IMPROVING LAND DATA ASSIMILATION PERFORMANCE WITH A WATER BUDGET CONSTRAINT

by

M. Tugrul Yilmaz
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
In Partial fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Earth Systems and Geoinformation Sciences

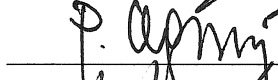Committee:

_____ Dr. Paul R. Houser, Dissertation Director

_____ Dr. Timothy DelSole, Committee Member

_____ Dr. Paul Dirmeyer, Committee Member

_____ Dr. Zafer Boybeyi, Committee Member

_____ Dr. Peggy Agouris, Department Chair

_____ Dr. Richard Diecchio, Associate
Dean for Academic and Student Affairs,
College of Science

_____ Dr. Vikas Chandhoke, Dean,
College of Science

Date: __10/19/2010__ Fall Semester 2010
George Mason University
Fairfax, VA

Improving Land Data Assimilation Performance with
a Water Budget Constraint

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at George Mason University

By

M. Tugrul Yilmaz
Master of Science
Vrije University Amsterdam, 2005
Bachelor of Science
Middle East Technical University, 2003

Director: Dr. Paul R. Houser, Professor
Department of Earth Systems and Geoinformation Sciences

Fall Semester 2010
George Mason University
Fairfax, VA

Dedication

I dedicate this dissertation to my mum Gulseren for her unconditional love, support, and dedication to me and my siblings.

Acknowledgments

Table of Contents

List of Tables

List of Figures

Abstract

IMPROVING LAND DATA ASSIMILATION PERFORMANCE WITH
A WATER BUDGET CONSTRAINT

M. Tugrul Yilmaz, PhD

George Mason University, 2010

Dissertation Director: Dr. Paul R. Houser

A weak constraint solution was introduced to reduce the water budget imbalance that appears in land data assimilation as a result of state updates. Constrained Kalman Filter results were shown to be identical in single- or two-stages solutions for Ensemble Kalman Filter (EnKF) whereas constrained Ensemble Transform Kalman Filter (ETKF) single- and two-stage solutions form two different square root solutions. Weakly Constrained Ensemble Kalman Filter (WCEnKF) and Weakly Constrained Ensemble Transform Kalman Filter (WCETKF) were evaluated for 3-hourly and daily update frequencies with soil moisture only, or soil moisture and soil temperature assimilated together. Not perturbed observations in EnKF was revisited. Both constrained and standard solutions were performed for not perturbed observations and without the constraint anomalies. Sensitivity of the constraint error variance is analyzed by comparing the results from objectively estimating and by using tuned values. Simulations were performed using the Noah Land Surface Model (LSM) over Oklahoma, USA, using synthetic observations.

State errors of constrained and unconstrained solutions were found to be similar; neither type had significantly smaller errors for most experiments. Constrained filters had smaller water balance residuals than unconstrained standard filters for all tested scenarios. The water balance residual of the ETKF and EnKF were similar for both 3-hourly and daily update experiments. The majority of the total column water change for daily updated filters resulted from the assimilation update. Not perturbing the observations and not using the constraint anomalies affected the state prediction skill only slightly where the residuals are significantly reduced when compared to the standard filters. Tuned constraint variances gave similar performance with objective variance estimation from the ensemble for WCEnKF but the tuned variances were better than objective estimation for WCETKF.

Chapter 1: Introduction

Data assimilation is a technique for optimally combining observations and model forecasts into a single best estimate of the state, while taking into account the accuracy of the two independent estimates. Data assimilation systems are optimum only in so far as certain underlying assumptions are fulfilled, namely that the forecast model is perfect, observations and forecasts are unbiased, observation errors are independent of the state, and all the distributions are Gaussian. However, available modeling and observing systems do not satisfy all these assumptions. In practice, the model is not perfect, observations and forecasts are biased, and the error covariances that are needed to solve the optimal solution are unknown.

In land surface, data assimilation methods have used satellite-, air-, and ground-based observations to improve estimates of soil moisture, soil skin temperature, snow water equivalent, and snow cover estimates {Houser et al. (1998); Lakshmi (2000); Crow and Reichle (2008); Reichle et al. (2008); De Lannoy et al. (2010)}. However, special problems occur when conserved quantities are assimilated. For instance, assimilation of hydrological observations (e.g. soil moisture) may improve estimates of hydrological variables, but generally degrade the water balance because the analysis increments do not conserve water since they are compensating for system biases or errors. Even if the dynamical model conserves water, the state update generally creates a water budget imbalance. If the degree of water imbalance is excessive, then it is reasonable to question whether an alternative data assimilation system should be employed, particularly one that reduces or removes the imbalance of water.

Skillful water estimation is important for hydrologists since it determines the location of the stored water on land, eg. for streamflow, agricultural, and water management applications [Alsdorf et al. (2007)]. Accurate water budgets are important for estimating runoff, because runoff is calculated as a residual of other water balance terms. Skillful estimations of the water and energy cycles are also important for developing and validating hydrological models [Wei et al. (2010)]; in particular in model skill assessment, facilitating model parameterization developments, calibrating model parameterizations, better understanding the hydrological processes, assessing the role of land over climate predictability [DelSole et al. (2009) and Dirmeyer (2003)], and predicting future changes. In fact obtaining a better energy and water balance has been focus of many scientific experiments, particularly World Climate Research Program (WCRP) Global Energy and Water-Cycle Experiment (GEWEX). It has been emphasized that the land-atmosphere interaction and the land water storage still remains as the future issues to be addressed [WCRP JSC Report (2010)], which are primarily linked with water and energy cycles.

Focusing on the energy and water balance variations analysis on a global scale plays a key role in change related studies and in determining the predictability of the climate, which are part of the primary goals of GEWEX. Land surface water regulates the climate through the memory of the land which is primarily associated with the soil moisture [Dirmeyer et al. (2009)]. Soil moisture controls the energy and the water exchange between land and atmosphere, hence an accurate energy and water balance at the land surface would translate into accurate climate predictions and perhaps better understanding the change through better soil moisture estimation. In fact obtaining a better soil moisture to infer the water storage in the land has been one of the primary goals of Global Soil Wetness Project [GSWP, Dirmeyer et al. (1999)] and GSWP2 [Dirmeyer et al. (2006)]. Furthermore, obtaining a "closed" water and

energy balance estimate on a continental scale is critical for many hydrology applications; in fact this was the primary goal of the Continental-scale International Project (GCIP)[Roads et al. (2003)] which is another contributing project to GEWEX.

However, obtaining a balanced or closed water budget is not trivial: observations are not temporally and spatially adequate to obtain useful closure information, or to estimate their sampling uncertainties. Models have the potential to completely cover the region of interest temporally and spatially, but, they may suffer from inaccurate parameterizations. Hence, correct closure information may not be obtained from models alone. Data assimilation combines both observations and models by taking into account their error structures; however, as described above, their corrections may lead to water budget imbalance due to the state updates that correct system bias or error.

Pan and Wood (2006) proposed a constraint in land-data assimilation to ensure that the data assimilation system conserved water. They have derived a two-stage constrained Kalman Filter in which the first stage is a traditional Kalman Filter and the second stage imposes a water balance constraint in an optimal manner. They have also included precipitation, evaporation, and runoff in their state vector and thereby used the filter to update these quantities. Pan and Wood (2006) showed that the constrained Kalman Filter gave estimates not far from the unconstrained filter, except that the water imbalance was removed.

In this study, it is shown that the constrained Kalman Filter can lead to very unrealistic state estimates. Specifically, if individual terms in the water budget have large errors, then imposing the budget to balance exactly requires these errors to be distributed among the state variables. If these errors are sufficiently large, then the budget constraint will cause some state variables to deviate beyond their natural range. There are at least two ways for dealing with large errors in the budget terms:

3

include forcing terms in the data assimilation procedure, as showed by Pan and Wood (2006), or to impose a weakly constrained in which the water budget derived from observed components is assumed to hold only approximately.

The purpose of this dissertation is to present a weakly constrained data assimilation system in which a water budget constraint is imposed on the conventional data assimilation systems while taking into account the uncertainties of the water balance elements. Weakly constrained solutions were introduced for both the Ensemble Kalman Filter (EnKF) and Ensemble Transform Kalman Filter (ETKF). The weakly constrained Kalman Filter is applied to idealized experiments and its performance was compared to the unconstrained Kalman Filter. It is shown that the weakly constrained solution improves the water budget imbalance without increasing the errors of the hydrological variables. In this study, not perturbation of the observations idea in standard filters is also revisited. Again using idealized setup perturbed observations and constraint anomalies were not used in the constrained solutions. It was shown that with little or no prediction skill loss, the water budget residuals were further reduced. In this study, the variance of the constraint was estimated through an objective way rather than through a tunable parameter. It was shown that for some filters tuning can be avoided through the objective estimation method where for some filters tuning resulted in better performance.

This study is organized as follows: chapter 2 briefly reviews the theoretical background of standard filters used in current land and atmospheric data assimilation; chapter 3 introduces the water budget constraint and its implementation; chapter 4 introduces new methods that further improve the constrained filter; chapter 5 summarizes the major outcomes of the results; and appendix section presents the detailed derivation of the constrained filter.

4

Chapter 2: Unconstrained Standard Filters

2.1 State Estimation in Least Squares Sense

Discovery of least squares has been a reason for a continuous debate in the scientific community. Gauss applied least squares in 1801 to find the location of the asteroid Ceres (that had been discovered by Giuseppe Piazzi but lost after tracking Ceres for only couple of days due to his illness). Later, Legendre printed the first publication on least squares [Legendre (1806)], where Gauss did not publish his solution until 1809 [Gauss (1963)] but claimed to have discovered this estimation theory in 1795. Today the debate over who discovered least squares first continues, although some historians attribute the discovery to Gauss [Sorenson (1970)].

After the discovery, least squares has gained its current form by Kalman (1960), where this estimation theory can be used to describe most (but not all) assimilation algorithms that have been used so far [Talagrand (1997)]. "One advantage of studying assimilation theory in the perspective of general estimation theory is that it forces one to explicitly formulate hypothesis which are necessarily made in one way or other "[Talagrand (1997)]. Hence, in this section state estimation in least squares sense will be briefly reviewed, before the full derivation of the assimilation methods introduced in later sections.

2.1.1 Least Squares Method

One of the most common applications of least squares is to fit a line to a group of points, where a linear relationship is expected. In this example, only a single realization (observed state estimate) is available for a given time step or space where the minimization is seeked. Let us assume there is a linear relation between a predictand (y) and a predictor (x) through some prediction parameters (a) with and observation error (ϵ) as

$$y_1 = a_1 + x_1 a_2 + \epsilon_1$$

$$y_2 = a_1 + x_2 a_2 + \epsilon_2$$

$$\vdots$$

$$y_N = a_1 + x_t a_2 + \epsilon_t$$

which can be conveniently grouped as

$$Y = XA + \epsilon$$

where $Y(N \text{ x } 1)$ is a vector holding the best estimates in time or space, $X(N \text{ x } 2)$ is a matrix holding the observations, $A(2 \text{ x } 1)$ is a vector holding the parameters, and $\epsilon(N \text{ x } 1)$ is a vector holding the noise in the estimation.

The noise ϵ is assumed to be independently and Gaussian distributed with 0 mean and σ_ϵ^2 variance. It is the final goal of Least Squares to minimize a squared estimate (namely a "cost function"), which in this case is the variance of this noise. This cost

function is defined as

$$J = \sigma_\epsilon^2 = E[\epsilon^T \epsilon] = E[(Y - X^T A)^T (Y - X^T A)]$$

$$J = E[Y^T Y - Y^T X^T A - A^T X Y + A^T X X^T A].$$

This cost function can be minimized by setting the first derivative to 0 and by solving. Using the matrix identities given in Appendix section (B.1)

$$\frac{\partial J}{\partial a} = -2XY + 2XX^T A = 0,$$

where the second derivation of this cost function $(2XX^T > 0)$ shows the solution is a minima. Hence solution for A that minimizes the above described cost function is

$$A = (XX^T)^{-1} XY.$$

2.1.2   Weighted Least Squares

Above derivation assumes equal weights for all residuals. If there is a reason we believe that the uncertainties of the predictions are not equal in time or space, then a new cost function can be formed by giving different weight to each residual. However, this would require advance knowledge of the weights before the least squares fit. The new cost function can be formed as

$$J = E[\epsilon^T R^{-1} \epsilon] = E[(Y - X^T A)^T R^{-1} (Y - X^T A)]$$

$$J = E[Y^T R^{-1} Y - Y^T R^{-1} X^T A - A^T X R^{-1} Y^T + A^T X R^{-1} X^T A], \qquad (2.1)$$

where R is a symmetric matrix holding the weights (error variances) for each prediction in time or space. Taking the first derivative of (2.1) and setting it to 0,

$$\frac{\partial J}{\partial a} = -2XR^{-1}Y + 2XR^{-1}X^T A = 0.$$

Hence solution for A can be found as

$$A = (XR^{-1}X^T)^{-1}XR^{-1}Y.$$

2.1.3 Recursive Least Squares

In above derived least squares examples, there is only a single realization of the state together with its uncertainty available; namely observations. Given the presence of another realization that is independent from the observations (could be from model simulations), least squares can be still used to optimally combine both information and obtain the best estimate. In above solutions, all data are given at once, prior to the solution, where the minimization of the cost function is performed only once to obtain a set of regression parameters. Rather than solving for all time steps together, above solutions could have been identically obtained by solving the problem recursively for each time step. On the other hand, with the presence of two realization, where the model simulations at any time step is dependent on the analysis of the previous time, a general solution for the best estimate requires a recursive estimation rather than solving for all time-steps at once. In such state estimation, the initial conditions may greatly effect the predictions, which is a property of nonlinear models (not necessarily chaotic). Perturbations in the initial conditions may also self-amplify in time, and could dominate over the signal in the initial conditions (especially in chaotic models), which is particularly true for models like in atmospheric science

8

applications [Talagrand (1997)].

In a recursive filtering framework, the cost function can be minimized separately for each time step where the uncertainties in the initial state can be also included in the state estimation. In fact, recursive estimation of the state is one of the advantages of Kalman's solution over the least squares solutions of Gauss (1963) and Legendre (1806) [Sorenson (1970)]. Following Talagrand (1997), below illustrates a least squares solution for the estimation of a state from independent sources of realizations with their relative initial state uncertainties.

Assume we have two different temperature realizations (T_1 and T_2) with their errors (ϵ_1 and ϵ_2) and error variances (σ_1 and σ_2).

$$T_1 = T_t + \epsilon_1$$

$$T_2 = T_t + \epsilon_2$$

where T_t is the true temperature. For simplicity, it is assumed that the errors have zero mean and are uncorrelated, $E[\epsilon_1 \epsilon_2] = 0$. It is desired to obtain an analysis (T_a) from these two estimates with error variance (σ_a) that is smaller than both σ_1 and σ_2. This analysis can be a linear combination of the two temperature estimate as

$$T_a = a_1 T_1 + a_2 T_2, \tag{2.2}$$

where a_1 and a_2 are the relative weights of T_1 and T_2 respectively and the error of the analysis is defined as $\epsilon_a = T_a - T_t$. In order for the analysis to be unbiased ($E[T_a - T_t] = 0$), it is required that the weights sum to 1 ($a_1 + a_2 = 1$). The error

variance of the analysis becomes

$$\sigma_a^2 = E[\epsilon_a^T \epsilon_a] = E[(T_a - T_t)^2]$$

$$= E[((a_1 T_1 + a_2 T_2) - (a_1 + a_2)T_t)^2]$$

$$= E[(a_1(T_1 - T_t) + a_2(T_2 - T_t))^2]$$

$$= E[(a_1\epsilon_1 + a_2\epsilon_2)^2]$$

$$= E[(a_1\epsilon_1)^2] + E[(a_2\epsilon_2)^2] + E[2a_1 a_2 \epsilon_1 \epsilon_2].$$

Since errors are uncorrelated, the last term vanishes and the analysis error variance becomes

$$\sigma_a^2 = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2.$$

The desired cost function can be chosen to minimize σ_a^2. The solution can be found by minimizing this cost function relative to a_1:

$$J = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 \tag{2.3}$$

$$J = \sigma_a^2 = a_1^2 \sigma_1^2 + (1 - a_1)^2 \sigma_2^2$$

$$J = a_1^2 \sigma_1^2 + (a_1^2 + 1 - 2a_1)\sigma_2^2$$

$$\frac{\partial J}{a_1} = 2a_1(\sigma_1^2 + \sigma_2^2) - 2\sigma_2^2 = 0$$

$$a_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \quad \& \quad a_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}.$$

The solution is intuitive that the "weights" are inversely proportional to the uncertainty of the realization itself. Above weights a_1 and a_2 provide the solution for the analysis itself, where the uncertainty of the analysis can be found after substituting these weights back into (2.2),

$$\sigma_a^2 = \left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right)^2 \sigma_1^2 + \left(\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right)^2 \sigma_2^2$$

$$\frac{1}{\sigma_a^2} = \frac{1}{\sigma_a^1} + \frac{1}{\sigma_a^2}. \tag{2.4}$$

It should be also emphasized that if the uncertainties in one of the estimate (say in model) is very large when compared to the other estimate (say observations), then the solution is reduced to the weighted least squares state estimate described above, where the uncertainty of the state is equivalent to the uncertainty of the more accurate estimate (in this case observations).

2.2 Kalman Filter

2.2.1 Standard Kalman Filter

Complete derivations of both KF and ETKF solutions can be found in numerous papers; here, these derivations are reviewed once more to emphasize the differences between the unconstrained and the constrained solutions.

The objective of data assimilation is to "optimally" estimate a set of quantities using all available observations, prior knowledge of the underlying model structure, and associated error statistics. In Kalman Filtering, the goal is to solve for the best state estimate and its uncertainty, where this best estimate and its error covariance information is propagated in time. Similar to (2.3), this optimal estimate can be

11

estimated by minimizing a cost function [Lorenc (1986)]

$$J = (o - Hx)^T R^{-1}(o - Hx) + (x - x_f)^T P_f^{-1}(x - x_f), \qquad (2.5)$$

where lower case letters represent vectors, capital letters represent matrices; o is the observations; x is the best estimate of the state to be found; H is a linear observation operator that maps the model state to observation space; superscript T is the transpose operator; R is the observation error covariance matrix; x_f is the prior estimate of the model state, usually obtained from a model forecast; and P_f is the model background error covariance matrix. The first term on the right side of (2.5) measures the distance between the state and the observations, and the second term measures the distance between the state and the background. Both distances are measured using a norm based on the appropriate error covariance matrix. The vector x that minimizes (2.5) gives the best estimate according to maximum likelihood or Bayesian derivation methods [Maybeck (1982)]. The minimization of (2.5) can be obtained by setting the derivative of J w.r.t. x equal to 0 and solving

$$\frac{\partial J}{\partial x} = 2(H^T R^{-1} H + P_f^{-1})x - 2(H^T R^{-1} o + P_f^{-1} x_f) = 0.$$

Hence, the best estimate of the state is

$$x_a = (H^T R^{-1} H + P_f^{-1})^{-1}(H^T R^{-1} o + P_f^{-1} x_f),$$

where this solution is shown in Appendix (B.6) to be equivalent to,

$$x_a = x_f + K(o - Hx_f) \qquad (2.6)$$

12

and

$$K = P_f H^T (H P_f H^T + R)^{-1} \qquad (2.7)$$

where x_a is the updated state vector and K is the Kalman gain matrix. Consistent with the above solution, the analysis error covariance can be found from the inverse of the second derivation of the cost function [shown in Appendix (B.5); and Lorenc (1986)],

$$P_a = (H^T R^{-1} H + P_f^{-1})^{-1}$$

which can be rewritten using the Sherman-Morrison-Woodbury formula as

$$P_a = P_f - P_f H^T (H P_f H^T + R)^{-1} H P_f. \qquad (2.8)$$

Following Ide et al. (1997), temporal evolution of the best estimate of the analysis is performed as

$$x_{f_t} = M x_{a_{t-1}} + \epsilon_{m_{t-1}} \qquad (2.9)$$

where M is a linear operator that performs the temporal evolution of the state and ϵ_t is the temporal evolution error. For atmospheric and hydrologic systems, M can be a linearized numerical model, and $\epsilon_{m_{t-1}}$ is the prediction error of the model. Error

covariance of this state estimate can be found

$$P_{f_t} = E[(x_t - x_{f_t})(x_t - x_{f_t})^T]$$

$$P_{f_t} = E[(Mx_{t_{t-1}} - Mx_{a_{t-1}} + \epsilon_{m_{t-1}})(Mx_{t_{t-1}} - Mx_{a_{t-1}} + \epsilon_{m_{t-1}})^T]$$

$$P_{f_t} = E[(M(x_{t_{t-1}} - x_{a_{t-1}}) + \epsilon_{m_{t-1}})(M(x_{t_{t-1}} - x_{a_{t-1}}) + \epsilon_{m_{t-1}})^T]$$

$$P_{f_t} = MPa_{t-1}M^T + E[\epsilon_{m_{t-1}}\epsilon_{m_{t-1}}^T] - E[M(x_{t_{t-1}} - x_{a_{t-1}})\epsilon_{t-1}^T] + E[\epsilon_{m_{t-1}}(x_{t_{t-1}} - x_{a_{t-1}})^T M^T].$$

Assuming the best estimate and the model prediction errors are independent, the expectation operators of the cross terms in the last line would vanish.

$$P_{f_t} = MPa_{t-1}M^T + Q_{m_{t-1}} \tag{2.10}$$

where $Q_{m_{t-1}}$ represent the unpredictable terms in the model prediction at time $t-1$. First term in (2.10) represents the error in the current time step which is a result of the propagation of the past analysis error, and the second term represents the modeling integration error particular to the time interval that the prediction is done. In general the second modeling error term is neglected with the perfect model assumption.

