Conditional Bias-Penalized Kalman Filter for Improved Estimation and Prediction of Extremes

by

MIAH MOHAMMAD SAIFUDDIN

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN CIVIL ENGINEERING

THE UNIVERSITY OF TEXAS AT ARLINGTON

JULY 2017

Acknowledgements

At the first, I would like to remember and express my gratitude to Almighty for giving me the inspiration, patience and capabilities to finish this work.

I want to express my gratitude to my supervising Professor Dr. Yu Zhang for keeping faith on me and helping me throughout the whole study. I want to thank my previous supervisor Dr. Dong Jun Seo for whom I have been able to start working as a researcher. I cannot but say that this work would not have been possible without the guidance and help of Dr. Dong Jun Seo.

I want to thank Dr. Habib Ahmari and Dr. Nick Fang, under whom I had taken some major courses towards completion of my degree program. I cannot but thank Dr. Noh who helped me whenever I asked for his help no matter how busy he was.

My special thanks go to Reza Ahmed, Babak Alizadeh, Behzad Nazari, Hamideh Habibi, Behzad Rouhanizadeh, Dr. Sunghee Kim, Fatemeh (Soona) Habibi Ardekani, Mohammad Nabatian and Biplop Dhakal for all their love and support.

My parents and my siblings are my ultimate source of motivation and everything I am, is because of them. I do not have enough words to thank them.

July 27, 2017

Abstract

Conditional Bias-Penalized Kalman Filter for Improved Estimation and Prediction of Extremes

Miah Mohammad Saifuddin, MS

The University of Texas at Arlington, 2017

Supervising Professor: Dr. Yu Zhang

Kalman filter (KF) and its variants are widely used for real-time updating of model states and prediction in environmental sciences and engineering. Whereas in many applications the most important performance criteria may be the fraction of the times when the filter performs satisfactorily under different conditions, in many other applications the performance for estimation and prediction of extremes, such as floods, droughts, algal blooms, etc., may be of primary importance. Because KF is essentially a least squares solution, it is subject to conditional biases (CB) which arise from the error-in-variable, attenuation, effects when the model dynamics are highly uncertain, the observations have large errors and/or the system is not very predictable. In this work, conditional bias-penalized Kalman filter is developed based on CB-penalized linear estimation which minimizes a weighted sum of error covariance and expectation of Type-II CB squared, and comparatively evaluate with KF through a set of synthetic experiments for one-dimensional state estimation under the idealized conditions of normality and linearity. The results show that CBPKF reduces root mean square error (RMSE) over KF by 10 to 20% or more over the tails of the distribution of the true state. For nonstationary cases, CBPKF performs comparably to KF in the unconditional sense in that CBPKF increased RMSE over all ranges of the true state only by 3% or less. With the ability to reduce CB explicitly, CBPKF provides a significant addition to the existing suite of filtering techniques toward improving analysis and prediction of extreme states of uncertain environmental systems.

Contents

List of Illustrations

List of Tables

Chapter 1 Introduction

1.1 Background

Kalman filter (KF, Kalman 1960) and its variants are arguably the most popular class of techniques for state estimation and prediction of dynamic systems. In environmental sciences and engineering, they are widely used for real-time updating of model states via data assimilation (DA), and prediction. Just to name a small subset, ensemble KF (EnKF, Evensen 2003) was used for hydrologic forecasting (Clark et al. 2008, Komma et al. 2008, Moradkhani et al. 2005, Neal et al. 2007, Rafieeinasab et al. 2014, Weerts and Serafy 2006, Xie and Zhang 2010), extended KF (EKF, Jazwinski 1970) and EnKF were used for water quality forecasting and ecological modeling (Cosby 1984, Ennola et al. 1998, Huang et al. 2013, Kim K et al. 2014, Kim S et al. 2014, Xue et al. 2012), and EnKF was used for marine ecosystem and ocean modeling (Eknes and Evensen 2002, Evensen 2003, 2009) and for weather forecasting (Houtekamer and Mitchell 1998, Zhang and Pu 2010). Whereas in most applications the most important performance criterion may be the fraction of the times when the filter performs satisfactorily under different conditions, in many applications the performance for estimation and prediction of extreme may be of primary importance. For example, in flood forecasting, accurate prediction of flood peaks or stages, which occur only for a very small fraction of the times, is far more important than that of average flows. In drought monitoring and prediction, accurate estimation of extremely dry soil moisture states is much more important than estimation of those above the 10th or even lower percentile. In water quality forecasting, being able to estimate accurately high concentrations of Chlorophyll a and other state variables is critical to skillful prediction of algal blooms (Guo et al. 2003, Mao et al. 2009, Pastres et al. 2003, Twigt et al. 2011, Whitehead and Hornberger 1984, Zingone and Enevoldsen 2000). Because KF is essentially a least squares solution (Jazwinski 1970, Schweppe 1973), it is subject to conditional biases (CB) if the model

dynamics are highly uncertain, the observations have large errors and/or the system is not very predictable. By “predictable”, it is meant that it is possible to make skillful predictions of the future state of the system. By “skillful”, it is meant that the prediction in question is superior to some clearly defined reference prediction such as climatological or simple statistical forecast. To illustrate the adverse effect of CB, Figure 1.1 shows an example of the KF estimates vs. the verifying observations from a synthetic experiment using a linear Gaussian state-space model (Kitagawa and Gersch 1996, Durbin and Koopman 2001; see the Evaluation Section for the model used). It is readily seen that, whereas the KF estimates are unbiased over the near-median range of the truth, they are biased everywhere else. This CB problem, where the larger and smaller values are systematically under- and overestimated, respectively, is referred to in the statistical literature as the error-in-variable, or attenuation, effects (Carroll et al. 1995, Fuller 1987, Seber 1989). These effects occur in least squares regression when the predictors are observed with significant error. In modeling and observation of environmental systems, such effects are the norm rather than the exception (see, e.g., Ciach et al. 2000). As such, the solutions obtained from KF or its variants often suffer from CB with potentially very large negative impact on estimation and prediction of extremes.

Type-I and –II CBs emerge from Type-I and Type-II errors which occur when falsely detecting an effect which does not exist, and when failing to detect an existing effect, respectively (Joliffe and Stephenson 2003). Whereas Type-I CB can be reduced via calibration, Type-II CB cannot. Therefore, reducing Type-II CB in optimal estimation addresses an important gap in estimation and prediction of extremes. As the resolution of environmental models continues to increase, it is expected that the model dynamics and observations will be subject to larger uncertainties and variabilities, and hence to larger CBs. As such, there is an increasingly large need for filtering techniques that can explicitly address CB. Recently, Seo (2013) developed a Fisher-like solution for optimal linear

estimation by minimizing a weighted sum of error variance and expectation of Type-II CB squared. The technique has since been applied successfully to rain gauge-only precipitation estimation in the form of conditional bias-penalized kriging (CBPK, Seo 2013, Seo et al. 2014), multi-sensor precipitation estimation using radar and rain gauge data in the form of conditional bias-penalized co-kriging (CBPCK, Kim et al. 2016), bias correction of forecast precipitation and streamflow in the form of conditional bias-penalized indicator co-kriging (CBP-ICK, Brown and Seo 2010), and high-resolution fusion of multiple radar-based quantitative precipitation estimation (QPE) products (Rafieeinasab et al. 2015).

In this work, CB-penalized optimal linear estimation is applied for dynamic filtering, and formulate and evaluate conditional bias-penalized Kalman filter, or CBPKF, which extends KF to explicitly minimize Type-II CB in addition to error variance. While much progress has been made since the introduction of KF in extending KF for nonlinear predictions and ensemble forecasting as referenced above, to the best of my knowledge no attempts have been reported to address CB explicitly until this paper. With the ability to reduce CB explicitly, CBPKF provides a significant addition to the existing suite of filtering techniques toward improving analysis and prediction of extreme states of uncertain environmental systems.

1.2 Objective

The main objective of this study is to develop conditional bias-penalized Kalman filter, or CBPKF, based on CB-penalized linear estimation which minimizes a weighted sum of error covariance and expectation of Type-II CB squared, and comparatively evaluate with KF through a set of synthetic experiments for one-dimensional state estimation under the idealized conditions of normality and linearity. It is expected that that CBPKF reduces root mean square error (RMSE) over KF over the tails of the distribution of the true state. For nonstationary cases, CBPKF performs comparably to KF in the unconditional sense in that

CBPKF increased RMSE over all ranges of the true state only by 3% or less. With the ability to reduce CB explicitly, CBPKF provides a significant addition to the existing suite of filtering techniques toward improving analysis and prediction of extreme states of uncertain environmental systems.

1.3 Outline of the Thesis

Chapter 1 Presents the background of the study and the statement of the problem, specific objectives of the study.

Chapter 2 Describes the CBPKF methodology and the algorithm.

Chapter 3 Describes comparative evaluation of CBPKF with KF.

Chapter 4 Presents the evaluation results.

Chapter 5 Presents the conclusion and suggested recommendations for further improvement of the proposed technique.

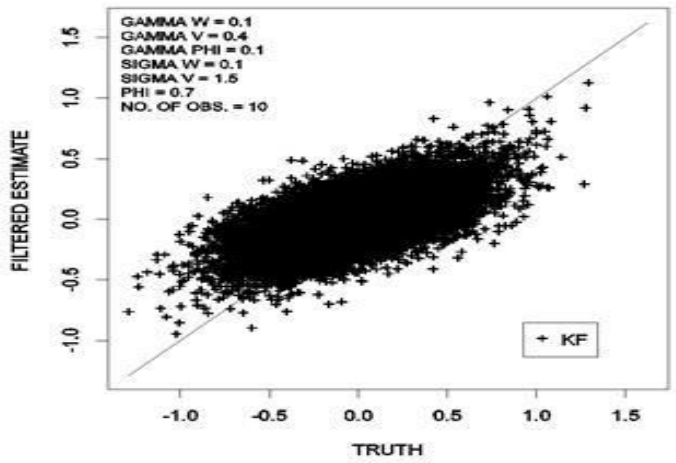

Figure 1.1 Example of the KF estimates vs. the truth from a synthetic experiment.

Chapter 2 Methodology

This chapter surveys the literature related to CB type and describes the algorithms related to CBPKF.

Conditional Bias:

Unbiased estimates having minimum error variance in the unconditional sense is produced by most of the conventional precipitation estimation techniques. But in the conditional (on the magnitude of the truth being estimated) sense, heavy precipitation is underestimated and light precipitation are overestimated by these techniques (Seo, 2012). There are two types of CB.

1) Type-I conditional bias (Type-I CB)
2) Type-II conditional bias (Type-II CB)

There is a widespread confusion about CB in the literature and among the practitioners of geo-statistics (Mclennan and Duestch, 2003). Isaak (2004) stated that CB is poorly understood even though it is a well-recognized problem in geo-statistics. Lack of distinction between Type-I and Type-II CB in the literature may be one of the main sources of such confusion (Seo, 2012).

Type-I CB:

Type-I CB is defined as:

$$E[X|X^*] - X^*$$

where X^* and X the estimate and the truth, respectively. Type-I CB occurs when the estimate is biased against the expected value of the true precipitation conditional on the estimate. (Siddique, R. 2014)

Type-II CB:

Type-II CB is defined as:

$$E[X^*|X] - X$$

Type-II CB exists when the expected value of the estimate given the truth differs from the

truth.

This study is concerned with reducing Type-II CB. Different scientists and geo-statisticians discuss the problem of Type-II CB in different contexts e.g. in those of mining and radar rainfall estimation (Siddique, R. 2014).

Brown and Seo (2012) proposed a new non-parametric technique to minimize Type-II CB in streamflow prediction. This technique is analogous to indicator co-kriging (ICK) and is called conditional bias-penalized indicator co-kriging (CBP-ICK). It is found that CBP-ICK successfully reduce Type-II CB and produce estimates that are more skillful than the estimates from other post processors used in hydrologic prediction (Siddique, R. 2014).

Seo (2012) proposed and described a new estimation technique, CBPK, which is an extension of SK. CBPK adds a penalty term for Type-II CB in addition to error variance. Seo (2012) evaluated CBPK using normal and lognormal synthetic experiments and found that CBPK successfully reduces Type-II CB for large precipitation amounts (Siddique, R. 2014). Seo (2012) also described a Fisher-like solution of CBPK

Seo (2013) has shown that, by minimizing the weighted sum, $J = \Sigma_{EV} + \alpha\Sigma_{CB}$, of the error variance of the true states, $\Sigma_{EV} = E_{X,X^*}[(X - X^*)(X - X^*)^T]$, and the expectation of the Type-II CB squared, $\Sigma_{CB} = E_X[(X - E_{X^*}[X^* \mid X])(X - E_{X^*}[X^* \mid X])^T$, where X, X* and α denote the true state of the system, optimal (in some sense of the word) estimate of the state, and scaler weight given to Σ_{CB}, and the subscripts to expectation operation signify that the expectation is with respect to the variables subscripted, one arrives at the following Fisher-like solution for linear estimation (see Equations. A20 and A21 in Appendix A for the derivation of Equations. (1) and (2)):

$$\Sigma = \mathrm{B}[\hat{H}^T \Lambda^{-1} H]^{-1} \tag{1}$$

$$X^* = WZ = [\hat{H}^T \Lambda^{-1} H]^{-1} \hat{H}^T \Lambda^{-1} Z \tag{2}$$

In the above, Σ denotes the (m×m) estimation error covariance matrix, B is the (m×m) scaling matrix, $\hat{H}^T$ is the (mx(n+m)) modified structure matrix, and Λ denotes the ((n+m)×(n+m)) modified observation error covariance matrix, X^* denotes the (m×1) vector of the estimates, W denotes the (m×(n+m)) weight matrix, and Z denotes the ((n+m)x1) observation vector. The matrices, $\hat{H}^T$ and Λ, are analogous to the structure and observation error covariance matrices, H and R, in Fisher estimation (see Appendix A) but modified according to Eqs.(3) and (4), respectively. The matrices, $\hat{H}^T$, Λ and B are given by (see Eqs.(A11), (A12) and (A18) in Appendix A):

$$\hat{H}^T = H^T + \alpha \Psi_{XX}^{-1} \Psi_{XZ} \tag{3}$$

$$\Lambda = R + \alpha(1-\alpha)\Psi_{ZX} \Psi_{XX}^{-1} \Psi_{XZ} - \alpha H \Psi_{XZ} - \alpha \Psi_{ZX} H^T \tag{4}$$

$$\mathrm{B} = \alpha \Psi_{XX} \hat{H}^T \Lambda^{-1} \hat{H} + I \tag{5}$$

where the (m×m) and ((n+m)xm) matrices, Ψ_{XX} and Ψ_{ZX}, denote $\mathrm{Cov}[X,X^T]$ and $\mathrm{Cov}[Z,X^T]$, respectively, R denotes the ((n+m)×(n+m)) observation error covariance matrix, $R=E[VV^T]$, and I denotes the (m×m) identity matrix. In this work, it is assumed for simplicity that α is a scaler rather than a matrix so that the penalty given to CB is applied proportionately across all state variables. It is possible to prescribe α, e.g., as a diagonal matrix which would allow state variable-specific weights for the CB penalty term. Such relaxation of α, however, is beyond the scope of this study and is left as a future endeavor.

