**Annales
Geophysicae**

Applications of Kalman filters based on non-linear functions to numerical weather predictions

G. Galanis[1,2]**, P. Louka**[1,3]**, P. Katsafados**[1]**, I. Pytharoulis**[1,4]**, and G. Kallos**[1]

[1]University of Athens, School of Physics, Division of Applied Physics, Atmospheric Modelling and Weather Forecasting Group, University Campus, Bldg. PHYS-V, 15784 Athens, Greece
[2]Naval Academy of Greece, Section of Mathematics, Xatzikyriakion, Piraeus 185 39, Greece
[3]Hellenic National Meteorological Service, El. Venizelou 14, Hellinikon 167 77, Athens, Greece
[4]Section of Meteorology and Climatology, Department of Geology, Aristotle University of Thessaloniki, University Campus, 54124 Thessaloniki, Greece.

Abstract. This paper investigates the use of *non-linear* functions in classical Kalman filter algorithms on the improvement of regional weather forecasts. The main aim is the implementation of non linear polynomial mappings in a usual linear Kalman filter in order to simulate better non linear problems in numerical weather prediction. In addition, the optimal order of the polynomials applied for such a filter is identified. This work is based on observations and corresponding numerical weather predictions of two meteorological parameters characterized by essential differences in their evolution in time, namely, air temperature and wind speed. It is shown that in both cases, a polynomial of low order is adequate for eliminating any systematic error, while higher order functions lead to instabilities in the filtered results having, at the same time, trivial contribution to the sensitivity of the filter. It is further demonstrated that the filter is independent of the time period and the geographic location of application.

Keywords. Meteorology and atmospheric dynamics (Mesoscale meterology; Instruments and techniques)

1 Introduction

It is well known that Numerical Weather Prediction (NWP) models usually exhibit systematic errors in the forecasts of certain meteorological parameters 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 handle successfully sub-grid scale phenomena. The model horizontal resolution associated with smoothing/averaging the orographic and landscape characteristics leads to weak representation of local effects on the airflow (e.g. systematic underestimations in the wind speed). Increasing the model resolution 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 limited area models improves the forecast skill considerably, while, even if this is true, it is still uncertain whether such improvement compensates for the usage of increasing computational resources required for these applications (Mass et al., 2002). In addition, inaccuracies in the NWP model forecasts may also be due to possible errors in the initial and lateral boundary conditions as well as to interpolations to areas that are not close to model levels.

In order to reduce the influence of the above mentioned drawbacks in the final output of a NWP model, a variety of approaches based on statistical methods has been used. Most of them are derived from Model Output Statistics (MOS), which are able to account for local effects and seasonal changes. However, discrepancies have been found in MOS applications in cases of short time local weather changes or updates of the atmospheric model in use (see e.g. Landberg, 1994; Joensen et al., 1999).

One of the most successful approaches to this problem is the use of Kalman filters (Kalman, 1960; Kalman and Bucy, 1961; Bossanyi, 1985; Persson, 1990; Dragulanescu, 1993; Kalnay, 2002; Galanis and Anadranistakis, 2002; Crochet, 2004; Giebel, 2000). They consist of a set of mathematical equations that provides an efficient computational solution of the least square method. In practice, the Kalman filter is the statistically optimal sequential estimation procedure for dynamic systems. Observations are recursively combined with recent forecasts using weights that minimize the corresponding biases.

The main advantage of this statistical methodology is the easy adaptation to any alteration of the observations as well as the fact that it may utilize short series of background information. Kalman filters can be widely used for pure meteorological purposes as well as for several other applications, e.g. wind power prediction (c.f. ANEMOS

Correspondence to: G. Kallos
(kallos@mg.uoa.gr)

project[1]). The structure of Kalman filter algorithms is more suitable to describe linear procedures. For this reason, their application on meteorological parameters that follow a non-linear or discontinuous behavior is always dubious.

The aim of this paper is twofold: At first, a new way for the encapsulation of non linear polynomial functions in a classical linear Kalman filter algorithm is proposed. In this way, a more convenient method for the description of the major part of meteorological parameters is obtained. Thereafter, a detailed study of Kalman filtering application to NWP data, leading to the clarification of the optimum polynomial for such type of simulations is illustrated. The results are based on NWPs and observations of temperature and wind speed obtained at two different locations in south Europe and for different time periods. Further, the obtained optimal filter is applied to a longer data set (of one year) for which the improvement of the direct model output is revealed.

2 Description of the methodology

This section initially describes the general form of a Kalman filter using the unified notation proposed by Ide et al., 1997. The main goal is the simulation of the evolution in time of an unknown process (state vector), whose "true" value at time t_i is denoted here by $x^t(t_i)$. A relevant known array (odbervations) y_i^O at the same time is also utilized. The change of x in time is described by:

$$x^t(t_{i+1}) = M_i[x^t(t_i)] + \eta(t_i) \tag{1}$$

and the relation between the observation vector and the unknown one is:

$$y_i^O = H_i[x^t(t_i)] + \varepsilon_i. \tag{2}$$

The matrices M_i (system operator) and H_i (observation operator) have to be determined before the application of the filter. This is also the case for the covariance matrices $Q(t_i)$, $R(t_i)$ of the Gaussian with zero mean (by assumption) and independent random vectors $\eta(t_i)$ (noise process) and ε_i.