Multiplication of one vector with M is equivalent to running the model from time $t-1$ to t. Hence, (2.10) requires model integration twice the number of state variables. This temporal evolution of the forecast error can be computationally expensive, even prohibitive, for atmospheric or oceanographic data assimilation systems. As a remedy, to avoid the computational cost, forecast error covariances were obtained by inflating the analysis error covariances with constant numbers [constant inflation factor of 1.5 for 6-hour forecast, Talagrand (1997)] were used in the past operational numerical weather predictions.

2.2.2 Extended Kalman Filter

Standard Kalman Filter solution requires the linear operator M (2.9) to evolve the temporal evolution of the state error covariance. Since Kalman Filter is optimum for linear systems, for non-linear models (like atmospheric, ocean, and land models) linearization of the temporal transition model is required.

In Extended Kalman Filter (EKF), a linearized version of the nonlinear model is used to propagate the error statistics of the state [Miller et al. (1994)]. Taking into account the modeling errors, forecast and the true model predictions can be formulated as

$$x_{f_t} = M_{t-1}(x_{a_{t-1}})$$

where the true state is an unbiased estimate of the model with random modeling errors,

$$x_{t_t} = M_{t-1}(x_{t_{t-1}}) + \epsilon_{m_{t-1}}.$$

Subtracting the first line from the second

$$x_{t_t} - x_{f_t} = M_{t-1}(x_{t_{t-1}}) - M_{t-1}(x_{a_{t-1}}) + \epsilon_{m_{t-1}}.$$

Using the Taylor expansion of r.h.s. of the above equation [reminder for the reader $f(x) = f(\hat{x}) + (x - \hat{x})f'(\hat{x}) + (x - \hat{x})^2 f''(\hat{x}) + ..$ where $\hat{x}$ is the expected value of x]

$$x_{t_t} - x_{f_t} = (x_{t_{t-1}} - x_{a_{t-1}})M'(x_{a_{t-1}}) + \frac{1}{2}(x_{t_{t-1}} - x_{a_{t-1}})^2 M''(x_{a_{t-1}}) +$$

$$\frac{1}{6}(x_{t_{t-1}} - x_{a_{t-1}})^3 M'''(x_{a_{t-1}}) + \frac{1}{24}(x_{t_{t-1}} - x_{a_{t-1}})^4 M''''(x_{a_{t-1}}) + \cdots + \epsilon_{m_{t-1}}.$$

Taking the square and the expected value of both sides

$$P_{f_t} = P_{a_{t-1}}(M'_{t-1})^2 + E[(\epsilon_{a_{t-1}})^3]M'_{t-1}M''_{t-1}$$

$$+ E[(\epsilon_{a_{t-1}})^4][\frac{1}{3}M'_{t-1}M''_{t-1} + \frac{1}{4}M''_{t-1}M''_{t-1}] + \cdots + Q_{m_{t-1}}$$

where $\epsilon_{a_{t-1}}$ is the error of the analysis and M'_{t-1} is the first order derivative of the analysis at time $t-1$. Discarding the third and higher order moments assuming their contributions are negligible [Evensen (2009)], above simplifies to the scaler form of (2.10). Similarly the error covariance evolution for matrices can be shown to be

$$P_{f_t} = [DM_{t-1}]P_{a_{t-1}}[DM_{t-1}]^T + Q_{m_{t-1}}$$

where DM holds the partial derivatives of M w.r.t. the analysis state at time $t-1$. At the same time, the best estimate of the state, Kalman gain, and the analysis error covariance of EKF for time t can be found with the same standard Kalman Filter solution equations as (2.6),(2.7), and (2.8) respectively.

2.2.3 Ensemble Kalman Filter

In typical geophysical data assimilation, the KF is prohibitively expensive. Moreover, the background error covariance P_f is often unavailable due to its large dimension and/or the underlying model is nonlinear. Above described closure scheme of EKF neglects the higher order derivatives, which results in unbounded error variance growth due to the linearization of the nonlinear model [Evensen (1992)]. To circumvent these problems, Evensen (1994) introduced the EnKF, whereby ensembles of realizations are created by Monte Carlo methods and carry the error covariance information.

Evensen (1994) proposed updating the individual ensemble members using

$$x_{ai} = x_{fi} + K(o - Hx_{fi})$$

where an i-index is included to identify the ensemble member. It proves convenient to collect the ensemble members into a single matrix as

$$X_f = \frac{1}{\sqrt{(N-1)}}[x_{f1} - \mu_f, x_{f2} - \mu_f, \ldots, x_{fN} - \mu_f].$$

where μ_f denotes the ensemble mean state vector, N is the ensemble size, and similarly for the update X_a. In this notation, the analysis anomaly and the best estimate of the state update equations become

$$X_a = X_f + K(O - HX_f) \tag{2.11}$$

$$\mu_a = \mu_f + K(o - H\mu_f) \tag{2.12}$$

where $P_f = X_f X_f^T$ is substituted in (2.7) and (2.8); and O is a matrix of perturbed observations in which each column is of the form $o + \epsilon_i$, where ϵ_i is drawn from a normal distribution with 0 mean and covariance R [Burgers et al. (1998)].

2.2.4   Ensemble Transform Kalman Filter

Bishop et al. (2001), Anderson (2001), and Whitaker and Hamill (2002) proposed alternative ensemble filtering method that avoided perturbed observations. These filters were shown to belong to a single family of filters called square root filters [Tippett et al. (2003)]. Just as a square root is not unique due to an ambiguity in

sign, square root filters are not unique due an ambiguity in a unitary transformation. Bishop et al. (2001) derived the analysis error covariance matrix shown in (2.8) as

$$P_a = X_f X_f^T - X_f X_f^T H^T (H X_f X_f^T H^T + R)^{-1} H X_f X_f^T$$

$$P_a = X_f \big[I - X_f^T H^T (H X_f X_f^T H^T + R)^{-1} H X_f \big] X_f^T$$

using the inverse transformation of Sherman-Morrison-Woodbury formula, above equation becomes

$$P_a = X_f \big[(X_f^T H^T * R^{-1} * H X_f + I)^{-1} \big] X_f^T$$

$$P_a = X_f D X_f^T \tag{2.13}$$

where
$$D = (I + X_f^T H^T R^{-1} H X_f)^{-1} \tag{2.14}$$

and (2.13) is also consistent with $P_a = X_a X_a^T$.

The square root of D can be derived from the eigenvectors of $X_f^T H^T R^{-1} H X_f$. Specifically, if the eigenvector decomposition of this matrix is expressed as

$$X_f^T H^T R^{-1} H X_f = U S U^T$$

where U is unitary and S is a real positive diagonal matrix, then the most general square root of D is

$$A = U (I + S)^{-1/2} V^T \tag{2.15}$$

where V is <u>any</u> unitary matrix and $AA^T = D$. This expression allows us to write the

updated analysis anomaly matrix as

$$X_a = X_f A \tag{2.16}$$

It is also shown in Appendix (B.7) that the Kalman Gain can be obtained once the eigenvalue decomposition is performed

$$K = X_f D X_f^T H^T R^{-1}.$$

It should be recognized that the square root matrix A depends on the choice of V^T. In contrast, the matrices P_a, D, and K are independent of V^T and hence unique. Choosing $V = U$ makes the square root matrix A symmetric. Ott et al. (2004) show that the quadratic form $(X_a - X_f)^T P_a^{-1}(X_a - X_f)$, which is a measure for the magnitude of the analysis update, is also minimized if A is selected as the symmetric square root of D (which is unique). Accordingly, in the present study V^T is chosen to be U^T.

Although both EnKF and ETKF have the same solution for P_a, D, K, and μ_a when starting with the same ensemble, they produce different ensemble anomalies — the EnKF produces the anomalies X_a as defined in (2.11), while the ETKF produces anomalies given in (2.16). The EnKF requires inverting the matrix $(HP_fH^T + R)$, which is expensive for meteorological data assimilation applications, but relatively cheaper for land data assimilation applications when the simulations at different pixels are assumed uncoupled. In contrast, the ETKF requires calculating the eigenvector decomposition $(X_f^T H^T R^{-1} H X_f)$ and inverting the matrix R, both of which are feasible for moderate ensemble sizes and diagonal R.

2.3 State Estimation from Bayesian Approach

Estimation of the optimum state in least squares is performed through a cost function by minimizing analysis error variances. Alternatively, the same solution can be obtained by maximizing the probability of the likelihood of the analysis state from Baye's Theorem. Although, theoretically these are two different problems (finding optimum weights to minimize the analysis error variance in observation space and finding the optimum state that maximizes the likelihood of the observations using error variances in state space), the solutions for these optimization problems are equivalent [Talagrand (1997) and Kalnay (2003)].

Given an estimate of a multivariate observations are available, it is our goal to obtain the best estimate of the state based on the available observations and the background forecasts. This best estimate is obtained through the probability distribution of the state conditioned on the available observations, $P(x_t|O_t)$. Following DelSole and Tippett (2010), using the Bayes theorem

$$P(x_t|O_t) = P(x_t|O_{t-1}o_t)$$

$$= \frac{P(x_t, O_{t-1}, o_t)}{P(o_t, O_{t-1})}$$

$$= \frac{P(o_t|x_t, O_{t-1})P(x_t, O_{t-1})}{P(o_t, O_{t-1})}$$

$$= \frac{P(o_t|x_t, O_{t-1})P(x_t|O_{t-1})P(O_{t-1})}{P(o_t, O_{t-1})}$$

$$P(x_t|O_t) = \frac{P(o_t|x_t, O_{t-1})P(x_t|O_{t-1})}{P(o_t|O_{t-1})} \tag{2.17}$$

where O_t denotes all the observations available until time t, and o_t denotes the observations only at time t. Assuming the observations are unbiased estimates of the true

state through operator H, the observations at time t can be related to the state with an observation model $o_t = Hx_t + r_t$. Also assuming the observational errors are temporally uncorrelated ($< r_t r_{t-1} >= 0$), then equality $P(o_t|x_t, O_{t-1}) = P(o_t|x_t)$ would hold; where from the Markovian property x_t also holds the signal of the observations at previous time steps. Accordingly, (2.17) would simplify to

$$P(x_t|O_t) = \frac{P(o_t|x_t)P(x_t|O_{t-1})}{P(o_t|O_{t-1})} \tag{2.18}$$

The term in the denominator of (2.18) is independent of x_t; hence, the best estimate of the state is related to

$$P(x_t|O_t) \propto P(o_t|x_t)P(x_t|O_{t-1}) \tag{2.19}$$

Both terms on r.h.s. of (2.19) have Gaussian distributions. The above model of observations implies that

$$P(o_t|x_t) = \frac{1}{\sqrt{2\pi|R|}}e^{-\frac{(o-Hx)^T R^{-1}(o-Hx)}{2}}$$

and following the modeling assumptions of the state

$$P(x_t|O_{t-1}) = \frac{1}{\sqrt{2\pi|\Sigma_f|}}e^{-\frac{(x-\mu_f)^T \Sigma_f^{-1}(x-\mu_f)}{2}},$$

where R and Σ_f are the error covariances of the observations and the background states. Hence, best estimate of the state is

$$P(x_t|O_t) \propto \frac{1}{2\pi|R||\Sigma_f|}e^{-\frac{(o-Hx)^T R^{-1}(o-Hx)+(x-\mu_f)^T \Sigma_f^{-1}(x-\mu_f)}{2}}. \tag{2.20}$$

21

Taking the natural logarithm of both sides

$$ln(P(x_t|O_t)) \propto constant - \frac{(o - Hx)^T R^{-1}(o - Hx) + (x - \mu_f)^T \Sigma_f^{-1}(x - \mu_f)}{2}.$$

(2.21)

To maximize the likelihood of the state estimate, the second term on the r.h.s of (2.21) should be minimized; this term can be defined as a cost function,

$$J = (o - Hx)^T R^{-1}(o - Hx) + (x - \mu_f)^T \Sigma_f^{-1}(x - \mu_f).$$

(2.22)

Hence, the maximum likelihood estimator of the state maximizes the posterior distribution derived from a Bayesian framework, and yields the same cost function that was used to obtain the standard Kalman Filter solution in a Least Squares sense. It can be also shown that the best estimate of the state can be obtained as the same solution as the Least Squares solution [(2.8) and (2.6)] by deriving (2.20).

2.4 Ensemble Based Data Assimilation studies in Atmospheric, Oceanography, and Hydrology Sciences

A detailed chronological listing of EnKF based publications can be found in Evensen (2009). Here, some milestones in ensemble based studies that are used in Atmospheric, Oceanographic, and Hydrology studies are summarized.

Kalman (1960) paved the road to recursive state estimation on which many current data assimilation applications are based. However, the standard Kalman Filter solution requires a linear model to temporally evolve the state and its uncertainty.

Accordingly, early Kalman Filter applications in meteorology [Petersen (1968)] and oceanography [Barbieri and Schopf (1982)] were performed using linear processes. Later, some applications performed linearizations of nonlinear models in EKF framework [Budgell (1986) and Lacarra and Talagrand (1988)], where the temporal evolution of the error covariances requires linearization of the model while discarding the higher order moments. Evensen (1992) and Miller et al. (1994) emphasized that this linearization of the nonlinear error propagation assumption may lead to growing unbounded errors. In addition to the closure problems, EKF is also computationally very expensive for high dimensional models, as adjoints of the nonlinear model are required to be calculated.

Evensen (1994) introduced EnKF as a stochastic alternative to the deterministic EKF [Evensen (2009)]. Ensemble of realizations, that are created in a Monte Carlo implementation, are propagated in time where the ensemble mean is selected as the best estimate and the error covariance is obtained from the ensemble covariance. This approach immediately solved the closure problem in EKF related with the propagation of the covariance. However, later Houtekamer and Mitchell (1998) used perturbed observations; and Burgers et al. (1998) emphasized that without the perturbation of observations the analysis error covariance would be underestimated, hence observations should be introduced as random variables through random perturbation. On the other hand, the earlier applications of EnKF did not have any perturbation, yet they did not have problems [Evensen (1994),Evensen and van Leeuwen (1996), and Evensen (1997)].

Divergence of filters has been a major problem in filtering techniques; the model error covariance is underestimated due to sampling issues and as a result observations are weighted too little [Anderson and Anderson (1999)]. These underestimations may lead the filter to collapse. However, it should be emphasized that the optimality of

a filter does not imply stability [Jazwinski (1970)]; unstable Kalman filters are the result of the simplifications or the nonlinearities that exist in the system but not because of the underlying equations of the filter itself. Filters with lower number of ensemble members are more prone to this underestimation than with higher ensemble members [Hamill et al. (2001)]. Heuristic and adaptive methods are used to inflate the underestimated error covariance: Anderson and Anderson (1999) and Anderson (2007) inflated the error covariance with a tuned constant λ, Whitaker et al. (2008) used an additive inflation method, Wang and Bishop (2003) used the expected statistics of the innovation to find an adaptive inflation method, and Li et al. (2009) used an adaptive method to estimate both an inflation factor and the observation errors simultaneously.

Methods to improve the error covariance estimation particularly for larger scale applications were investigated through limiting the effect of the neighboring grids over the covariance estimation of a point depending on the spatial distance of the grid. Houtekamer and Mitchell (1998) emphasized an inbreeding problem in EnKF exists that the weights to update the forecast is calculated using the same forecast. To handle this inbreeding effect, they have used pair of ensembles where one of the pair is used to calculate the weights and the other is used to calculate the analysis. They have also a cutoff radius to limit the effect of neighboring pixels on error covariances, which also reduce the computational demand of the assimilation system. Houtekamer and Mitchell (2001) have weighted the error covariance again with a distance dependent method to reduce the spurious error covariance from distant grid points, which in turn also reduces the computational power demand of the assimilation system. Ott et al. (2004) has introduced Local Ensemble Kalman Filter (LEKF) where Earth's surface is divided into local regions, so that large numbers of observations can be processes simultaneously on a much lower subspace than the number of ensemble members.

Although many applications concluded that determining the analysis ensemble in a stochastic way (with EnKF) performed well, deterministic ensemble filtering methods were introduced that avoided perturbation of the observations: Ensemble Adjustment Kalman Filter (EAKF) by Bishop et al. (2001), Ensemble Transform Kalman Filter (ETKF) by Anderson (2001), and Ensemble Square Root Filter (EnSRF) by Whitaker and Hamill (2002). These three deterministic filters were later shown to belong to a single family of square root filters [Tippett et al. (2003)]. These filters generate the same analysis error covariance from the same observation and forecast ensemble, but generate ensembles that differ from each other by unitary transformations. Among the infinite different square root solutions, Ott et al. (2004) showed selecting the ensemble that gives a symmetric square root also minimizes the analysis update.

Many of the hydrologic data assimilation studies have benefited from the theoretical background built previously in atmospheric and oceanographic data assimilation studies. Among the hydrologic data assimilation studies, soil moisture attracted the most attention, perhaps owing to its role in the predictability of the climate [Koster et al. (2004), Dirmeyer (2006), and Dirmeyer et al. (2009)]. Early applications have evaluated variational methods [Houser et al. (1998), Reichle et al. (2001)]. Walker and Houser (2001) assimilated near surface soil moisture observations to investigate the initialization problems that may occur in soil moisture profile predictions. Reichle et al. (2002) assessed the performances of EKF and EnKF in a twin experiment setup. Crow (2003) investigated a method to correct for the impact of poorly sampled rainfall over the root zone soil moisture using TOPLATS model where the Microwave radiative transfer parameters were taken from the study of Jackson et al. (1999). Zhan et al. (2006) assimilated brightness temperature observations in an observation system simulation experiment (OSSE) to investigate the potential of obtaining different resolution soil moisture products. Houser (2003) summarized the theoretical

background and the applications of streamflow, soil moisture, snow, and temperature land data assimilation studies. Walker and Houser (2004) addressed the soil moisture measurement mission requirements in a twin experiment study. Following Friedland (1969), De Lannoy et al. (2007) estimated the bias for the soil moisture profile by assimilating ground measurements. Ryu et al. (2009) investigated the unintended biases in hydrologic data assimilation framework. Reichle (2008) reviewed the data assimilation methods used in hydrologic sciences. Recently, adaptive filtering techniques were applied to improve the error estimation in land data assimilation studies [Crow and Reichle (2008) and De Lannoy et al. (2009)]. However, these methods require temporally uncorrelated observation errors, where the innovations are found to have correlations [Crow and van den Berg (2010)]. An alternative approach is proposed by Crow and van den Berg (2010) for tuning a surface soil moisture data assimilation system.

Soil moisture observations are also used to estimate the errors of other hydrological variables, like precipitation and runoff. Crow et al. (2005) assimilated satellite born soil moisture data into precipitation based soil moisture proxy to improve runoff/precipitation ratio predictions. Using the same soil moisture proxy, Crow and Bolten (2007) assimilated soil moisture data to estimate the errors of different daily precipitation products. Crow and Ryu (2009) assimilated soil moisture observations to improve both the pre-storm soil moisture conditions and the storm scale cumulative rainfall estimations.

