Review of the Kalman type hydrological data assimilation

Article *in* Hydrological Sciences Journal/Journal des Sciences Hydrologiques · December 2015

DOI: 10.1080/02626667.2015.1127376

4 authors:

Leqiang Sun
University of Ottawa

Ousmane Seidou
University of Ottawa

Ioan Nistor
University of Ottawa

Kailei Liu
Hohai University

Hydrological Sciences Journal

ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage: http://www.tandfonline.com/loi/thsj20

Review of the Kalman type hydrological data assimilation

Leqiang Sun, Ousmane Seidou, Ioan Nistor & Kailei Liu

Publisher: Taylor & Francis & IAHS

Journal: *Hydrological Sciences Journal*

DOI: 10.1080/02626667.2015.1127376

Review of The Kalman Type Hydrological Data Assimilation

Leqiang Sun[1] Ousmane Seidou[1] Ioan Nistor[1] Kailei Liu[2]

[1]Civil Engineering Department, Univeristy of Ottawa, Ottawa ON, K1N 6N5, CANADA

[2]College of Hydrology and Water Resources, Hohai University, Nanjing, Jiangsu Province, 210098, CHINA

Abstract

There is great potential in Data Assimilation (DA) for the purposes of uncertainty identification, reduction and real time correction of the hydrological models. This paper reviews the latest development of Kalman filters, particularly the Extended Kalman Filter (EKF) and the Ensemble Kalman Filter (EnKF) in hydrological DA. The hydrological DA targets, methodologies and their applicability are examined. The recent applications of EKF and EnKF in hydrological DA are summarized and assessed critically. Furthermore, this review highlights the existing challenges in the implementation of EKF and EnKF, especially error determination and joint parameter estimation. A detailed review of these issues would not only benefit the Kalman type DA but also provide an important reference to other type hydrological DA.

Keywords

Data assimilation; Extended Kalman filter; Ensemble Kalman filter; model error; parameter estimation

1 Introduction

In the past few decades, the hydrological models have significantly benefited from the improvement of computation capacity and the availability of multi-source measurement data to simulate and forecast the hydrological processes. However, due to the deepening involvement of the uncertainties stemming from initial conditions, inputs and outputs, operational hydrological models will remain inherently imprecise.

Data Assimilation (DA) is a procedure developed to optimally merge information from the model simulations and the independent observations with appropriate modeling (Liu *et al.* 2012b). The DA could provide optimized initial conditions, updated parameters and even improved structures for the dynamic model. Originally used in areas like atmospheric and oceanographic science (Dee 1995, Derber and Rosati 1989), the DA has drawn more and more attention of hydrologists for its convincing performances in the real time correction of the hydrological models (Robinson and Lermusiaux 2000).

Walker and Houser (2005) and Reichle (2008) summarized the basic hydrological DA methods; Moradkhani (2008) reviewed the remote sensing measurement techniques and their applications in data assimilation; Liu and Gupta (2007) discussed the role of DA in addressing the uncertainties in hydrological models and discussed the application of the hydrological DA in operational scenarios (Liu *et al.* 2012b). Montzka *et al.* (2012) examined the joint assimilation of observational data precedents from different spatial scales and different data types. Reviews that focus on a specific method are relatively rare compared to reviews with a wide scope (e.g., Van Leeuwen 2009).

The Kalman filter is a classic sequential method that has been widely used in hydrological DA for more than two decades (Evensen 1994b, McLaughlin 1995). Compared to other methods such as particle filters and variational methods, the Kalman filter is easier to implement and could produce comparable or even better results with a lower computation demanding (Weerts and El Serafy 2006, Abaza *et al.* 2014b). Furthermore, the Kalman filter is very flexible to couple with the hydrological models and has more derivative variants than any other methods. Among these variants, the Extended Kalman Filter (EKF) and the Ensemble Kalman Filter

(EnKF) were both developed to extend the application of the Kalman filter to nonlinear systems; also, they are the two major descendants of the linear Kalman filter. EKF applies a straightforward Taylor extension scheme to linearize the nonlinear system. As logical as it is, EKF has endured a 'notorious' reputation for being unstable when applied to complex nonlinear hydrological models. EnKF avoids direct linearization by statistically analyzing the ensemble members. Although it increases the computational cost, EnKF is one of the most widely used hydrological DA methods. Despite some comparison case studies (Reichle *et al.* 2002b, El Serafy and Mynett 2004, Dumedah and Coulibaly 2012), there lacks a critical review that specifically focuses on the Kalman filter. The objective of this review is to fill this gap by assessing the latest developments and analyzing the challenges of the Kalman type hydrological DA, especially EKF and EnKF. Nevertheless, this review is not intended to judge which method is better than the other, but rather regards them as different solutions to the same problem. This is not only because they are both branched from the linear Kalman filter, but also they face some similar issues when applied in hydrological DA, of which some are also faced by methods other than Kalman filters.

