Journal of Wind Engineering & Industrial Aerodynamics 167 (2017) 1–22

Contents lists available at ScienceDirect

Journal of Wind Engineering and Industrial Aerodynamics

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

A hybrid Bayesian Kalman filter and applications to numerical wind speed modeling

George Galanis*, Evgenia Papageorgiou, Aristotelis Liakatas

Hellenic Naval Academy, Section of Mathematics, Mathematical Modeling and Applications Laboratory, Hatzikiriakion, Piraeus 18539, Greece

ARTICLE INFO

Keywords:
Bayesian inference
Kalman filtering
Numerical atmospheric modeling

ABSTRACT

A new hybrid optimization technique for numerical environmental simulation models is proposed and tested in this work. Bayesian modeling is utilized in conjunction with a nonlinear Kalman filter towards a novel post process algorithm applied to numerical wind speed simulations. The new model is tested on idealized data as well as on numerical model forecasts leading to promising results and supporting both the reduction of systematic biases but also the significant limitation of the error variability and the associated forecast uncertainty, a point where classical Kalman filters usually fail to contribute.

1. Introduction

The added value of accurate local environmental forecasts is today generally recognized since a variety of activities are directly associated with them: Wind and wave power forecasts over land and sea areas and the corresponding uncertainty hold a critical role in the smooth integration of available renewable resources in the general power grid; harbor management and coastline safety is closely related with the accurate knowledge of the local weather and wave conditions while early warning systems for natural hazards utilize localized weather forecasts.

Numerical Wind and Wave Predictions Systems (NWPs) provide a great support in this framework. In particular, high resolution simulation systems, developed the last decades in many operational and research centers, have increased the forecasting accuracy. Despite this general progress, local environmental predictions, needed for the applications discussed above, are still revealing the limitations of NWPs with systematic or not errors appearing especially for near surface forecasting. Several parameters contribute to such problems relevant to the physics of atmospheric and wave modeling systems, e.g. parameterization, averaging in landscape characteristics, technical inabilities in simulating successfully sub-scale processes as well as potential problems in the utilized boundary and initial conditions (see e.g. Galanis et al., 2006 and 2009; Kalnay, E, 2002).

These problems make the use of local adaptation post processes and bias removal techniques a critical and indispensable component of integrated forecasting systems. In this framework, a number of approaches have been proposed varying from stable MOS (Model Output Statistics) algorithms to dynamically adjustable filters. The former is a good option when dealing with constant behavior of errors but proved to be expensive on the amount of history data needed and to exhibit discrepancies in short time local weather changes (see e.g. Landberg, 1994; Joensen et al., 1999). On the other hand, Kalman filters (Galanis and Anadranistakis, 2002; Galanis et al., 2006, 2009; Kalman, 1960; Kalman and Bucy, 1961; Kalnay, 2002; Pelland et al., 2011) provide a useful tool for systematic bias removal being the statistically optimal sequential estimation procedure that combine recursively observations with recent forecasts using weights that minimize the corresponding biases. A critical advantage of Kalman filters is the reduced CPU memory needed. However, in many cases, such filters are not able to reduce the error variability but only the systematic mean biases (Louka et al., 2008). This is particularly the case for wind speed predictions that have a variable and discontinuous in some cases time evolution. Such forecasts are, however, of critical importance for a number of activities and applications.

The main objective of this work is to propose new techniques for dealing with the above issues by developing hybrid models which combine Kalman filters with Bayesian inference. The latter is an approach to statistics in which all forms of uncertainty are expressed in terms of probability based on the likelihood function and the prior distribution, deriving information for both the evidence contained in the data and background uncertainty. Bayesian techniques revolves around the posterior probability, which represents total knowledge of parameters after the data have been observed (Box and Tiao, 1992; Bernardo and Smith, 2000) and have been used to deal with a wide range of problems in many scientific areas: Economics and Finance (Beck et al., 2012), Ecology (McCathy, 2007), Data Analysis (Gelman et al., 2013), Image Processing (Sabuncu and Leeinput, 2011), Signal

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

http://dx.doi.org/10.1016/j.jweia.2017.04.007
Received 24 November 2015; Received in revised form 12 March 2017; Accepted 15 April 2017

Processing (Mohammad-Djafari, 2006), Medicine (Etzioni and Kanade, 1995), Mechanical Engineering (Aye et al., 2014).