The other hydrological observations that are also assimilated in land data assimilation framework include discharge, temperature, and snow variables. Pauwels and De Lannoy (2006) assimilated discharge rates to improve the performance of the hydrologic models and also analyzed different methods for their optimality in ensemble

based discharge assimilation. Lakshmi (2000) assimilated surface temperature observations to improve the soil moisture errors. van den Hurk et al. (2002) assimilated land surface temperature to improve numerical weather prediction (NWP) model. Rodell and Houser (2004) assimilated snow cover observations to update snow water storage. Andreadis and Lettenmaier (2006) assimilated MODIS snow cover observations to update snow water equivalent in a hydrologic model. De Lannoy et al. (2010) downscaled coarse-scale SWE observations and assimilated in a Noah land surface model. Some of the few studies focused on the water storage and its cycling include: Zaitchik et al. (2008) assimilated Grace gravity observations to improve water storage estimation in Mississippi River basin; following Simon and Chia (2002), Pan and Wood (2006) performed a constrained solution to eliminate the water budget residual; and Yilmaz et al. (2010) introduced the weakly constrained solution to decrease the water balance residual (discussed in chapter 3).

Chapter 3: Constrained Kalman Filter

3.1   Water Budget Constraint

In land data assimilation, assimilation of soil moisture (SM) results in an analysis update that does not conserve water. In this section, a water budget constraint is introduced to reduce the water imbalance. The water balance residual at time step t is

$$r_t = c_{sm}^T(SM_{at-1} - SM_{at}) + c_{cmc}(CMC_{at-1} - CMC_{at})+$$

$$c_{swe}(SWE_{at-1} - SWE_{at}) + c_pPr_t - c_eEv_t - c_rRn_t \qquad (3.1)$$

where SM is a 4-dimensional vector specifying the soil moisture in each of the 4 layers; the scalar CMC specifies canopy moisture content; the scalar SWE specifies the snow water equivalent; the scalars Pr, Ev, and Rn specify the integrated precipitation, evapotranspiration, and runoff respectively, during the data assimilation window; prefactors c_{cmc}, c_{swe}, c_p, c_e, and c_r are constants for unit conversion; and subscript a denotes the analysis. Note that SM, CMC, and SWE are prognostic variables; Pr is a forcing variable; Ev, and Rn are diagnostic variables. It is of interest to write the residual equation as combination of state and non-state variables. For a given time step all terms in (3.1), except for the analysis states, are known. Hence, these water balance terms can be condensed into the form

$$r_t = \beta_t - c_x^T x_t \qquad (3.2)$$

where

$$\beta_t = c_r Pr_t - c_e Ev_t - c_r Rn_t + c_{sm}^T SM_{at-1} + c_{cmc} CMC_{t-1} + c_{swe} SWE_{t-1} \quad (3.3)$$

$$x_t = [SM1_{at}, SM2_{at}, SM3_{at}, SM4_{at}, ST1_{at}, ST2_{at}, ST3_{at}, ST4_{at},$$

$$SkT_{at}, CMC_{at}, SWE_{at}]^T \quad (3.4)$$

where β_t is a known constant that holds the residual terms involving non-prognostic variables; where $SM1_{at}$, $SM2_{at}$, $SM3_{at}$, and $SM4_{at}$ are the soil moistures in the four layers, $ST1_{at}$, $ST2_{at}$, $ST3_{at}$, and $ST4_{at}$ are the soil temperatures in four soil layers, and SkT_{at} is the skin temperature; and c_x^T is the unit conversion vector, where temperature terms that are not part of water balance are weighted as zero (eg. assuming the units of SM, CMC, and SWE are same, then $c_x^T = [1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1]$) in order (3.2) to be consistent with (3.1).

Residuals (r_t) and the total column water content change ($c_{sm}^T(SM_{at-1} - SM_{at})$) were estimated for standard EnKF and ETKF filters (Fig. 3.1; see section 3.3.1 for more details about the model, input datasets, and the data assimilation setup). Residuals, particularly when observations are assimilated once a day, are very high when compared to the natural variability of the true total column content change. Hence, the aim of this chapter is to reduce the residuals through constrained filters.

Applying a strong constraint (i.e. forcing $r_t = 0$) would preserve the total amount of water in the water storage terms (soil moisture at different soil layers, canopy moisture content, and snow water content). In a system where precipitation, runoff, and evapotranspiration are not updated, the strongly constrained solution would redistribute the water between the storages and would preserve the total amount of water in the storage terms. However, the problem with enforcing a strong constraint

29
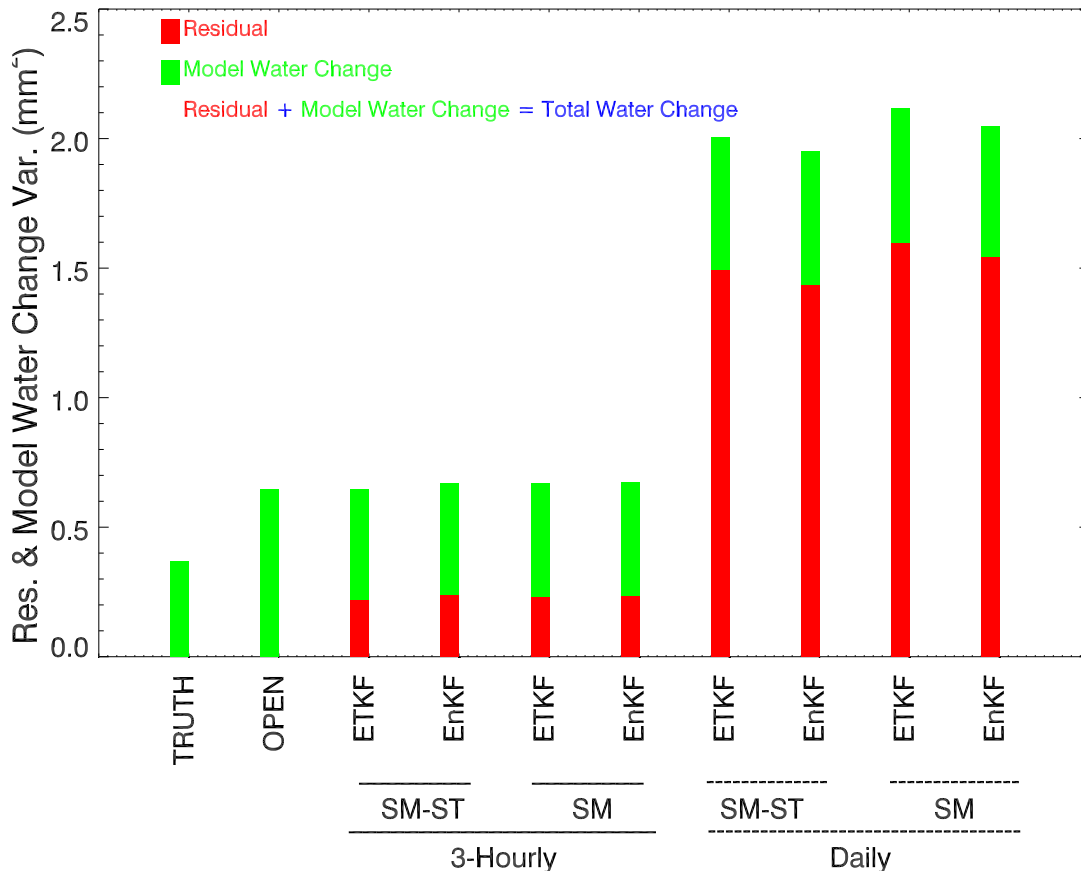
Figure 3.1: Residual and total column water chance variance of unconstrained EnKF and ETKF for 3-hourly and daily assimilation frequencies. Green, black, and blue bars indicate the total column water content change of truth, open loop, and data assimilation simulations; and red bars indicate the residuals of data assimilation simulations.

is that the individual terms (including the non-storage terms) in the water budget have error, and the errors themselves are not conserved. Thus, it is inappropriate to force an imperfectly observed budget to be held exactly. One approach is to correct the forcing terms, as described by Pan and Wood (2006). Here a weak constraint is imposed, which accounts for uncertainty in the water budget itself.

One way to impose a residual constraint is to add another term to the cost function (2.5) of the form $\lambda * f(r)$, where λ is a Lagrange multiplier and $f(r)$ is a positive definite function of r. For a strongly constrained solution [Simon and Chia (2002)]

the weighting factor λ can be determined by setting the derivation of the chosen cost function w.r.t. to λ to 0 and solving. However, for a weakly constrained system, it is not clear how this λ should be selected. Here the penalizing function $f(r)$ is set to be $(\beta - c_x^T x)^2$, and the Lagrange multiplier λ is chosen as φ^{-1}, where φ is the error variance of β (more details on how φ is calculated are given below section 3.3). Note that the Lagrange multiplier is objectively estimated. Hence the imposed constraint is of the form $(\beta - c_x^T x)^T \varphi^{-1} (\beta - c_x^T x)$, and the cost function to be minimized is of the form

$$
J_c = (o - Hx)^T R^{-1}(o - Hx) + (x - \mu_f)^T P_f^{-1}(x - \mu_f) + (\beta - c_x^T x)^T \varphi^{-1}(\beta - c_x^T x)
$$

$$(3.5)$$

where the constraint is conceived as a third penalization function which measures the degree of water imbalance $(r = 0)$.

In the standard cost function (2.5), uncertainty of the observations and the forecast are represented with error covariance matrices of R and P_f respectively, that can be obtained from the ensemble of their anomalies. Analogously, the error variance (φ) of β in (3.5) can be obtained in the form

$$
\varphi = \beta' \beta'^T / (N - 1) \tag{3.6}
$$

where β' is a vector with dimension $(1, N)$ that holds the ensemble anomaly of β (3.3), and it is trivially calculated from the ensemble of variables that are known.

Although the governing equations of the filters are same for land, atmopsheric, and oceanic data assimilation studies, in general the nature of land data assimilation systems is much different than the atmospheric and oceanographic data assimilation systems. It is common to assume in land data assimilation studies that the errors

in neighboring grid boxes are independent from each other and observations in each grid are assimilated separately; hence there is no "lateral" information exchange exists as far as the error covariance matrices are concerned. Furthermore, for many hydrological applications, the available observations are limited to the surface layer only, if not there are only 3-5 soil layers for many land surface models; hence there is practically no "depth" dimension. Briefly, in land data assimilation studies, the ensemble size is larger than the state variable size, hence sampling size issues (hence underestimation of P_a matrix) that apply to atmospheric and oceanographic studies may apply on a much smaller magnitude. Perhaps this could be the reason why error covariance inflation may not be needed, yet the ensemble analysis would not collapse in land studies. Accordingly, in this study an analysis error covariance inflation is not performed.

3.2   Constrained Kalman Filter

The vector x that minimizes (3.5) can be found by setting the derivative of J_c with respect to x equal to 0 and solving. It is shown in the appendix (A.9) that the constrained Kalman Filter solution is

$$\mu_{aa} = \mu_f + P_{aa}H^T R^{-1}(o - H\mu_f) + P_{aa}c_x\varphi^{-1}(\beta - c_x{}^T\mu_f) \tag{3.7}$$

where P_{aa} is the analysis error covariance of the constrained filter which is given in Appendix (A.5).

The Weakly Constrained Ensemble Kalman Filter (WCEnKF) solution is obtained by updating the ith ensemble member using (3.7) where perturbed observations (o') are used instead of the observations (o) to update the ith ensemble member.

Also, it is shown in the appendix (A.14) that the Weakly Constrained Ensemble

Transform Kalman Filter (WCETKF) is of the form $X_{aa} = X_f A_{aa}$ where A_{aa} is the symmetric square root of

$$D = (I + X_f^T (H^T R^{-1} H + c_x \varphi^{-1} c_x^T) X_f)^{-1}$$

The square root can be obtained from the eigenvector decomposition of $X_f^T (H^T R^{-1} H + c_x \varphi^{-1} c_x^T) X_f$.

The above constrained Kalman Filter solution can also be shown to approach the unconstrained standard Kalman Filter solution as $\varphi \to \infty$ [see appendix, (A.18)]. Moreover, the residual of the constrained filter is shown to be smaller than the residual of the standard filter [see appendix, (A.19)]. It is also shown in the appendix (A.18) that the constrained Kalman Filter solution can be solved equivalently in two recursive filters:

$$\mu_{aa} = \mu_a + P_a c_x (\varphi + c_x^T P_a c_x)^{-1} (\beta - c_x^T \mu_a)$$

where $\mu_a = \mu_f + P_a H^T R^{-1} (o - H\mu_f) = \mu_f + K(o - H\mu_f)$ is the solution of the standard KF without the constraint. This solution implies that the constrained solution can be obtained by first calculating the solution (μ_a) for the standard KF, and then adjusting this solution to take into account the constraint by adding $(P_a c_x (\varphi + c_x^T P_a c_x)^{-1} (\beta - c_x^T \mu_a))$. The single-stage and two-stage solutions yield identical WCEnKF updates, but generally different WCETKF analysis anomaly updates due to the fact that the single- and two-stage WCETKF equations are solved using two different matrix square roots for the same analysis error covariances.

A strongly constrained Kalman Filter solution (A.20) can be estimated by taking

the limit $\varphi \to 0$.

$$\mu_{aa} = \mu_a + P_a c_x (c_x^T P_a c_x)^{-1} (\beta - c_x^T \mu_a)$$

where this solution is identical to the strongly constrained solution of Simon and Chia (2002) [(25) in their paper]. Similar to the weakly constrained solution, a strongly constrained Ensemble Kalman Filter (SCEnKF) can be estimated by updating the each ensemble using the above equation with perturbed observations. Similarly, a strongly constrained Ensemble Transform Kalman Filter (SCETKF) can be obtained by taking the limit $\varphi \to 0$ in (A.23) as

$$X_{aa} = X_a - P_a c_x c_x^T X_a / c_x^T P_a c_x$$

which implies the adjustment term for the constraint in the second state is $P_a c_x c_x^T X_a / c_x^T P_a c_x$.

3.3   Sample Simulations

3.3.1   Experiment Setup

To illustrate the weakly constrained filters, synthetic experiments were performed using the Noah land surface model [Ek et al. (2003)] version 2.7. Noah model uses the lower atmospheric boundary fields (precipitation, humidity, air temperature at a reference level, short–wave and long–wave, and pressure), provided by an atmospheric model or an offline forcing data, and evaluates the hydrological variables at the surface like soil moisture, soil temperature, evapotranspiration, runoff, etc. Soil and land cover types in Noah are selected fixed, while the canopy and vegetation parameters (e.g. greenness and LAI) vary in time. Noah model heritages the surface albedo,

roughness, and surface resistance parameters from Simple Biosphere Model [SIB, Sellers et al. (1986)] parameterization (Dorman and Sellers (1989)). Noah model includes a frozen soil scheme following Koren et al. (1999). The Noah model states (with memory) includes soil moisture, soil temperature, skin temperature, canopy moisture content, and snow water equivalent. Number of soil layers is defined by the user, which in general is chosen as 4 layers. For each layer separate water and energy balance is calculated, where the total balances are preserved (input water equals to output water and input energy equals output energy).

The study area was chosen to be Red Arkansas River Basin, US (between 32.0°N - 37.0°N and 96.0°W - 91.0°W) with 0.125°spatial resolution (Fig. 3.2). There are total of 1521 pixels (39*39). The pixels are assumed to have uncorrelated errors. Simulations were performed between April - October 2006 (total 4500 hourly time steps) using hourly North America Land Data Assimilation [NLDAS; Cosgrove et al. (2003)] forcing data (precipitation, pressure, relative humidity, wind speed, short wave and long wave radiation, and air temperature) which have 0.125°spatial resolutions. Model grid spatial resolutions were selected consistent with the NLDAS data native resolution, so that no averaging or downscaling was needed. The initial states were generated by running the land model for 10 years, but with repeating 2006 NLDAS forcing data in each of the 10 years where the state obtained after each year of simulation is used as an initial condition for the following year. The state obtained at the end of the 10th year were selected as the initial states for all simulations. Assimilation of observations are performed in warm climate, where the ensemble of model realizations are simulated starting from January to provide a smooth transition before the assimilation of observations. All initial states and the forcing data (air temperature, short and long wave radiations, and precipitation) were perturbed (as described below) to create the ensembles for all simulations. The truth run is identified as a single
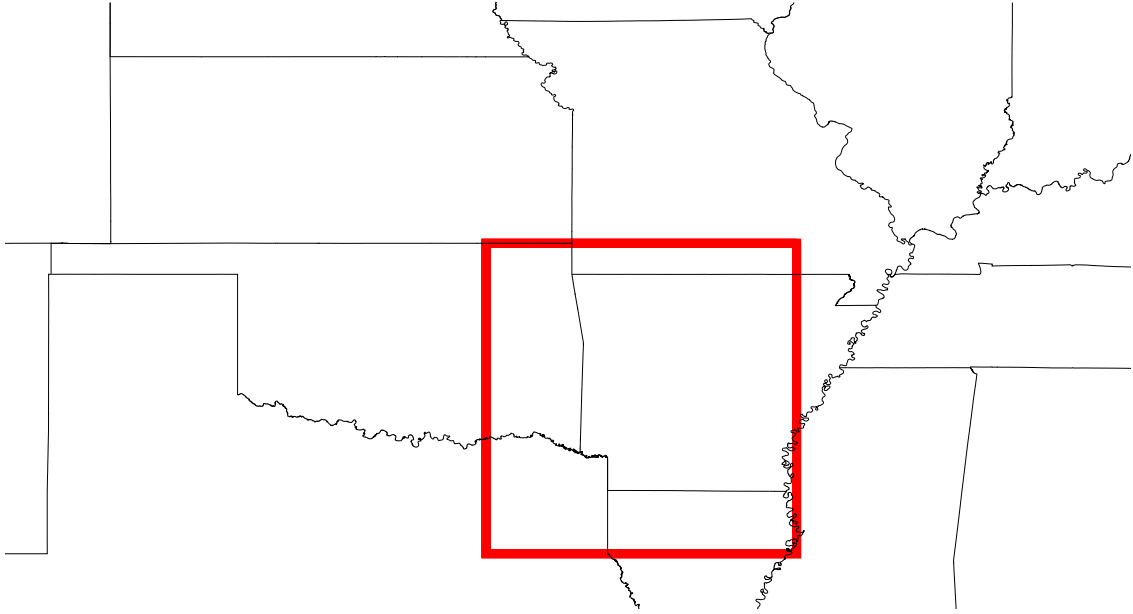
Figure 3.2: Location of the Study Area shown in red box.

run of the model with unperturbed initial condition and forcing.

The experiments were based on a 'perfect model'assumption in which the same model that generated the "truth"was used to generate the prior distribution. The observation operator H equals to the identity matrix. Initial states were perturbed using additive Gaussian noise [selected from normal distribution with (μ=0,σ=1K°) and (μ=0,σ=0.02%) for ST and SM respectively]. Precipitation forcing was perturbed using multiplicative noise with a log-normal distribution (μ=1,σ=0.7); short-wave radiative forcing was perturbed using multiplicative noise with normal distribution [N(μ=1, σ=0.25)]; air temperature forcing and long-wave forcing data were perturbed using additive noises with normal distribution [N(μ=0, σ=2.5 K°) and N(μ=0, σ=10 $W.m^{-2}$) respectively]. The above perturbations are independent. The precipitation perturbation multiplication factor was limited between 0 and 4 where the actual precipitation value was further prevented to exceed the true precipitation value with

5mm/hour in ensemble generation. The short-wave perturbation multiplication factor was limited between 0.2 and 1.8. Temperature and long-wave radiation perturbations were limited to 4 times their respective standard deviations.

All forecasts were performed for an ensemble size of 50. Ensembles of Open loop simulations (through an ensemble of model simulations without the assimilation of observations) were simulated using the same perturbed initial states and forcings as the assimilation experiments. Although it is not possible to remotely sense the full SM and ST profiles with the current observation systems, there are many monitoring stations that provide in-situ deep soil layer variables (i.e. Oklahoma Mesonet Network). Hence, for the proof of concept, observations through the entire soil column were assimilated (not only the top layer). After open loop simulations were performed and their errors were calculated, observation perturbation variances were selected based on these open loop error variances in order to have comparable open loop and observation realizations. Accordingly, observations were created by adding zero mean Gaussian noise to the truth states for all four soil layers (ST perturbation standard deviations were 0.40K°, 0.20K°, 0.20K°, 0.10K°for the four layers from top to bottom respectively; SM perturbation standard deviations were 0.004%, 0.004%, 0.004%, 0.004% for the four layers from top to bottom respectively). Unconstrained and constrained simulations had the same forcing and initial state perturbations as the open loop.

