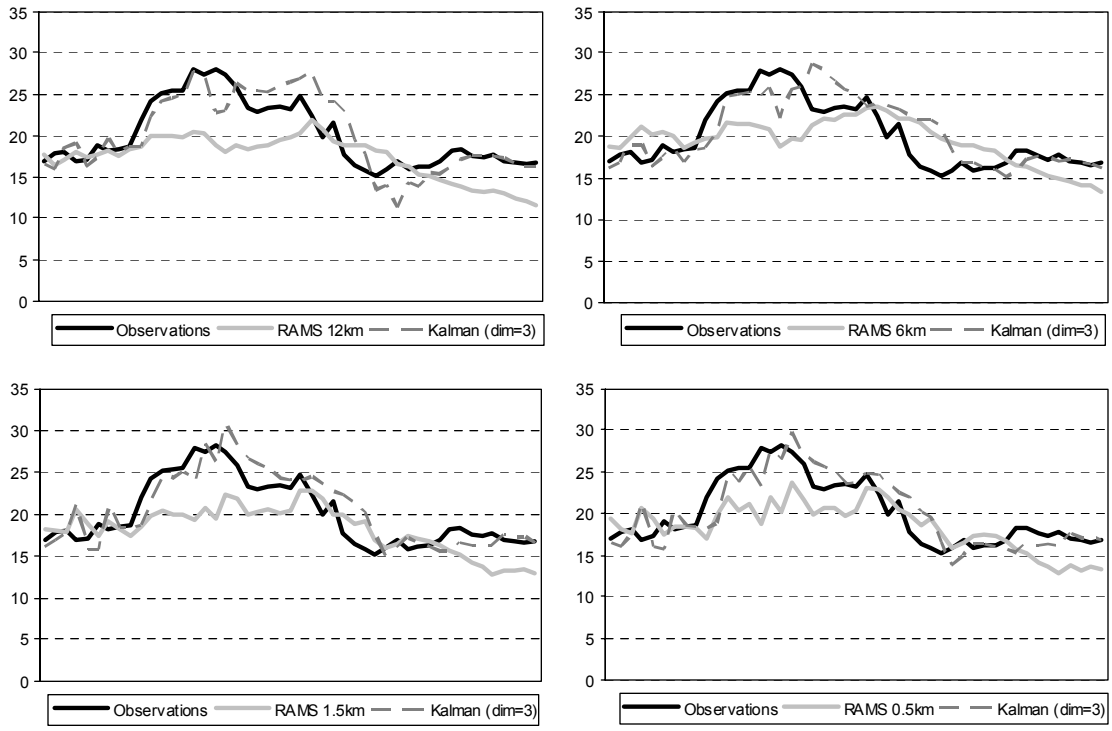

35
30
25
20
15
10
5
0
Observations
RAMS 12km
Kalman (dim=3)
Observations
RAMS 6km
Kalman (dim=3)
Observations
RAMS 1.5km
Kalman (dim=3)
Observations
RAMS 0.5km
Kalman (dim=3)

Figure 8

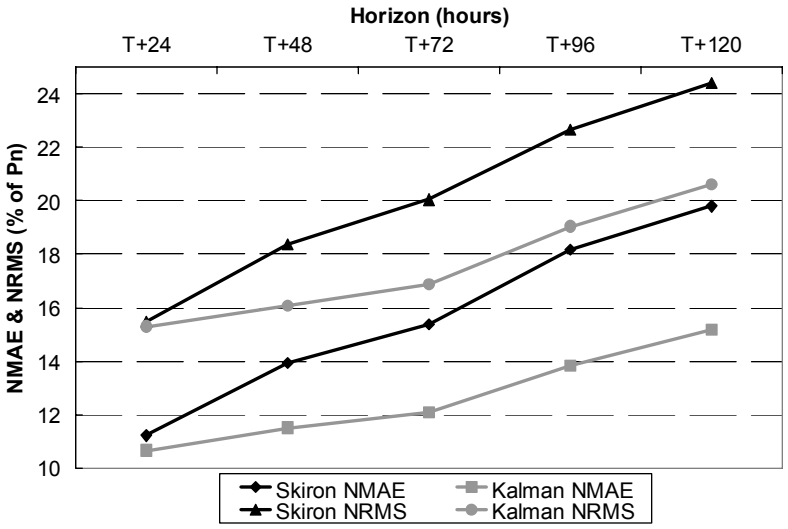

Horizon (hours)
T+24
T+48
T+72
T+96
T+120
NMAE & NRMS (% of Pn)
24
22
20
18
16
14
12
10
Skiron NMAE
Skiron NRMS
Kalman NMAE
Kalman NRMS