In this work, we propose a new post processing model to Numerical Weather Predictions for wind speed based on the combination of a nonlinear Kalman filter with a Bayesian model. The Kalman component aims to the elimination of any systematic error in the direct NWP outputs by utilizing nonlinear functions. On the other hand, a subsequent Bayesian algorithm targets to the reduction of the variability of the remaining nonsystematic noise. More precisely, the new system is an extension of the well-established Kalman filter approaches proposed and used for local adaptation of NWPs in the recent years (see, e.g. Galanis and Anadranistakis, 2002; Galanis et al., 2006, 2009; Kalnay, 2002; Louka et al., 2008; Landberg, 1994; Pelland et al., 2011; Wilks, 1995). The novelty of the extended system presented here comes from the enhanced capabilities supported by the combination with the Bayesian approach.

A detailed evaluation has been performed and discussed for the proposed system: Firstly, the new post process technique is checked over idealized data produced randomly by specific probability distribution functions that optimally describe wind speed time series. In this way, the good performance of the proposed model is ensured under ideal conditions. Moreover, tests for "real world" operational forecasts based on a variety of statistical measures is performed over a number of sites covering the major part of the Greek region, where the results of a high resolution numerical model are post-processed by the proposed system. Namely, SKIRON model (Kallos et al., 1997; Papadopoulos et al., 2001) an Eta/NCEP (Janjic, 1994) based non-hydrostatic model has been employed for wind speed forecasting over different local climatological conditions and time periods. The obtained results are more than promising ensuring not only the elimination of any possible systematic biases in the direct NWP outputs but also the significant reduction of the variability of the remaining white noise.

The work presented here is organized as follows: In Section 2 the numerical weather prediction model, the data sets as well as the proposed methodology and algorithms are described. The test cases both for idealized data and operational forecasts are presented, statistically analyzed and discussed in Section 3. The main conclusions are summarized in Section 4.

2. Models and methodology

The modeling systems utilized in this work are presented in this Section. A discussion on the background methodology and the specific configuration adopted aims to the clarification of the capability of the modeling tools that have been utilized.

2.1. The atmospheric model SKIRON

The regional atmospheric system SKIRON (Kallos, 1997; Papadopoulos et al., 2001) has been employed for providing the first level-primordial data under study that is the direct model wind speed forecasts. The SKIRON model is able to support both hydrostatic and non-hydrostatic dynamics by means of explicit time finite-difference arithmetic schemes designed to optimally support computational economy requirements as well as physical parameterization.

The domain of the model covered the major part of Europe, North Africa, the Mediterranean and the Black Sea (see Fig. 1). The horizontal resolution used, based on an Arakawa E-grid, was 0.05×0.05° while 45 vertical levels where employed from surface to 20 km altitude using the step-mountain Eta coordinate grid (Janjic, 1990). Initial and boundary conditions were employed from the NCEP/GFS atmospheric system at 0.5×0.5° horizontal resolution while sea surface boundary conditions are interpolated from the 0.5×0.5° degrees Sea Surface Temperature NCEP analysis field. Recent updates of the model that include the Rapid Radiative Transfer Model RRTMG (Barker et al., 2003; Clough et al., 2005) have been also utilized.

SKIRON model has been successfully evaluated and utilized in a number of major European projects and by many operational and research centers worldwide (see, e.g., Nickovic et al., 2001; Papadopoulos et al., 2002; Papadopoulos and Katsafados, 2009; Zodiatis et al., 2003, 2008; Galanis et al., 2012; Janeiro et al., 2012; Correia et al., 2013). However, as all regional numerical weather prediction systems, is exposed to systematic or not biases, especially when focusing on local area predictions.