In the rest of the paper, we introduce of hydrological DA, including the DA targets and methods in section 2. The Kalman filter (including EKF and EnKF) theories and its state of the art applications in hydrological DA are discussed in section 3. Section 4 discussed the issues regarding the implementation of Kalman filter in hydrological models. The summary and conclusions are given in section 5.

2 Hydrological Data Assimilation

2.1 Hydrological DA targets

DA was firstly used in the 1950s for the numerical weather forecast model; however, hydrologists did not pay enough attention to it until the 1990s (Evensen 1994b, McLaughlin 1995). Proper use of DA may help to handle the uncertainties from the model inputs, the initialization and propagation of the states, the model structures and even the model parameters (Vrugt *et al.* 2006, Liu and Gupta 2007, He *et al.* 2012). Meanwhile, global DA may improve the regional field estimation by achieving more accurate external boundary condition estimations (Robinson and Lermusiaux 2000), and local DA could also lead to improved global estimations (Clark *et al.* 2008).

The development of remote sensing has promoted the application of DA in hydrological models. The remotely sensed hydrological data that has had, or currently has the potential to be applied in hydrological models includes (Walker and Houser 2005, Houser *et al.* 2005, Xu *et al.* 2014):

1) overland parameters (e.g., topography, land cover, albedo)

2) forcing inputs (e.g., precipitation, air humidity and temperature)

3) states (e.g., soil moisture, snow cover)

4) fluxes (e.g., carbon flux)

The overland parameters are usually regarded as 'static' in the model, even though it is not true in long-term forecasts. Forcing inputs have the potential to replace the traditional ground-based observations with the development of the remote sensing instruments and more accurate retrieval algorithms. Neither overland parameters nor forcing inputs are the main interests of DA for hydrological modelers at this stage. The states and fluxes provide validation to the intermediate processes of the models. This was not available before, and hence they drew extra attention.

Snow cover has a high albedo and thermal properties, as well as a medium-term water storage capacity; therefore, the assimilation of snow observations could improve hydrological prediction (Walker *et al.* 2003). Andreadis and Lettenmaier (2006) used EnKF to assimilate remotely sensed snow observations MODIS SCE data into the VIC hydrologic model to update SWE estimates. From these estimates, a simple snow depletion curve scheme from SNOTEL station SWE data and MODIS imagery (Moradkhani 2008) was used to form the observation operator. Clark *et al.* (2006) assimilated the observations of a snow covered area (SCA) to update the hydrologic model, and they found that the assimilation of SCA information results in minor improvements in the accuracy of stream flow simulations near the end of the snowmelt season. Both of them found that the snow cover update works better during the snowmelt season than the snow accumulating season. Further research with regards to snow cover assimilation can be found in literatures (Parajka and Blöschl 2008, Rodell and Houser 2004, Sheffield *et al.* 2003, Pan *et al.* 2003, Kumar *et al.* 2008).

Galantowicz *et al.* (1999) demonstrated the Kalman filter retrieval of the soil moisture profile and temperature from L-band radio brightness observations, while Crow and Wood (2003) applied EnKF to assimilate the airborne measurements of surface brightness temperature into a TOPMODEL-based Land–Atmosphere Transfer Scheme (TOPLATS). Due to the short memory the land surface skin temperature holds, it is suggested to combine other state variables with longer memories such as deeper soil temperature or moisture to obtain longer DA effectiveness (Walker *et al.* 2003). The surface temperature products are also sensitive to terrain, vegetation and cloud contamination; hence, multiple sets of products and better retrieval algorithms are required for continuous DA operations (Huang *et al.* 2008).