The Kalman filter provides a method for the recursive estimation of the unknown state x^t based on all observation values y^O up to time t_i. A first (forecast) step of the state vector and its error covariance matrix P, based only on the previous' time step analysis values, is given by:

$$x^f(t_i) = M_{i-1}[x^a(t_{i-1})], \tag{3a}$$

$$P^f(t_i) = M_{i-1}P^a(t_{i-1})M_{i-1}^T + Q(t_{i-1}). \tag{3b}$$

This is followed up by an update (analysis) step in which the observation available at time t_i is blended with the previous information:

$$x^a(t_i) = x^f(t_i) + K_i(y_i^O - H_i[x^f(t_i)]), \tag{4a}$$

$$P^a(t_i) = (I - K_i H_i)P^f(t_i) \tag{4b}$$

where

$$K_i = P^f(t_i)H_i^T[H_i P^f(t_i)H_i^T + R_i]^{-1} \tag{5}$$

is the Kalman gain that arranges how easily the filter adjusts to possible new conditions. Note that the superscripts o, t, f, a denote observations, true, forecast and analysis value correspondingly. Moreover, T, -1 denote the transpose and the inverse matrix while I stands for the unitary matrix.

Equations (1)–(5) update the Kalman algorithm from time t_{i-1} to t_i.

During the last years, Kalman filters of the above type have been used for meteorological purposes and especially for improving weather forecasts. However, the linear form of the algorithm as well as the special type (non continuous in some cases) of wind speed time series, are important drawbacks for such applications on the prediction of wind parameters, affecting significantly the final outcome. Therefore, while the application of Kalman filters for the improvement of local air temperature predictions seems, in most cases, to be successful, analogous work for wind speed predictions may lead to poor results. Some previous work on this subject can be found in Bossanyi, 1985; Giebel, 2000.

Our approach is based on the non linear correction of forecast bias using 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, let m_i denoting the direct output of the model at time t_i referring on one parameter (temperature or wind speed) that we estimate at each case. Let also y_i^O be the bias of this forecast. We realize it by means of m_i as a polynomial:

$$y_i^O = a_{0,i} + a_{1,i} \cdot m_i + a_{2,i} \cdot m_i^2 + \ldots + a_{n-1,i} \cdot m_i^{n-1} + \varepsilon_i, \tag{6}$$

where the coefficients $(a_{j,i})$, j=0,1,...,n-1, are the parameters that have to be estimated by the filter and ε_i the Gaussian, non systematic, error of the previous procedure. Note here that y_i^O, ε_i are scalar variables contrary to the general case Eq. (2) where they are vectors.

We consider, in other words, as state vector the one formed by the coefficients $(a_{j,i})$, namely, $\mathbf{x}(t_i) = \begin{bmatrix} a_{0,i} & a_{1,i} & a_{2,i} & \ldots & a_{n-1,i} \end{bmatrix}^T$, as observation the (scalar here) bias y_i^O, as observation matrix $H_i = \begin{bmatrix} 1 & m_i & m_i^2 & \ldots & m_i^{n-1} \end{bmatrix}$ and as system matrix the identity matrix I_n. In this way, the system and observation equations take the following form:

$$x^t(t_{i+1}) = x^t(t_i) + \eta(t_i), \quad y_i^O = H_i[x^t(t_i)] + \varepsilon_i. \tag{7}$$

This estimation may be of a linear form by choosing as order n=2, or a polynomial one of arbitrary degree. We succeed in this way to study non linear procedures, and therefore to improve the numerical predictions of several meteorological parameters of non-linear behavior.

The covariance matrices $Q(t_i)$, $R(t_i)$ are defined explicitly in Sect. 3.

3   The optimal filter

The algorithm presented in the previous section offers the advantage of employing non-linear functions for the elimination of the systematic error in the forecast of a specific meteorological parameter. Moreover, one can use any natural number (at least theoretically) as the order of non-linearity, increasing in this way the sensitivity of the filter. The question then rising is: "To which extent the increase in the order of the polynomial in use influences positively the performance of the filter?", in other words, "Is there an "optimum" polynomial that ensures the maximum credibility of the filter?"

In this section the application of different order polynomials is examined in order to extract the optimum order based on the best performance of the filter in improving the NWP data, in combination with the minimization of the required CPU time. The NWP data used in this study were extracted from SKIRON model (see Kallos, 1997; Papadopoulos et al., 2001). SKIRON, based on the Eta model (Janjic, 1994), runs operationally at the University of Athens providing 5-day forecasts[2]. It is a full physics non-hydrostatic model with sophisticated convective, turbulence and surface energy budget scheme. It has several unique capabilities making it appropriate for regional/mesoscale simulations in regions with varying physiographic characteristics. SKIRON uses NCEP/GFS meteorological data for initial and lateral boundary conditions for operational purposes at a resolution of $1°$, 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 variance of the sub-grid scale topography is also taken into account. The domain of the model covers the entire Mediterranean region with a horizontal increment of $0.1° \times 0.1°$ (i.e. $10\,\text{km} \times 10\,\text{km}$ approximately), while 38 Eta levels are used in the vertical.

SKIRON NWP data of long time periods, i.e. at least one month forecast of 5-days horizon, have been provided for wind energy production purposes in the framework of ANEMOS project. In particular, SKIRON forecasts, namely wind speed and direction, air temperature and mean sea level pressure, have been supplied for different locations in the Mediterranean where wind farms operate. For the specific case study, SKIRON NWP data, namely, air temperature at $2\,\text{m}$ and wind speed at $10\,\text{m}$ above the ground for the area of Alaiz, Spain (grid point used: W $1.6°$, N $42.7°$, $672\,\text{m}$ above mean sea level), for December 2003 and for Rokas, Crete (grid point used: E $26.2°$, N $35.2°$, $480\,\text{m}$ above mean sea level) for the whole year 2003 were available. At the same time observations of temperature at $2\,\text{m}$ and wind speed

---

at $55\,\text{m}$ at Alaiz and at $40\,\text{m}$ at Rokas were provided in ANEMOS project. Kalman filtering was applied to the wind speed at $55\,\text{m}$ and $40\,\text{m}$ as it is calculated from the model forecasts at the nearest model levels (weighted interpolation). Hereafter, Alaiz will refer to as Case I and Rokas as Case II.