For complete derivation, the reader is referred to Appendix A which provides the technical context for CBPKF and also includes additional details omitted in Seo (2013) in notations that are consistent with this thesis to avoid confusion. Note in Eqs.(1) through (5) that, if α=0 (i.e., no penalty for Type-II CB), Eqs.(1) and (2) reduce to the Fisher solution. Because Eqs.(3) through (5) require a priori knowledge of Ψ_{XX} and Ψ_{ZX}, Eqs.(1) and (2) do not represent a Fisher solution but a hybrid solution that combines Bayesian and Fisher estimation (see Appendix A). The CB-penalized formulation above may also be considered as a form of regularization in which the a priori information of Ψ_{XX} and Ψ_{ZX} is added to the objective function through a quadratic penalty for Type-II CB.

Noting strong resemblance of Eqs.(1) and (2) to the Fisher solution of $\Sigma_{Fisher} = [H^T R^{-1} H]^{-1}$ and $X^*_{Fisher} = [H^T R^{-1} H]^{-1} H^T R^{-1} Z$, one may arrive at CBPKF by analogy with KF. Appendix B provides the derivation which is largely of technical nature. Appendix C provides an alternative expression for CBPKF based on factorization of Λ^{-1} which juxtaposes with KF for direct comparison and to obtain additional positive semidefinite conditions for specific elements of the filter. In the following, is present the resulting CBPKF algorithm in the context of the following state-space representation of the dynamical model and observation equation in direct analogy with the standard KF algorithm (Schweppe 1973, Bras and Rodriguez-Iturbe 1985):

$$X_k = \Phi_{k-1} X_{k-1} + G_{k-1} w_{k-1} \tag{6}$$

$$Z_k = H_k X_k + V_k \tag{7}$$

where X_k and X_{k-1} denote the (mx1) state vectors at time steps k and k-1, respectively, Φ_{k-1} denotes the (m×m) state transition matrix between time steps k-1 and k,

G_{k-1} denotes the (m×m) scaling matrix for the (mx1) random error vector, w_{k-1} with the (m×m) model error covariance matrix, $Q_{k-1} = E[w_{k-1}w_{k-1}^T]$, Z_k denotes the (nx1) observation vector, H_k denotes the (n×m) structure matrix and V_k denotes the (nx1) observation error vector with the (n×n) observation error covariance matrix, $R_k = E[V_k V_k^T]$. In the above, it is assumed that X_{k-1} and w_{k-1} as well as X_k and V_k are independent. From Appendices B and C, then have the following CBPKF algorithm for estimation of the state at time step k and its error variance:

1) At $k=1$, prescribe the initial conditions for the state and its error covariance, $\hat{X}_{k-1|k-1}$ and $\Sigma_{k-1|k-1}$.

2) Make one step-ahead predictions for the state and its error covariance:

$$\hat{X}_{k|k-1} = \Phi_{k-1}\hat{X}_{k-1|k-1} \tag{8}$$

$$\Sigma_{k|k-1} = \Phi_{k-1}\Sigma_{k-1|k-1}\Phi_{k-1}^T + G_{k-1}Q_{k-1}G_{k-1}^T \tag{9}$$

3) Set initial α, or α0, that satisfies $0 \le \alpha_0 \le (\sqrt{5}-1)/2 \approx 0.618$ (see (B7)).

4) Evaluate the (m×n), (n×n) and (m×m) modified covariance matrices, Λ_{21}, Λ_{11} and Λ_{22}, and invert Λ_{11}:

$$\Lambda_{21} = -\alpha(\alpha+1)\Sigma_{k|k-1}H_k^T \tag{10}$$

$$\Lambda_{11} = R_k - \alpha(\alpha+1)H_k\Sigma_{k|k-1}H_k^T \tag{11}$$

$$\Lambda_{22} = \{1-\alpha(\alpha+1)\}\Sigma_{k|k-1} \tag{12}$$

5) Check for positive semi-definiteness of Λ_{11}. If α does not satisfy Eq.(13) below, reduce α according to $\alpha_i = c\alpha_{i-1,\ i=1,2,3,\ldots}$ where $0 < c < 1$ and α_i denotes the value of α at the i-th iteration, and go back to Step 4. If α satisfies Eq.(13), proceed with Step 6. $0 \le \alpha \le \sqrt{\frac{Tr[R_k]}{Tr[H_k \Sigma_{k|k-1} H_k^T]} + \frac{1}{4}} - \frac{1}{2}$ (13)

6) Evaluate the (m×m), (n×m) and (n×n) matrices, Γ_{22}^{-1}, Γ_{12} and Γ_{11}, and invert Γ_{22}^{-1}:

$$\Gamma_{22}^{-1} = \Lambda_{22} - \Lambda_{21}\Lambda_{11}^{-1}\Lambda_{12} \quad (14)$$

$$\Gamma_{12} = -\Lambda_{11}^{-1}\Lambda_{12}\Gamma_{22} \quad (15)$$

$$\Gamma_{11} = \Lambda_{11}^{-1} + \Lambda_{11}^{-1}\Lambda_{12}\Gamma_{22}\Lambda_{21}\Lambda_{11}^{-1} \quad (16)$$

7) Check for positive semi-definiteness of $(H_k^T\Gamma_{11} + \Gamma_{21})H_k$ and $\Gamma_{21}H_k + \Gamma_{22}$. If either constraint is violated, reduce α according to $\alpha_i = c\alpha_{i-1,\ i=1,2,3,\ldots}$ and go back to Step 4. If both constraints are met, proceed to Step 8.

$$Tr[(H_k^T\Gamma_{11} + \Gamma_{21})H_k] \ge 0 \quad (17)$$

$$Tr[\Gamma_{22} + H_k^T\Gamma_{12}] \ge 0 \quad (18)$$

8) Evaluate the (m×m) updated error covariance matrix, $\Sigma_{k|k}$:

$$\Sigma_{k|k} = (1+\alpha)\alpha\Sigma_{k|k-1} + \{(1+\alpha)[(H_k^T\Gamma_{11} + \Gamma_{21})H_k + \Gamma_{21}H_k + \Gamma_{22}]\}^{-1} \quad (19)$$

9) Evaluate the (m×n) conditional bias-penalized Kalman (CBPK) gain matrix, K_k:

$$K_k = [(H_k^T \Gamma_{11} + \Gamma_{21}) H_k + \Gamma_{21} H_k + \Gamma_{22}]^{-1} [H_k^T \Gamma_{11} + \Gamma_{21}] \quad (20)$$

10) Evaluate the (mx1) updated state vector, $\hat{X}_{k|k}$:

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k [Z_k - H_k \hat{X}_{k|k-1}] \quad (21)$$

11) Go to Step 2 for the next time step and repeat.

Computationally, CBPKF can be significantly more expensive than KF if the size of the state and/or the observation vector is large. Note that, in addition to solving an (m×n) linear system for the gain matrix in Step 9 as in KF, it is also necessary in CBPKF to solve (n×n) and (m×m) linear systems for inversion of Λ_{11} in Step 4 and Γ_{22}^{-1} in Step 6, respectively. An obvious strategy to minimize the number of iterations while satisfying the positive semi-definiteness conditions is to start in Step 3 with a value of α near the upper bound of the feasible region of [0, 0.618] and reduce α incrementally until all constraints are satisfied as close to the equality conditions as possible. In this way, the resulting α represents the largest possible weight for the CB penalty that yields a valid CBPKF solution. The above strategy is guaranteed to succeed in that, if α has to be reduced all the way to zero, CBPKF simply becomes KF. For incremental reduction of α, different strategies of varying complexity are possible. In this work, it is employed that a very simple iterative procedure in which α is reduced geometrically according to $\alpha_i = c\alpha_{i-1},\ i=1,2,3,\ldots$ where $0 < c < 1$ and α_i denotes the value of α at the i-th iteration. A smaller c would satisfy the constraints in fewer iterations but at the expense of potentially over-reducing α. A larger c would produce a larger α but at the expense of increasing the number of iterations. Depending on the size of the DA problem,

some experimentation may be necessary to choose a satisfactory c. Development of a more efficient scheme for iteration is left as a future endeavor.

For stationary processes, i.e., Φ_{k-1}, G_{k-1}, Q_{k-1}, H_k and R_k in Eqs.(6) and (7) are time-invariant, there may exist a better α_0 than ~0.6 in Step 3. For example, if all iteratively-reduced α which satisfies the inequality conditions of Eqs.(13), (17) and (18) is smaller than α_0 at all time steps, one should use the new smaller value of α_0 in Step 3 to avoid unnecessary iterations in real-time implementation. If the combination of model prediction and observation is very informative so that CB is not very large, α may have to be reduced to a level well below ~0.6. In such cases, it is necessary to optimize α to avoid overcorrecting CB. Numerical experiments with one-dimensional examples (see the Results Section) suggest that one may chose α such that the quantile-quantile plot (qqplot) of the CBPKF estimates vs. the verifying observations closely approaches the diagonal line, i.e., the CBPKF estimates have a similar marginal probability distribution as the verifying observations. If the qqplot lies above the diagonal line over the range of the truth greater than its median, it is an indication that α is too large. Numerical experiments of this study also indicate that an optimized α may range between 0.25 and 0.50 for a wide range of conditions that may be encountered in real-world applications. If the KF estimates exhibit only a very small CB but reduction of CB is still desired, it may be necessary to reduce α to below 0.15. The information content of an observation and the number of observations available, n, also impact the choice of α. If n is very small, the observational information content may be too small for CBPKF to offer significant benefits. In such cases, the CBPKF results will be very similar to the KF results. If the individual observations are informative and the number of available observations increases, one may expect the performance of both KF and CBPKF to improve. Numerical experiments of this study suggest that CBPKF in such

cases very often provides significant improvement over KF for estimation of extremes if the processes modeled are not very predictable. Because the qqplots of the filtered estimates do not in general form straight lines, it may not be readily possible to determine whether the optimization of α is completely satisfactory or not based solely on the qqplot. For this reason, it may be necessary to examine additionally conditional error statistics such as the conditional root mean square error (RMSE) as well as relative performance measures such as the relative percent reduction in RMSE in optimizing α. In the Results Section, is provided the examples of the above cases.

In this work, CBPKF is described in the context of state space estimation involving linear dynamical and observation models only. In most real-world applications, the dynamical model and possibly the observation model as well would be nonlinear. It is hence expected that CBPKF would be applied in most applications in the form of an ensemble filter, or ensemble CBPKF (EnCBPKF), to nonlinear dynamical and linear observation models analogously to ensemble KF (EnKF). Description and evaluation of EnCBPKF, however, is beyond the scope of this paper. This work focus on comparative evaluation of CBPKF under idealized conditions via synthetic experiments. The motivation for such evaluation is two-fold. The first is that real-world applications would reflect not only the comparative performance between CBPKF and KF but also other factors such as the degree and extent to which the assumptions of linearity and normality may be met, the performance of the dynamical model and the quality of the uncertainty modeling involved. As such, with real-time applications it would be very difficult to isolate the comparative performance due solely to the filter formulations. The second is that, without applying to many diverse state estimation and prediction problems, it would not be possible to identify easily from real-world applications the conditions under which CBPKF may offer significant benefits over KF. The evaluation carried out in this work, on the other hand, allows performance comparisons

under the idealized conditions of linear dynamical and observation models, normally distributed observation and model errors, and perfectly known uncertain parameters so that we may attribute the difference in performance solely to that in the filters. Also, the wide-ranging conditions considered in this work allow the user to assess the potential benefits of CBPKF to diverse applications that may be encountered in environmental science and engineering.