Soil moisture is the key variable to control the runoff generation process. It is logical to assimilate soil moisture directly, because it is a continuous storage indication as well as a natural bridge to transfer the hydrological model into state space equations. The correct estimation of antecedent soil moisture content is critical to stream flow simulation when the soil is neither too dry nor over saturated (Reichle *et al.* 2002a). Soil moisture assimilation has been proved beneficial to stream flow, especially that of low flow prediction (Chen *et al.* 2014, Wanders *et al.* 2014b). The application of soil moisture requires the ground based measurements to be dense enough (Chen *et al.* 2011) otherwise the spatial interpolation error would have to be considered. Satellite remote sensing can provide an economic and more spatial-temporal continuous observation of soil moisture.

Parrens *et al.* (2014) assimilated the in situ soil moisture observations into a soil model at local scale. Reichle *et al.* (2002a) assimilated the L-band microwave brightness temperature into a Land Surface Model to estimate the near surface soil moisture. Crow and Ryu (2009) used the remotely sensed soil moisture retrievals to correct both the soil moisture state and satellite rainfall products. Alvarez-Garreton *et al.* (2014) assimilated the surface soil moisture and the soil wetness index derived from the passive microwave AMSR-E. Brocca *et al.* (2012) assimilated the surface and root-zone soil moisture products derived from the active miscrowave ASCAT into a rainfall runoff model using EnKF. Recently, the soil moisture assimilation has also been used to assist the parameter identification of hydrological models (Wanders *et al.* 2014a, Tran *et al.* 2014).

Despite the existence of some root zone soil moisture byproducts from the surface soil moisture observations (Wagner *et al.* 1999, Das *et al.* 2008), one of the obvious drawbacks of remote sensing soil moisture is that only surface or near surface soil moisture is available (Moradkhani 2008, Han *et al.* 2012). Preliminary experiments showed that the surface soil moisture assimilation has minimal effect on the simulation of the deep layer soil moisture (Chen *et al.* 2011), although it is the latter that has a more significant impact on the runoff simulations (Brocca *et al.* 2012, Han *et al.* 2012). Houser *et al.* (1998) argued that the remotely sensed soil moisture must be supplemented by in situ surface and root zone observations to specify error correlation, calibrate parameters, and validate the model-calculated fields. Draper *et al.* (2011) found that it might be more effective to address the cause of model bias instead of relying on soil moisture assimilation to correct it. Some research combined soil moisture assimilation with the correction of other rainfall runoff model forcings, such as precipitation, to improve the stream flow predictions (Chen *et al.* 2014, Massari *et al.* 2014). The application of remote sensing soil moisture may also involve issues like rescaling (Sahoo *et al.* 2013, Kaheil *et al.* 2008), error evaluation (Alvarez-Garreton *et al.*, Doubková *et al.* 2012), and radiative transfer modelling (Verhoef and Bach 2003, Reichle 2008), among others. A more detailed explanation of the limitations of soil moisture assimilation can be found in literature (Vereecken *et al.* 2015).

Stream flow is the most commonly used and is sometimes the only available prognostic observation variable (Clark *et al.* 2008, Abaza *et al.* 2014a, Trudel *et al.* 2014, Samuel *et al.* 2014, Randrianasolo et al. 2014). Great efforts have been made to improve the stream flow forecast using output assimilation/error assimilation over the past two decades (Sene 2008, Anctil *et al.* 2003, Broersen 2007, Yu and Chen 2005). The output assimilation/error assimilation methods treat stream flow forecast as a pure model output and update it by adding errors calculated with another independent procedure or model. Such procedures/models could either be nonlinear like ANNs (Anctil *et al.* 2003) or linear such as ARMA (Broersen 2007, Chen *et al.* 2015). Output/error assimilation of stream flow is relatively simple to implement, as there is no feedback to the original rainfall runoff model from the manipulation of model outputs.

Besides being the direct assimilation goal (Liu *et al.* 2012a), stream flow is also the prevalent observation to assimilate other state variables and parameters (Coustau *et al.* 2013). In hydrological DA, stream flow is often treated as a diagnostic variable, and is therefore not updated directly (Clark *et al.* 2008). In the case of the nonlinear measurement operator, it is probably the only choice to augment the state vector with the stream flow (Evensen 2003). By doing so, the nonlinear measurement operator would be reduced to a linear matrix, otherwise one would have to linearize the measurement operator. Pauwels and De Lannoy (2009) attempted to linearize the nonlinear discharge-watershed storage relationship (observation operator) in a 'brutal' way. Linearization was undertaken within both a simple time series model and a conceptual model HBV; subsequently, the results found that direct linearization should be bypassed to obtain a better assimilation result.