3.3.2   Filter Performance Analysis

The simulations were performed for four filters (ETKF, EnKF, WCETKF, and WCEnKF), for two types of assimilated observations (all 4 layers of SM, or all 4 layers of SM and ST together), and for two state update frequencies (3-hourly or once a day) giving a total of 16 sets of experiments. Soil temperature affects the water balance through

37

evaporation term, hence the effect of assimilating also soil temperature is analyzed. Only the single-stage solutions were used for the constrained filters. State error and the water balance residual statistics were calculated for all 16 sets of experiments. The state error statistics were also calculated for the open loop simulations (open loop simulations have no water balance residual).

Innovation Statistics

If the assumptions on which the Kalman Filter equations were derived are true, then the quadratic form $[(o - Hx)^T (HP_f H^T + R)^{-1} (o - Hx)]$ should have chi-squared distribution with d.o.f. equal to the size of the observation vector. This chi-squared statistic was calculated at each time step for each pixel and each experiment separately. The percentage of pixels that were within the 2.5 and 97.5 percentiles was calculated for each experiment separately. The 2.5 and 97.5 percentiles of a chi-square distribution are 0.484 and 11.14 for 4 d.of. (for SM only updated scenario); and 2.180 and 17.535 for 8 d.o.f. (for both SM and ST updated scenario).

State Errors

Updated states during the assimilation are SM (all 4 layers), ST (all 4 layers), SkT, CMC, and SWE, regardless of the observed variable that is assimilated (SM, or SM and ST). Due to the time interval selection (April-October, no snow), snow related variables were effectively not updated; hence snow related results were not investigated or presented in this study. Mean square error of ensemble means (MSE) for each of 10 states and for each of 16 experiments per pixel were calculated as

$$MSE_{s\ i\ lon\ lat} = \sum_t \left(\mu_{s\ i\ lon\ lat\ t} - TS_{s\ i\ lon\ lat\ t} \right)^2 / (ts - 1)$$

where μ is the ensemble mean state, TS is the true state, s is each state (total 11), i is each experiment (defined above, total 16 sets), lon is longitude pixel number (total 39), lat is latitude pixel number (total 39), t is each time step, and ts denotes the number of time steps (total 4501) respectively. Resulting MSE values for each pixel and for all 4 soil layers were then averaged to a single number separately for ST and SM variables and for each experiment.

$$RMSE.SM_i = \sqrt{\sum_{sm}^{4}\sum_{lat}^{39}\sum_{lon}^{39} MSE_{sm\ i\ lon\ lat}/(4*39*39))}$$

$$RMSE.ST_i = \sqrt{\sum_{st}^{4}\sum_{lat}^{39}\sum_{lon}^{39} MSE_{st\ i\ lon\ lat}/(4*39*39))}$$

Water Balance Residual

The water balance residual was calculated for each ensemble member, at each time step, at each pixel in the study area, and each set of experiments (total 16, defined above). The variance and the mean of the residuals were calculated using all time step and ensemble member values for each set of experiment and for each pixel in the study area as:

$$r_{i\ lon\ lat\ .\ t} = \sum_{n} r_{i\ lon\ lat\ n\ t}/N$$

$$r_{i\ lon\ lat\ .\ .} = \sum_{t} r_{i\ lon\ lat\ .\ t}/ats$$

$$\sigma^2 r_{i\ lon\ lat} = \sum_{t}(r_{i\ lon\ lat\ .\ t} - r_{i\ lon\ lat\ .\ .})/(ats-1)$$

where the "dot"denotes an index that is averaged out, $\sigma^2 r$ is the residual variance, n denotes ensemble member, and ats is the total number of time steps that the observations are assimilated (1500 and 187 for 3-hourly and daily update scenarios respectively), where only the residuals due to assimilation were included in the statistics. Then $\sigma^2 r_{i\ lon\ lat}$ and $r_{i\ lon\ lat\ ..}$ values were averaged over the study area into single number ($\sigma^2 r_{i\ ..}$ and $r_{i\ ....}$) for each experiment separately.

Degree of Residual Improvement

The degree of performance change between unconstrained and constrained solutions was assessed by a series of F-tests (simply the ratio of constrained filter residual variance over the unconstrained filter residual variance). F-tests were performed for 8 experiments separately: for both ETKF and EnKF filters; for both SM alone, and SM and ST together assimilated cases; and for both daily and 3-hourly assimilation frequency scenarios. For each scenario, the unconstrained residual variance to constrained residual variance ratio was calculated seperately for each individual pixel in the study area. These ratios were compared to the critical F-test values at 5% significance level (upper and lower critical values are 1.33 and 0.75 for daily, and 1.11 and 0.90 for 3-hourly assimilation frequency scenarios with 187 and 1500 residual time-series values respectively). For each pixel and experiment an F-test value of -1, 0, or 1 were assigned for when unconstrained residual was significantly smaller than constrained residual, when they are not significantly different, and when constrained residual was significantly smaller than unconstrained residual respectively. All pixel F-test values were then averaged into a single number and multiplied by 100 for each experiment separately. Accordingly, a positive 100% represents smaller residuals for the constrained filter for the entire area, and -100% represents smaller residuals for the unconstrained filter for the entire area.

Total Column Water Change

Total column water content is defined as the summation of the total soil moisture content (mm) for all 4 soil layers at any given time where its change is defined as,

$$\Delta WC_{i\ lon\ lat\ t} = \sum_N \sum_d (SM_{i\ lon\ lat\ n\ t-1\ d} - SM_{i\ lon\ lat\ n\ t\ d}) * Depth_d/N$$

$$\Delta WC_{i\ lon\ lat\ .} = \sum_t \Delta WC_{i\ lon\ lat\ t}/ats$$

$$\sigma^2 Wat_{i\ lon\ lat} = \sum_t (\Delta WC_{i\ lon\ lat\ t} - \Delta WC_{i\ lon\ lat\ .})^2/(ats-1)$$

where ΔWC is the total column water content change (mm), d is the soil layer identifier, and $\Delta WC_{i\ lon\ lat\ .}$ and $\sigma^2 Wat_{i\ lon\ lat}$ are the mean and the variance of the total column water change. Calculated $\sigma^2 Wat_{i\ lon\ lat}$ values are then averaged over the study area into a single variance for each experiment ($\sigma^2 Wat_{i\ .\ .}$). For daily update scenarios $\sigma^2 Wat_{i\ .\ .}$ variances, similar to residual variances, were calculated only for the time-steps of the assimilation updates.

3.3.3 Results

State Error

The result of applying a strongly constrained EnKF for a single pixel located at 34.63N and 94.75W between May-Oct, 2006 with 3-hourly observations is shown in Fig. 3.3. This figure shows that the strong-constrained filter produces very unrealistic soil temperatures, in the sense that the estimates are well beyond the range of variability of the truth. This result suggests that forcing data has large uncertainties that should be taken into account in the filter. Hence, the remaining constraint experiments were
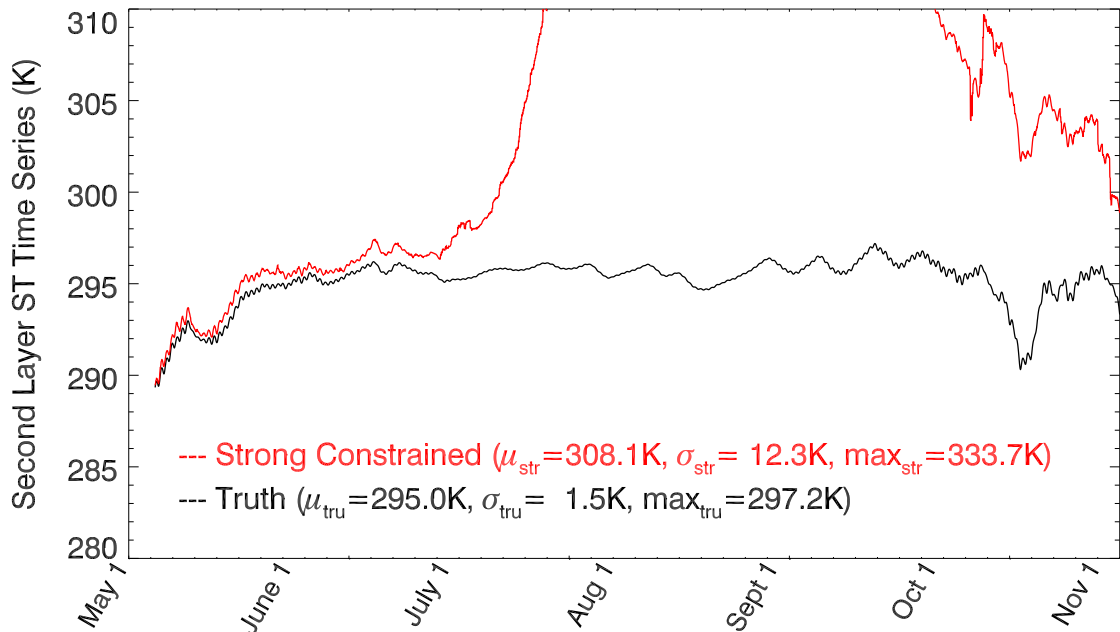
Figure 3.3: Second soil layer temperature errors of strongly constrained EnKF simulations

performed using weak constrained filters (WCEnKF or WCETKF).

The RMSE of all assimilation experiments, observations, and the open loop runs are shown in Fig 3.4. In most cases, the RMSEs for the constrained filter were close (within 2%) to the RMSEs for the unconstrained filter. The RMSEs for the constrained filter can be larger than for the unconstrained filter, but in these cases the RMSEs still were much smaller than the RMSEs in observations or the open loop. Not surprisingly, the RMSEs of a variable were much smaller than those of the corresponding observations or the open loop, when observations of that variable were assimilated. However, if the observations of a variable were not assimilated, then the RMSE of that variable can be comparable to that of the open loop, indicating very little benefit from the filter. Three-hourly assimilation of observations has smaller RMSEs than the corresponding daily assimilation, but not by an order of magnitude

(even though 3-hourly assimilation was 8 times more expensive than the daily assimilation). In general, the RMSEs for the EnKF, ETKF, WCEnKF, and WCETKF were comparable to each other.

Innovation statistics were analyzed for the filter performance. Observed innovations fell within the 95% confidence interval 92% to 95% of the time, suggesting consistency with the underlying assumptions of the Kalman Filter.

Water Balance Residual

Water balance residual variances for the 16 experiments using the single-stage filters are shown in Fig. 3.5a. In general, the time mean of the residuals differed only slightly between the 16 sets of experiments (results not shown), where the annual water budget is not conserved on average. The magnitude of the residual bias was orders magnitude smaller (2-3% for daily simulations) than the magnitude of the residual variance for all experiments.

Constrained filter residual variances were smaller than unconstrained filter residual variances over all pixels in the study area regardless of the update variable (SM alone, or SM and ST together), filter (WCETKF vs ETKF, or WCEnKF vs EnKF), or update frequency (3-hourly or daily) selection (Fig. 3.5a). The residual variances of the constrained filters were 14% to 44% less than those for the unconstrained filters.

Two-stage WCETKF was performed using the unconstrained ETKF square root for the first stage. Two-stage WCETKF has consistent tendency to have higher (but not significant) state errors than the single-stage WCETKF errors, whereas two-stage WCETKF residuals were almost identical with the single-stage WCETKF residuals (results not shown).
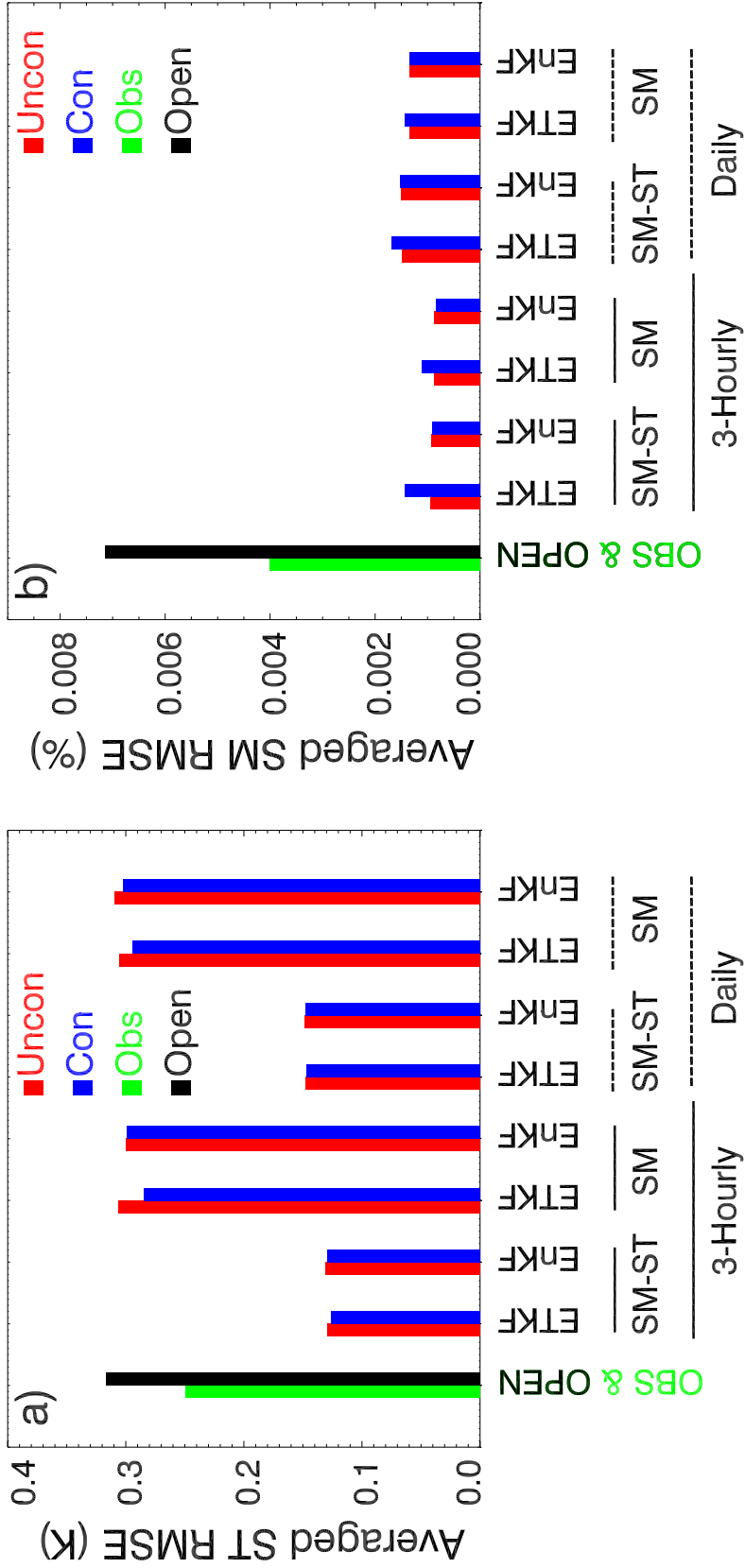
Figure 3.4: (a) Soil temperature and (b) soil moisture errors averaged across soil layers. Horizontal axis: OBS refers to observation errors (in green color); OPEN refer to open loop errors (in black color); and Uncon and Con refer to unconstrained and constrained filters respectively.
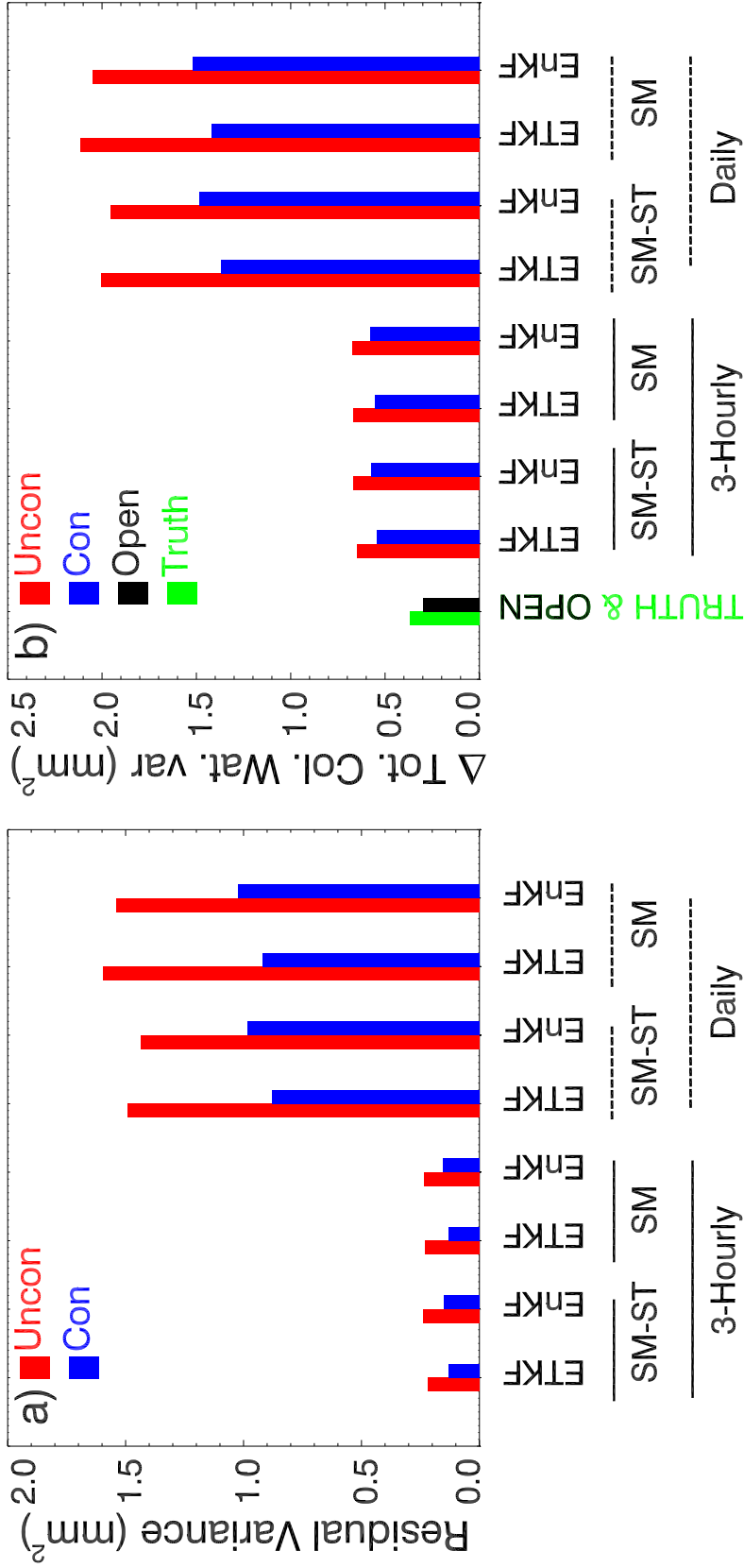
Figure 3.5: (a) Water budget residual variance and (b) Total Column Water Content Change variance of constrained and unconstrained experiments. True total column water content change is shown in green bar, open loop water content change in black, unconstrained filter results in red, and constrained filter results in blue.

Degree of Residual Improvement

Signifance of the residual differences between unconstrained and constrained filters were checked for various scenarios: for assimilated variables SM alone, and SM and ST together; for ETKF and EnKF filters; for 3-hourly and daily assimilation frequencies. Constrained residuals were significantly smaller than the unconstrained residuals for all the pixels (except for a small percentage of the study area for ETKF when SM was assimilated) regardless of the assimilated frequency, assimilated variable, and the filter selection (Table 3.1).

Table 3.1: F-Test results for residual variances of constrained and unconstrained solutions. AF-24 and AF-03 represent the daily and 3-hourly assimilation experiments. Values represent the percentage of pixels that the constrained residual variances were significantly smaller than the unconstrained filter residual variances.

	AF-24	AF-03
ETKF SM-ST	100.	100.
ETKF SM	98.	100.
EnKF SM-ST	100.	100.
EnKF SM	100.	100.

Total Column Water Content Change

Cross comparisons of the variances of the total column water content were performed for the 16 sets of assimilation experiments, the truth, and the open loop simulations (Fig. 3.5b).

The water content change variance of the open loop simulations was slightly lower than that of truth simulations. The constrained assimilation experiments had 14%-33% smaller total column water change than the unconstrained experiments regardless of the assimilation frequency, observed variable, or the filter selection (Fig. 3.5b), supporting the above discussed residual results that the constrained filters were closer to

46

the truth simulations with respect to their closure of water cycling than unconstrained filters.