Chapter 3 Evaluation

To comparatively evaluate CBPKF with KF, the study carried out a set of one-dimensional (1D) numerical experiments in which the dynamics of the true state is given by the state space model in Eq.(22) and the states are observed via the linear observation equation in Eq.(23):

$$X_k = \varphi_{k-1} X_{k-1} + \sigma_{w,k-1} w_{k-1} \tag{22}$$

$$Z_k = UX_k + V_k \tag{23}$$

where X_k and X_{k-1} denote the true states at time steps k and k-1, respectively, φ_{k-1} denotes the state transition coefficient at time step k-1, $\sigma_{w,k-1}$ denotes the input coefficient at time step k-1 for the serially uncorrelated random error, $w_{k-1} \sim N(0,1)$, Z_k denotes the (nx1) observation vector, U denotes the (nx1) unit vector and V_k denotes the (nx1) observation error vector, $V_k = [v_{1,k}, v_{2,k}, ..., v_{n,k}]^T$, with $v_{i,k} \sim N(0, \sigma_{v,k}^2)$, $i = 1, ..., n$. Note that, if the process in Eq.(22) is stationary, φ_{k-1} denotes the constant lag-1 serial correlation of the state variable and $\sigma_{w,k-1}$ denotes the constant standard deviation of the random error, $\sigma_{w,k-1} w_{k-1}$. The number of observations, n, is assumed to be time-invariant for simplicity. The

observation errors are assumed to be independent among themselves and of the true state. Then is applied KF and CBPKF to obtain $\hat{X}_{k|k}$ and $\Sigma_{k|k}$, and verify them against the assumed truth. Under the above experimental design, the impact of CB on filtering and prediction depends on the choices of φ_{k-1}, $\sigma_{w,k-1}$, $\sigma_{v,k}$ and n, where φ_{k-1} represents the predictability of the process, $\sigma_{w,k-1}$ represents the magnitude of the model error, and $\sigma_{v,k}$ and n represent the information content of the observations. The one-step prediction equations for Eq.(22) are given by $\hat{X}_{k|k-1} = \varphi_{k-1}\hat{X}_{k-1|k-1}$ and $\Sigma_{k|k-1} = \varphi_{k-1}^2\Sigma_{k-1|k-1} + \sigma_{w,k-1}^2$ where $\hat{X}_{k|k-1}$ and $\Sigma_{k|k-1}$ denote the predicted or estimated state and its error variance, respectively, valid at time step k given all available information through time step k-1. To assess comparative performance of CBPKF under widely varying conditions, two types of experiments were carried out, stationary and nonstationary. In the stationary experiment, different values of time-invariant φ_{k-1} is used and n while time-invariant $\sigma_{w,k-1}$ and $\sigma_{v,k}$ remain fixed. In the two nonstationary experiments, φ_{k-1}, $\sigma_{w,k-1}$ and $\sigma_{v,k}$ are randomly perturbed according to Eqs.(24) through (26) and used only those deviates that satisfy the bounds imposed below:

$$\varphi_{k-1}^p = \varphi_{k-1} + \gamma_\varphi \varepsilon_\varphi, \qquad 0.5 \le \varphi_{k-1}^p \le 0.95 \tag{24}$$

$$\sigma_{w,k-1}^p = \sigma_{w,k-1} + \gamma_w \varepsilon_w, \qquad \sigma_{w,k-1}^p \ge 0.01 \tag{25}$$

$$\sigma_{v,k}^p = \sigma_{v,k} + \gamma_v \varepsilon_v, \qquad \sigma_{v,k}^p \ge 0.01 \tag{26}$$

In the above, the superscript p signifies that the variable superscripted is a perturbation, ε_φ, ε_w and ε_v denote the standard normal random deviates, and γ_φ, γ_w and

γ_v denote the standard deviation of the normally-distributed random perturbations added to φ_{k-1}, $\sigma_{w,k-1}$ and $\sigma_{v,k}$, respectively. Table 3.1 summarizes the parameter settings used in all experiments. The ranges of values in Table 3.1 are chosen to encompass less predictable (small φ_{k-1}) to more predictable (large φ_{k-1}) processes, certain (small $\sigma_{w,k-1}$) to uncertain (large $\sigma_{w,k-1}$) model dynamics, and informative (small $\sigma_{v,k}$) to non-informative (large $\sigma_{v,k}$) observations. The above parameters may be explained in real-world terms using real-time streamflow forecasting as an example, in which autoregressive-1 and state-space models are used for stationary and nonstationary cases, respectively. A small/large φ_{k-1} would represent the predictability of streamflow for a small/large catchment (and hence of short/large system memory) or that for any catchment in high/low flow conditions. A large/small $\sigma_{w,k-1}$ would represent a large/small collective model error, or hydrologic uncertainty (Krzysztofowicz 1999, Seo et al. 2006), associated with imperfect model structures and parameters, initial conditions and input. The number of observations, n, would represent the number of available real-time streamflow observations valid at the time of filtering in which the observations are made repeatedly at the same location to reduce sampling uncertainty. A large/small $\sigma_{v,k}$ would represent a large/small collective measurement or estimation error in streamflow observation due to lack of precision in the instrument, representativeness (e.g., point vs. cross-sectional area) errors, errors in the rating curve if estimated from stage observations, etc. The bounds for φ_{k-1}^p in Eq. (24) is based on the range of lag-1 serial correlation that represents moderate to high predictability (25 to ~90% of variance explained) where CBPKF and KF are most likely to differ in performance. Note that, if the process is not very predictable, no filter may be expected to perform well, and that, if it is extremely predictable, any reasonable filter would perform well. The purpose of bounding the perturbed values $\sigma_{w,k-1}^p$ and $\sigma_{v,k}^p$ in Eqs. (25) and (26),

respectively, in the nonstationary experiments is to avoid the observational or model prediction uncertainty becoming near zero which is not possible in reality. An unrealistically small $\sigma^p_{w,k-1}$ and $\sigma^p_{v,k}$ would render the information content in the model prediction, $\Sigma_{k|k-1}$, and the observations, Z_k, respectively, extremely large which would keep the filters operating in extremely favorable conditions for extended periods of time, thereby inflating performance.

To provide additional context for the range of parameter values used in this work, is used below the KF solution under stationarity to relate the parameter settings to the weight given to the observations in the KF process of optimally combining the observations, Z_k, with the model prediction, $\hat{X}_{k|k-1}$. Under stationarity, the KF solution for Eqs.(22) and (23) is given by (Schweppe 1973, Bras and Rodriguez-Iturbe 1985):

$$\Sigma_{k|k} = [U^T R^{-1} U + \Sigma^{-1}_{k|k-1}]^{-1} = \frac{1}{n/\sigma_v^2 + 1/\left(\sigma_w^2/(1-\varphi^2) + \sigma_w^2\right)} \tag{27}$$

$$\hat{X}_{k|k} = \Sigma_{k|k}(U^T R^{-1} Z_k + \Sigma^{-1}_{k|k-1}\hat{X}_{k|k-1}) = w_{obs}(1/n)\sum_{i=1}^{n} z_{k,i} + (1-w_{obs})\hat{X}_{k|k-1} \tag{28}$$

$$w_{obs} = \frac{n/\sigma_v^2}{n/\sigma_v^2 + 1/\left(\sigma_w^2/(1-\varphi^2) + \sigma_w^2\right)} \tag{29}$$

where the time indices have been dropped from the time-invariant parameters and $z_{k,i}$ denotes the i-th observation, i=1,…,n, in Z_k. In Eq.(27), the first and second terms in the denominator represent the information content in the observations, Z_k, and the model prediction, $\hat{X}_{k|k-1}$. Note that the uncertainty in $\hat{X}_{k|k}$, $\Sigma_{k|k}$, decreases as the number of observations, n, increases, the observation uncertainty, σ_v^2 decreases, the predictability of

the process, φ, increases, or the model uncertainty, σ_w^2, decreases. The rightmost column of Table 3.1 shows the range of values for w_{obs} for all cases considered based on the parameter settings in Table 3.1 under the assumption of stationarity and the bounds in Eqs.(24) through (26). One may interpret w_{obs} in Eq.(29) as the contribution of the observations relative to the model prediction in reducing the uncertainty in the filtered estimate. Table 3.1 shows that the stationary experiment allows variations of w_{obs} up to about 0.60, and that the first nonstationary experiment with only φ_{k-1} perturbed allows only a rather limited range of w_{obs} whereas all other nonstationary cases encompass effectively all possible ranges of w_{obs}. Because not all cases encompass all possible ranges of w_{obs}, it might seem that the above experiments may not be sufficiently realistic. It is easy to see in Eq.(29), however, that the reduced ranges of w_{obs} arise from the realistic ranges of the parameter settings employed. Note that φ=1, $\sigma_v^2 = 0$ or $\sigma_w^2 = 0$ would yield $w_{obs} = 1$ but none of them is achievable in reality. For all cases, the simulation horizon was set at 1,000,000 time steps which produced 1,000,000 data points for each case. For the scatter- and qqplots in the Results Section, it is only displayed the first 100,000 to limit the size of the plots. To evaluate the relative performance between CBPKF and KF, then is calculated the RMSE conditional on the true state exceeding some threshold between 0 and the largest truth, and percent reduction in conditional RMSE by CBPKF over KF. To identify the specific attributes that CBPKF improves or does not improve over KF, also was carried out mean square error (MSE) decomposition (Murphy and Winkler 1987, Nelson et al. 2010). Lastly, the accuracy of uncertainty estimates was assessed by comparing the CBPKF and KF error variances with the observed error squared in the mean sense.

Table 3.1 Parameter settings used in the synthetic experiments

Experiment type	φ_{k-1}	$\sigma_{w,k-1}$	$\sigma_{v,k}$	n	Υ_{φ}	Υ_{w}	Υ_{v}	$w_{obs}{}^{1}$
Stationary	0.5, 0.7, 0.8, 0.9, 0.95	0.1	1.5	1, 10, 20, 30	0	0	0	[0.01,0.60]
Non-stationary 1	0.7	0.1	1.5	10	0.1, 0.2, 0.4, 0.8	0	0	[0.09,0.33]
					0	0.05, 0.1, 0.15	0	[0.00,1.00]
					0	0	0.4, 0.8, 1.2, 1.6	[0.00,1.00]
Non-stationary 2					0.1, 0.8	0.01, 0.1, 0.2	0.4, 1.2	[0.00,1.00]

Chapter 4 Results

In this section, is presented the results from the three experiments listed in Table 4.1. here is presented only a limited number of figures due to space limitations, but are summarized the results for all cases considered so that the reader may gauge performance under the full range of the parameter settings shown in Table 4.1. Figure 4.1 shows the KF (in black) vs. CBPKF (in red) results for the stationary cases of φ_{k-1} of 0.8 and 0.9, respectively, while all other parameters are kept constant as indicated in the plots. Being stationary, the initial conditions have an impact only until steady-state error variance is reached. The optimized values of α are 0.54, 0.54, 0.50 and 0.44 for φ_{k-1} of 0.7, 0.8, 0.9 and 0.95 which represent predictability levels (i.e., $100\,\varphi_{k-1}^2$) of 49, 64, 81 and 90% of the variance explained, respectively. Note that, as φ_{k-1} increases, the optimized α tends to decrease, a reflection of the fact that CB in the KF estimates decreases as predictability increases. At φ_{k-1} of 0.7 (or ~50% of variance explained), the CBPKF estimates are not very much different from the KF estimates, suggesting that, at this level of predictability, CBPKF is not able to reduce CB very significantly. For larger φ_{k-1}, it is readily seen in Figure 4.1 Scatter plots of the KF (in black) and CBPK (in red) estimates vs. the truth for the stationary cases of a) φk-1=0.8 and b) φk-1=0.9, while all other parameters are kept constant (see Table 4.1). that CBPKF significantly reduces CB, particularly over the tail ends of the distribution of the truth. For φ_{k-1} of 0.7 and 0.8, α was found to be 0.54 as obtained by running CBPKF initially with α~0.6 and identifying the lower bound above which all iteratively-reduced α values lie at all time steps as described in the Methodology Section. For φ_{k-1} of 0.9 and 0.95, the lower bound for iteratively-reduced α was zero, which meant that α had to be optimized by matching the qqplot of the CBPKF estimates with the diagonal line (see Figure 4.3 QQ plots of the KF (in black) and CBPK (in red) estimates vs. the truth for the stationary case of ϕ_{k-1}=0.9, while all other parameters are kept constant (see Table

4.1). for φ_{k-1} of 0.9). The optimal α values obtained in this way for φ_{k-1} of 0.9 and 0.95 were 0.5 and 0.44, respectively. At φ_{k-1} of 0.95, the KF solutions are of very high quality and exhibit rather small CB. Accordingly, the optimal α for CBPKF is smaller than that at φ_{k-1} of 0.9. Figure 4.2 QQ plots of the KF (in black) and CBPK (in red) estimates vs. the truth for the stationary case of ϕ_{k-1}=0.8, while all other parameters are kept constant (see Table 4.1). and Figure 4.3 QQ plots of the KF (in black) and CBPK (in red) estimates vs. the truth for the stationary case of ϕ_{k-1}=0.9, while all other parameters are kept constant (see Table 4.1). show the qqplots corresponding to Figure 4.1 Scatter plots of the KF (in black) and CBPK (in red) estimates vs. the truth for the stationary cases of a) φk-1=0.8 and b) φk-1=0.9, while all other parameters are kept constant (see Table 4.1)., respectively. For φ_{k-1} of 0.7, CBPK is not able to reduce CB very significantly. With increased predictability in Figure 4.2 QQ plots of the KF (in black) and CBPK (in red) estimates vs. the truth for the stationary case of ϕ_{k-1}=0.8, while all other parameters are kept constant (see Table 4.1)., CBPKF significantly reduces CB. For φ_{k-1} of 0.9 and 0.95, CBPKF is able to effectively eliminate CB. Figure 4.4 Percent reduction in RMSE by CBPK over KF conditioned on the truth exceeding the value on the x-axis for the stationary cases of φk-1 of 0.7, 0.8 (Figs 4.1a, 4.2), 0.9 (Fig 4.1b, 4.3) and 0.95. shows the percent reduction in RMSE by CBPKF over KF conditioned on the truth exceeding the value shown on the x-axis. All results for percent reduction in RMSE presented in this paper is for truth $\geq$ 0 only. The minimum number of pairs of the estimates and the verifying observations used for calculation of percent reduction in RMSE presented in this paper is 10. At x=0 where x denotes the verifying truth, the y-axis in Figure 4.4 Percent reduction in RMSE by CBPK over KF conditioned on the truth exceeding the value on the x-axis for the stationary cases of φk-1 of 0.7, 0.8 (Figs 4.1a, 4.2), 0.9 (Fig 4.1b, 4.3) and 0.95.represents the unconditional RMSE. That the percent reduction at x=0 is negative for all cases, i.e., the KF estimates have smaller unconditional RMSE than the CBPKF

estimates, is fully expected given that KF provides the minimum error variance solution. For x > 0.2, however, CBPKF improves over KF substantially for φ_{k-1}=0.8 or larger with a maximum reduction approaching 40%.