Stream flow assimilation is different from other state variables because it involves the issue of routing. When the runoff is assimilated into a hydrological model at the current time step, not only does the current state of the watershed at a given location (e.g., outlet) need to be updated and propagated forward, but also the state at a number of different locations at multiple previous time steps (Pauwels and De Lannoy 2006). This is the case especially concerning large and distributed hydrological models. Many authors actually choose not to get into the details of the complicated channel network to control the complexity of the question (Clark *et al.* 2008, Weerts and El Serafy 2006).

Despite the unique challenges faced by these targets, they share some common issues in the implementation, most notably in the quantification of observation errors and the determination of observation operators that connect them with the model output (Andreadis and Lettenmaier 2006). It is also worth pointing out that the assimilation of one single variable does not necessarily improve the estimation of the other variables. Trudel (2014) reported that the assimilation of the stream flow at the outlet of the watershed would distort the estimate of soil moisture. The combination of multiple observation variables seems to be superior to the single variable assimilation (Trudel *et al*. 2014, Xie *et al*. 2008).

The last decade saw the development of various global and regional land assimilation systems (Rodell *et al.* 2004, Mitchell *et al.* 2004, Cosgrove *et al.* 2003, Kumar *et al.* 2008). These systems enable the incorporations of multi-source observations, multi models and multi assimilation schemes in creating an optimal land surface states output. Due to the possible water and energy balance issues, a major concern in land assimilation systems is that if they should be coupled to the atmospheric models (Walker *et al.* 2003). Betts (2003) compared the water budget of ECMWF 40-year reanalysis and NASA DAO fvGCM with the hydrological balance of VIC model and the radiative fluxes with the basin averages derived from ISCCP. They found that the runoff from both atmospheric models was significantly underestimated compared to VIC runoff simulation, which is consistent with the observed stream flow. Large bias are also observed for radiation fluxes and surface temperature between ECMWF 40, fvGCM and the ISCCP data. Pan (2006) suggested a constrained EnKF to maintain the benefit of DA without violating the water balance principle. Boulet (2000) and Bøgh (2004) developed a simple water and energy balance model that allows the direct application of the remote sensing data.

2.2 DA methods

The DA methods can be divided into different categories based on different standards (Rakovec *et al.* 2015). According to the dimension they focus on, these approaches can be classified as objective analysis methods and time dependent methods (Wang and Kou 2009).

2.2.1. Objective analysis

Objective analysis aims to minimize the error between the observation field and background field (usually the output of the numerical models) by 'fusing' the new observations into the background with spatial dimensions. Typical objective analysis methods include successive correction (Cressman 1959, Barnes 1964), optimal correction (Lorenc 1981, Gandin 1963), statistical bias correction (Piani *et al.* 2010), Newtonian nudging (Houser *et al.* 1998, Paniconi *et al.* 2003), variational methods (Reichle *et al.* 2001, Seo *et al.* 2003, Seo *et al.* 2009, Navon 2009), etc. Detailed descriptions about the theory and development of the above methods can be found in literature (Navon 2009).

Since objective analysis methods mostly take 'snapshots' of the background filed at a given time (McLaughlin 1995), the temporal dimension is usually not well incorporated with the spatial dimensions. Even though some of them do consider the temporal dimension, such as 4D VAR and Newton Nudging, time-variant objective methods are most simply viewed as a dynamic extension of the time-invariant version of the objective methods (McLaughlin 2002).

Among various objective analysis methods, the successive correction and Nudging methods fail to consider the errors in the observations, while 3D VAR and 4D VAR ignore the uncertainties in the models. Objective analysis methods usually involve huge computations (Clark *et al.* 2008). Houser *et al.* (2005) suggested that the adjoints should be calculated as the model is developed, but this is by no means an easy task for distributed hydrological models.