The use of these extended time periods, apart from increasing the credibility of the statistical analysis, ensures also that the final performance of the proposed filters will not be seriously affected by the choice of the initial conditions. This issue is also supported by the definition of all initial values for filter parameters. More precisely, the initial value of the state vector x is zero, assuming, in this way, that the initial bias of the forecasting model in use is non-systematic: $y_0^O = \varepsilon_0$ Eqs. (6, 7). On the other hand, the covariance matrix P (Eq. 3b) is considered initially diagonal, indicating trivial correlations between different coordinates of the state vector x. The diagonal elements have an initially relatively large value, here we propose $P(t_0) = \begin{pmatrix} 4 & 0 & ... & 0 \\ 0 & 4 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & & 4 \end{pmatrix}$, that declares low credibility of our first guess. Finally, the initial values of the variances $Q(t_i)$, $R(t_i)$ (Eqs. 3b, 5) are $Q(t_0) = I_n$ (i.e. the identity matrix with dimension equal to the order of the filter in use), $R(t_0) = 6$ (a sufficiently large estimation leading to quick independence from initial conditions). The selection of these values for the variables, leads also to an initial Kalman gain that contributes to fast adaptability of the filter to "any possible new conditions Eq. (5)".

There is a long history of successful Kalman filter applications for air temperature (see e.g. Persson, 1990; Dragulanescu, 1993). On the other hand, the non linear evolution in time of wind speed sets under question the possibility of successful application of classical Kalman filters. In the present study the implementation of non linear functions in the classical (linear) Kalman algorithm made possible the simulation of such type of time series.

A wide range of different forecast times is used in order to avoid any possible forecasting time dependences. Our statistical analysis was based on the:

– Bias of forecasted (filtered or not) values:

$$\text{Bias} = \frac{1}{k} \cdot \sum_{i=1}^{k} (\text{for}(i) - \text{obs}(i)), \qquad (8)$$

where $\text{obs}(i)$ denotes the recorded (observed) value at time i, $\text{for}(i)$ the respective forecasted value (direct model output or improved forecast via the proposed filter) and k the size of our sample. Bias is a crucial parameter for Kalman filtering. Any type of Kalman methods aims at eliminating the standard error and, thus, at vanishing the corresponding biases. The fulfillment of this requirement is the main criterion ensuring the credibility of the filter.

– Absolute Bias:
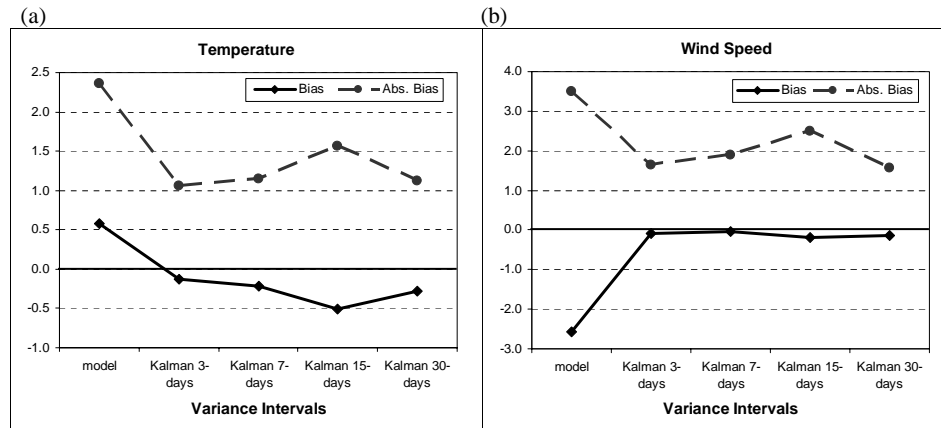
Fig. 1. Bias and Absolute Bias of direct model outputs and Kalman filtered results for **(a)** temperature and **(b)** wind speed, based on different time intervals for the estimation of variance matrices (3, 7, 15 and 30 days)

Table 1. Biases of direct model output and Kalman filters – Case I.

Temperature										
Forecast period	Model	Kal 1	Kal 2	Kal 3	Kal 4	Kal 5	Kal 6	Kal 7	Kal 8	Kal 9	Kal 10
T+24 h	2.47	0.32	**-0.01**	**-0.05**	**0.10**	0.19	0.56	0.46	0.66	0.37	1.06
T+48 h	2.30	0.29	**0.12**	**0.08**	**0.24**	0.38	0.32	0.15	0.43	1.29	0.68
T+72 h	2.55	0.33	**0.07**	**-0.19**	**0.05**	-0.02	0.54	0.39	1.12	0.54	1.13
T+96 h	2.68	0.41	**-0.09**	**-0.07**	**0.30**	0.11	0.18	0.26	0.15	0.35	1.24
T+120 h	3.41	0.23	**0.00**	**-0.05**	**0.08**	0.09	0.71	0.74	0.01	0.82	0.45
Average	**2.68**	**0.32**	**0.02**	**-0.06**	**0.15**	**0.15**	**0.46**	**0.40**	**0.47**	**0.67**	**0.91**
Wind Speed										
Forecast period	Model	Kal 1	Kal 2	Kal 3	Kal 4	Kal 5	Kal 6	Kal 7	Kal 8	Kal 9	Kal 10
T+24 h	-0.47	-0.64	**-0.40**	**-0.53**	**-0.43**	-0.27	-1.10	-0.25	-0.77	-1.04	-1.35
T+48 h	-0.59	-0.23	**-0.22**	**-0.06**	**0.44**	0.52	0.62	0.49	0.48	1.02	0.92
T+72 h	-0.97	0.08	**-0.07**	**0.23**	**0.50**	0.69	0.59	-0.08	-0.11	-0.34	-0.85
T+96 h	-0.98	0.36	**0.16**	**0.29**	**0.12**	0.04	0.20	-0.90	-0.45	-2.94	-2.49
T+120 h	-1.50	-0.29	**-0.58**	**-0.28**	**-0.31**	-0.51	-0.75	-1.25	-1.90	-0.90	-1.90
Average	**-0.90**	**-0.14**	**-0.22**	**-0.07**	**0.06**	**0.09**	**-0.09**	**-0.40**	**-0.55**	**-0.84**	**-1.13**

Table 2. Absolute Biases of direct model output and Kalman filters – Case I.