Performance for stationary cases, however, is not a good indicator of how CBPKF may perform as a dynamic filter when the predictability of the processes modeled, accuracy of the model and/or the information content of the observations varies in time. Figure 4.5 Example scatter plot of the KF (in black) and CBPKF (in red) estimates vs. the truth when only $\sigma_{w,k-1}$ is assumed to vary in time., Figure 4.6 Example scatter plot of the KF (in black) and CBPKF (in red) estimates vs. the truth when only $\sigma_{v,k}$ is assumed to vary in time. and Figure 4.7 Example scatter plot of the KF (in black) and CBPKF (in red) estimates vs. the truth when only φ_{k-1} is assumed to vary in time show examples of the scatter plots of the KF (in black) and CBPKF (in red) estimates vs. the verifying observations when only one of the three parameters, $\sigma_{w,k-1}$, $\sigma_{v,k}$ and φ_{k-1}, respectively, is assumed to vary in time. The figures show that the impact of the variations in $\sigma_{w,k-1}$, $\sigma_{v,k}$ and φ_{k-1} to the CBPKF estimates varies from one parameter to another. Examination of similar figures for all ranges of the parameter settings in Table 4.1 indicates that the variations in $\sigma_{w,k-1}$ and $\sigma_{v,k}$ have a larger impact than those in φ_{k-1} (see Table 4.1). The figures also suggest that both KF and CBPKF benefit greatly from the filtering results at the preceding time steps when the model error and/or the observation error is very small; they produce highly accurate filtered estimates with very small error variances which effectively reinitialize the filter with very accurate initial conditions. It is interesting to see in the figures, in particular in Figure 4.6 Example scatter plot of the KF (in black) and CBPKF (in red) estimates vs. the truth when only $\sigma_{v,k}$ is assumed to vary in time., how the CBPKF estimates differ from the KF estimates; CBPKF

estimates are progressively larger and smaller than the KF estimates as the truth increases and decreases, respectively, while tipping the cluster of the estimates to align more closely with the diagonal. The resulting CBPKF estimates above and below the diagonal line show much stronger tendencies to cancel each other out than the KF estimates regardless of the magnitude of the truth, thereby reducing CB. Figure 4.8 Percent reduction in RMSE by CBPKF over KF for γ_w =0.05, 0.10 (Figure 4.5 Example scatter plot of the KF (in black) and CBPKF (in red) estimates vs. the truth when only $\sigma_{w,k-1}$ is assumed to vary in time.), 0.15 and 0.20, Figure 4.9 Percent reduction in RMSE by CBPKF over KF for γ_v =0.4 (Figure 4.6 Example scatter plot of the KF (in black) and CBPKF (in red) estimates vs. the truth when only $\sigma_{v,k}$ is assumed to vary in time.), 0.8, 1.2 and 2.4. and Figure 4.10 Percent reduction in RMSE by CBPKF over KF for γ_φ =0.1, 0.2, 0.4 and 0.8 (Figure 4.7 Example scatter plot of the KF (in black) and CBPKF (in red) estimates vs. the truth when only φ_{k-1} is assumed to vary in time). show the percent reduction in RMSE by CBPKF over KF for the scatter plots for all parameters settings of $\sigma_{w,k-1}$, $\sigma_{v,k}$ and φ_{k-1} in Table 4.1, respectively. The figures indicate that CBPKF reduces significantly to substantially conditional RMSE over KF whether the time-varying changes occur in predictability, model uncertainty or observational uncertainty, and that the percent improvement by CBPKF is the largest under the time-varying model uncertainty and the smallest under the time-varying observational uncertainty. Note also that, whereas CBPKF increases unconditional RMSE by up to about 3% under varying observational uncertainty (Figure 4.9 Percent reduction in RMSE by CBPKF over KF for γ_v =0.4 (Figure 4.6 Example scatter plot of the KF (in black) and CBPKF (in red) estimates vs. the truth when only $\sigma_{v,k}$ is assumed to vary in time.), 0.8, 1.2 and 2.4.), the increase is less than 1.5% under varying model uncertainty or predictability, and that CBPKF

is inferior to KF only over an extremely small region around the median of the true state. The above indicates that, for the filtering problems for which performance over non-median regions in the state space is important, CBPKF is clearly superior to KF.

Figure 4.11 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for first case)., Figure 4.12 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for second case). and Figure 4.13 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for third case). show selected results from the 2nd nonstationary experiment. The parameter values used are shown in the respective figures. The positive impact of CBPKF is readily seen particularly for estimation of extreme values. It is revealed here that Figure 4.1 Scatter plots of the KF (in black) and CBPK (in red) estimates vs. the truth for the stationary cases of a) φk-1=0.8 and b) φk-1=0.9, while all other parameters are kept constant (see Table 4.1)., shown in the Introduction Section to illustrate CB, are the KF estimates shown in Figure 4.12 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for second case).. For Figure 4.12 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for second case). and Figure 4.13 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for third case)., it was necessary to optimize α. Optimization was carried out by visually matching the qqplots closely to the diagonal as described above. Figure 4.14 QQ plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for first case)., Figure 4.15 QQ plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for

second case). and Figure 4.16 QQ plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for third case). show the qqplots associated with Fig 4.11, fig 4.12 and fig 4.13 respectively. In practice, the qqplots may be too irregularly shaped to readily assess closeness to the diagonal line. Various synthetic and real-world experiments carried out thus far for the CBPK family of algorithms (Brown and Seo 2010, Seo 2013, Seo et al. 2014, Kim et al. 2016) suggests that a reasonable match, in which the main body of the qqplot of the CBPKF estimates lies close to the diagonal line, generally suffices, and elaborate optimization is usually not necessary. The above strategy, however, may not work well if the state variables are highly skewed. In such cases, a somewhat smaller α may be necessary to avoid over-correcting CB which may produce excessively large variability in the CBPKF estimates (see Seo 2013 for examples for stationary cases). If computational requirements are not an issue, one may choose to explicitly optimize α under the user-chosen performance criteria. Such an effort, however, is beyond the scope of this work and is left as a future endeavor. Figure 4.17 Percent reduction in RMSE by CBPKF over KF conditional on the true state exceeding the value on the x-axis for all 12 cases in the 2nd nonstationary experiment. shows the percent reduction in RMSE by CBPKF over KF conditional on the truth exceeding the value on the x-axis for all 12 cases in the 2nd nonstationary experiment (see Table 4.1). In the figure, one may divide the 12 cases into Groups 1 (Cases 1 to 4), 2 (Cases 5 to 8) and 3 (Cases 9 to 12), from the nearest to the origin (least skillful) to the farthest (most skillful). Groups 1 through 3 are associated with $\gamma_w = 0.01,\ 0.1,\ 0.2$, respectively. It might seem counter-intuitive that adding largest perturbations to $\sigma_{w,k-1}$ (Group 3) is associated with the largest skill. This is because, with large perturbations, $\sigma^p_{w,k-1}$ often hits the lower bound (see Eq.(25)) which reinitializes the filter with very accurate state and error covariance. Note in Figure 4.17 Percent reduction in RMSE by CBPKF over KF conditional on the true state

exceeding the value on the x-axis for all 12 cases in the 2nd nonstationary experiment. that, for most cases, CBPKF is able to reduce conditional RMSE over KF by 10% or more, that, for the majority of the cases, the reduction is 20% or larger, and that the increase in unconditional RMSE by CBPKF is only about 3% or less.

Figure 4.18 MSE and MSE decomposition of the errors in the CBPK and KF estimates for a selected case (see text) in the 2nd nonstationary experiment shows the MSE and MSE decomposition of the CBPKF and KF estimates for the cases shown in Figure 4.13 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for third case). and Figure 4.16 QQ plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for third case).. MSE decomposition is based on the following identity (Murphy and Winkler, 1987, Nelson et al. 2010):

$$MSE = \frac{1}{N}\sum_{j=1}^{N}(f_j - o_j)^2 \tag{30a}$$

$$= (m_f - m_o)^2 + (\sigma_f - \sigma_o)^2 + 2\sigma_f\sigma_o(1-\rho) \tag{30b}$$

where N denotes the total number of pairs of the estimates and verifying observations, f_j and o_j denote the j-th estimate and truth, respectively, m_f and m_o denote the mean of the estimate and truth, respectively, ($_f$ and ($_o$ denote the standard deviation of the estimate and truth, respectively, and ρ denotes the correlation between the estimate and the truth. In Eq.(30b), the first and second terms measure biases in the mean and in the univariate variability of the estimate, respectively, and the third term measures the strength

of covariation between the estimate and the truth (the smaller, the stronger). Figure 4.18 MSE and MSE decomposition of the errors in the CBPK and KF estimates for a selected case (see text) in the 2nd nonstationary experiment. indicates that CBPKF significantly reduces conditional RMSE, that the reduction in RMSE is due mostly to the reduction in conditional bias in the mean, but that the CBPKF estimates are slightly more conditionally biased in standard deviation and have a slightly smaller strength of covariation. The slightly increased conditional bias in standard deviation may seem odd in that in general the CBPKF estimates represent the variability of the truth significantly better than the KF estimates in the unconditional sense. If the system is uncertain and/or the observations are not very informative, however, the filtered estimates often cannot capture the peaks and valleys in the variations of the true state. The CBPKF estimates, which generally have larger variability, hence may make excursions well below the conditioning threshold for the true state, thereby introducing additional biases in the standard deviation conditional on the threshold. The above picture, however, varies among different cases and may not be generalized. Figure 4.19 Filtered variance from CBPKF vs. that from KF for the nonstationary case shown in Figure 4.12 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for second case).. shows the filtered variance from CBPKF vs. that from KF for the nonstationary case shown in Figure 4.11 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for first case).. Also, shown for reference is the one-to-one line. Note that the CBPKF error variance is very close to the KF error variance when the latter is small, but is progressively larger than the KF error variance as the latter increases. To assess the accuracy of the error variance estimates, is shown in Fig 4.20, fig 4.21 and fig 4.22 the box-and-whisker plots, from left to right in each figure, of the absolute error of the KF estimate, the KF error standard deviation, the absolute error of the CBPKF estimate, and the CBPKF error standard deviation conditional on the true state

exceeding the 99th percentile (i.e., the largest 1%) for the three nonstationary cases shown in Figure 4.11 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for first case)., Figure 4.12 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for second case). and Figure 4.13 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for third case)., respectively. Similar plots for all ranges of the true state show little difference between KF and CBPKF and are not shown. If the error variance estimates are unbiased, one should see in Fig 4.20, fig 4.21 and fig 4.22 the mean of the absolute error of the filtered estimate match the mean of the estimated error standard deviation. Fig 4.20, fig 4.21 and fig 4.22 indicates that, for the filtering results for the largest 1% of truth, the CBPKF error variances are more accurate than the KF error variances for all three cases of Fig 4.20, fig 4.21 and fig 4.22, but that, for the less than very skillful cases of Fig 4.20, fig 4.21, both KF and CBPKF significantly underestimate error variance for the largest 1% of the events.

Table 4.1 Parameter settings used in the synthetic experiments

Experiment type	φ_{k-1}	$\sigma_{w,k-1}$	$\sigma_{v,k}$	n	Υ_{φ}	Υ_{w}	Υ_{v}	$w_{obs}{}^{1}$
Stationary	0.5, 0.7, 0.8, 0.9, 0.95	0.1	1.5	1, 10, 20, 30	0	0	0	[0.01,0.60]
Non-stationary 1	0.7	0.1	1.5	10	0.1, 0.2, 0.4, 0.8	0	0	[0.09,0.33]
					0	0.05, 0.1, 0.15	0	[0.00,1.00]
					0	0	0.4, 0.8, 1.2, 1.6	[0.00,1.00]
Non-stationary 2					0.1, 0.8	0.01, 0.1, 0.2	0.4, 1.2	[0.00,1.00]

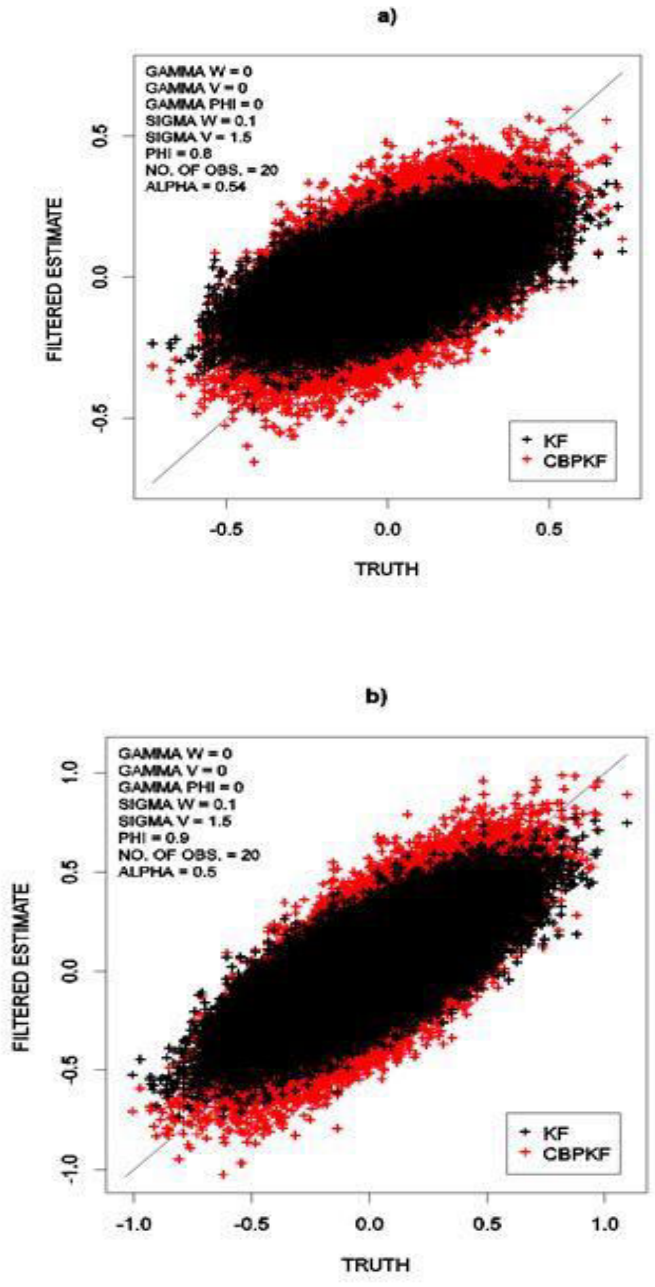

Figure 4.1 Scatter plots of the KF (in black) and CBPK (in red) estimates vs. the truth for the stationary cases of a) φk-1=0.8 and b) φk-1=0.9, while all other parameters are kept constant (see Table 4.1).