2.2.2. Time dependent methods

Time dependent methods use a probabilistic framework and estimate the system state sequentially by propagating information forward in time (Bertino *et al.* 2003). The strength of the time dependent methods is in time series analysis. The time dependent methods work on a fixed but moving time window and only the most recent observations that fall into this window are incorporated into the final estimation results. Typical time dependent methods that are frequently used in hydrological DA include the linear Kalman filter (Kalman 1960) and it's variants, such as the Extended Kalman Filter (EKF) (Puente and Bras 1987), the Ensemble Kalman filter (EnKF) (Evensen 1994b), the Unscented Kalman Filter (UKF) (Wan and Van Der Merwe 2000), the Particle Filter (Weerts and El Serafy 2006, Pham 2001, Moradkhani *et al.* 2005a), and the H-infinity filter (Lü *et al.* 2010, Wang and Cai 2008, Moradkhani *et al.* 2005a), among others. There are also some alternative approaches to solving specific issues in hydrological DA, most notably in the application of a genetic algorithm (GA) in the estimate of the pixel-based soil hydraulic parameters for hydroclimatic modeling (Ines and Mohanty 2008) and the evolutionary based assimilation in the stream flow simulations in ungauged watersheds (Dumedah and Coulibaly 2012). The spatial variation of the relationship between the background field and observation field of the variable in question is of less concern in time dependent methods. Meanwhile, time dependent methods are capable of handling more uncertainties (Moradkhani 2008) and are less complex to implement compared to some objective analysis methods that require an inverse or a joint model (Bertino *et al.* 2003).

Time dependent methods can be well explained with Wiener-Kolmogorov estimation theory (Wiener 1949, Kolmogorov *et al.* 1941). Suppose $s(t)$ is the original signal, and $\widehat{s(t)}$ is the estimated signal; the estimation error is defined as:

$$e(t) = s(t+\alpha) - \widehat{s(t)}$$

Where α is the delay of the estimation. In other words, the error is the difference between the estimated signal and the true signal shifted by α. Depending on the value of α, the estimation problem can be described as in Fig. 1:

i) if $\alpha>0$, the estimation is a prediction problem (error is reduced when $\widehat{s(t)}$ is similar to a later value of *s(t)*);

ii) if $\alpha=0$, the estimation is a filtering problem (error is reduced when $\widehat{s(t)}$ is similar to *s(t)*) ;

iii) if $\alpha<0$, the estimation is a smoothing problem (error is reduced when $\widehat{s(t)}$ is similar to an earlier value of *s(t)*).

In a broad sense, both smoothing and filtering are time dependent DA methods as they both combine the advantages of the measurements and model outputs. In a narrower sense, only filtering problems count as time dependent DA methods, as DA only deals with 'real-time' measurements instead of historical ones. In the case that the information in the time series is only propagated forward without any backward loop or window, the time dependent methods are also termed as sequential methods. The prediction problems usually serve as the validation of the smoothing and filtering problems. The traditional batch calibration technologies for hydrological models can be categorized as a smoothing problem (McLaughlin 2002). Filtering technologies usually play an important role to improve the prediction from the calibration; hence, they are widely used in the real time modules of operational models (McMillan *et al.* 2013, Divac *et al.* 2009).

2.2.3 Applicability of DA methods to hydrological models

Many approaches have been applied in hydrological DA (Houser *et al.* 2005, Walker and Houser 2005). The most commonly used methods include the variational method, the Particle Filters and the Kalman filters. Although traditionally dominant in numerical weather forecasts, variational methods have not been widely used in hydrological DA. Variational methods such as 3D VAR assume the forecast error statistics are isotropic and largely homogeneous with little variation in time (Houtekamer and Mitchell 1998), yet the consideration of time dimension would overwhelmingly increase the computation burden. It is also complicated to develop the adjoint model for the distributed hydrological models (Clark *et al.* 2008). Successful cases using variational methods in hydrological DA are mostly based on simpler lumped models (Seo *et al.* 2009, Seo *et al.* 2003, Abaza *et al.* 2014b). Abaza *et al.* (2014b) compared the classic variational method with EnKF, and they found that the latter is more stable in stream flow assimilation.

The Particle Filters are also widely used in hydrology DA (Weerts and El Serafy 2006, Pham 2001, Moradkhani *et al.* 2005a). One of the major advantages of particle filters is that the system does not need to be Gaussian (Liu *et al.* 2012b). Particle filters are also better at handling model nonlinearities compared to other sequential methods (Moradkhani *et al.* 2005a). The fact that the Particle Filters use full prior density means this method is more computationally intensive (Weerts and El Serafy 2006). The operational applications of particle filters in distributed hydrological models are also limited due to the setup of the particle numbers (Liu *et al.* 2012b).