2.2. The new Bayesian Kalman filter

The main development presented in this work is a new Kalman filtering model which employs nonlinear functions for reducing possible systematic errors, exhibited by the Numerical Weather Prediction (NWP) system discussed in the previous section, in conjunction with a Bayesian inference approach targeting to the limitation of the remaining white noise and the associated variability in the non-systematic part of the model error.

More precisely, the estimation of the bias y_t in time for wind speed forecasts is pursuit by means of a nonlinear Kalman filter as a polynomial of the direct NWP model output m_t:

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

The above equation forms the observation equation of Kalman filter with state vector $\mathbf{x_t} = [x_{0,t} \;\; x_{1,t} \;\; x_{2,t} \cdots \;\; x_{n,t}]^T$ and observation matrix $\mathbf{H_t} = [1 \;\; m_t \;\; m_t^2 \ldots \;\; m_t^n]$. The evolution of the former in time is described by the system equation:

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

where the system equation matrix F_t coincides with the identity matrix I in the proposed filter, $\boldsymbol{v_t}$ and $\boldsymbol{w_t}$ give the Gaussian nonsystematic parts of the model error with covariance matrices $\mathbf{V_t}$, $\mathbf{W_t}$ respectively which should be estimated before the application of the filter. More precisely, in the proposed approach, $\mathrm{V_t}$ and $\mathrm{W_t}$ are estimated by utilizing the last 7 values of the observed and modeled data, a number that sensitivity tests proved as optimal in order to reach a dynamically flexible but also credible filter (see Galanis et al., 2006, 2009, 2011):

$$V_{t_i} \equiv \frac{1}{6}\cdot\sum_{i=0}^{6}(((y_{t_i} - H_{t_i}x_{t_i}) - (\frac{\sum_{i=0}^{6}(y_{t_i} - H_{t_i}x_{t_i})}{7})))^2,$$

$$W_{t_i} = \frac{1}{6}\cdot\sum_{i=0}^{6}(((x_{t_i+1} - x_{t_i}) - (\frac{\sum_{i=0}^{6}(x_{t_i+1} - x_{t_i})}{7})))^2.$$

Kalman filters in general provide the optimum least square method for the recursive estimation of the unknown state vector $\mathbf{x_t}$ based on recorded "known" values $\mathbf{y}$ up to time t by means of the following main steps:

- An initial estimate of $\mathbf{x_t}$ is provided by $\mathbf{x_{t/t-1}} = \mathbf{F_t}\cdot\mathbf{x_{t-1}}$
- The associated covariance matrix $\mathbf{P_t}$ is given by the relation $\mathbf{P_{t/t-1}} = \mathbf{F_t}\cdot\mathbf{P_{t-1}}\cdot\mathbf{F_t}^\mathbf{T} + \mathbf{W_t}$
- As soon as a new observation value $\mathbf{y_t}$ is available, the estimation of $\mathbf{x}$ at time t takes the form: $\mathbf{x_t} = \mathbf{x_{t/t-1}} + \mathbf{K_t}\cdot(\mathbf{y_t}-\mathbf{H_t}\cdot\mathbf{x_{t/t-1}})$

with the quantity $\mathbf{K_t} = \mathbf{P_{t/t-1}}\cdot\mathbf{H_t}^\mathbf{T}\cdot(\mathbf{H_t}\cdot\mathbf{P_{t/t-1}}\cdot\mathbf{H_t}^\mathbf{T}+\mathbf{V_t})^{-1}$ giving the *Kalman gain*, which is a critical parameter that controls the flexibility of the filter to adjust to any data alterations.

The optimum order of the polynomial utilized is time and space sensitive: Different options may prove to be the best choices depending on the local area and time period characteristics. More details on the Kalman filtering theory, application and sensitivity tests can be found in (Kalman, 1960; Kalman and Bucy, 1961; Kalnay, 2002; Galanis et al., 2006; Louka et al., 2008). The filter has been tested successfully for the local adaptation of the NWP model outputs for different environmental parameters: wind speed (Galanis et al., 2006; Louka