Total column water change in an assimilation experiment can be conceived as the summation of the true change plus the residual added due to the assimilation update. Comparison of the residual variances against total water change for the assimilation experiments indicates 70% of the total water change was due to the residual for daily assimilations where this ratio was around 30% for 3-hourly assimilation experiments (Fig. 3.5a and Fig. 3.5b); suggesting that in the absence of frequent observations the obtained total soil moisture content change is heavily affected from the residuals along with the true soil moisture change.

Sensitivity of φ

Estimation of φ in an objective way from the ensemble of realizations with the above described methodology (3.6) improves the residuals with little effect on the state errors. The effect of inflating (or deflating) the φ values and using constant φ values rather than being objectively estimated (Fig. 3.6) scenarios were investigated. These simulations were performed for 117 pixels located between 32.0°N 32.375°N and 96.0°W 91.0°W with 3-hourly observations. An apparent trade-off was found between the state errors and the residuals: the more the φ values were deflated (constraint was applied stronger), the more the state errors were increased and the more the residuals were decreased (Fig. 3.6). Applying the constraint too strongly (with inflation factor of 0.001 or using constant 0.001 φ values) resulted in state errors equal to observation errors, suggesting no additional benefit from the filter, whereas applying the constraint too weakly (by inflating φ 5 times or using constant φ values of 5) resulted in residuals that are very close to residuals of the unconstrained simulations.

In this sensitivity study, the range of constant (tuned) φ values were chosen based

on a priori information obtained from objective estimation. SM error–residual trade off performance of WCEnKF was better than the performance of WCETKF. Objective estimation of φ had same performance with the estimation through tuned φ for WCEnKF; whereas for WCETKF using tunable parameter gave better performance than objective estimation. Hence, in this study we conclude there is no universal solution in selection of tuning or inflating φ; for some filters tuning gives better, for some inflation avoids tuning φ.

Optimality of the constrained filter depends on the goal of the specific application; depending on the priority given to the state error or the residual error, φ can be inflated or deflated to improve one error while degrading the other one at a different magnitude (Fig. 3.6). In general, in hydrological studies, having smaller state error is generally preferred. From this point of view, smaller residuals can be obtained without degrading the state errors noticeably. For example, inflating φ values with factors of 0.50–0.75 gave almost the same state errors with the standard EnKF, while it reduces the residuals to less than half of the standard EnKF. Objectiveness of how a constant φ value can be selected is still questionable; however similar results can be obtained by tuning the φ values prior to the simulations. The objective selection of a tuned φ value or an inflation factor could be less of a problem for reanalysis type of studies; whereas for an operational platform, particularly in a changing system, the selection of tuned φ could be more critical.
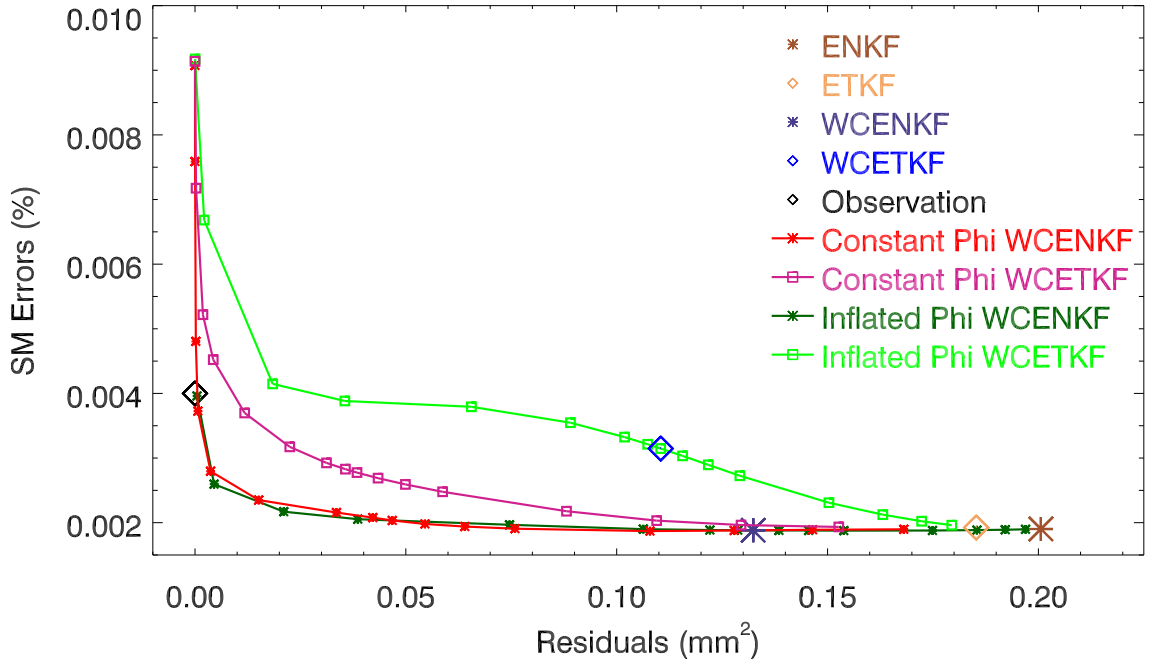
Figure 3.6: SM error and residual relation for varying Phi values, where both SM and ST observations are assimilated using 50 ensemble members. Each line represents series of simulations using 16 different constant values or 16 different inflation values (Both inflation and constant values were selected as 0.001, 0.01, 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 1.00, 1.20, 1.50, 2.0, 5.0, 10.0, 20.0, and 50.0). Single points represent single simulations of constrained filters with un-inflated values or of unconstrained filters. For both constant and inflated experiments, higher residuals are result of higher φ values and lower residuals are result of lower φ values (Inflation or constant φ values increase from left to right for green and red lines). Observation error is also marked with a black diamond. The residuals and the errors asymptotically approached to those of unconstrained simulations or strongly constrained simulations as the inflation factors for or the constant value was increased to or decreased to 0 respectively.

49

Chapter 4: No Perturbed Observations and No Constraint Anomalies

In section 2.4, we showed why perturbation of observations is necessary and without the perturbations the analysis error covariance is underestimated. On the other hand, before this underestimation was noted by Burgers et al. (1998), several studies had already implemented the EnKF without the perturbation of the observations [Evensen (1994); Evensen and van Leeuwen (1996); and Evensen (1997)]. Whitaker and Hamill (2002) has presented some of the shortcomings of perturbations of observations. However, since then this issue has not been studied. Here we revisit perturbing the observations in EnKF and compare EnKF with other filters. We have also studies the affect of eleminating the perturbations of observations and the constraint ensemble anomalies over the water balance residuals.

4.1 No Perturbed Observations and No Constraint Anomalies

The constrained Kalman Filter is derived in the Appendix section by minimizing a cost function. The solution for the state is shown in (A.9). In an ensemble aproach (WCEnKF or WCETKF), this solution can be performed in two pieces, by estimating the analysis mean and by estimating the analysis anomaly separately; the analysis mean represents the best estimate of the state and the analysis anomaly represents the uncertainty of this best estimate. Hence the mean state is fit to the mean of

observations, forecast, and β [(3.3), holds all the water water balance terms except for the states at the current time step], where the state anomaly is fit to the anomaly of observations (namely the perturbations), forecast, and β'.

Although, it has been emphasized by Burgers et al. (1998) that taking $O' = 0$ underestimates the analysis error covariance, it has been also pointed out that this underestimation would not affect the updated analysis mean [Burgers et al. (1998)]. Hence, in this study, the effect of not perturbing the observations ($O' = 0$) and removing the constraint anomalies ($B' = 0$) together or individually is investigated for different simulations.

Starting from the same initial conditions, the solutions of the EnKF and ETKF have the same P_a, D, K, and μ_a; and the solutions of the WCEnKF and WCETKF have the same P_{aa}, D, K, and μ_{aa}. However, the analysis ensemble anomalies for EnKF (2.11) and ETKF (2.16), as well as for the WCEnKF (A.9) and WCETKF (A.14), are not the same. The analysis anomaly solution of EnKF requires the perturbed observations, and the solution of WCEnKF requires both perturbed observations and the constraint anomalies. On the other hand, the analysis anomaly solutions of ETKF and WCETKF are obtained through square roots which do not require either the perturbed observations or the constraint anomalies. Hence, the above proposed removal of O' and B' affect only the EnKF and WCEnKF solutions but not the ETKF and WCETKF solutions. Accordingly, EnKF and WCEnKF solutions are modified into additional 4 filters. Including the ETKF and WCETKF, total 8 filters are analyzed in this chapter, which all have the same analysis mean but different analysis anomalies.

Analysis anomaly solution for the standard EnKF (with perturbed observations) is

$$X_a = X_f + K(O' - HX_f). \tag{4.1}$$

Analysis anomaly solution for the Ensemble Kalman Filter with no perturbed observations [EnKF-noPO; Whitaker and Hamill (2002)] is

$$X_a = X_f + KHX_f. \tag{4.2}$$

Analysis anomaly solution for WCEnKF is given in Appendix (A.10) as

$$X_{aa} = X_f + P_{aa}H^T R^{-1}(O' - HX_f) + P_{aa}c_x\varphi^{-1}(B' - c_x^T X_f). \tag{4.3}$$

Analysis anomaly solution for Weakly Constrained Ensemble Kalman Filter with no perturbed observations (WCEnKF-noPO) is

$$X_{aa} = X_f - P_{aa}H^T R^{-1}(HX_f) + P_{aa}c_x\varphi^{-1}(B' - c_x^T X_f) \tag{4.4}$$

where B' holds the ensemble anomalies of β. Analysis anomaly solution for Weakly Constrained Ensemble Kalman Filter with no constraint anomalies (WCEnKF-noCA) is

$$X_{aa} = X_f + P_{aa}H^T R^{-1}(O - HX_f) - P_{aa}c_x\varphi^{-1}(c_x^T X_f) \tag{4.5}$$

and analysis anomaly solution for Weakly Constrained Ensemble Kalman Filter with no perturbed observations and no constraint anomalies (WCEnKF-noPO-noCA) is

$$X_{aa} = X_f - P_{aa}H^T R^{-1}(HX_f) - P_{aa}c_x\varphi^{-1}(c_x^T X_f). \tag{4.6}$$

Above equations (4.1)-(4.6) are only used to derive anomalies, whereas the solutions for the means are obtained from (B.6) and (A.9) for the unconstrained and constrained filters respectively. The characteristic differences between these filters are summarized below in Table 4.1.

Table 4.1: Summary of filters and their distinctive analysis anomaly properties. O'
denotes perturbed observations and B' denotes the presence of constraint anomalies.
A dash means the particular anomaly does not apply, ✓ means the anomaly exists in
the solution, and X means the anomaly is not used in the solution.

	O'	B'
ETKF	-	-
WCETKF	-	-
EnKF	✓	-
EnKF-noPO	X	-
WCEnKF	✓	✓
WCEnKF-noPO	X	✓
WCEnKF-noCA	✓	X
WCEnKF-noPO-noCA	X	X

4.2   Sample Simulations

4.2.1   Experiment Setups

Separate simulations were performed for each of the above described filters (ETKF,
WCETKF, EnKF, EnKF-noPO, WCEnKF, WCEnKF-noPO, WCEnKF-noCA, and
WCEnKF-noPO-con). The setups of the experiment in this chapter were identical to
the setups described in chapter 3. Synthetic experiments were performed using Noah
model version 2.7. The study area was chosen as Oklahoma, US (between $32.0°N -$
$37.0°N$ and $96.0°W - 91.0°W$) with $0.125°$spatial resolution between April $-$ October
2006. North America Land Data Assimilation [NLDAS; Cosgrove et al. (2003)] data
were used as the atmospheric forcing. Model grid spatial resolutions were selected
consistent with the NLDAS data, so that no averaging or downscaling was needed.
Initial states were obtained after spinning the model for 10 years.

The "truth"run is identified as a single run of the model with unperturbed initial
conditions and forcing. The initial state and the forcing perturbation for both the
assimilation and the open loop experiments are described in chapter 3. Open loop

is defined as ensemble of simulations without the assimilation of observations, where the ensemble members are generated in the same way as in the assimilation experiments. All forecasts were performed for an ensemble size of 50. Four layers of Soil Moisture (SM) and Soil Temperature (ST) observations were assimilated once a day. Observations were created by adding zero mean Gaussian noise to the truth states as described in chapter 3. State error, residual, and total column water content change statistics were also calculated similar to the methodology described in chapter 3. Sensitivity analysis were also performed for all filters where the assimilation frequencies change between hourly to daily and ensemble sizes change from 10 to 150.

4.2.2 Results

In this study variety of filters were compared, which also includes not perturbing the observations; acknowledging eliminating the perturbations results the assumptions behind the Kalman Filter are not fulfilled (hence filter is not optimum). SM RMS error and residual sensitivities to the ensemble size and the assimilation frequency were tested for EnKF, ETKF, EnKF-noPO, WCEnKF, WCEnKF-noPO, WCEnKF-noCA, and WCEnKF-noPO-noCA filters [Table 4.1] with assimilating both SM and ST observations.

Similar to Burgers et al. (1998), in this study not perturbing the observations also resulted in similar state errors as the filters with the perturbed observations. In general all filters (ETKF, WCETKF, EnKF, EnKF-noPO, WCEnKF, WCEnKF-noPO, WCEnKF-noCA, and WCEnKF-noPO-noCA) had similar SM and ST errors (Fig. 4.1). On the other hand, standard filters (EnKF and ETKF) had the largest residual and total water content errors (Fig. 4.2). Not perturbing the observations or taking the constraint anomaly as zero improved both the residual and the total column water content change errors. In fact, WCEnKF-noPO-noCA had the smallest
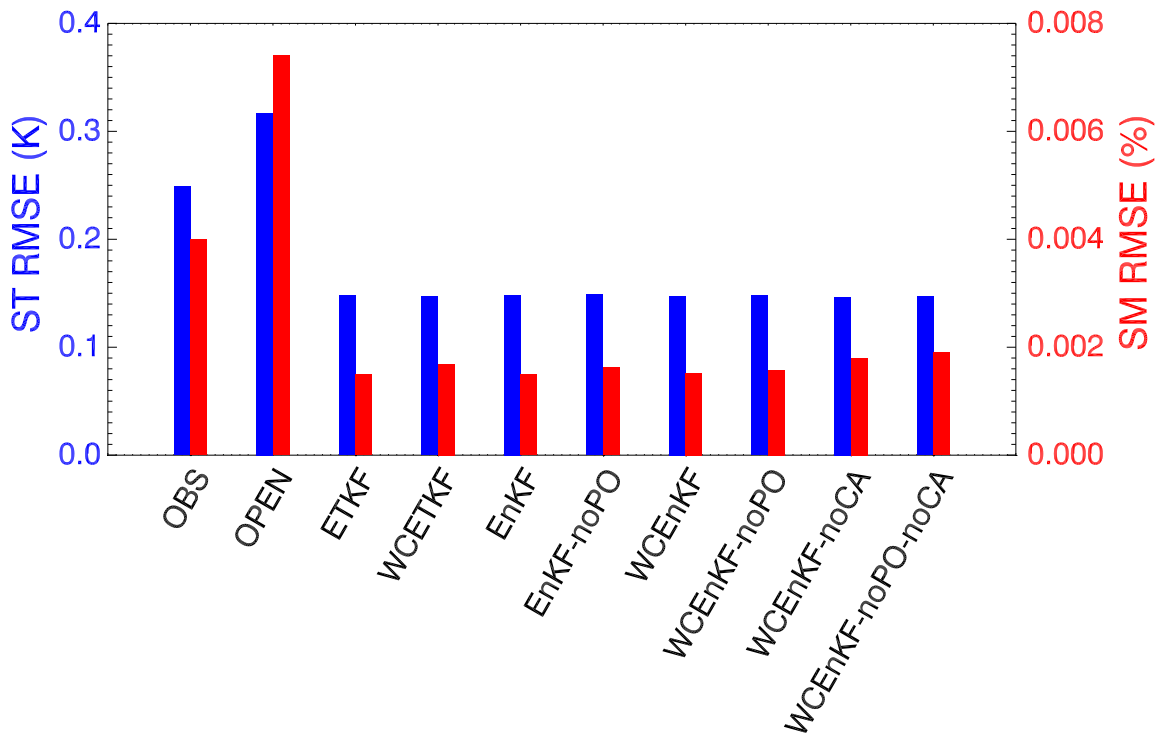
Figure 4.1: SM and ST RMSE of ETKF, WCETKF, EnKF, EnKF-noPO, WCEnKF, WCEnKF-noCA, WCEnKF-noPO-noCA filters for assimilation of daily SM and ST observations with 50 ensemble members. ST errors are shown in blue color on left y-axis and SM errors are shown in red color on the right y-axis.

residual errors.

As expected from a consistent filter, in general more frequent observations results in smaller residuals and SM errors. Increased ensemble size did not improve SM errors. Although the simulations with the smallest ensemble size (10 or 15 members) in general have the highest SM errors, simulations with largest ensemble sizes (120 or 150 members) did not always have the smallest errors (Fig. 4.3).

The effect of more frequent observations on residuals was more dramatic than the effect of ensemble member size selection (Fig. 4.3). In fact the magnitude of the residuals remain almost unaffected from the ensemble size selection (Fig. 4.3). In
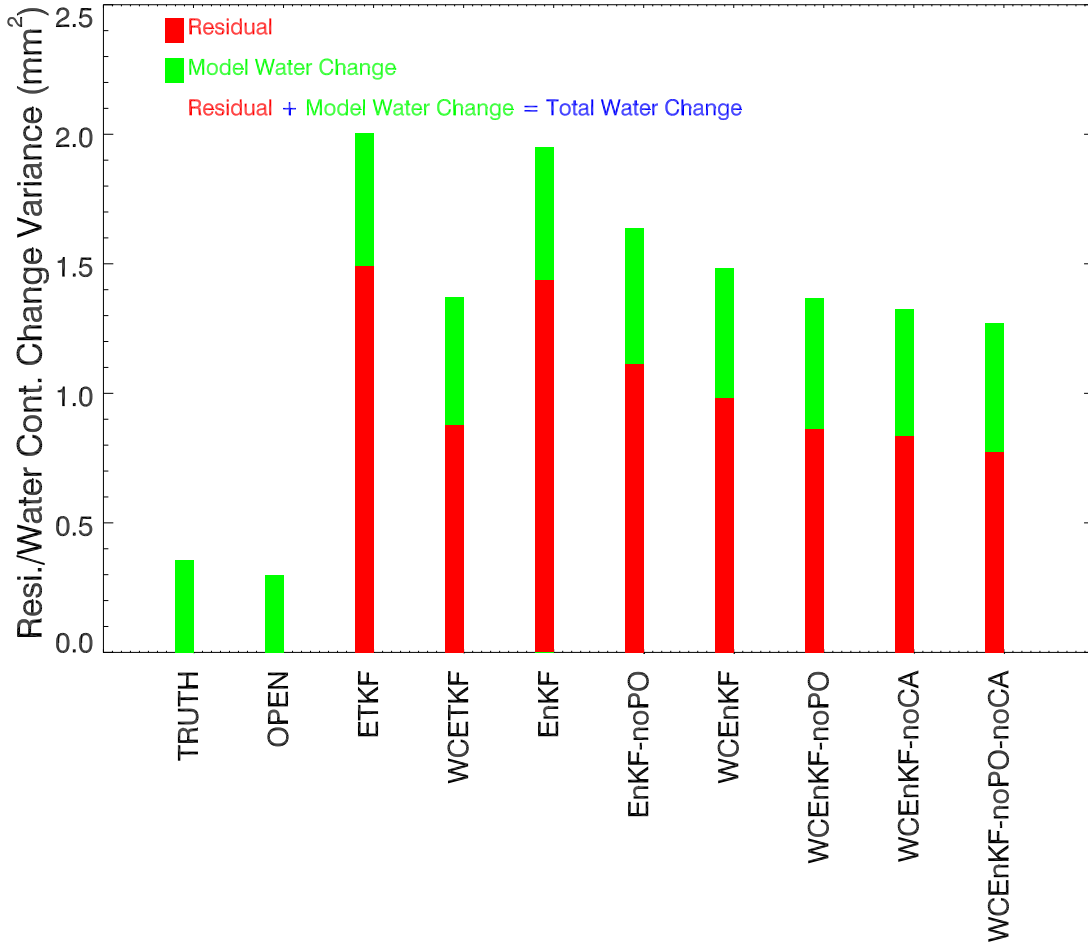
Figure 4.2: Residual variances for various filters (in blue bars); and total column water change variances for various filters (shown in red columns), and for truth and open loop simulations (shown in green and black columns respectively).

general the magnitude of constrained filter residuals for a given observation assimilation frequency were comparable to the residuals of unconstrained filter with higher observation frequency (Fig. 4.3).