Despite the prosperous development of DA methods in 'interpolation', 'smoothing' or 'filtering', not many of them are actually extensively validated in 'forecast' mode, not to mention in operational hydrological forecasts (Liu *et al.* 2012b). One of the problems that hydrological DA faces is the short 'efficient period'. The updating of model structures and storages generally has a major impact on the forecast only within shorter lead times (El Serafy and Mynett 2004, Knight and Shamseldin 2006). A possible reason for this is that DA deals with uncertainties from different data sources in a statistical way, rather than in a physical or mechanical way (McLaughlin 2002). It somehow 'detours' the imperfect model structure and parameters in hydrologic models by adding more weight to the model uncertainties, thereby emphasizing observations in the final results.

Numerical Weather Prediction (NWP) is fundamentally an initial problem that is very sensitive to state variation versus the hydrological models, which lean toward a process problem that relies on model structures and parameters. For this reason, DA might not work as well in hydrology as in NWP. However, because of the 'conceptual' nature of most hydrological models, DA does have the potential to improve the model forecast either by updating the initial condition or modifying the model parameters.

A potential development direction of hydrological DA is the combination of objective analysis methods and sequential methods. 4D VAR expands the strength of 3D VAR by considering the temporal evolution of variables within the fixed time window. However, depending on the model structure, it could be expensive to calculate the gradients of the cost function (Bin and Ying 2005). The methods that combine the merits of both objective analysis and sequential methods, such as 4D VAR and EnKF, are generally regarded as the most promising DA technology in Numerical Weather Prediction (Lorenc 2003, Kalnay *et al.* 2007). However, as far as the authors are aware of, such a dominant DA method still does not exist in the operational hydrological simulation and forecast.

3 EKF and EnKF

The linear Kalman Filter is a classic sequential method. Together with its multiple nonlinear variants, such as EKF and EnKF, the Kalman filter has become a very promising method cluster used in hydrological DA.

3.1 Linear Kalman filter

The Kalman filter is an optimal estimator that recursively couples the most recent measurements into the linear model to update the model state output (Kalman 1960). Under the assumption that the linear system is a stochastic process with Gaussian noises, the Kalman filter produces the best estimation with minimum mean square error.

The Kalman filter works on a stochastic system in the form of state space equations (Hamilton 1994):

$$x_{k+1} = M_k x_k + B_k u_k + \eta_k \tag{1}$$

$$y_k^o = H_k x_k + \varepsilon_k \tag{2}$$

Equation (1) is the model function that propagates state *x* from step *k* to step *k+1*. x_k is the *a posteriori* state vector and x_{k+1} is the *a priori* state. M_k is the model function (also known as the dynamic function or dynamic operator). B_k is a linear matrix to convert the dimension of residual vector u_k to state vector x_k. u_k is the 'forcing' term of the model in the form of linear residuals. η_k is the model estimation error (covariance matrix=Q_k). Sometimes, to simplify the problem, u_k is regarded as part of M_k and its estimation error is included in η_k, and in this case B_k=0.

Equation (2) is the observation function that relates the state vector x_k to the observation y_k^o, with observation error ε_k (covariance matrix=R_k). To guarantee the linearity of the system, the observation operator H_k should be linear too. Normally, the number of the observation variables is limited while the state variable combinations could be infinite. It is probable that the observation vector dimension (degree of freedom) is much smaller than that of the state vector.

A typical implementation of the Kalman filter is described below (Drécourt 2003). At time step 0, set up the initial estimation error covariance matrix P_0^a of the initial state x_0^a. The model propagation without an explicit form of u_k is given by:

$$x_k^f = M_{k-1}x_{k-1}^a \tag{3}$$

$$P_k^f = M_{k-1}P_{k-1}^a M_{k-1}^T + Q_{k-1} \tag{4}$$

The superscript 'a' represents *a posteriori* estimation and the superscript 'f' means *a priori* estimation.
The state is then updated via:

$$x_k^a = x_k^f + K_k \cdot (y_k^o - H_k x_k^f) \tag{5}$$

Where $y_k^o - H_k x_k^f$ is the innovation and the Kalman gain K_k is calculated with:

$$K_k = P_k^f H_k^T [R_k + H_k P_k^f H_k^T]^{-1} \tag{6}$$

The estimation error is finally updated as:

$$P_k^a = (I - K_k H_k)P_k^f \tag{7}$$

With the updated estimation error, the filter can restart from equation (3) and (4) to begin another loop recursively.