Temperature										
Forecast period	model	Kal 1	Kal 2	Kal 3	Kal 4	Kal 5	Kal 6	Kal 7	Kal 8	Kal 9	Kal 10
T+24 h	3.02	1.54	**1.19**	**1.33**	**1.35**	1.52	1.77	1.97	2.20	2.31	3.03
T+48 h	2.92	1.40	**1.40**	**1.42**	**1.51**	1.55	1.73	2.34	2.36	2.92	3.53
T+72 h	3.11	1.43	**1.21**	**1.41**	**1.80**	1.69	2.15	1.88	2.57	2.57	2.46
T+96 h	3.05	1.56	**1.21**	**1.56**	**1.54**	1.50	1.48	1.79	1.86	1.90	2.67
T+120 h	3.72	1.48	**1.19**	**1.28**	**1.30**	1.38	1.98	2.10	1.52	2.23	1.90
Average	**3.16**	**1.48**	**1.24**	**1.40**	**1.50**	**1.53**	**1.82**	**2.02**	**2.10**	**2.39**	**2.72**
Wind Speed										
Forecast period	model	Kal 1	Kal 2	Kal 3	Kal 4	Kal 5	Kal 6	Kal 7	Kal 8	Kal 9	Kal 10
T+24 h	2.36	2.43	**2.48**	**2.53**	**2.81**	2.77	3.08	2.91	3.84	3.48	4.49
T+48 h	2.79	2.86	**2.68**	**2.67**	**2.80**	2.67	2.62	2.84	3.22	3.50	4.66
T+72 h	3.22	3.18	**2.94**	**3.05**	**2.83**	3.02	2.97	3.08	3.86	3.88	4.80
T+96 h	3.41	2.96	**2.77**	**2.74**	**2.71**	2.64	2.95	3.33	3.27	5.44	4.65
T+120 h	3.80	3.67	**3.57**	**3.74**	**3.76**	4.11	4.20	4.84	6.93	5.62	4.63
Average	**3.12**	**3.02**	**2.89**	**2.95**	**2.98**	**3.04**	**3.16**	**3.40**	**4.22**	**4.38**	**4.65**

$$\text{Absolute Bias} = \frac{1}{k} \cdot \sum_{i=1}^{k} |\text{for}(i) - \text{obs}(i)| \qquad (9)$$

where | | denotes the absolute value. This parameter assesses the real improvement in the model predictions despite any possible changes in the type of errors.

– Standard deviation of the bias:

$$\text{St. Dev. of Bias} =$$

$$\sqrt{\frac{1}{k} \cdot \sum_{i=1}^{k} ((\text{for}(i) - \text{obs}(i)) - \text{Bias})^2} \qquad (10)$$

and of the absolute bias:

St. Dev. of Absolute Bias=

$$\sqrt{\frac{1}{k} \cdot \sum_{i=1}^{k} (|\text{for}(i) - \text{obs}(i)| - (\text{Absolute Bias}))^2} \quad (11)$$

are used as indexes that estimate the variability of the results.

The variances $Q(t_i)$, of the system equation, and $R(t_i)$, of the observation equation, are based on the sample of the last 7 values of $\eta(t_i) = x^t(t_{i+1}) - x^t(t_i)$ and $\varepsilon_i = y_i^O - H_i[x^t(t_i)]$ respectively :

$$Q(t_i) \equiv$$

$$\frac{1}{6} \cdot \sum_{i=0}^{6} (((x^t(t_{i+1}) - x^t(t_i)) - (\frac{\sum_{i=0}^{6}(x^t(t_{i+1}) - x^t(t_i))}{7})))^2, \quad (12)$$

$$R(t_i) \equiv$$

$$\frac{1}{6} \cdot \sum_{i=0}^{6} (((y_i^O - H_i[x^t(t_i)]) - (\frac{\sum_{i=0}^{6}(y_i^O - H_i[x^t(t_i)])}{7})))^2. \quad (13)$$

The latter are objective estimators of $Q(t_i)$, $R(t_i)$ respectively since the variables $\eta(t_i)$ and ε_i, denoting the non-systematic part of errors in Eq. 7, follow the normal distribution by assumption.

The time period for the estimation of variances was determined to 7 days after a series of tests which led to the conclusion that this short time interval is adequate to obtain significantly improved forecasts with the application of the filter. On the other hand, this choice allows fast adaptability to possible data alternations and, at the same time, does not create needs for extended data storage.

A part of these tests is presented in Fig. 1, where the filter has been used for a period of 1 year based on different time intervals for the estimation of variance matrices: 3, 7, 15 and 30 days for both meteorological parameters in study (temperature and wind speed). The results indicate that no other choice leads to significantly improved results.