Among the filters, the sensitivity of EnKF was clearly different than the other filters when higher ensemble size and high assimilation frequencies were applied together (Fig. 4.3). More specifically, the residuals of hourly EnKF increased dramatically with the ensemble size increase (top left panel of Fig. 4.3). Although SM errors
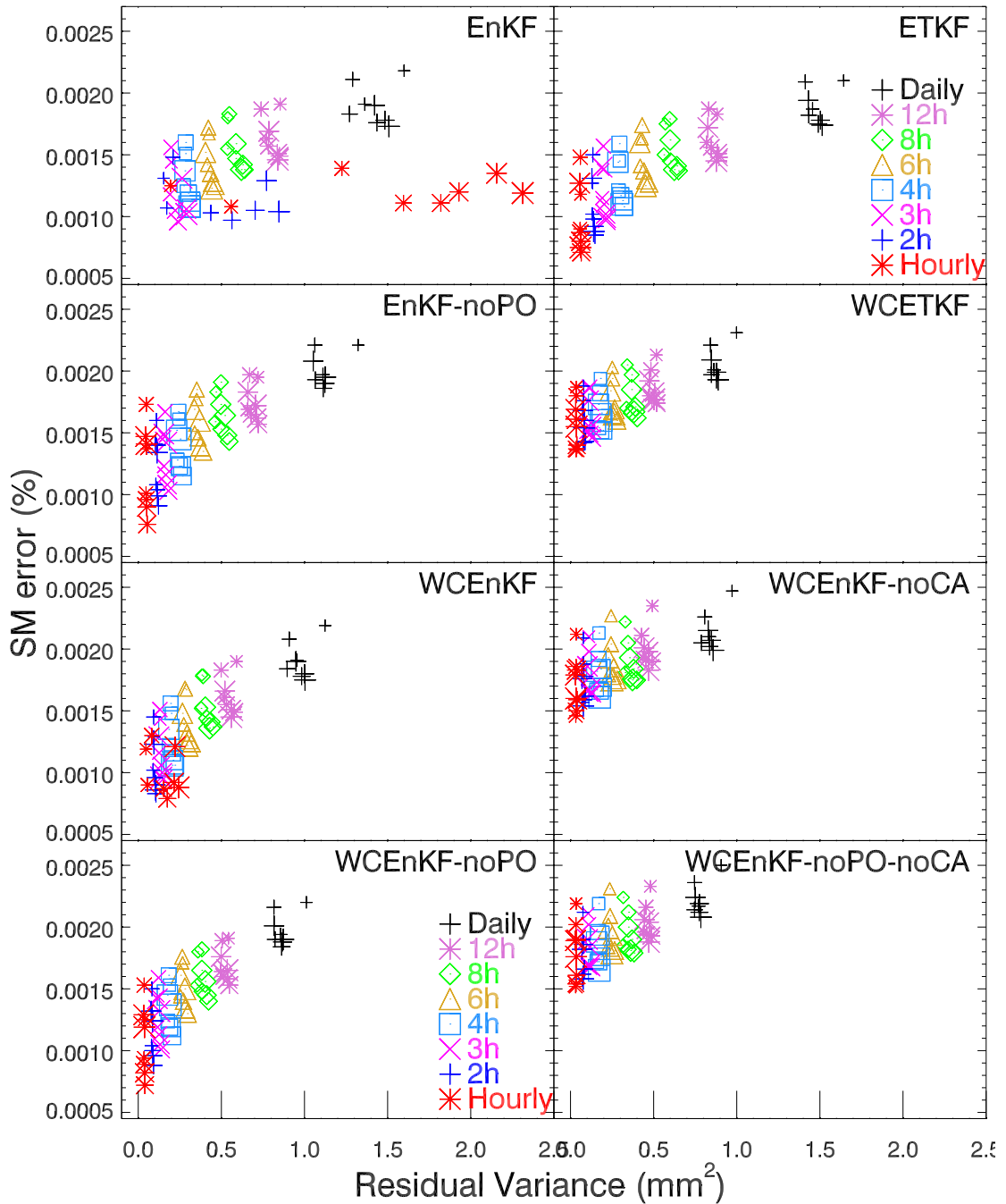
Figure 4.3: Sensitivities of the ETKF, EnKF, EnKF-noPO, WCETKF, WCEnKF, WCEnKF-noPO, WCEnKF-noCA, and WCEnKF-noPO-noCA filter residuals and SM errors to the ensemble size and assimilation frequency. Each panel represent a different filter; each color in each panel represent a different assimilation frequency varying from hourly to daily; and points with the same color and symbol represent a simulation with a different ensemble size (10, 15, 20, 30, 50, 80, 120, and 150), where the ensemble size increases with the increasing symbol size. All simulations are performed with both SM and ST observations were assimilated.

57

of the high frequency/high ensemble size scenario were not higher than the SM errors of low frequency/low ensemble size scenario, the residuals of these two scenarios differ dramatically for EnKF, which is not seen in any other filter (top left panel in Fig. 4.3). In fact the EnKF residuals of simulations with hourly assimilation frequencies and high ensemble size were higher than the EnKF residuals with daily assimilation frequencies and same number of ensemble size (Fig. 4.3). This behavior was unique to the standard EnKF with perturbed observations and hourly simulations (Fig. 4.3).

More frequent assimilation of observations reduces P_a (P_a and P_f will be used, in this chapter only, to refer analysis and forecast error covariance matrices respectively for both the standard and the constrained filters) more, than less frequent assimilation scenarios, which results in smaller Kalman gain (R and H are time invariant, and $K = P_a H^T R^{-1}$). When observations were assimilated every 1–hour, this was true for ETKF, EnKF-noPO, and WCEnKF but was not true for standard EnKF (Fig. 4.4). As a result of this higher than expected P_a for 1-hourly assimilation, Kalman gain of EnKF is also higher than other filters (4.5).

The P_a of the constrained filter is smaller than P_a of the standard filter with the term of $c_x \varphi^{-1} c_x^T$ (A.5), which is consistent with the results of 1-hourly and 6-hourly simulations (Fig. 4.4). Perturbation of observations inflates P_a which is partly balanced by the deflation of P_a with the constrained filter. Perhaps this is why WCEnKF did not have the drastic residual increase as EnKF did (Fig. 4.3).

Not perturbing observations, eliminating constraint anomalies, and using the constrained filter all have the same effect: to shrink P_a when compared to the standard filter. Constraining the filter with a water budget constraint does not have any effect on the optimality of the filter. However, not using the perturbed observations and not including the constraint anomalies does (because P_a is not consistent with its expected distribution). It can be expected from a non-optimal filter to have less
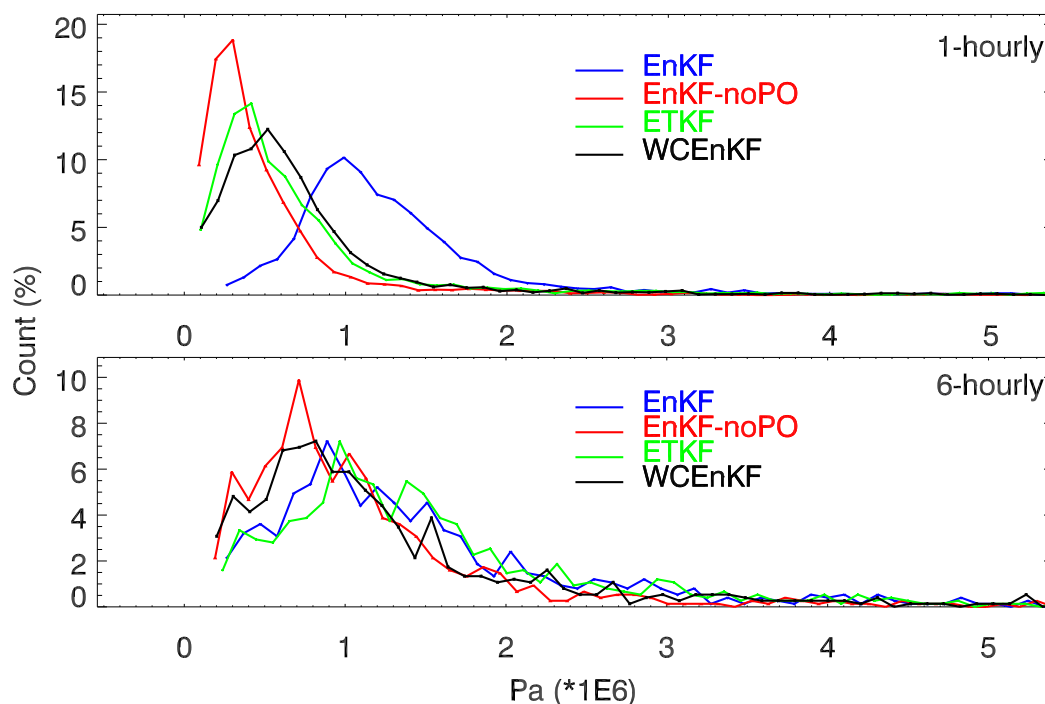
Figure 4.4: First layer soil moisture analysis error variance histogram for EnKF, ETKF, EnKF-noPO, and WCEnKF filters.

state prediction skill than the skill of the standard optimum filter. This could be the reason why SM errors of WCEnKF-noPO-noCA has slightly higher errors than the other filters (lower right panel of Fig. 4.3). In general, smaller residuals are favored by WCEnKF-noPO-noCA whereas smaller errors are favored by WCEnKF and EnKF-noPO (which is almost indistinguishable with the skill of standard EnKF); meaning depending on the purpose of the study (smaller state errors or residual errors), a different filter could be selected.

Due to the nature how K is calculated (nonlinear dependence on P_f), *on average* K is underestimated even though P_f may not be biased (assuming time invariant R). Whitaker and Hamill (2002) recognized this and showed the mean absolute error of P_a decreases with increasing ensemble size. Here we elaborate on the effect of this ensemble size effect over this underestimation of K. The variance of P_f around its
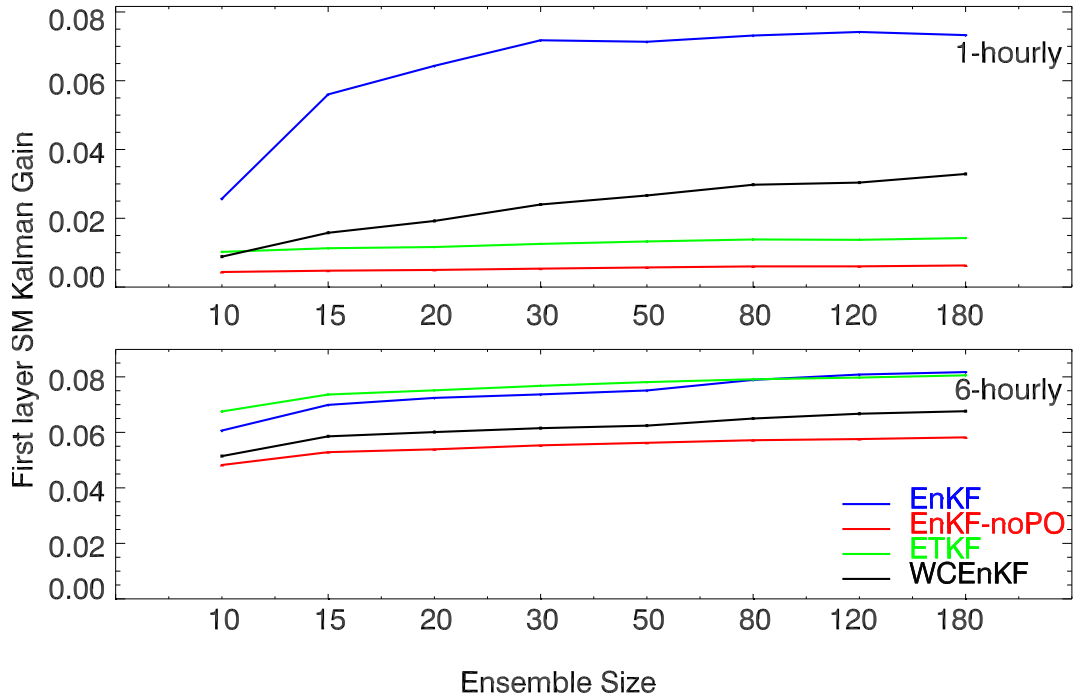
Figure 4.5: First layer soil moistrure Kalman gain change with the ensemble size change.

expected (or "true") value is larger for smaller ensemble sizes than than for larger ensemble sizes, where P_f should converge to its "true" value for infinite numbers of ensemble members. The magnitude of the underestimation of K depends on the degree of fluctuation of P_f around its expected value: for smaller ensemble sizes the degree of underestimation is larger than for larger ensemble sizes. It is expected that the residuals could be more reduced for smaller ensemble sizes than for higher ensemble sizes (smaller K = smaller residual). This is consistent with the results that the residuals increase with the increasing ensemble sizes (Fig. 4.5). This was particularly prominent for 1-hourly EnKF simulations (Fig. 4.5). Exact reasoning for this different behavior is unknown, but it is perhaps related with perturbation of observations (as only EnKF has it but not ETKF) when observations are assimilated too frequently (1–hourly EnKF has it but not 3– to 24–hourly assimilation simulations).

Chapter 5: Conclusions and Final Directions

5.1   Conclusions

In land data assimilation systems, the state updates produce a water budget imbalance, called a residual. In this study, a weakly constrained data assimilation solution was introduced to reduce the residual of standard EnKF [Evensen (1994)] and ETKF [Bishop et al. (2001)]. The solutions of these filters for the optimum state estimation can be found by minimizing a cost function which penalizes both the model forecast and the observation errors weighted by their error uncertainty. Similarly, constrained filter solutions (WCEnKF and WCETKF) were derived by minimizing a cost function that is the summation of three terms that represent the forecast errors, observation errors, and the water budget imbalance. These solutions were shown to be obtained in a single stage or in two stages where the first stage is the standard solution and the second stage is the constrained filter update. Two stage solution was shown to be identical to the single stage solution for WCEnKF where the analysis anomaly solutions of WCETKF differ for single and two stage solutions.

The minimization of the constraint cost function requires uncertainty estimates for the water balance elements (φ). This φ term was estimated through a flow dependent way from the ensembles. Optimality of φ was analyzed by inflating, deflating, and using constant values of φ and comparing the results of these analysis with the flow dependent method.

In standard EnKF, observations are perturbed to satisfy the equality of the expected value of analysis error covariance. In an attempt to further reduce the residuals, observations were not perturbed (acknowledging the filter is not optimum anymore) and the ensemble constraint anomalies were removed in the solutions. Accordingly additional new filters were obtained {EnKF-noPO [Whitaker and Hamill (2002)], WCEnKF-noPO, WCEnKF-noCA, and WCEnKF-noPO-noCA}. The role that the model forecast, perturbation of observations, and the Kalman gain play in the residual and state error performances of all these different filters were analyzed.

Major outcomes of this study can be summarized as follows:

- In general, the constrained solution affected the state RMSE only slightly when compared to unconstrained solution: constrained filter errors were indistinguishable from the unconstrained filter errors for the majority of the experiments.

- There is little-to-no improvement in ST errors when only SM observations are assimilated. There is also no improvement in SM errors and residuals when ST observations are also assimilated along with SM observations.

- Water balance residual variances of weakly constrained filters (WCEnKF and WCETKF) are smaller than that of unconstrained filters (ETKF or EnKF) regardless of the update frequency (daily or 3-hourly) or the assimilated variable (SM only, or SM and ST together) selection.

- There is no major difference found between single-stage WCETKF (with a symmetric square-root) and two-stage WCETKF (with symmetric square-root only in the first stage) when state errors and residuals are compared.

- Residuals remain unaffected from the ensemble size selection. Residuals can be decreased either by assimilating more frequent observations or by using the

constrained filter introduced in this study.

- Estimation of φ in a flow dependent way (3.6) did not give smaller SM errors and residuals when φ values were selected as a constant.

- Using constrained filter compensated for the lack of observations when the residual errors are concerned.

- Not perturbing the observations reduced the residuals while preserving the SM state prediction skill. In fact, the state errors of the scheme with non-perturbed observations were indistinguishable from the ones with the perturbed observations.

- Removing the anomalies of the constraints reduced the residuals, however increased the state errors, where the errors still remained well below the errors of the open loop and observation errors.

- Hourly EnKF residuals using large ensemble size were much higher than the residual of any other filter. This was primarily linked to the perturbation of observations, high assimilation frequency, and the ensemble size. Constraining the filter reduces this effect, whereas not perturbing the observations totally diminishes it.

Overall, it is suggested that the constrained filter should be used if the purpose of a particular study is concerned about the residuals. Furthermore, it is also suggested not to perturb the observations but keeping the constraint anomalies to reduce the residuals further without loosing the state prediction skill. Although the latter one is obtained through a non-optimal filter, this did not cause any prediction problems yet improved the simulations.

5.2 Future Directions and Applications

As with the water balance, land surface models also conserve the energy balance, but an imbalance occurs during assimilation as a result of the temperature state update. Although an energy balance constraint was not performed, the solution implemented in this study for water balance residuals also can be used to reduce the energy balance residuals. In general, data assimilation of hydrological states results in an inconsistency between the predicted diagnostic variables (i.e. evapotranspiration, runoff) and the updated prognostic variables. Diagnostic variables remain unaffected from the prognostic variable update in current hydrologic data assimilation schemes; unaffected diagnostic variables and the updated prognostic variables are model predictions for two different initial conditions. A remedy can be obtained by also updating the diagnostic variables (eg. evapotranspiration and runoff) along with the prognostic variables, where the error covariances for the diagnostic variables are estimated from the ensembles [Pan and Wood (2006)]. In this study, an idealized setup was used, where the model errors and the model parameterization errors were not taken into account. An alternative approach could have been a fraternal twin experiment, where the truths are generated in one model and the experiments are performed in another. In this study, flow dependent estimated φ did not give superior results over a constant value for φ. An alternative flow dependent methodology can be obtained where φ can be treated as a parameter to be optimized inside the Kalman Filter and be solved simultaneously with the estimated state.

In this study the residuals of the standard data assimilation techniques were reduced with a constrained filter. Building on this result, a better evapotranspiration and runoff estimations could be seeked using the presented methodology. Obtaining better fluxes through this method could be particularly important to modelers in

parameter calibration. The constrained solution introduced in this study could be very valuable to GEWEX community to obtain a better water and energy cycling information as this study lays a solution to reduce the uncertainty of the water and potentially energy budgets. In general, reanalysis data are used to obtain better analysis of historical data that were not available in the past; NCEP reanalysis [Kalnay et al. (1996)] is one of the early examples that produced 40 years of global atmospheric data. Data assimilation offers the ideal platform for reanalysis type of studies as new methods emerge. The introduced weakly constrained filter in this study could be used in reanalysis type of studies to acquire improved water and energy cycles. Weakly constrained assimilation can make the reanalysis products more valuable to the same community without making it less valuable to another community.

Appendix A: Constrained Filter

A.1 Single-Stage Constrained Filter

A.1.1 Single-Stage Constrained Kalman Filter

Similar to the traditional Kalman Filter, a constrained filter solution can be also obtained through minimizing the cost function in (3.5)

$$J_c = (o - Hx)^T R^{-1}(o - Hx) + (x - \mu_f)^T P_f^{-1}(x - \mu_f) + (\beta - c_x^T x)^T \varphi^{-1}(\beta - c_x^T x) \quad \text{(A.1)}$$

$$\frac{\partial J}{\partial x} = 2(H^T R^{-1} H + P_f^{-1} + c_x \varphi^{-1} c_x^T)x - 2(H^T R^{-1} o + P_f^{-1}\mu_f + c_x\varphi^{-1}\beta) \quad \text{(A.2)}$$

Setting derivation (A.2) to 0, the solution for the constrained filter can be expressed as

$$\mu_{aa} = (H^T R^{-1} H + P_f^{-1} + c_x \varphi^{-1} c_x^T)^{-1}(H^T R^{-1} o + P_f^{-1}\mu_f + c_x\varphi^{-1}\beta) \quad \text{(A.3)}$$

This equation can be used as a final solution to the constrained KF. However, the analogy with the standard KF is not obvious. Below, a constrained KF filter solution analogous to the standard solution was derived.