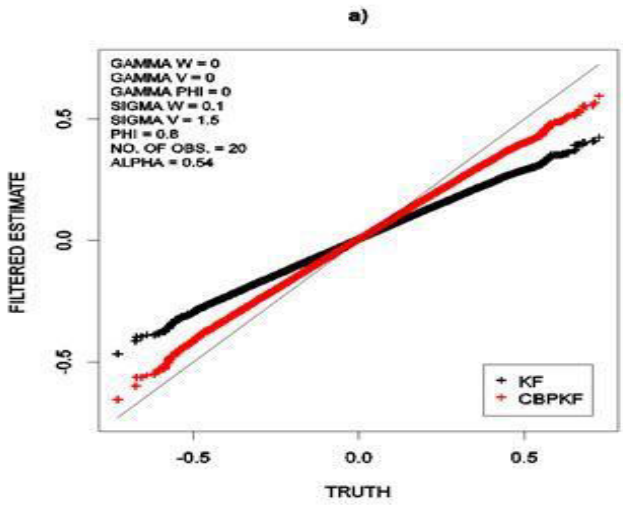

Figure 4.2 QQ plots of the KF (in black) and CBPK (in red) estimates vs. the truth for the stationary case of $\phi_{k-1}=0.8$, while all other parameters are kept constant (see Table 4.1).

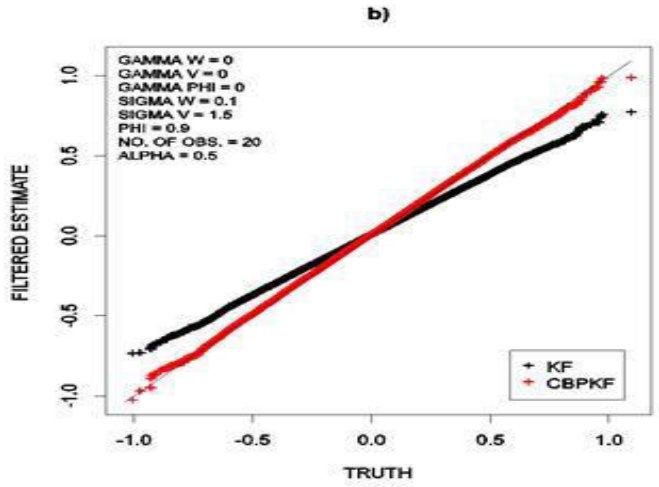

Figure 4.3 QQ plots of the KF (in black) and CBPK (in red) estimates vs. the truth for the stationary case of $\phi_{k-1}=0.9$, while all other parameters are kept constant (see Table 4.1).

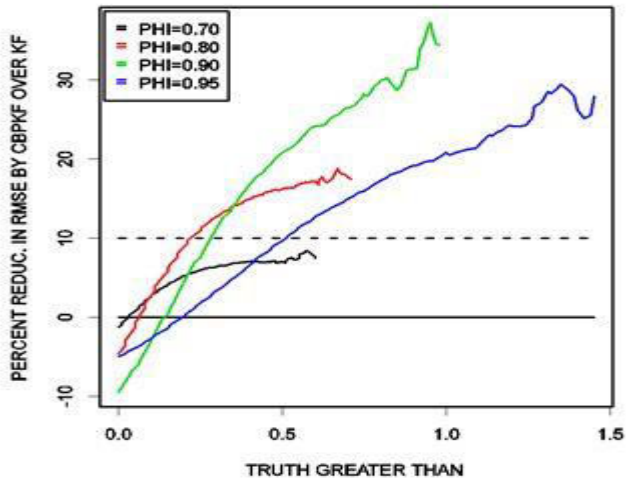

Figure 4.4 Percent reduction in RMSE by CBPK over KF conditioned on the truth exceeding the value on the x-axis for the stationary cases of φk-1 of 0.7, 0.8 (Figs 5.1a, 5.2), 0.9 (Fig 5.1b, 5.3) and 0.95.

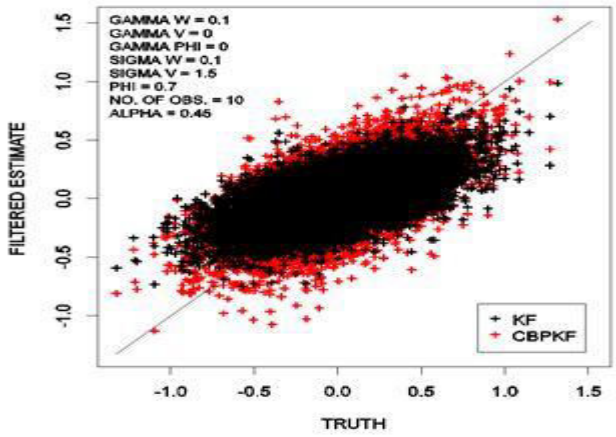

Figure 4.5 Example scatter plot of the KF (in black) and CBPKF (in red) estimates vs. the truth when only $\sigma_{w,k-1}$ is assumed to vary in time.

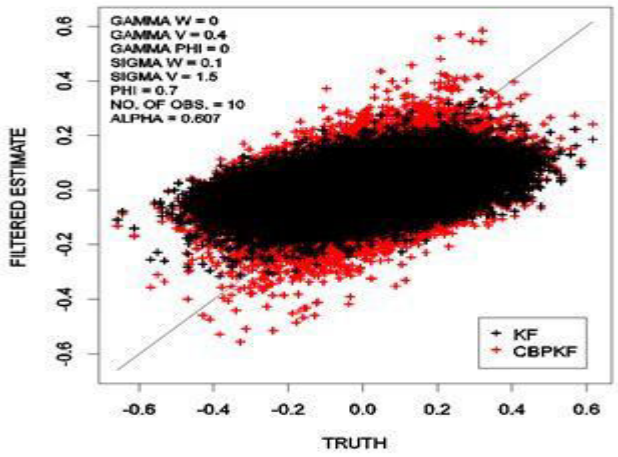

Figure 4.6 Example scatter plot of the KF (in black) and CBPKF (in red) estimates vs. the truth when only $\sigma_{v,k}$ is assumed to vary in time.

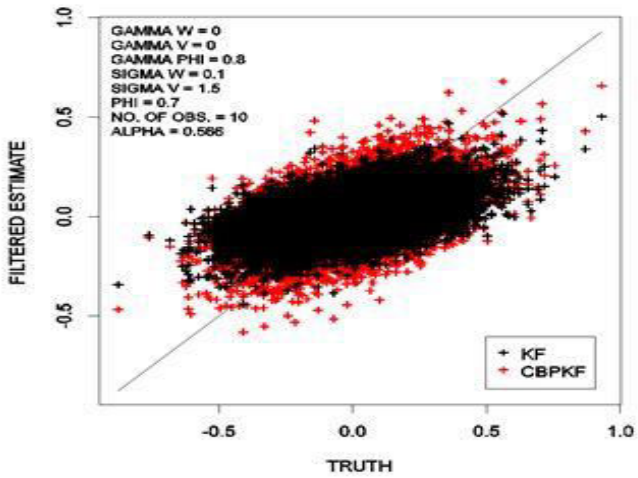

Figure 4.7 Example scatter plot of the KF (in black) and CBPKF (in red) estimates vs. the truth when only φ_{k-1} is assumed to vary in time

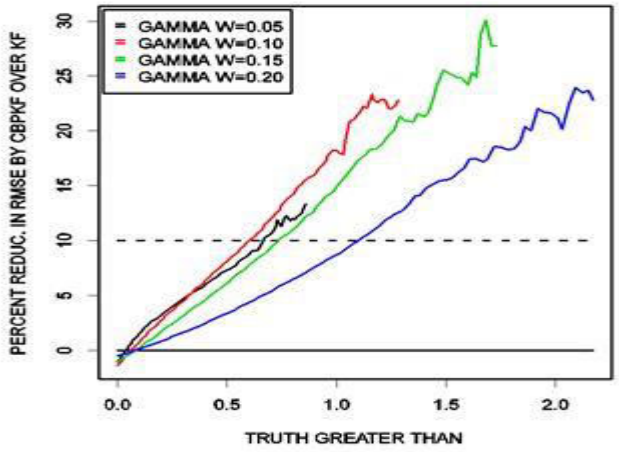

Figure 4.8 Percent reduction in RMSE by CBPKF over KF for γ_w =0.05, 0.10 (Figure 5.5), 0.15 and 0.20

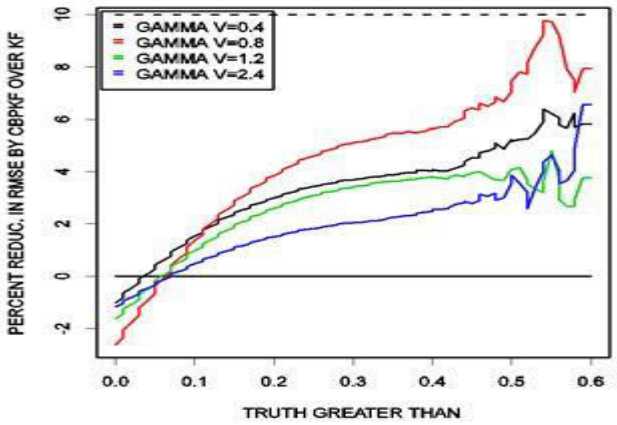

Figure 4.9 Percent reduction in RMSE by CBPKF over KF for γ_v =0.4 (Figure 5.6), 0.8, 1.2 and 2.4.

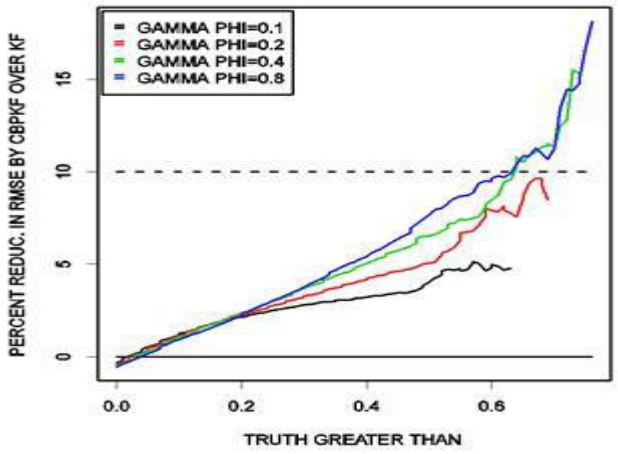

Figure 4.10 Percent reduction in RMSE by CBPKF over KF for γ_{φ} =0.1, 0.2, 0.4 and 0.8 (Figure 5.7).

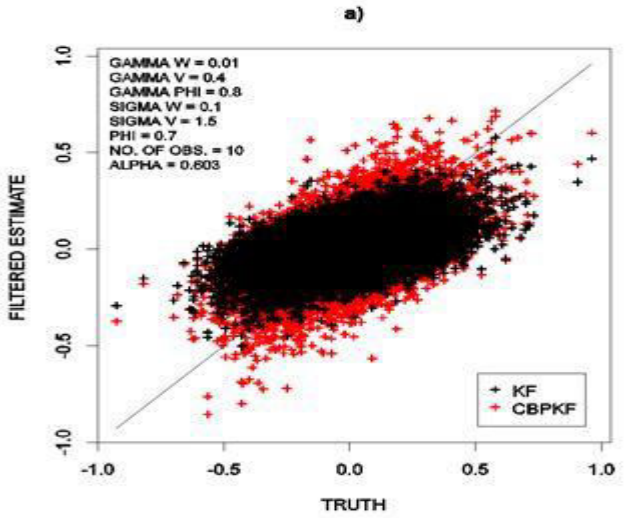

Figure 4.11 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for first case).

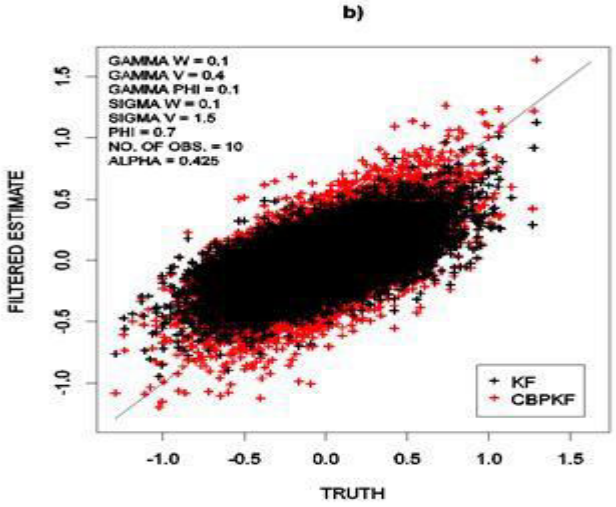

Figure 4.12 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for second case).

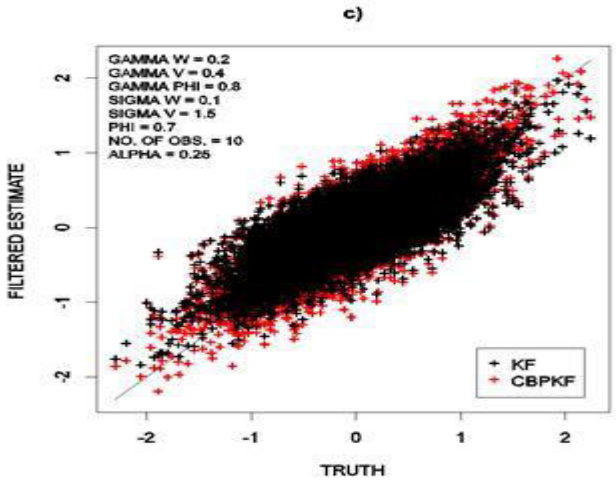

Figure 4.13 Scatter plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for third case).