Figure 9

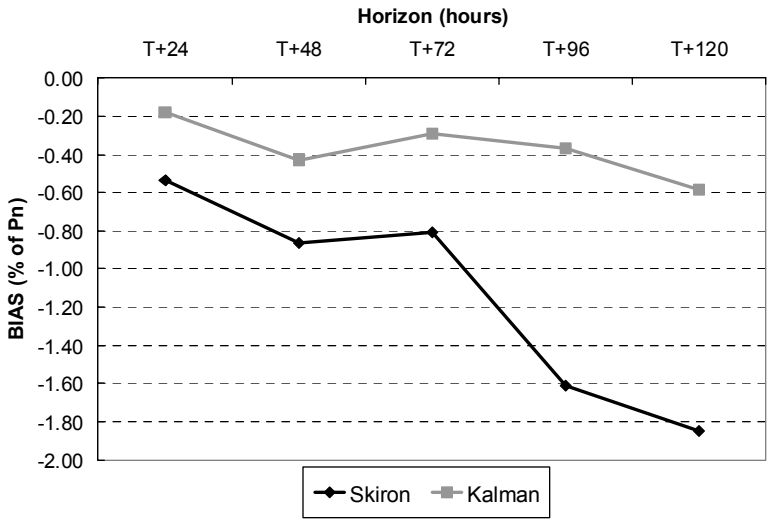
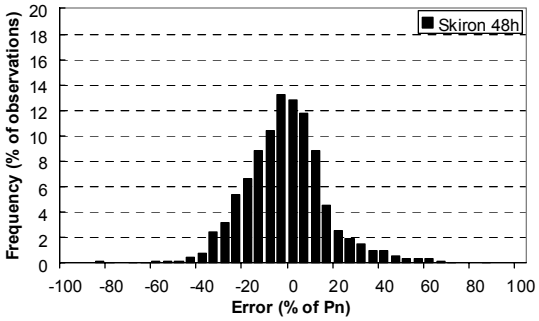

Horizon (hours)
T+24
T+48
T+72
T+96
T+120
0.00
-0.20
-0.40
-0.60
-0.80
-1.00
-1.20
-1.40
-1.60
-1.80
-2.00
BIAS (% of Pn)
Skiron
Kalman

Figure 10


Skiron 48h
Frequency (% of observations)
Error (% of Pn)
0
2
4
6
8
10
12
14
16
18
20
-100
-80
-60
-40
-20
0
20
40
60
80
100

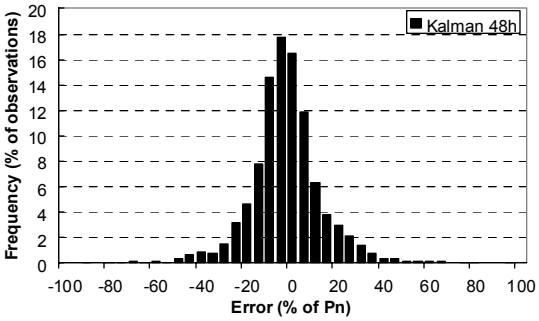
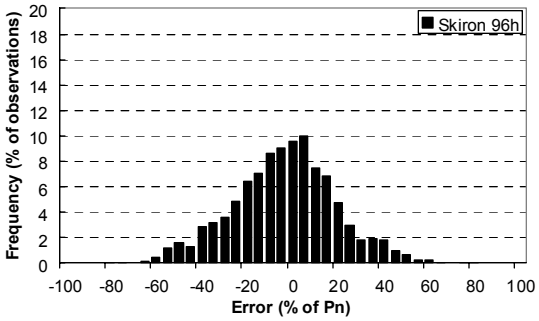

Kalman 48h
Frequency (% of observations)
Error (% of Pn)
0
2
4
6
8
10
12
14
16
18
20
-100
-80
-60
-40
-20
0
20
40
60
80
100

Figure 11


Skiron 96h
Frequency (% of observations)
Error (% of Pn)
0
2
4
6
8
10
12
14
16
18
20
-100
-80
-60
-40
-20
0
20
40
60
80
100