To ease the notation, we define $S^{-1} = H^T R^{-1} H + c_x \varphi^{-1} c_x^T$, then (A.3) becomes:

$$\mu_{aa} = (P_f^{-1} + S^{-1})^{-1}(H^T R^{-1} o + P_f^{-1}\mu_f + c_x\varphi^{-1}\beta) \quad \text{(A.4)}$$

The notation was eased further by using the second derivation of the cost function, which is equal to the inverse of the analysis error covariance matrix [Lorenc (1986),

and shown in (B.5)] of the constrained filter.

$$\frac{\partial^2 J}{\partial x^2} = P_{aa}^{-1} = P_f^{-1} + S^{-1}$$

$$P_{aa} = (P_f^{-1} + S^{-1})^{-1} \tag{A.5}$$

Hence, above equation (A.4) can be rewritten as

$$\mu_{aa} = P_{aa}(H^T R^{-1} o + P_f^{-1} \mu_f + c_x \varphi^{-1} \beta)$$

$$\mu_{aa} = P_{aa} H^T R^{-1} o + P_{aa} P_f^{-1} \mu_f + P_{aa} c_x \varphi^{-1} \beta \tag{A.6}$$

Before continuing the derivation from (A.6), another equality is introduced

$$P_{aa} = (P_f^{-1} + S^{-1})^{-1}$$

$$P_{aa}(P_f^{-1} + S^{-1}) = I$$

$$P_{aa} P_f^{-1} = I - P_{aa} S^{-1} \tag{A.7}$$

Using this equality in (A.7), (A.6) can be rewritten as

$$\mu_{aa} = P_{aa} H^T R^{-1} o + (I - P_{aa} S^{-1}) \mu_f + P_{aa} c_x \varphi^{-1} \beta \tag{A.8}$$

$$= P_{aa} H^T R^{-1} o + \mu_f - P_{aa} S^{-1} \mu_f + P_{aa} c_x \varphi^{-1} \beta$$

$$= \mu_f + P_{aa} H^T R^{-1} o - P_{aa}(H^T R^{-1} H + c_x \varphi^{-1} c_x^T) \mu_f + P_{aa} c_x \varphi^{-1} \beta$$

$$= \mu_f + P_{aa} H^T R^{-1}(o - H \mu_f) + P_{aa}(c_x \varphi^{-1} \beta - c_x \varphi^{-1} c_x^T \mu_f)$$

and the final constrained KF equation is obtained as

$$\mu_{aa} = \mu_f + P_{aa}H^T R^{-1}(o - H\mu_f) + P_{aa}c_x\varphi^{-1}(\beta - c_x^T\mu_f). \tag{A.9}$$

Similarly the analysis anomaly in an EnKF framework can be found as

$$X_{aa} = X_f + P_{aa}H^T R^{-1}(O' - HX_f) + P_{aa}c_x\varphi^{-1}(B' - c_x^T X_f). \tag{A.10}$$

where O' and B' are matrices holding the observation anomalies (namely random numbers used for the perturbations) and the constraint anomalies respectively.

Solution of the standard KF requires computation of a single inverse, $P_a H^T R^{-1}$, which can be equivalently derived as Kalman gain. Similarly, the solution of the constrained filter can be obtained through a single inverse, $P_{aa}H^T R^{-1}$, which also can be obtained through a single inverse,

$$P_{aa}H^T R^{-1} = (H^T R^{-1}H + P_f^{-1} + c_x\varphi^{-1}c_x^T)^{-1}H^T R^{-1}$$

$$P_{aa}H^T R^{-1} = P_f\big(I + (H^T R^{-1}H + c_x\varphi^{-1}c_x^T)P_f\big)^{-1}H^T R^{-1}$$

provided that the observation error covariance matrix (R) is assumed diagonal, hence its inverse is trivial.

Whitaker and Hamill (2002) showed that without the perturbation of observations, the analysis error covariance of EnKF is underestimated by a term of KRK^T. The term β holds the prognostic variables of the previous time-step analysis, fluxes, and the forcing data, where β is obtained from ensembles $(B' \neq 0)$. Hence, construction of perturbed constraints is not needed for the constrained filters.

A.1.2 Single-Stage Constrained Ensemble Transform Kalman Filter

Similar to the traditional ETKF solutions, WCETKF solution can also be obtained by using analysis error covariance matrix of the constrained filter.

$$P_{aa} = (P_f^{-1} + S^{-1})^{-1}$$

$$= P_f P_f^{-1} (P_f^{-1} + S^{-1})^{-1}$$

$$= P_f (P_f^{-1} P_f + S^{-1} P_f)^{-1}$$

$$= P_f (I + S^{-1} P_f)^{-1}$$

$$= X_f X_f^T (I + S^{-1} X_f * I * X_f^T)^{-1} \tag{A.11}$$

Using the Sherman-Morrison-Woodbury formula [a reminder for the reader $(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$], (A.11) can be rewritten as

$$P_{aa} = X_f X_f^T[I - I * S^{-1}X_f(I + X_f^T * I * S^{-1}X_f)^{-1}X_f^T * I]$$

$$= X_f[X_f^T - X_f^T * S^{-1}X_f(I + X_f^T S^{-1}X_f)^{-1}X_f^T]$$

$$= X_f[I - X_f^T * S^{-1}X_f(I + X_f^T S^{-1}X_f)^{-1}]X_f^T$$

$$= X_f[(I + X_f^T S^{-1}X_f)(I + X_f^T S^{-1}X_f)^{-1} - X_f^T * S^{-1}X_f(I + X_f^T S^{-1}X_f)^{-1}]X_f^T$$

$$= X_f[(I + X_f^T S^{-1}X_f - X_f^T S^{-1}X_f)(I + X_f^T S^{-1}X_f)^{-1}]X_f^T$$

$$P_{aa} = X_f(I + X_f^T S^{-1}X_f)^{-1}X_f^T \tag{A.12}$$

$$P_{aa} = X_f D X_f^T \tag{A.13}$$

where $D = (I + X_f^T S^{-1}X_f)^{-1}$. Using eigenvalue decomposition of $X_f^T S^{-1}X_f$ (U is eigenvectors and Λ is diagonal) and defining its square root as $D = A_{aa}A_{aa}^T$, this square root can be found $A_{aa} = U(I+\Lambda)^{-1/2}V^T$, where V^T is unitary. These equalities imply the solution for the anomaly of the analysis for the constrained filter can be rewritten as

$$X_{aa} = X_f A_{aa} \tag{A.14}$$

where this solution is also consistent with (A.13). After analysis anomaly is estimated, analysis mean of the WCETKF can be estimated also from (A.9).

A.2   Two-Stage Constrained Filter

In this section it is shown that the single-stage solution in (A.9) can equivalently be performed in two-recursive stages where the first stage is the standard KF equations and the second stage is the constrained filter adjustment.

A.2.1   Two-Stage Constrained Kalman Filter

Recalling (A.9)

$$\mu_{aa} = \mu_f + P_{aa} H^T R^{-1}(o - H\mu_f) + P_{aa} c_x \varphi^{-1}(\beta - c_x{}^T \mu_f)$$

and expending the terms

$$\mu_{aa} = \mu_f + (P_f^{-1} + H^T R^{-1}H + c_x \varphi^{-1}c_x^T)^{-1} H^T R^{-1}(o - H\mu_f) + (P_f^{-1} +$$
$$H^T R^{-1}H + c_x \varphi^{-1}c_x^T)^{-1} c_x \varphi^{-1}(\beta - c_x{}^T \mu_f) \quad \text{(A.15)}$$

Substituting inverse of the standard KF analysis error covariance $P_a^{-1} = P_f^{-1} + H^T R^{-1}H$, (A.15) becomes

$$\mu_{aa} = \mu_f + (P_a{}^{-1} + c_x \varphi^{-1}c_x^T)^{-1} H^T R^{-1}(o - H\mu_f) +$$
$$(P_a{}^{-1} + c_x \varphi^{-1}c_x^T)^{-1} c_x \varphi^{-1}(\beta - c_x{}^T \mu_f) \quad \text{(A.16)}$$

Using the Sherman-Morrison-Woodbury formula, $(P_a^{-1} + c_x \varphi^{-1} c_x^T)^{-1} = P_a - P_a c_x (\varphi + c_x^T P_a c_x)^{-1} c_x^T P_a$, (A.16) becomes

$$\mu_{aa} = \mu_f + \left(P_a - P_a c_x (\varphi + c_x^T P_a c_x)^{-1} c_x^T P_a\right) H^T R^{-1}(o - H\mu_f) +$$

$$\left(P_a - P_a c_x (\varphi + c_x^T P_a c_x)^{-1} c_x^T P_a\right) c_x \varphi^{-1}(\beta - c_x^T \mu_f)$$

$$\mu_{aa} = \mu_f + \left[P_a H^T R^{-1}(o - H\mu_f) - P_a c_x (\varphi + c_x^T P_a c_x)^{-1} c_x^T P_a H^T R^{-1}(o - H\mu_f)\right] +$$

$$\left[P_a c_x \varphi^{-1}(\beta - c_x^T \mu_f) - P_a c_x (\varphi + c_x^T P_a c_x)^{-1} c_x^T P_a c_x \varphi^{-1}(\beta - c_x^T \mu_f)\right]$$

$$\mu_{aa} = \mu_f + P_a H^T R^{-1}(o - H\mu_f) + P_a c_x \left[\varphi^{-1}(\beta - c_x^T \mu_f) - (\varphi + c_x^T P_a c_x)^{-1}\right.$$

$$\left. c_x^T P_a H^T R^{-1}(o - H\mu_f) - (\varphi + c_x^T P_a c_x)^{-1} c_x^T P_a c_x \varphi^{-1}(\beta - c_x^T \mu_f)\right]$$

$$\mu_{aa} = \mu_f + P_a H^T R^{-1}(o - H\mu_f) + P_a c_x (\varphi + c_x^T P_a c_x)^{-1}$$

$$\left[(\varphi + c_x^T P_a c_x)\varphi^{-1}(\beta - c_x^T \mu_f) - c_x^T P_a H^T R^{-1}(o - H\mu_f) - c_x^T P_a c_x \varphi^{-1}(\beta - c_x^T \mu_f)\right]$$

$$\mu_{aa} = \mu_f + P_a H^T R^{-1}(o - H\mu_f) + P_a c_x (\varphi + c_x^T P_a c_x)^{-1}$$

$$\left[(\beta - c_x^T \mu_f) + c_x^T P_a c_x \varphi^{-1}(\beta - c_x^T \mu_f) - c_x^T P_a H^T R^{-1}(o - H\mu_f) - c_x^T P_a c_x \varphi^{-1}(\beta - c_x^T \mu_f)\right]$$

$$\mu_{aa} = \mu_f + P_a H^T R^{-1}(o - H\mu_f) + P_a c_x (\varphi + c_x^T P_a c_x)^{-1}$$

$$\left[(\beta - c_x^T \mu_f) - c_x^T P_a H^T R^{-1}(o - H\mu_f)\right]$$

$$\mu_{aa} = \mu_f + P_a H^T R^{-1}(o - H\mu_f) + P_a c_x (\varphi + c_x^T P_a c_x)^{-1}$$

$$\left[\beta - c_x^T \left(\mu_f - P_a H^T R^{-1}(o - H\mu_f)\right)\right] \tag{A.17}$$

The two-stage solution can be written as

$$\mu_{aa} = \mu_a + P_a c_x (\varphi + c_x^T P_a c_x)^{-1} (\beta - c_x^T \mu_a) \tag{A.18}$$

where $\mu_a = \mu_f + P_a H^T R^{-1}(o - H\mu_f)$ is the standard KF solution without the constraint.

This solution implies that the constraint can be performed in two sequential stages: the first stage is the standard KF (μ_a) without the constraint and the second stage is the constrained filter $P_a c_x (\varphi + c_x^T P_a c_x)^{-1} (\beta - c_x^T \mu_a)$.

A comparison of the residual terms $(\beta - c_x^T x)$ of the constrained and standard filters can be performed using the two-stage solution in (A.18).

$$\mu_{aa} = \mu_a + P_a c_x (\varphi + c_x^T P_a c_x)^{-1} (\beta - c_x^T \mu_a)$$

$$-c_x^T \mu_{aa} = -c_x^T \mu_a - c_x^T P_a c_x (\varphi + c_x^T P_a c_x)^{-1} (\beta - c_x^T \mu_a)$$

$$\beta - c_x^T \mu_{aa} = \beta - c_x^T \mu_a - c_x^T P_a c_x (\varphi + c_x^T P_a c_x)^{-1} (\beta - c_x^T \mu_a)$$

$$\beta - c_x^T \mu_{aa} = [I - c_x^T P_a c_x (\varphi + c_x^T P_a c_x)^{-1}](\beta - c_x^T \mu_a)$$

$$\beta - c_x^T \mu_{aa} = \left[(\varphi + c_x^T P_a c_x) - c_x^T P_a c_x \right] (\varphi + c_x^T P_a c_x)^{-1} (\beta - c_x^T \mu_a)$$

$$\beta - c_x^T \mu_{aa} = \varphi (\varphi + c_x^T P_a c_x)^{-1} (\beta - c_x^T \mu_a) \tag{A.19}$$

For scalar $c_x^T P_a c_x > 0$, $\varphi(\varphi + c_x^T P_a c_x)^{-1} < 1$; hence the constraint shrinks the residual of the standard filter toward zero by a rate that depends on φ.

It is noted that for $(lim_\varphi \to \infty)$, the second term in (A.18) vanishes, and the constrained filter solution equals to the standard KF solution.

Moreover, setting $\varphi = 0$ in (A.18), strongly constrained Kalman Filter solution is

obtained as

$$\mu_{aa} = \mu_a + K_s(\beta - c_x^T \mu_a) \tag{A.20}$$

where $K_s = P_a c_x (c_x^T P_a c_x)^{-1}$. This strongly constrained solution in (A.20) is identical with the maximum probability method constrained solution of Simon and Chia (2002) [(25) in their paper].

A.2.2 Two-stage Constrained Ensemble Transform Kalman Filter

Two stage solution of the state anomalies for the WCETKF can be found using the inverse of the analysis error covariance of the constrained filter:

$$P_{aa}^{-1} = P_f^{-1} + H^T R^{-1} H + c_x \varphi^{-1} c_x^T$$

$$P_{aa}^{-1} = P_a^{-1} + c_x \varphi^{-1} c_x^T$$

Taking the inverse of both sides

$$P_{aa} = (P_a^{-1} + c_x \varphi^{-1} c_x^T)^{-1}$$

Using the Sherman-Morrison-Woodbury formula, this can be written as

$$P_{aa} = P_a - P_a c_x (\varphi + c_x^T P_a c_x)^{-1} c_x^T P_a$$

$$P_{aa} = X_a (I - X_a^T c_x (\varphi + c_x^T P_a c_x)^{-1} c_x^T X_a) X_a^T$$

$$X_{aa} X_{aa}^T = X_a (I - z \alpha z^T) X_a^T$$

where X_{aa} is the analysis anomaly of the constrained filter, $\alpha = (\varphi + c_x^T P_a c_x)^{-1}$ is a scalar, and $z = X_a^T c_x$. A square root can be found analytically by finding a scalar (δ) such that

$$(I + \delta z z^T)(I + \delta z z^T)^T = I - z \alpha z^T$$

and rearranging the terms on both sides as

$$(\gamma \delta^2 + 2\delta + \alpha) z z^T = 0$$

where $\gamma = z^T z$ is a scalar and the solution is found as

$$\delta_\pm = \frac{-1 \pm \sqrt{1 - \alpha \gamma}}{\gamma}$$

This quadratic form gives two solutions, but only one of them produces a positive definite square root. To determine the correct choice, we choose the solution that renders

$$z^T (I + \delta_\pm z z^T) z > 0$$

where the above quadratic form checks for the positive definiteness of $(I + \delta_\pm z z^T)$ for vector z. Rearranging the above equation,

$$z^T z + \delta_\pm z^T z z^T z > 0$$

Replacing $\delta_\pm$

$$\gamma(1 - 1 \pm \sqrt{1 - \alpha\gamma}) > 0$$

$$\pm\sqrt{1 - \alpha\gamma} > 0$$

Hence the positive root is selected:

$$X_{aa} = X_a(I + \delta_+ zz^T)$$

$$= X_a\left[+\frac{-1 + \sqrt{1 - (\varphi + c_x^T P_a c_x)^{-1} c_x^T P_a c_x}}{c_x^T P_a c_x} X_a{}^T c_x c_x^T X_a\right]$$

The final solution for the constrained filter analysis anomaly can be found as

$$X_{aa} = X_a\left[I + X_a{}^T c_x c_x^T X_a\left(-1 + \sqrt{\varphi(\varphi + c_x^T P_a c_x)^{-1}}\right)/c_x^T P_a c_x\right] \tag{A.21}$$

which can be also rewritten as

$$X_{aa} = X_a E = X_f A E \tag{A.22}$$

where A is the square-root multiplier matrix that is estimated from the standard ETKF equations and E is the matrix obtained from the operations within the square-brackets on the rhs of (A.21). This equation implies the two-stage analysis anomaly of WCETKF (X_{aa}) can be obtained by first solving for the analysis anomaly of the standard ETKF (X_a) and then multiplying it by the matrix E.

Similar to the strongly constrained KF solution, a strongly constrained ETKF

solution can be estimated by setting φ in (A.21) into 0 as

$$X_{aa} = X_a - P_a c_x c_x^T X_a / c_x^T P_a c_x \tag{A.23}$$

WCETKF analysis anomaly of single-stage (A.14) and two-stage (A.21) solutions differ, although they have identical solutions for the analysis error covariance matrix P_{aa}. In fact, these single-stage and two-stage solutions are two different square root filters with the same error covariance matrixes but with different state analysis anomalies. It is fairly easy to make the single-stage WCETKF square-root A_{aa} (A.14) symmetric with the selection of $V^T = U$; whereas for the two-stage filter, it is not immediately clear which selection for the V^T matrix would make the AE term in (A.22) symmetric. On the other hand, it is stressed that WCEnKF solutions are identical for both single-stage (A.9) and two-stage (A.18) constrained filters.

Computationally, both standard (B.6) and two-stage constrained (A.18) KF solutions require single inverse $(HP_f H^T + R)$, where the single-stage constrained KF solution (A.9) requires two inverses $[(I + S^{-1}P_f)$ and $R]$. Although the inverse of R can be avoided by a diagonal observation error covariance matrix assumption, the dimension of the term to be inverted is higher for the single-stage constrained KF solution than it is for other two solutions (assuming not all state variables are observed). Hence computationally, the two-stage solution is similar to the standard KF whereas the single-stage KF solution is more expensive. The load for the square root filters is the same for all standard ETKF, single-stage WCETKF, and two-stage WCETKF solutions. They all require single inverse (R) and single eigenvalue decomposition. Standard ETKF and two-stage WCETKF solutions require the eigenvalue decomposition of the term $X_f^T H^T R^{-1} H X_f$ (2.14); this term for the single-stage WCETKF solution is $X_f^T (H^T R^{-1} H + c_x \varphi^{-1} c_x^T) X_f$ (A.12).

Appendix B: Useful matrix identities and matrix equalities

B.1   Matrix Derivation Identities

Although complete list of matrix idetities can be found in numerous of references, it is useful to briefly list some of them as they form the basis of the above derivations in the first part of the appendix. For a random variable X and constants vectors c,d, and e,

$$\frac{\partial(X^T c)}{\partial x} = \frac{\partial(c^T X)}{\partial x} = c$$

$$\frac{\partial(X^T X)}{\partial x} = 2X$$

$$\frac{\partial(c^T X d)}{\partial x} = cd^T$$

$$\frac{\partial(X^T c X)}{\partial x} = (c + c^T)X$$

$$\frac{\partial(c^T X^T d)}{\partial x} = dc^T$$

$$\frac{\partial(c^T X^T X d)}{\partial x} = Xcd^T + Xdc^T$$

$$\frac{\partial(c^T X^T e X d)}{\partial x} = eXcd^T + eXdc^T \tag{B.1}$$

B.2 Best Guess in Mean Square Sense

For a random variable x, the best estimate of x that minimizes $E[(x-k)^2]$ can be found

$$E[(x-k)^2] = E[(x - E[x] + E[x] - k)^2]$$

$$= E[(x - E[x])^2 + (E[x] - k)^2 + 2(x - E[x])(E[x] - k)]$$

$$= E[(x - E[x])^2] + E[(E[x] - k)^2] + E[2(x - E[x])E[x]] - E[2(x - E[x])k]$$

$$= E[(x - E[x])^2] + E[(E[x] - k)^2] \tag{B.2}$$

The first term is the variance of x and the best guess can only affect the sum through the second term. Hence we choose the best estimate (k) of x as $E[x]$, which vanishes the second term and minimizes $E[(x-k)^2]$.