Additionally, this is supported by the results of absolute bias presented in Tables 2 and 4 for Cases I and II, respectively. The significant reduction of the absolute bias for the Kalman filters of order of two to four ensures that the discrepancy between the two time series (observed and forecasted) has been considerably decreased despite any possible changes in the type of error.
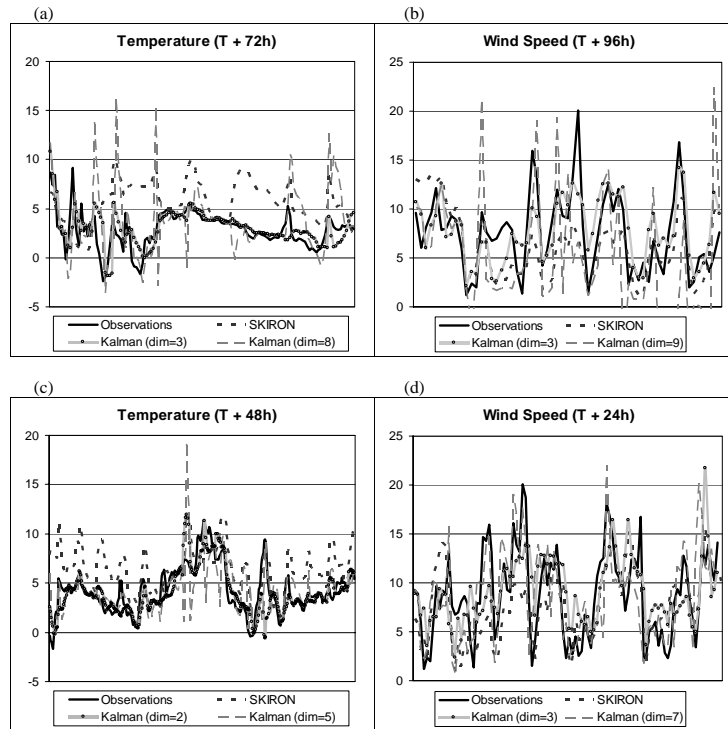
Tables 1 and 3 present the overall performance of the Kalman filter based on the bias using polynomials of first to tenth order against the model direct output for both areas of application. More precisely, Table 1 refers to Case I for December 2003 and Table 3 contains the results for Case II for the period March to May 2003 (this 3-month period

Table 3. Biases of direct model output and Kalman filters – Case II.
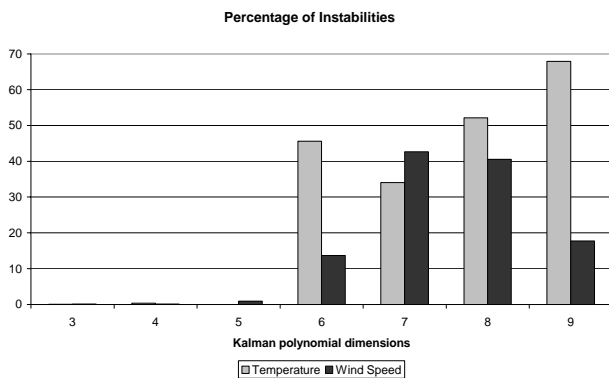
Temperature											
Forecast period	model	Kal 1	Kal 2	Kal 3	Kal 4	Kal 5	Kal 6	Kal 7	Kal 8	Kal 9	Kal 10
T+24 h	0.37	0.39	**0.08**	**-0.05**	**-0.03**	-0.08	-0.01	0.78	-0.36	0.74	0.17
T+48 h	0.46	0.22	**-0.08**	**-0.13**	**-0.09**	-0.07	-0.99	0.58	0.82	1.01	0.45
T+72 h	0.56	0.29	**-0.08**	**-0.10**	**-0.02**	-0.23	0.16	-0.24	-0.25	0.15	-0.14
T+96 h	0.31	0.33	**0.00**	**0.03**	**0.07**	-1.49	-0.27	0.08	0.29	0.77	-0.55
T+120 h	0.28	0.39	**-0.18**	**-0.23**	**-0.06**	0.23	0.01	-0.95	0.20	-0.71	0.15
Average	**0.40**	**0.32**	**-0.05**	**-0.10**	**-0.03**	**-0.33**	**-0.22**	**0.05**	**0.14**	**0.39**	**0.02**

Wind Speed											
Forecast period	model	Kal 1	Kal 2	Kal 3	Kal 4	Kal 5	Kal 6	Kal 7	Kal 8	Kal 9	Kal 10
T+24 h	-2.03	-0.12	**-0.14**	**-0.26**	**-0.23**	-0.20	-0.25	-0.07	-0.08	-0.36	-0.31
T+48 h	-2.23	-0.07	**-0.09**	**-0.13**	**-0.14**	-0.01	0.33	0.02	-0.63	-0.62	-0.60
T+72 h	-2.33	-0.07	**-0.10**	**-0.05**	**-0.04**	-0.24	-0.63	-0.19	-0.44	-0.20	-0.60
T+96 h	-2.64	0.03	**-0.14**	**-0.19**	**-0.12**	-0.23	-0.64	-0.07	-0.46	-0.73	-1.16
T+120 h	-3.17	-0.17	**-0.26**	**-0.32**	**-0.34**	-0.38	-0.25	0.06	-0.46	-0.98	-0.90
Average	**-2.48**	**-0.08**	**-0.15**	**-0.19**	**-0.17**	**-0.21**	**-0.29**	**-0.05**	**-0.41**	**-0.58**	**-0.71**

Table 4. Absolute Biases of direct model output and Kalman filters – Case II.