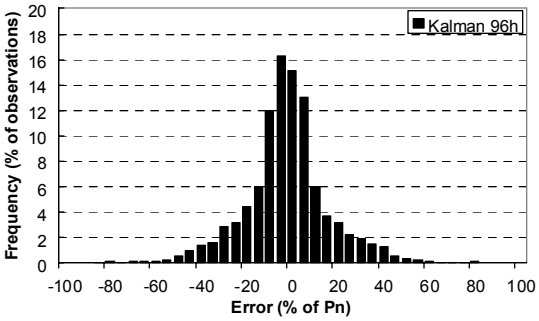

Kalman 96h
Frequency (% of observations)
Error (% of Pn)
0
2
4
6
8
10
12
14
16
18
20
-100
-80
-60
-40
-20
0
20
40
60
80
100

Figure 12

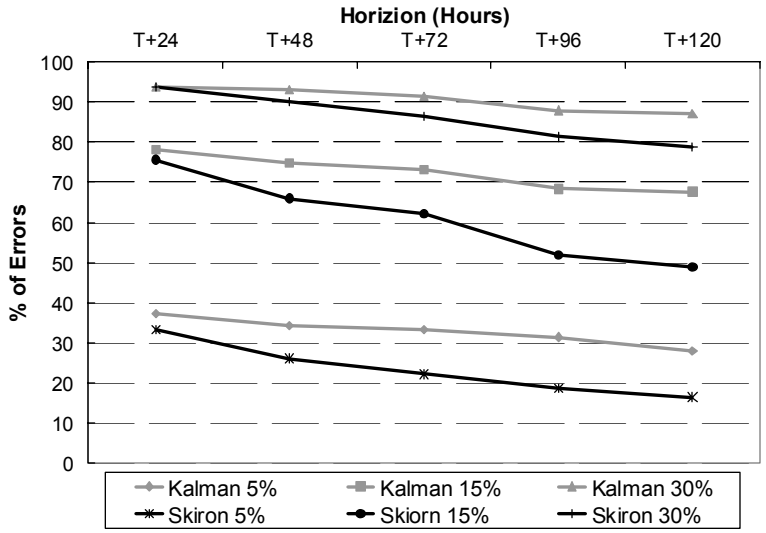
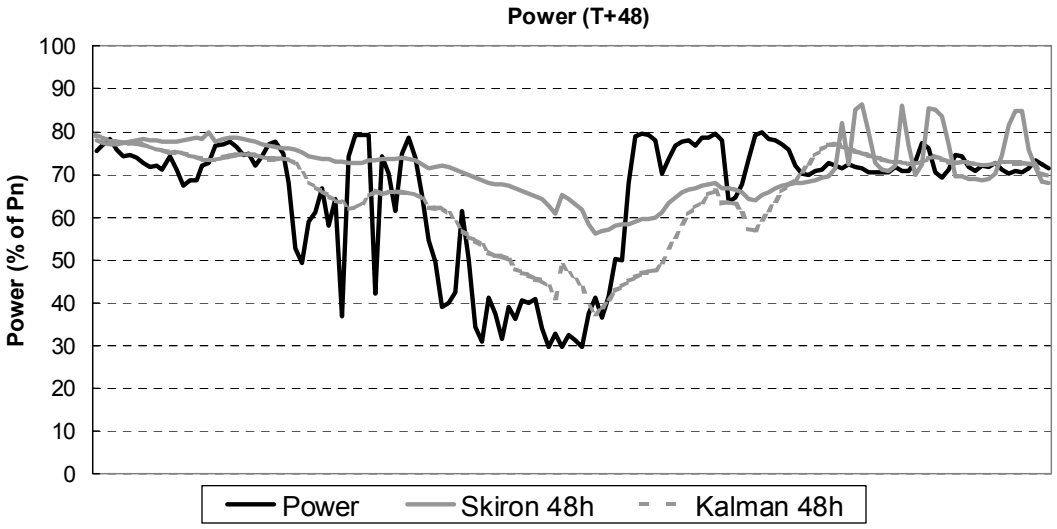

Horizion (Hours)
T+24
T+48
T+72
T+96
T+120
100
90
80
70
60
50
40
30
20
10
0
% of Errors
Kalman 5%
Kalman 15%
Kalman 30%
Skiron 5%
Skiorn 15%
Skiron 30%

Figure 13


Power (T+48)
Power (% of Pn)
100
90
80
70
60
50
40
30
20
10
0
Power
Skiron 48h
Kalman 48h