B.3 Hessian and Analysis Covariance Inverse

Below, it is shown that the Hessian (second derivative of the cost function) is equal to the inverse of the analysis error covariance matrix.

$$J = (o - H\mu_a)^T R^{-1}(o - H\mu_a) + (\mu_a - \mu_f)^T P_f^{-1}(\mu_a - \mu_f)$$

$$\frac{\partial J}{\partial \mu} = -H^T R^{-1}(o - H\mu_a) + P_f^{-1}(\mu_a - \mu_f) \tag{B.3}$$

$$\frac{\partial^2 J}{\partial \mu^2} = H^T R^{-1} H + P_f^{-1} \tag{B.4}$$

where the second derivative of the cost function (B.4) is called Hesssian. From (B.3),

$$- H^T R^{-1}(o - H\mu_t + H\mu_t - H\mu_a) + P_f^{-1}(\mu_a - \mu_t + \mu_t - \mu_f) = 0$$

$$- H^T R^{-1}(o - H\mu_t) - H^T R^{-1} H(\mu_t - \mu_a) + P_f^{-1}(\mu_a - \mu_t) + P_f^{-1}(\mu_t - \mu_f) = 0$$

$$- H^T R^{-1}(o - H\mu_t) + P_f^{-1}(\mu_t - \mu_f) = H^T R^{-1} H(\mu_t - \mu_a) + P_f^{-1}(\mu_a - \mu_t)$$

$$- H^T R^{-1}(o - H\mu_t) + P_f^{-1}(\mu_t - \mu_f) = (H^T R^{-1} H + P_f^{-1})(\mu_t - \mu_a)$$

Multiplying both sides by their transposes, and taking the expectation of both sides, the terms with $(o - H\mu_t)(\mu_t - \mu_f)$ multiplication would vanish,

$$(H^T R^{-1} H + P_f^{-1}) = (H^T R^{-1} H + P_f^{-1}) P_a (H^T R^{-1} H + P_f^{-1})$$

Hence, the Hessian (B.4) and the analysis error covariance equality is found

$$P_a = (H^T R^{-1} H + P_f^{-1})^{-1} \tag{B.5}$$

B.4 Standard Kalman Filter Solution

Solution of Kalman Filter equations can be found by setting the first derivation of the cost function (2.5)

$$\frac{\partial J}{\partial x} = 2(H^T R^{-1} H + P_f^{-1})x - 2(H^T R^{-1} o + P_f^{-1} \mu_f) = 0$$

Using the Sherman-Morrison-Woodbury formula,

$$\mu_a = \left[P_f - P_f H^T (R + H P_f H^T)^{-1} H P_f\right](H^T R^{-1} o + P_f^{-1} \mu_f)$$

$$\mu_a = \left[P_f - P_f H^T (R + H P_f H^T)^{-1} H P_f\right] H^T R^{-1} o +$$

$$\left[P_f - P_f H^T (R + H P_f H^T)^{-1} H P_f\right] P_f^{-1} \mu_f$$

$$\mu_a = P_f H^T \left[I - (R + H P_f H^T)^{-1} H P_f H^T\right] R^{-1} o + \left[I - P_f H^T (R + H P_f H^T)^{-1} H P_f\right]\mu_f$$

$$\mu_a = P_f H^T \left[(R + H P_f H^T)^{-1}\left((R + H P_f H^T) - H P_f H^T\right)\right] R^{-1} o +$$

$$\mu_f - P_f H^T (R + H P_f H^T)^{-1} H \mu_f$$

$$\mu_a = P_f H^T (R + H P_f H^T)^{-1} o + \mu_f - P_f H^T (R + H P_f H^T)^{-1} H \mu_f$$

$$\mu_a = \mu_f + P_f H^T (R + H P_f H^T)^{-1} (o - H \mu_f)$$

$$\mu_a = \mu_f + K(o - H \mu_f) \tag{B.6}$$

B.5 Kalman Gain in Square Root Filters

In square root filters, Kalman gain can be computed without the need of an extra inverse once the eigenvalue decomposition (to find the square root) is performed,

$$K = P_f H^T (R + H P_f H^T)^{-1}$$

$$K = P_f H^T (R + H P_f H^T)^{-1} R R^{-1}$$

$$K = P_f H^T \left[(R + H P_f H^T)^{-1} (R + H P_f H^T - H P_f H^T) \right] R^{-1}$$

$$K = P_f H^T \left[(R + H P_f H^T)^{-1} (R + H P_f H^T) - (R + H P_f H^T)^{-1} H P_f H^T \right] R^{-1}$$

$$K = P_f H^T \left[I - (R + H P_f H^T)^{-1} H P_f H^T \right] R^{-1}$$

$$K = \left[P_f H^T - P H^T (R + H P_f H^T)^{-1} H P_f H^T \right] R^{-1}$$

$$K = \left[P_f - P H^T (R + H P_f H^T)^{-1} H P_f \right] H^T R^{-1}$$

$$K = \left[X_f X_f^T - X_f X_f^T H^T (R + H X_f X_f^T H^T)^{-1} H X_f X_f^T \right] H^T R^{-1}$$

$$K = X_f \left[I - X_f^T H^T (R + H X_f * I * X_f^T H^T)^{-1} H X_f \right] X_f^T H^T R^{-1}$$

$$K = X_f \left[(I + X_f^T H^T R^{-1} H X_f)^{-1} \right] X_f^T H^T R^{-1}$$

$$K = X_f D X_f^T H^T R^{-1} \tag{B.7}$$

Hence, once the eigenvalue decomposition of the term $X_f^T H^T R^{-1} H X_f$ is performed, the Kalman gain can be calculated without an extra cost.

Bibliography

Bibliography

Alsdorf, D. E., E. Rodriguez, and D. P. Lettenmaier, 2007: Measuring surface water from space. *Rev. Geophys.*, **45 (2)**, RG2002.

Anderson, J. L., 2001: An ensemble adjustment filter for data assimilation. *Monthly Weather Review*, **129**, 2884–2903.

Anderson, J. L., 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus*, **59 (2)**, 210–224.

Anderson, J. L. and S. L. Anderson, 1999: A monte carlo implementation of the non-linear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, **127 (12)**, 2741–2758.

Andreadis, K. M. and D. P. Lettenmaier, 2006: Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources*, **29 (6)**, 872–886.

Barbieri, R. and P. Schopf, 1982: Oceanic applications of kalman filter. *Tech. Memo. TM83993, NASA Goddard Space Flight Center, Greenbelt, MD.*

Bishop, C. H., B. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform kalman filter. part i: Theoretical aspects. *Monthly Weather Review*, **129**, 420–436.

Budgell, W. P., 1986: Nonlinear data assimilation for shallow water equations in branched channels. *Journal of Geophysical Research*, **91 (C9)**, 10 633–10 644.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review*, **126**, 1719–1724.

Cosgrove, B. A., et al., 2003: Real-time and retrospective forcing in the north american land data assimilation system (NLDAS) project. *Journal of Geophysical Research*, **108 (D22)**, 8842.

Crow, W. T., 2003: Correcting land surface model predictions for the impact of temporally sparse rainfall rate measurements using an ensemble kalman filter and surface brightness temperature observations. *Journal of Hydrometeorology*, **4 (5)**, 960–973.

Crow, W. T., R. Bindlish, and T. J. Jackson, 2005: The added value of spaceborne passive microwave soil moisture retrievals for forecasting rainfall-runoff partitioning. *Journal of Geophysical Research*, **32**, L18 401.

Crow, W. T. and J. D. Bolten, 2007: Estimating precipitation errors using spaceborne surface soil moisture retrievals. *Geophysical Research Letters*, **34**, L08 403.

Crow, W. T. and R. H. Reichle, 2008: Comparison of adaptive filtering techniques for land surface data assimilation. *Water Resources Research*, **44**, W08 423.

Crow, W. T. and D. Ryu, 2009: A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals. *Hydrol. Earth Syst. Sci.*, **13**, 1–16.

Crow, W. T. and M. J. van den Berg, 2010: Estimating observation and model error parameters for soil moisture data assimilation (in press). *Water Resources Research*.

De Lannoy, G. J. M., P. R. Houser, N. E. C. Verhoest, and V. R. N. Pauwels, 2009: Adaptive soil moisture profile filtering for horizontal information propagation in the independent column-based clm2.0. *Journal of Hydrometeorology*, **10 (3)**, 766–779.

De Lannoy, G. J. M., R. H. Reichle, P. R. Houser, K. R. Arsenault, N. E. C. Verhoest, and V. R. N. Pauwels, 2010: Satellite-scale snow water equivalent assimilation into a high-resolution land surface model. *Journal of Hydrometeorology*, **11 (2)**, 352–369.

De Lannoy, G. J. M., R. H. Reichle, P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest, 2007: Correcting for forecast bias in soil moisture assimilation with the ensemble kalman filter. *Water Resources Research*, **43**, W09 410.

DelSole, T. and M. K. Tippett, 2010: *Statistical Methods in Climate Science*. unpublished book.

DelSole, T., M. Zhao, and P. Dirmeyer, 2009: A new method for exploring coupled land-atmosphere dynamics. *Journal of Hydrometeorology*, **10 (4)**, 1040–1050.

Dirmeyer, P. A., 2003: The role of the land surface background state in climate predictability. *Journal of Hydrometeorology*, **4**, 599–610.

Dirmeyer, P. A., 2006: The hydrologic feedback pathway for landclimate coupling. *Journal of Hydrometeorology*, **7 (5)**, 857–867.

Dirmeyer, P. A., A. J. Dolman, and N. Sato, 1999: The pilot phase of the global soil wetness project. *Bulletin of the American Meteorological Society*, **80 (5)**, 851–878.

Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki, 2006: Gswp-2: Multimodel analysis and implications for our perception of the land surface. *Bulletin of the American Meteorological Society*, **87 (10)**, 1381–1397.

Dirmeyer, P. A., C. A. Schlosser, and K. L. Brubaker, 2009: Precipitation, recycling, and land memory: An integrated analysis. *Journal of Hydrometeorology*, **10 (1)**, 278–288.

Dorman, J. L. and P. J. Sellers, 1989: A global climatology of albedo, roughness length and stomatal resistance for atmospheric general circulation models as represented by the simple biosphere model (sib). *Journal of Applied Meteorology*, **28 (9)**, 833–855.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayand, and J. D. Tarpley, 2003: Implementation of noah land surface model advances in the national centers for environmental prediction operational mesoscale eta model. *Journal of Geophysical Research*, **108 (D22)**, 8851.

Evensen, G., 1992: Using the extended kalman filter with a multilayer quasi-geostrophic ocean model. *Journal of Geophysical Research*, **97 (C11)**, 905–924.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research*, **99 (C5)**, 10 143–10 162.

Evensen, G., 1997: Advanced data assimilation for strongly nonlinear dynamics. *Monthly Weather Review*, **125 (6)**, 1342–1354.

Evensen, G., 2004: Sampling strategies and square root analysis schemes for the enkf. *Ocean Dynamics*, **54**, 539–560.

Evensen, G., 2009: *Data Assimilation: The Ensemble Kalman Filter*. Springer-Verlag.

Evensen, G. and P. J. van Leeuwen, 1996: Assimilation of geosat altimeter data for

the agulhas current using the ensemble kalman filter with a quasigeostrophic model. *Monthly Weather Review*, **124 (1)**, 85–96.

Friedland, B., 1969: Treatment of bias in recursive filtering. *IEEE Transactions on Automatic Control*, **14 (4)**, 359 – 367.

Gauss, C., 1963: *Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections*. Dover, 326 pp.

Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter. *Monthly Weather Review*, **129 (11)**, 2776–2790.

Houser, P. R., 2003: *Assimilation of Land Surface Data in Data Assimilation for the Earth Systems*. Kluwer Academic Publishers, 331-343 pp.

Houser, P. R., W. J. Shuttleworth, J. S. Famiglietti, H. V. Gupta, K. H. Syed, and D. C. Goodrich, 1998: Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Resources Research*, **34 (12)**, 3405–3420.

Houtekamer, P. L. and H. L. Mitchell, 1998: Data Assimilation Using an Ensemble Kalman Filter Technique. *Monthly Weather Review*, **126**, 796–811.

Houtekamer, P. L. and H. L. Mitchell, 2001: A sequential ensemble kalman filter for atmospheric data assimilation. *Monthly Weather Review*, **129 (1)**, 123–137.

Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997: Unified notation for data assimilation. *Journal of the Meteorological Society of Japan*, **75 (1B)**, 181–189.

Jackson, T. J., et al., 1999: Soil moisture mapping at regional scales using microwave radiometry: The southern great plains hydrology experiment. *IEEE Transactions on Geoscience and Remote Sensing*, **37**, 2136–2150.

Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory*, Vol. 64. Academic Press, 376 pp.

Kalman, R., 1960: A new approach to linear filtering and prediction problems. *Trans. ASME, J. Basic Eng.*, **82D**, 35–45.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.

Kalnay, E., et al., 1996: The ncep/ncar 40-year reanalysis project. *Bulletin of the American Meteorological Society*, **77 (3)**, 437–471.

Koren, V., J. Schaake, K. Mitchell, Q.-Y. Duan, F. Chen, and J. Baker, 1999: A parameterization of snowpack and frozen ground intended for ncep weather and climate models. *Journal of Geophysical Research*, **104 (D16)**, 19 569–19 585.

Koster, R. D., et al., 2004: Regions of strong coupling between soil moisture and precipitation. *Science*, **305 (5687)**, 1138–1140.

Lacarra, J.-F. and O. Talagrand, 1988: Short-range evolution of small perturbations in a barotropic model. *Tellus A*, **40 (2)**, 81–95.

Lakshmi, V., 2000: A simple surface temperature assimilation scheme for use in land surface models. *Water Resources Research*, **36 (12)**, 3687–3700.

Legendre, A., 1806: *Nouvelles Methodes pour la determination des orbites des cometes*. Courcier.

Li, H., E. Kalnay, and T. Miyoshi, 2009: Simultaneous estimation of covariance inflation and observation errors within an ensemble kalman filter. *Quarterly Journal of the Royal Meteorological Society*, **135**, 523–533.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, **112**, 1177–1194.

Maybeck, P. S., 1982: *Stochastic models, estimation, and control*, Vol. 2. Academic Press.

Miller, R. N., M. Ghil, and F. Gauthiez, 1994: Advanced data assimilation in strongly nonlinear dynamical systems. *Journal of the Atmospheric Sciences*, **51 (8)**, 1037–1056.

Ott, E., et al., 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **A56**, 415–428.

Pan, M. and E. F. Wood, 2006: Data assimilation for estimating the terrestrial water budget using a constrained ensemble kalman filter. *Journal of Hydrometeorology*, **7 (3)**, 534–547.

Pauwels, V. R. N. and G. J. M. De Lannoy, 2006: Improvement of modeled soil wetness conditions and turbulent fluxes through the assimilation of observed discharge. *Journal of Hydrometeorology*, **7 (3)**, 458–477.

Petersen, D. P., 1968: On the concept and implementation of sequential analysis for linear random fields. *Tellus*, **20 (4)**, 673–686.

Reichle, R., D. McLaughlin, and D. Entekhabi, 2001: Variational data assimilation of microwave radiobrightness observations for land surface hydrology applications. *IEEE Transactions on Geoscience and Remote Sensing*, **39 (8)**, 1708 – 1718.

Reichle, R. H., 2008: Data assimilation methods in the earth sciences. *Advances in Water Resources*, **31 (11)**, 1411–1418.

Reichle, R. H., W. T. Crow, and C. L. Keppenne, 2008: An adaptive ensemble kalman filter for soil moisture data assimilation. *Water Resources Research*, **44 (3)**, W03 423.

Reichle, R. H., J. P. Walker, R. D. Koster, and P. R. Houser, 2002: Extended versus ensemble kalman filtering for land data assimilation. *Journal of Hydrometeorology*, **3 (6)**, 728–740.

Roads, J., et al., 2003: GCIP water and energy budget synthesis (WEBS). *Journal of Geophysical Research*, **108 (D16)**, 8609.

Rodell, M. and P. R. Houser, 2004: Updating a land surface model with modis-derived snow cover. *Journal of Hydrometeorology*, **5 (6)**, 1064–1075.

Ryu, D., W. T. Crow, X. Zhan, and T. J. Jackson, 2009: Correcting unintended perturbation biases in hydrologic data assimilation. *Journal of Hydrometeorology*, **10 (3)**, 734–750.

Sellers, P. J., Y. Mintz, Y. C. Sud, and A. Dalcher, 1986: A simple biosphere model (sib) for use within general circulation models. *Journal of the Atmospheric Sciences*, **43 (6)**, 505–531.

Simon, D. and T. L. Chia, 2002: Kalman filtering with state equality constraints. *IEEE Transactions on Aerospace and Electronic Systems*, **38 (2)**, 128–136.

Sorenson, H. W., 1970: Least-squares estimation: from gauss to kalman. *IEEE spectrum*, **7**, 63–68.

Talagrand, O., 1997: Assimilation of observations: An introduction. *Journal of the Meteorological Society of Japan*, **75**, 191–209.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square-root filters. *Monthly Weather Review*, **131**, 1485–1490.

van den Hurk, B. J. J. M., L. Jia, C. Jacobs, M. Menenti, and Z.-L. Li, 2002: Assimilation of land surface temperature data from atsr in an nwp environment - a case study. *International Journal of Remote Sensing*, **23 (24)**, 5193–5209.

Walker, J. P. and P. R. Houser, 2001: A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations. *Journal of Geophysical Research*, **106 (D11)**, 11 761–11 774.

Walker, J. P. and P. R. Houser, 2004: Requirements of a global near-surface soil moisture satellite mission: accuracy, repeat time, and spatial resolution. *Advances in Water Resources*, **27**, 785–801.

Wang, X. and C. H. Bishop, 2003: A comparison of breeding and ensemble transform kalman filter ensemble forecast schemes. *Journal of the Atmospheric Sciences*, **60 (9)**, 1140–1158.

WCRP JSC Report, 2010: WCRP joint scientific committee 31st session report. Tech. rep., WCRP.

Wei, J., P. A. Dirmeyer, and Z. Guo, 2010: How much do different land models matter for climate simulation? part ii: A decomposed view of the landatmosphere coupling strength. *Journal of Climate*, **23 (11)**, 3135–3145.

Whitaker, J. and T. M. Hamill, 2002: Ensemble Data Assimilation Without Perturbed Observations. *Monthly Weather Review*, **130**, 1913–1924.

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the ncep global forecast system. *Monthly Weather Review*, **136 (2)**, 463–482.

Yilmaz, M. T., T. Delsole, and P. R. Houser, 2010: Improving land data assimilation performance with a water budget constraint (submitted). *Journal of Hydrometeorology.*

Zaitchik, B. F., M. Rodell, and R. H. Reichle, 2008: Assimilation of grace terrestrial water storage data into a land surface model: Results for the mississippi river basin. *Journal of Hydrometeorology*, **9 (3)**, 535–548.

Zhan, X., P. Houser, J. Walker, and W. Crow, 2006: A method for retrieving high-resolution surface soil moisture from hydros l-band radiometer and radar observations. *IEEE Transactions on Geoscience and Remote Sensing*, **44 (6)**, 1534 – 1544.

Curriculum Vitae

M. Tugrul Yilmaz graduated from Kayseri Science High School, Kayseri, Turkey, in 1998. He received his Bachelor of Science from Middle East Technical University, Turkey in 2003. He received his Master of Science in Hydrogeology from Vrije University Amsterdam, Netherlands in 2005. He has worked as a visiting scientist at USDA Hydrology and Remote Sensing lab in Beltsville for 18 months from 2004 until 2006. He has started his PhD in Earth Systems and Geoinformation Sciences in George Mason University, Fairfax, Virginia in 2006. He has continued to work at USDA Hydrology and Remote Sensing lab for another year from 2006 until 2007. During 2007 until 2008, he has worked on several projects at Center for Research on Environment and Water. He has completed his PhD work and graduated from George Mason University in 2010.