Temperature											
Forecast period	model	Kal 1	Kal 2	Kal 3	Kal 4	Kal 5	Kal 6	Kal 7	Kal 8	Kal 9	Kal 10
T+24 h	1.62	1.34	**0.86**	**0.90**	**0.88**	0.91	2.54	3.58	6.52	5.74	5.25
T+48 h	1.77	1.42	**0.91**	**0.99**	**1.02**	1.25	2.66	3.58	6.88	7.37	4.29
T+72 h	1.81	1.46	**0.90**	**0.98**	**0.99**	1.19	1.99	3.53	5.01	6.31	6.82
T+96 h	1.78	1.37	**0.94**	**0.99**	**1.07**	3.12	1.67	4.04	4.48	5.84	5.34
T+120 h	2.12	1.49	**1.07**	**1.07**	**1.07**	1.36	1.89	4.82	5.36	6.23	6.34
Average	**1.82**	**1.42**	**0.94**	**0.99**	**1.01**	**1.57**	**2.15**	**3.91**	**5.65**	**6.30**	**5.61**

Wind Speed											
Forecast period	model	Kal 1	Kal 2	Kal 3	Kal 4	Kal 5	Kal 6	Kal 7	Kal 8	Kal 9	Kal 10
T+24 h	2.76	1.74	**1.72**	**1.80**	**1.83**	1.84	2.00	2.55	5.03	4.44	7.72
T+48 h	3.11	1.91	**1.82**	**1.87**	**2.11**	2.03	2.72	2.88	4.62	4.33	5.05
T+72 h	3.31	1.99	**1.91**	**2.04**	**2.05**	2.08	2.77	2.97	4.85	3.86	5.28
T+96 h	4.11	2.49	**2.45**	**2.48**	**2.59**	2.53	2.77	3.52	3.17	5.21	6.94
T+120 h	4.43	2.50	**2.52**	**2.55**	**2.65**	2.76	3.25	4.52	5.06	5.21	6.52
Average	**3.54**	**2.13**	**2.08**	**2.15**	**2.25**	**2.25**	**2.70**	**3.29**	**4.55**	**4.61**	**6.30**

(a) Temperature (T + 72h)   (b) Wind Speed (T + 96h)

(c) Temperature (T + 48h)   (d) Wind Speed (T + 24h)

Fig. 2. Time series of the **(a)**, **(c)** air temperature and **(b)**, **(d)** wind speed observed, forecasted and Kalman filtered for different forecasting hours with different polynomials.

Fig. 3. Percentage of instabilities resulting by remarkably increased values of coefficients for different Kalman filter polynomials in a time series of one year for air temperature and wind speed.

have been selected for investigating which the optimum filter is. This period was covering the spring time which is characterized by variable weather giving, thus, the opportunity to study the behaviour of the filter under different forecast situations). The bias concerns the data from all five different forecasting periods of SKIRON. The results imply that a Kalman filter of order of *two* to *four* for both air temperature and wind speed is the optimal choice for all forecasting

periods. In all cases the corresponding bias is close to zero suggesting that the main goal of a Kalman-type filter is fulfilled.

Any attempt to use higher order polynomials, even by restricting the corresponding Kalman gain, does not lead to an additional improvement of the sensitivity of the filter. On the contrary, several instabilities in the corresponding time series emerge (see Fig. 2), and the overall performance of the filter seems to deviate from its optimum values. These instabilities, resulting in remarkably increased coefficients of the models in use (i.e. $a_{j,i}$ of Eq. (6) obtain values greater than 100), are presented in Fig. 3 as a percentage of the total number of outputs. This figure shows the analysis of the 120 h forecast cycle for temperature and the 72 h forecast cycle for wind speed. It is obvious that any polynomial of degree greater than 5 leads to a large percentage of instabilities. This is further supported in Figs. 4 and 5 where the bias and absolute bias refer to all available data (covering any forecasting period) and all different types of Kalman filter. Moreover, the standard deviation of the bias is significantly increased in the cases of filters of order higher than 4 (see Fig. 6), a fact providing further support to the above remarks.

Based on these results, a third order polynomial Kalman filter is selected as the optimum one which was proven to eliminate the systematic error from both air temperature and wind speed predictions. The consistency of the performance
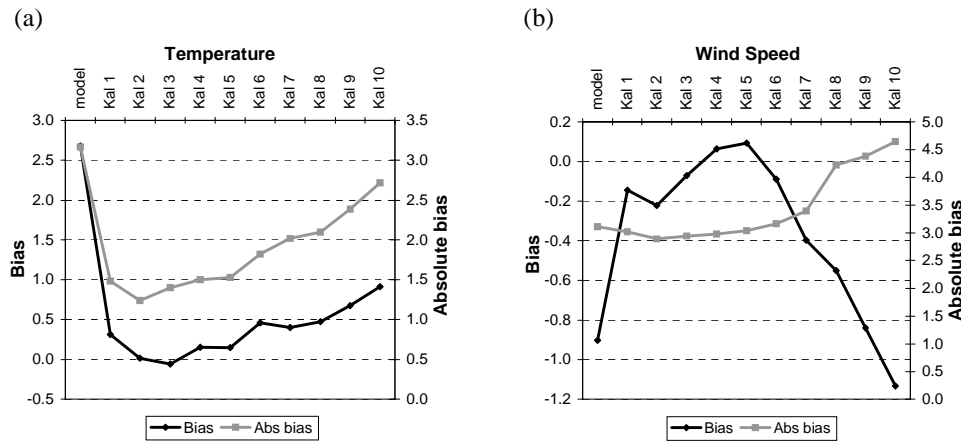
(a)

(b)