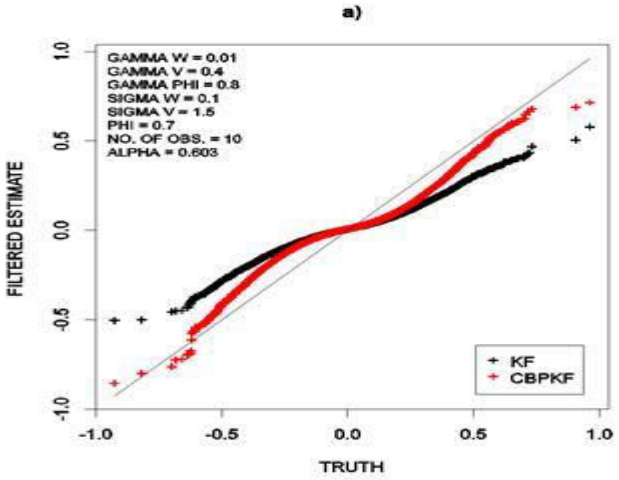

Figure 4.14 QQ plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for first case).

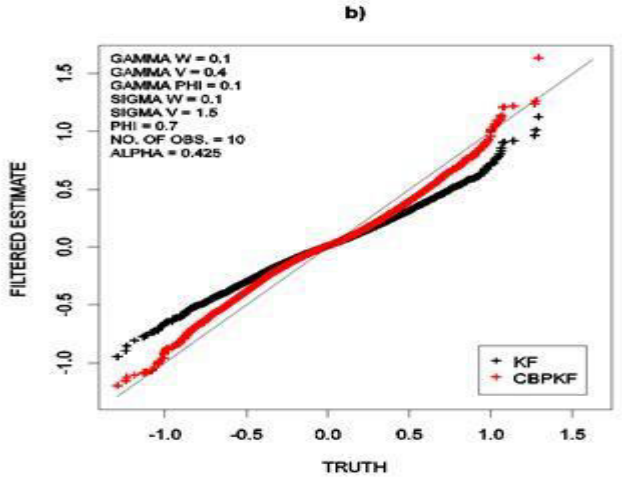

Figure 4.15 QQ plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for second case).

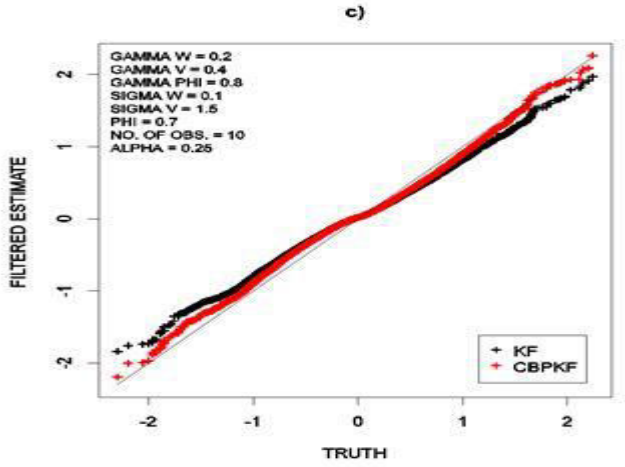

Figure 4.16 QQ plots of the CBPKF and KF estimates vs. the verifying truth for three selected cases from the 2nd nonstationary experiment (this is for third case).

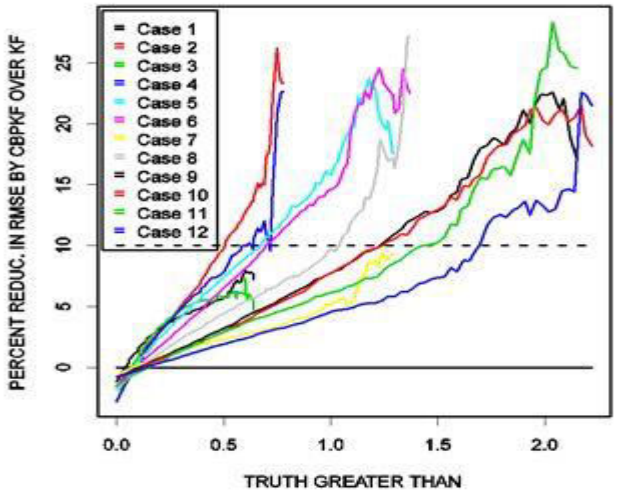

Figure 4.17 Percent reduction in RMSE by CBPKF over KF conditional on the true state exceeding the value on the x-axis for all 12 cases in the 2nd nonstationary experiment.

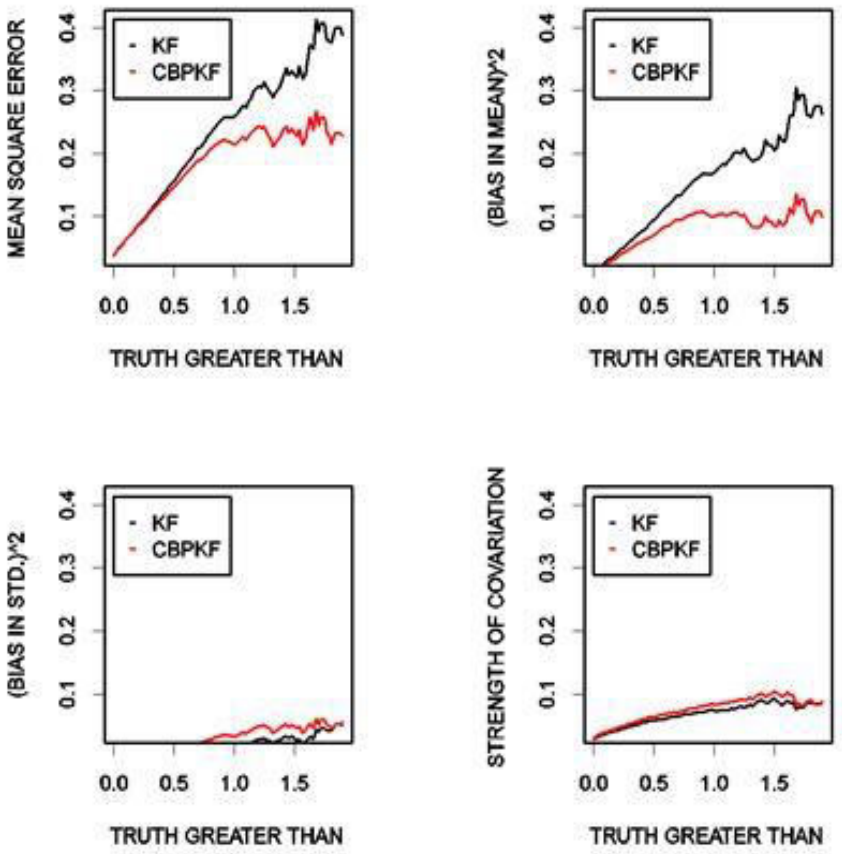

Figure 4.18 MSE and MSE decomposition of the errors in the CBPK and KF estimates for a selected case (see text) in the 2nd nonstationary experiment.

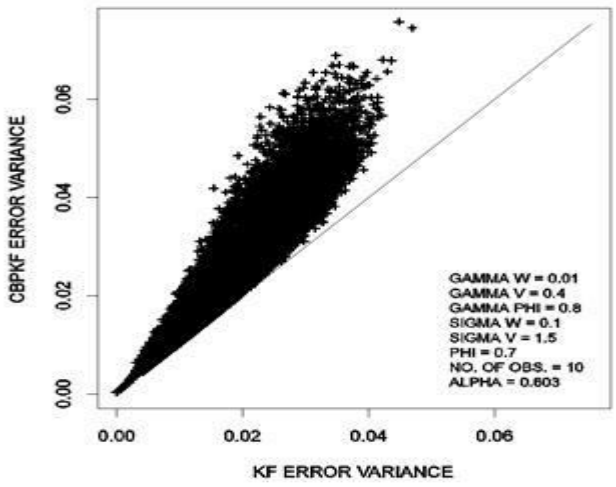

Figure 4.19 Filtered variance from CBPKF vs. that from KF for the nonstationary case shown in Figure 5.12.

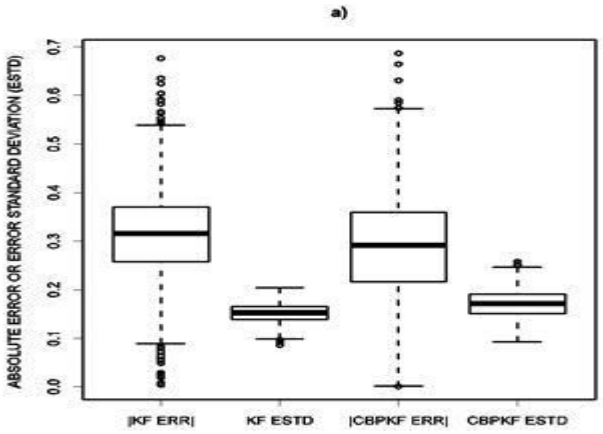

Figure 4.20 Box-and-whisker plots of the absolute error of the KF estimate (black left), KF error standard deviation (black right), absolute error of the CBPKF estimate (red left) and CBPKF error standard deviation (red right) for the case in Fig 5.11.

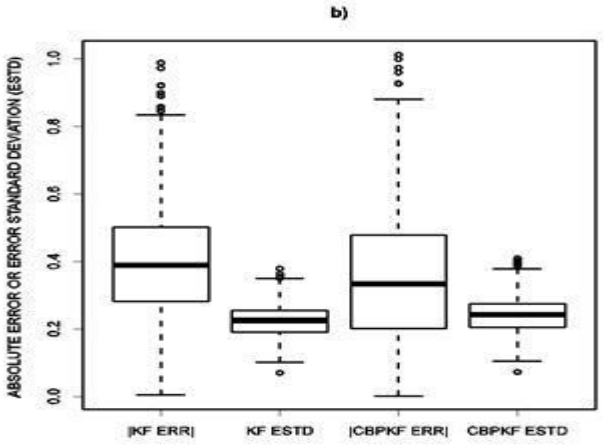

Figure 4.21 Box-and-whisker plots of the absolute error of the KF estimate (black left), KF error standard deviation (black right), absolute error of the CBPKF estimate (red left) and CBPKF error standard deviation (red right) for the case in Fig 5.12.

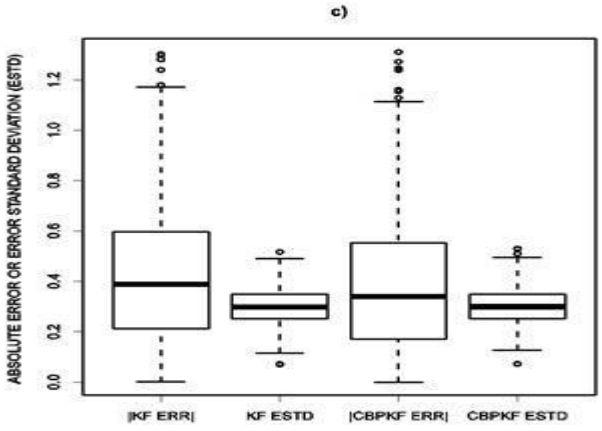

Figure 4.22 Box-and-whisker plots of the absolute error of the KF estimate (black left), KF error standard deviation (black right), absolute error of the CBPKF estimate (red left) and CBPKF error standard deviation (red right) for the case in Fig 5.13.

Chapter 5 Conclusion and Future Recommendation

Being a least squares solution, Kalman filter (KF) is subject to conditional biases (CB) that arise from the error-in-variable, or attenuation, effects. These effects occur if the model dynamics are highly uncertain, the observations have large errors and/or the system is not very predictable. In modeling and observation of environmental systems, the above effects are the norm rather than the exception. As such, KF or its variants often suffer from CB with potentially very large negative impact on estimation and prediction of extremes. In this work, CB-penalized Kalman filter is introduced, or CBPKF, for improved state estimation and prediction of extremes. CBPKF results from CB-penalized linear estimation which minimizes a weighted sum of error covariance and expectation of Type-II CB squared, $J = \Sigma_{EV} + \alpha\Sigma_{CB}$, where Σ_{EV}, Σ_{CB} and α denote the error covariance, the quadratic penalty for Type-II CB, and the weight for the latter, respectively (Seo 2013). One may consider CB-penalized linear estimation as an extension of classical Fisher estimation (Schweppe 1973) from which KF results in that for minimization of Σ_{EV}, the a priori covariance of the state vector is assumed non-informative as in Fisher estimation, but for minimization of Σ_{CB}, the prior is assumed informative. To comparatively evaluate CBPKF with KF, we designed and carried out a set of synthetic experiments for one-dimensional (1D) state estimation under the idealized conditions of normality and linearity.

The results show that CBPKF reduces root mean square error (RMSE) over KF by 10 to 20% or more over the tails of the distribution of the true state, and that, as expected, the improvement comes from reduced CB. For dynamical cases, it was found that CBPKF performs comparably to KF in the unconditional sense; CBPKF increases RMSE over all ranges of the true state only by 3% or less. The results indicate that CBPKF may be expected to significantly improve analysis and prediction of extreme states in uncertain

systems with little deterioration in unconditional performance, and should be favored over KF if improved performance over the tails of the distribution of the true state is desired. CBPKF are not without disadvantages, however. In its current form, CBPKF additionally requires inversion of (mxm) and (nxn) matrices, which KF does not, where m and n denote the number of state variables and observations, respectively. Additional research is needed to explore computationally more efficient reformulation or approximation. For maximum performance, it is necessary in CBPKF to optimize the weight, α, which requires hindcasting. If computational requirements are not an issue, one may optimize α explicitly under the performance criteria desired by the user. In this work, is used a very simple procedure for iterative reduction of α. Additional research is needed to develop efficient procedures for iterative reduction and optimization of α. It may also be possible to avoid hindcasting and optimization by specifying α based on (near) real-time assessment of CB using real-time observations for which additional research is needed. In the current formulation of CBPKF, α is assumed to be a scaler. Additional work is needed to generalize the formulation to allow α in a matrix form. The evaluation of CBPKF in this work was limited only to 1D synthetic experiments under the idealized conditions of linearity and normality. Multi-dimensional synthetic and real-world experiments are needed to assess performance for a wide range of higher-dimensional problems and under real-world conditions. Lastly, as with KF, CBPKF may be cast into extended and ensemble formulations.