Fig. 4. Bias and absolute bias of **(a)** temperature and **(b)** wind speed for model and different Kalman filters outputs (Case I).
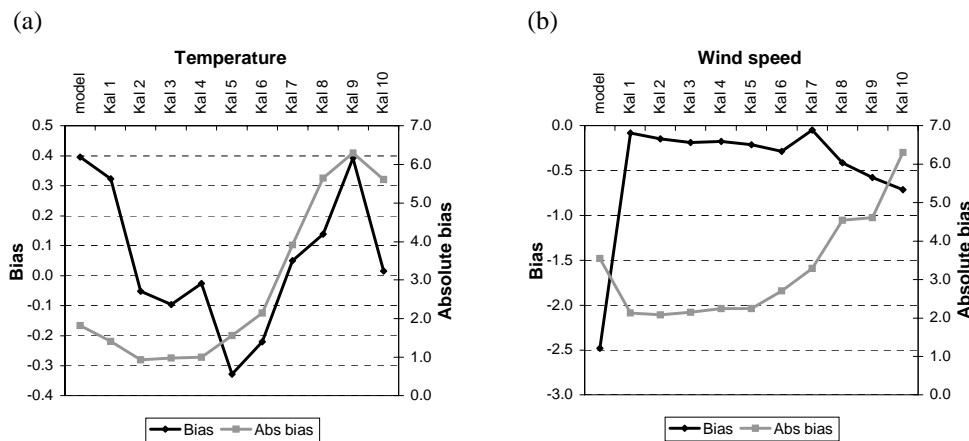
(a)

(b)

Fig. 5. Bias and absolute bias of **(a)** temperature and **(b)** wind speed for model and different Kalman filters outputs (Case II).

of this optimum filter is assessed for a longer time period of one year (2003) using the same statistical analysis.

Figures 7 to 9 focus on the variation of bias, absolute bias and standard deviation of bias of the direct model output and the Kalman filtered output (of order three) with the forecasting period for both air temperature and wind speed. Apart from the obvious improvement in the final forecast, it is worth noticing that this positive influence remains invariant, or even improves, with the forecast time.

The time series illustrated in Fig. 2 further support the significant improvement in the model output due to the application of Kalman filtering based on low order polynomials for both air temperature and wind speed. It is obvious that the systematic error is eliminated regardless its type (underestimation or overestimation). On the other hand, the use of a higher order Kalman filter gives better results only in isolated cases, while it creates considerable instabilities when the type of bias is alternated. Note that the cases presented in

Fig. 2 are characteristic and representative of the general case since similar results appeared in several different selections of the polynomial order.

4 Conclusions

A new methodology of implementing non-linear polynomial functions in the classical linear Kalman filter algorithms is proposed. Different order polynomials were examined, based on time series of two meteorological parameters with different type of behavior, namely the air temperature and the wind speed. The former has a long history of successful compatibility with Kalman filters, while the non continuous evolution in time of the latter may lead to important drawbacks. Our methodology showed high performance for both cases at low computational cost. More precisely, this study suggests that:
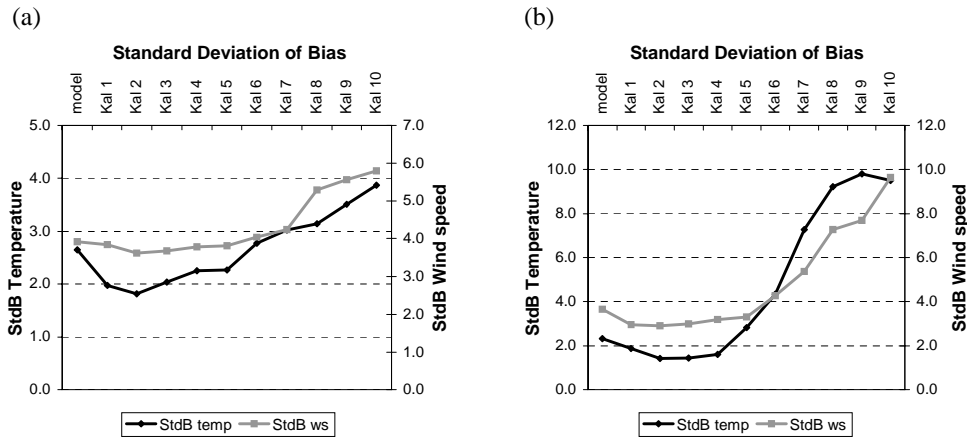
Fig. 6. Standard deviation of bias for temperature and wind speed for model and different Kalman filters outputs in **(a)** Case I and **(b)** Case II.
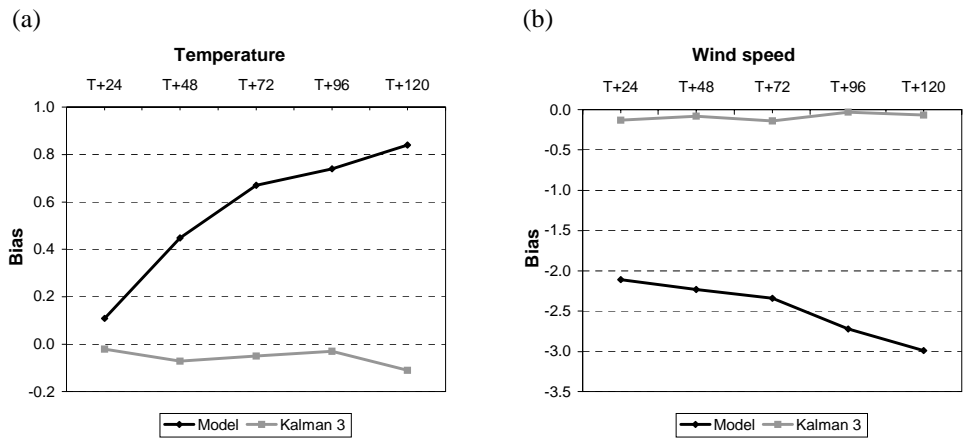
Fig. 7. Variation of **(a)** temperature and **(b)** wind speed bias with forecast time for the direct model output and after the application of order-3 Kalman filter, for the year 2003.

- A low order polynomial, of second to fourth order, seems to be the optimal choice that guarantees the successful elimination of any type of standard bias (linear or not) strongly contributing, in this way, to a successful final forecast.

- Higher order polynomials do not increase the sensitivity of the Kalman filter in use, while they require increased CPU time. Moreover, they create, in several cases, intense discontinuities in the filtered forecast.

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 as a post-processing method or it can be activated in certain time intervals during the integration, smoothing any possible temporary discontinuity that a rapid change in the time series could produce. It

could be also exploited as a pre-processing method provided that an operational observation network is available.

On the other hand, there is a wide range of potential applications beyond the traditional meteorological use, e.g. wind energy predictions and forest-fire fighting operations as well as image processing.

(a)

(b)
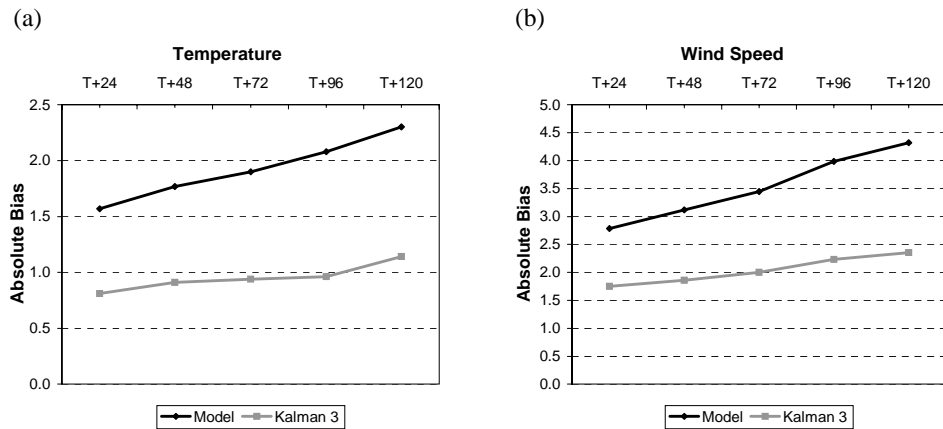
Fig. 8. Variation of **(a)** temperature and **(b)** wind speed Absolute Bias with forecast time for the direct model output and after the application of order-3 Kalman filter, for the year 2003.

(a)

(b)
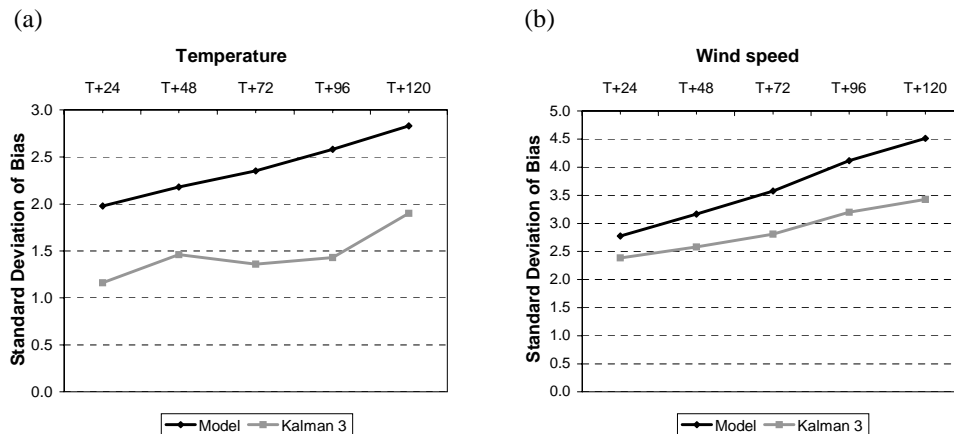
Fig. 9. Variation of **(a)** temperature and **(b)** wind speed Standard Deviation of Bias with forecast time for the direct model output and after the application of order-3 Kalman filter, for the year 2003.

References

Bossanyi, E. A.: Short-Term Wind Prediction Using Kalman Filters, Wind Engineering 9, 1–8, 1985.

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

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

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

Giebel, G.: On the Benefits of Distributed Generation of Wind Energy in Europe, PhD thesis from the Carl von Ossietzky Universität Oldenburg, Düsseldorf, 2001.

Ide, K., Courtier, P., Ghil, M., and Lorenc, A. C.: Unified Notation for Data Assimilation: Operational, Sequential and Variational, J. Meteor. Soc., 75, 181–189, 1997.

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

Joensen, A., Giebel, G., Landberg, L., Madsen, H., and Nielsen, H.: Model Output Statistics applied to wind power prediction, Proceedings of the European Wind Energy Conference, 1177–1180, Nice, France, 1999.

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

Kalman, R. E.: A new approach to linear filtering and prediction

problems, Trans. ASME, Ser. D, 82, 35–45, 1960.

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

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

Landberg, L.: Short-term predictions of local wind conditions, Ph.D. thesis, Riso National Laboratory, Denmark, 1994.

Mass, C. F., Ovens, D., Westrick, K., and Colle, B. A.: Does increasing horizontal resolution produce more skillful forecasts?, Bull. Amer. Meteor. Soc., 83, 3, 407–430, 2002.

Papadopoulos, A., Katsafados, P., and Kallos G.: Regional weather forecasting for marine application, Global Atmos. Ocean Syst., 8, 2–3, 219–237, 2001.

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

Pielke, Sr., R. A.: Further Comments on "The Differentiation between Grid Spacing and Resolution and Their Application to Numerical Modeling", Bull. Amer. Meteor. Soc., 82, 4, 699–700, 2001.