APPENDIX A

Derivation of Conditional Bias-Penalized Fisher-Like Linear Estimator

Here is derived the Fisher-like CB-penalized linear estimator in the context of CBPKF (see Appendix B) with additional details that were not presented in Seo (2013). The estimator sought is of the form, $X^* = WZ$ where X^* denotes the (mx1) vector of the estimated states, W denotes the (mx(n+m)) weight matrix and Z denotes the ((n+m)x1) observation vector. The particular choice of the dimensionality of Z is to relate to CBPKF. Here we assumed the following linear observation equation:

$$Z=HX+V \tag{A1}$$

where X denotes the (mx1) vector of the true state with $E[X]=M_X$ and $Cov[X,X^T]=\Psi_{XX}$, H denotes the ((n+m)xm) linear observation equation matrix, and V denotes the ((n+m)x1) zero-mean measurement error vector with Cov[V,V]=R. Assuming Cov[X,V]=0, the Bayesian estimator for X, or X^*, is given by (Schweppe 1973):

$$X^* = M_X + W(Z - M_Z) \tag{A2}$$

where W denotes the (mx(n+m)) weight matrix that minimizes the error covariance. The error covariance matrix for X^*, $\Sigma_{EV} = E_{X,X^*}[(X - X^*)(X - X^*)^T]$, where the variables subscripted denote the random variables on which the expectations operate, is given by:

$$\Sigma_{EV} = \Psi_{XX} - WH\Psi_{XX} - \Psi_{XX}H^T W^T + W(H\Psi_{XX}H^T + R)W^T \quad \text{(A3)}$$

The quadratic penalty due to Type-II CB, $X - E_{X^*}[X^* \mid X]$, is given by:

$$\Sigma_{CB} = E_X[(X - E_{X^*}[X^* \mid X])(X - E_{X^*}[X^* \mid X])^T] \quad \text{(A4)}$$

Using (A2), we may rewrite the CB in (A4) as:

$$X - E[X^* \mid X] = (X - M_X) - WE[(Z - HM_X) \mid X] \quad \text{(A5)}$$

We model $E[(Z - HM_X) \mid X]$ in (A5) using the Bayesian estimator again as:

$$E[(Z - HM_X) \mid X] = \Psi_{ZX}\Psi_{XX}^{-1}(X - M_X) \quad \text{(A6)}$$

where $\Psi_{ZX} = Cov(Z, X)$. With the above, we may now write Σ_{CB} as:

$$\Sigma_{CB} = \Psi_{XX} - WA - A^T W^T + WCW^T \quad \text{(A7)}$$

where $C \equiv \Psi_{ZX}\Psi_{XX}^{-1}\Psi_{XZ}$ and $A \equiv \Psi_{ZX}$. In CB-penalized estimation, we minimize $\Sigma = \Sigma_{EV} + \alpha\Sigma_{CB}$ where α is some positive weighting coefficient:

$$\Sigma = (1+\alpha)\Psi_{XX} - W(\alpha A + H\Psi_{XX}) - (\Psi_{XX}H^T + \alpha A^T)W^T + W(H\Psi_{XX}H^T + R + \alpha C)W^T \quad \text{(A8)}$$

The weighting coefficient, α, may be made into an (mxm) matrix if it is necessary to give different weights to the CB penalty for different state variables. In this work, it is

assumed for simplicity that the weights are the same for all state variables. Differentiating Σ with respect to W and setting it to 0, we have:

$$W = (\Psi_{XX} H^T + \alpha A^T)[H\Psi_{XX} H^T + R + \alpha C]^{-1} \quad \text{(A9)}$$

Replacing W in (A8) with (A9), we have for the estimation variance for the CB-penalized estimate:

$$\Sigma = (1+\alpha)\Psi_{XX} - \Psi_{XX}\hat{H}^T[\hat{H}\Psi_{XX}\hat{H}^T + \Lambda]^{-1}\hat{H}\Psi_{XX} \quad \text{(A10)}$$

where

$$\hat{H}^T = H^T + \alpha\Psi_{XX}^{-1}\Psi_{XZ} \quad \text{(A11)}$$

$$\Lambda = R + \alpha(1-\alpha)\Psi_{ZX}\Psi_{XX}^{-1}\Psi_{XZ} - \alpha H\Psi_{XZ} - \alpha\Psi_{ZX}H^T \quad \text{(A12)}$$

Using the matrix inversion lemma, we may rewrite Σ in (A10) as:

$$\Sigma = \alpha\Psi_{XX} + [\hat{H}^T\Lambda^{-1}\hat{H} + \Psi_{XX}^{-1}]^{-1} \quad \text{(A13)}$$

Replacing W in (A2) with (A9) and after some matrix manipulations, we have for X^*:

$$X^* = [\hat{H}^T\Lambda^{-1}\hat{H} + \Psi_{XX}^{-1}]^{-1}\{\hat{H}^T\Lambda^{-1}Z + \Psi_{XX}^{-1}M_X\} + \Delta \quad \text{(A14)}$$

where

$$\Delta = \Psi_{XX}\hat{H}^T[\hat{H}\Psi_{XX}\hat{H}^T + \Lambda]^{-1}\Psi_{ZX}\Psi_{XX}{}^{-1}M_X \qquad \text{(A15)}$$

To render the Bayesian solution to a Fisher solution, we let Ψ_{XX}^{-1} vanish in (A12) and (A14) in the brackets only, which are associated exclusively with the penalty for error covariance in (A3), and arrive at the following intermediate solution for CB-penalized Fisher-like linear estimation:

$$\Sigma = B[\hat{H}^T\Lambda^{-1}\hat{H}]^{-1} \qquad \text{(A16)}$$

$$X^* = [\hat{H}^T\Lambda^{-1}\hat{H}]^{-1}\hat{H}^T\Lambda^{-1}Z + \Delta \qquad \text{(A17)}$$

where

$$\mathrm{B} = \alpha\Psi_{XX}\hat{H}^T\Lambda^{-1}\hat{H} + I \qquad \text{(A18)}$$

To obtain the estimator of the form, $X^* = WZ$, we impose the unbiasedness condition, $E[X^*] = X$, or equivalently:

$$WH = I \qquad \text{(A19)}$$

It is readily seen in (A17) that the above unbiasedness condition is satisfied by replacing $[\hat{H}^T\Lambda^{-1}\hat{H}]^{-1}$ with $[\hat{H}^T\Lambda^{-1}H]^{-1}$ and dropping Δ. The Fisher-like solution for CB-penalized linear estimation is hence given by:

$$\Sigma = \mathrm{B}[\hat{H}^T\Lambda^{-1}H]^{-1} \qquad \text{(A20)}$$

$$X^* = [\hat{H}^T \Lambda^{-1} H]^{-1} \hat{H}^T \Lambda^{-1} Z \quad \text{(A21)}$$

The development above indicates that CB-penalized Fisher-like linear estimation is analogous to Fisher estimation in which the observation matrix, H, and the measurement error covariance matrix, R, are modified by the a priori knowledge of Ψ_{XX} and Ψ_{XZ}, and the estimation variance is scaled by a factor of B. Because (A20) and (A21) are not based on explicit constrained minimization, they may not represent the optimal solution in the least squares sense. It can be shown, however, that for m=1 and α=0.5 with perfect observations (A21) is identical to the conditional bias-penalized kriging (CBPK) estimate which is based on explicit constrained minimization (Seo 2013, Seo et al. 2014, Kim et al. 2016), and that (A20) converges to the CBPK estimation variance as $n \to \infty$.

APPENDIX B

Derivation of Conditional Bias-Penalized Kalman filter (CBPKF)

Here we derive CBPKF from the Fisher-like solution of Appendix A for estimation of the (mx1) true state, X_k, using the (nx1) observation, Z_k, and (mx1) model prediction, $\hat{X}_{k|k-1}$, and their (nxn) and (mxm) error covariances, $R_k = E[V_k V_k^T]$, and $\Sigma_{k|k-1}$, respectively. The observation equation is given by $Z_k = H_k X_k + V_k$ where it is assumed that the true state, X_k, is independent of the measurement error, V_k, or the model prediction error, $\hat{X}_{k|k-1} - X_k$, so that we may write, e.g., $\Psi_{Z_k X_k} = H_k \Psi_{X_k X_k}$. Decomposing the structure matrix H in (A1) into the first submatrix that relates the observations to the true states, $H_1=H_k$, and the second submatrix that relates the model-predicted states to the true

states, H_2=I, we have for the mx(n+m) modified structure matrix, $\hat{H}^T = [\hat{H}_1^T \quad \hat{H}_2^T]$ in (A11):

$$\hat{H}_1^T = H_k^T + \alpha \Psi_{X_k X_k}^{-1} \Psi_{X_k Z_k} = (1+\alpha) H_k^T \tag{B1}$$

$$\hat{H}_2^T = I + \alpha \Psi_{X_k X_k}^{-1} \Psi_{X_k X_{k|k-1}} = (1+\alpha) I \tag{B2}$$

where the (mxn) and (mxm) covariance matrices, $\Psi_{X_k Z_k}$ and $\Psi_{X_k X_{k|k-1}}$, denote $Cov[X_k, Z_k^T]$ and $Cov[X_k, X_{k|k-1}^T]$, respectively. An obvious choice for Ψ_{XX} in practice is $\Sigma_{k|k-1}$ obtained from propagating $\Sigma_{k-1|k-1}$ using the dynamical model with model errors as appropriate. With $\Psi_{XX} = \Sigma_{k|k-1}$, the (nxn), (nxm), (mxn) and (mxm) submatrices, Λ_{11}, Λ_{12}, Λ_{21} and Λ_{22}, of the (n+m)x(n+m) revised error covariance matrix, Λ in (A12), are given by:

$$\Lambda_{11} = R_k - \alpha(\alpha+1) H_k \Sigma_{k|k-1} H_k^T \tag{B3}$$

$$\Lambda_{12} = -\alpha(\alpha+1) H_k \Sigma_{k|k-1} \tag{B4}$$

$$\Lambda_{21} = -\alpha(\alpha+1) \Sigma_{k|k-1} H_k^T \tag{B5}$$

$$\Lambda_{22} = \{1 - \alpha(\alpha+1)\} \Sigma_{k|k-1} \tag{B6}$$

CBPKF requires that Λ_{11} and Λ_{22} to be positive semidefinite (see Appendix C) which yields the following constraint for α from (B3) and (B6):

$$0 \le \alpha \le \min\left\{\sqrt{\frac{Tr[R_k]}{Tr[H_k \Sigma_{k|k-1} H_k^T]} + \frac{1}{4}} - \frac{1}{2}, \frac{\sqrt{5}-1}{2}\right\} \qquad \text{(B7)}$$

where Tr[ ] denotes the trace of the symmetric matrix bracketed and the second term in the upper bound is ~0.618. Note in (B7) that, if the states are perfectly observed so that we have $Tr[R_k] = 0$, α is reduced to zero and hence CBPKF becomes KF. Similarly, if the model forecast is diffuse so that we have $\Sigma_{k|k-1} \to \infty$, CBPKF is again reduced to KF. The (mx(n+m)) non-normalized weight matrix, $\varpi = [\varpi_1 \quad \varpi_2] = \hat{H}^T \Lambda^{-1}$ in (A21), where ϖ_1 and ϖ_2 are the (mxn) and (mxm) non-normalized weight submatrices for Z_k and $\hat{X}_{k|k-1}$, respectively, may be evaluated by:

$$\varpi_1 = \hat{H}_1^T \Gamma_{11} + \hat{H}_2^T \Gamma_{21} = (1+\alpha)[H_k^T \Gamma_{11} + \Gamma_{21}] \qquad \text{(B8)}$$

$$\omega_2 = \hat{H}_1^T \Gamma_{12} + \hat{H}_2^T \Gamma_{22} = (1+\alpha)[H_k^T \Gamma_{12} + \Gamma_{22}] \qquad \text{(B9)}$$

In the above, the inverse of the (n+m)x(n+m) modified error covariance matrix Γ is given by:

$$\Lambda^{-1} = \Gamma = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix} = \begin{bmatrix} \Lambda_{11}^{-1} + \Lambda_{11}^{-1}\Lambda_{12}\Gamma_{22}\Lambda_{21}\Lambda_{11}^{-1} & -\Lambda_{11}^{-1}\Lambda_{12}\Gamma_{22} \\ -\Gamma_{22}\Lambda_{21}\Lambda_{11}^{-1} & \Gamma_{22} \end{bmatrix} \qquad \text{(B10)}$$

where

$$\Gamma_{22}^{-1} = \Lambda_{22} - \Lambda_{21}\Lambda_{11}^{-1}\Lambda_{12} \qquad \text{(B11)}$$

The (mxm) matrix, $\hat{H}^T \Lambda^{-1} H$ in (A20) and (A21), is given by:

$$\hat{H}^T \Lambda^{-1} H = \varpi_1 H_k + \varpi_2 = (1+\alpha)[(H_k^T \Gamma_{11} + \Gamma_{21}) H_k + \Gamma_{21} H_k + \Gamma_{22}] \quad \text{(B12)}$$

Positive semidefiniteness of $\hat{H}^T \Lambda^{-1} H$ and nonnegativity of the CBPK gain (see B20) require:

$$Tr[(H_k^T \Gamma_{11} + \Gamma_{21}) H_k] \geq 0 \quad \text{(B13)}$$

$$Tr[\Gamma_{22} + H_k^T \Gamma_{12}] \geq 0 \quad \text{(B14)}$$

From (A20), we then have for the filtered variance:

$$\Sigma_{k|k} = B[\hat{H}^T \Lambda^{-1} H]^{-1} = (1+\alpha)\alpha \Sigma_{k|k-1} + \{(1+\alpha)[(H_k^T \Gamma_{11} + \Gamma_{21}) H_k + \Gamma_{21} H_k + \Gamma_{22}]\}^{-1} \quad \text{(B15)}$$

where

$$\mathrm{B} = \alpha \Sigma_{k|k-1} \hat{H}^T \Lambda^{-1} \hat{H} + I = \alpha \Sigma_{k|k-1} (1+\alpha)^2 [(H_k^T \Gamma_{11} + \Gamma_{21}) H_k + \Gamma_{21} H_k + \Gamma_{22}] + I \quad \text{(B16)}$$

In (B15), $\Sigma_{k|k-1} - \Sigma_{k|k}$ is positive semidefinite if the following holds:

$$(1+\alpha)(\alpha^2 + \alpha - 1) \leq Tr([(H_k^T \Gamma_{11} + \Gamma_{21}) H_k + \Gamma_{21} H_k + \Gamma_{22}]^{-1}) / Tr(\Sigma_{k|k-1}) \quad \text{(B17)}$$

Noting that the minimum for the right-hand side of (B20) is zero, we may reduce (B20) to $(\alpha^2 + \alpha - 1) \leq 0$ which is identical to the positive semidefiniteness conditions for (B6). As such, (B7) suffices.

From (A21), we have for the filtered estimate:

$$\hat{X}_{k|k} = [\hat{H}^T \Lambda^{-1} H]^{-1} [\varpi_1 Z_k + \varpi_2 \hat{X}_{k|k-1}]$$
$$= [(H_k^T \Gamma_{11} + \Gamma_{21}) H_k + \Gamma_{21} H_k + \Gamma_{22}]^{-1} \{ [H_k^T \Gamma_{11} + \Gamma_{21}] Z_k + [H_k^T \Gamma_{12} + \Gamma_{22}] \hat{X}_{k|k-1} \} \quad \text{(B18)}$$

It can be easily shown using the matrix inversion lemma that (B18) can be rewritten in the more familiar form:

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k [Z_k - H_k \hat{X}_{k|k-1}] \quad \text{(B19)}$$

In the above, the CBPK gain, K_k, is given by:

$$K_k = [(H_k^T \Gamma_{11} + \Gamma_{21}) H_k + \Gamma_{21} H_k + \Gamma_{22}]^{-1} [H_k^T \Gamma_{11} + \Gamma_{21}] \quad \text{(B20)}$$

APPENDIX C

Alternative Form of Conditional Bias-Penalized Kalman filter (CBPKF)

Here we express CBPKF in an alternative form for direct comparison with KF by factorizing Λ^{-1} in (A20) and (A21) as follows:

$$\Lambda^{-1} = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix}^{-1} = \left[\begin{pmatrix} I & 0 \\ \Lambda_{21}\Lambda_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} \Lambda_{11} & 0 \\ 0 & \Lambda_{22} - \Lambda_{21}\Lambda_{11}^{-1}\Lambda_{12} \end{pmatrix} \begin{pmatrix} I & \Lambda_{11}^{-1}\Lambda_{12} \\ 0 & I \end{pmatrix} \right]^{-1}$$

$$= \begin{pmatrix} I & -\Lambda_{11}^{-1}\Lambda_{12} \\ 0 & I \end{pmatrix} \begin{pmatrix} \Lambda_{11}^{-1} & 0 \\ 0 & (\Lambda_{22} - \Lambda_{21}\Lambda_{11}^{-1}\Lambda_{12})^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -\Lambda_{21}\Lambda_{11}^{-1} & I \end{pmatrix}$$

(C1)

In the above, we could have factorized Λ such that the first and third matrices are lower and upper matrices, respectively, which would have yielded an alternative but equivalent expression. With (C1), it can be easily shown that the CBPKF error covariance and estimate, $\Sigma_{k|k}$ and $\hat{X}_{k|k}$, respectively, are given by:

$$\Sigma_{k|k} = \frac{1}{1+\alpha}\tilde{\Sigma}_{k|k} + \alpha(1+\alpha)\Sigma_{k|k-1}$$

(C2)

$$\hat{X}_{k|k} = \tilde{X}_{k|k} - \tilde{\Sigma}_{k|k}\tilde{\Sigma}_{k|k-1}^{-1}\Lambda_{12}\Lambda_{11}^{-1}Z_k$$

(C3)

In the above, $\tilde{\Sigma}_{k|k}$, $\tilde{X}_{k|k}$ and $\tilde{\Sigma}_{k|k-1}$ denote the "pseudo" updated error covariance, updated states and forecast error covariance, respectively, defined solely to render the CBPKF solution to look like the familiar KF solution below:

$$\tilde{\Sigma}_{k|k} = [H_k^T\Lambda_{11}^{-1}H_k + \tilde{\Sigma}_{k|k-1}^{-1}]^{-1}$$

(C4)

$$\tilde{X}_{k|k} = \hat{X}_{k|k-1} + \tilde{K}_k[Z_k - H_k^T \hat{X}_{k|k-1}]$$

(C5)

$$\tilde{K}_k = \tilde{\Sigma}_{k|k-1} H_k^T [\Lambda_{11} + H_k \tilde{\Sigma}_{k|k-1} H_k^T]^{-1}$$

(C6)

In the above, the forecast error covariance is given by:

$$\tilde{\Sigma}_{k|k-1}^{-1} = (I - H_k^T \Lambda_{11}^{-1} \Lambda_{12})(\Lambda_{22} - \Lambda_{21} \Lambda_{11}^{-1} \Lambda_{12})^{-1} = (I - H_k^T \Lambda_{11}^{-1} \Lambda_{12}) \Gamma_{22} = \Gamma_{22} + H_k^T \Gamma_{12}$$

(C7)

Positivity of the CBPK gain in (C6) requires that Λ_{11} in (B3), Λ_{22} (from the alternative expression for (C1)) in (B6) and $\tilde{\Sigma}_{k|k-1}^{-1}$ in (C7), respectively, are positive semidefinite. The third condition is already identified in (B14). The first two conditions are used in (B7). The alternative development described above indicates that CBPKF is a combination of KF with modified measurement and model forecast error covariances as shown in (C4), (C6) and (C7), and adjustment to the resulting KF solution according to (C2) and (C3).

References:

Bras RL, Rodriguez-Iturbe I (1985) Random Functions and Hydrology. Addison-Wesley, 559 pp

Brown JD, Seo DJ (2010) A nonparametric post-processor for bias correcting ensemble forecasts of hydrometeorological and hydrologic variables. J. Hydrometeorol. 11(3): 642-665

Carroll RJ, Ruppert D, Stefanski LA (1995) Measurement Error in Nonlinear Models. Chapman and Hall, 305 pp

Ciach GJ, Morrissey ML, Krajewski WF (2000) Conditional bias in radar rainfall estimation. J Appl Meteor 39: 1941-1946

Clark MP, Rupp DE, Woods RA, Zheng X, Ibbitt RP, Slater AG, Schmidt J, Uddstrom MJ (2008) Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. Adv. Water Resour. 31(10): 1309-1324

Cosby BJ (1984) Dissolved oxygen dynamics of a stream: model discrimination and estimation of parameter variability using an extended Kalman filter. Water Sci. Technol. 16(5-7): 561-569

Durbin J, Koopman SJ (2001) Time series analysis by state space methods. Oxford University Press, Oxford

Eknes M, Evensen G (2002) An ensemble Kalman filter with a 1-D marine ecosystem model. J. Mar. Syst. 36(1): 75-100

Ennola K, Sarvala J, Devai G (1998) Modelling zooplankton population dynamics with the extended Kalman filtering technique. Ecol. Model. 110(2): 135-149

Evensen G (2003) The ensemble Kalman filter: Theoretical formulation and practical implementation. Ocean Dyn. 53(4): 343-367

Evensen G (2009) Data assimilation: the ensemble Kalman filter. Springer Science & Business Media

Fuller WA (1987) Measurement Error Models. John Wiley and Sons, 440 pp

Guo HC, Liu L, Huang GH (2003) A stochastic water quality forecasting system for the Yiluo River. J. Environ. Inform. 1(2): 18-32

Houtekamer PL, Mitchell HL (1998) Data assimilation using an ensemble Kalman filter technique. Mon. Weather Rev. 126(3): 796-811

Huang J, Gao J, Liu J, and Zhang Y (2013) State and parameter update of a hydrodynamic-phytoplankton model using ensemble Kalman filter. Ecol. Model. 263: 81-91

Jazwinski AH (1970) Stochastic Processes and Filtering Theory. Academic Press, New York, 376 pp

Jolliffe IT, Stephenson DB (eds) (2003) Forecast Verification: A Practitioner's Guide in Atmospheric Science. John Wiley and Sons, Chichester, 254 pp

Kalman RE (1960) A new approach to linear filtering and prediction problems. J. Fluids Eng. 82(1): 35-45

Kim B, Seo D-J, Noh S, Prat OP, Nelson BR (2016) Improving Multisensor Estimation of Heavy-to-Extreme Precipitation via Conditional Bias-Penalized Optimal Estimation. J. Hydrol. http://dx.doi.org/10.1016/j.jhydrol.2016.10.052

Kim K, Park M, Min JH, Ryu I, Kang MR, Park LJ (2014) Simulation of algal bloom dynamics in a river with the ensemble Kalman filter. J. Hydrol. 519: 2810-2821

Kim S, Seo D-J, Riazi H, Shin C (2014) Improving water quality forecasting via data assimilation–Application of maximum likelihood ensemble filter to HSPF. J. Hydrol. 519: 2797-2809

Kitagawa G, Gersch W (1996) Smoothness Priors Analysis of Time Series. Springer-Verlag New York, doi:10.1007/978-1-4612-0761-0, ISBN: 978-0-387-94819-5 (Print) 978-1-4612-0761-0 (Online)
Komma J, Blöschl G, Reszler C (2008) Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting. J. Hydrol. 357(3): 228-242
Krzysztofowicz R (1999) Bayesian theory of probabilistic forecasting via deterministic hydrologic model. Water Resour. Res. 35(8): 2739–2750
Mao JQ, Lee JH, Choi KW (2009) The extended Kalman filter for forecast of algal bloom dynamics. Water Res. 43(17): 4214-4224
Moradkhani H, Sorooshian S, Gupta HV, Houser PR (2005) Dual state–parameter estimation of hydrological models using ensemble Kalman filter. Adv. Water Resour. 28(2): 135-147
Murphy AH, Winkler RL (1987) A general framework for forecast verification. Mon. Weather Rev. 115: 1330-1338
Neal JC, Atkinson PM, Hutton CW (2007) Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements. J. Hydrol. 336(3): 401-415
Nelson B, Seo D-J, Kim D (2010) Multisensor Precipitation Reanalysis. J. Hydrometeorol. 11(3): 666-682
Pastres R, Ciavatta S, Solidoro C (2003) The Extended Kalman Filter (EKF) as a tool for the assimilation of high frequency water quality data. Ecol. Model. 170(2): 227-235
Rafieeinasab A, Norouzi A, Seo D-J, Nelson B (2015) Improving High-Resolution Quantitative Precipitation Estimation via Fusion of Multiple Radar-Based Precipitation Products. J. Hydrol. 531 Part 2: 320–336
Rafieeinasab A, Seo D-J, Lee H, Kim S (2014) Comparative evaluation of maximum likelihood ensemble filter and ensemble Kalman filter for real-time assimilation of streamflow data into operational hydrologic models. J. Hydrol. 519: 2663-2675
Schweppe FC (1973) Uncertain dynamic systems. Prentice Hall, 563 pp
Seber GAF (1989) Nonlinear Regression. John Wiley and Sons, 768 pp
Seo D-J (2013) Conditional bias-penalized kriging (CBPK). Stoch. Environ. Res. Risk Assess. 27(1): 43-58
Seo D-J, Herr H, Schaake JC (2006) A statistical post-processor for accounting of hydrologic uncertainty in short-range ensemble streamflow prediction. Hydrol. Earth Syst. Sci. Discuss. 3: 1987-2035
Seo D-J, Siddique R, Zhang Y, Kim D (2014) Improving real-time estimation of heavy-to-extreme precipitation using rain gauge data via conditional bias-penalized optimal estimation. J. Hydrol. 519: 1824-1835
Twigt D, Tyrrell D, Rego Lima J, and Troost T (2011) Water Quality Forecasting Systems: Advanced Warning of Harmful Events and Dissemination of Public Alerts. In Proceedings of the 8th International ISCRAM Conference–Lisbon, Portugal
Weerts AH, El Serafy GY (2006) Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models. Water Resour. Res. 42(9): W09403
Whitehead PG, Hornberger GM (1984) Modelling algal behaviour in the River Thames. Water Res. 18(8): 945-953
Xie X, Zhang D (2010) Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. Adv. Water Resour. 33(6): 678-690
Xue P, Chen C, Beardsley RC (2012) Observing system simulation experiments of dissolved oxygen monitoring in Massachusetts Bay. J. Geophys. Res.: Oceans (1978–2012), 117(C5)

Zhang H, Pu Z (2010) Beating the uncertainties: ensemble forecasting and ensemble-based data assimilation in modern numerical weather prediction. Adv. Meteorol. http://dx.doi.org/10.1155/2010/432160
Zingone A, Enevoldsen HO (2000) The diversity of harmful algal blooms: a challenge for science and management. Ocean Coast. Management 43(8): 725-748
Siddique, R. (2014). Improving Analysis Of Heavy To Extreme Precipitation Using Conditional Bias-penalized Optimal Estimation.

Biographical Information

Miah Mohammad Saifuddin is born on 3rd January, 1991. He had a knack for Mathematics and Science from a very early age. When he was asked that what he wants to be when he grows up, he always replied that he wanted to be an engineer. His grandfather was a Civil Engineer thus he was fascinated by the tools that were used by his grandfather. He saw that not anyone could use those engineering staffs without knowing what actually is to be done with those. He understood that not everyone can be an engineer and engineers are smart. Miah got himself a chance in one of the best Engineering schools in his country for earning a Bachelor of Science degree in the field of Civil Engineering and became even more impressed with how things work out in the world of Engineers. He understood that implementation along with innovation is a key factor in Engineering world. For getting himself more into the innovation, he got himself admitted for a Master of Science degree in Civil Engineering at the U.S. He was happy when he saw that the opportunities are more wide in U.S for Civil Engineers. He loved water resources and thus wanted to pursue his MS with a concentration in Water Resources. He was fortunate to get excellent faculty members who helped him towards pursuing such honorable degree. He completed his MS degree and shall be looking for some real life industrial experience in the field of Civil Engineering. He also has in mind to do a PhD after he gains enough industrial experience.