Equation (5) shows that the updated state is a linear combination of the observation and the model estimate (Drécourt 2003). The gain would change between 0 and H^{-1} depending on the uncertainty comparison between the model outputs and the observations. For example, if the dynamic model is 100% accurate (which is unrealistic), then the new observation would not be considered at all in the updated state, as the gain is 0, and vice versa.

The linear Kalman filter tends to be more commonly used in stochastic models (Bergman and Delleur 1985, Bolzern *et al.* 1980, Szöllősi-Nagy and Mekis 1988) and channel routing problems (Huang 1999), which are easier to linearize (Wu *et al.* 2008, Fan 1991, Georgakakos and Bras 1982, Sun *et al.* 2013). However, there are a few successful cases that the linear Kalman filter was applied in complex rainfall runoff models. Liu *et al.*

(2011) and Lee and Singh (1999) coupled a Tank model with a linear Kalman filter to estimate the model parameter and outflow respectively. Kim *et al.* (2005) applied the Kalman filter in a one dimension physically based distributed model CDRMV3 with the storage amount of whole watershed as state and outlet discharge as measurement, while a discharge-storage (Q-S) curve was used as the observation equation, in addition to a Monte Carlo ensemble being used as the dynamic equation. The drawback of such operation is that it could be difficult to find the linear relationship between discharge and storage, even at local scale.

3.2 Nonlinear Kalman Filter

The Kalman filter is only legitimate in a linear system where both the model function M_k and the observation function H_k are linear. In the case of a nonlinear system (Drécourt 2003):

$$x_k^f = M_k(x_{k-1}^a, u_k, \eta_k) \tag{8}$$

$$y_k^o = H_k(x_k^f, \varepsilon_k) \tag{9}$$

Where M_k is the model function and H_k is the observation function, the linear Kalman filter is not applicable even if only one of them is nonlinear.

The hydrological system is a highly complicated nonlinear system with gigantic dimentions. In cases like stochastic blackbox stream flow prediction models, a linear Kalman filter might apply. However, it is not feasible to count on a linear Kalman Filter in DA of distributed/semi distributed hydrological models.

To cope with the nonlinear problem, many different versions of modified Kalman filters are developed (Rakovec *et al.* 2015, Gharamti and Hoteit 2014, de Rosnay *et al.* 2013, Chen *et al.* 2013, Dumedah and Coulibaly 2012, Muluye 2011, Shamir *et al.* 2010). A large portion of these booming new methods or algorithms are based on EKF and EnKF.

3.2.1 EKF

3.2.1.1 Theory of EKF

A prerequisite to apply EKF is that the nonlinear system should be derivativable contineously. Under this condition, Taylor extension is applied at the estimated point. It is therefore possible to obtain the converted linear dynamic matrix by expanding the nonlinear functions at the estimated point. Apply this to both model function and observation function to get equation (10) and equation (11) (Walker and Houser 2005):

$$\dot{M}_k = \frac{\partial M}{\partial x_k}\Big|_{(x_{k-1}^a, u_k, 0)} \tag{10}$$

$$\dot{H}_k = \frac{\partial H}{\partial x_k}\Big|_{(x_k^f, 0)} \tag{11}$$

The linearized $\dot{M}_k$ and $\dot{H}_k$ are then used in the calculation of error covariance propagation; all subsequent steps are the same as the linear Kalman Filter (Lewis *et al.* 2008). The EKF may be applied at the latest observation back and forth as many times as needed to reduce the linearization errors. However, such recursive operations do not guarantee a better estimation than one time EKF while the computation burden is largely increased (Puente and Bras 1987).

In many cases, the nonlinear systems do not have explicit analytical solutions and the derivatives can only be calculated numerically. As the state is a vector with multiple variables, the linearized matrix can be expressed as a Jacobian matrix of the partial differential functions.

EKF keeps the first order of the Taylor expansion, while it ignores the higher order terms. Theoretically, a high order EKF that keeps more than the first order terms is more precise, as more information is kept (Tanizaki 1996, Simon 2006). However, they are still biased because higher orders terms are still ignored (Tanizaki 1996). In reality, there are rare cases supporting the superiority of higher order EKF over the first order EKF (Ermolaev and Volynsky 2014, Sadeghi and Moshiri 2007). This is especially true when the analytical solutions of the system are unavailable, because the simultaneous numerical solution of the Jacobian and Hessian matrix may introduce more errors (Vittaldev *et al.* 2012, Roth and Gustafsson 2011).

3.2.1.2 Application of EKF

EKF has been widely used in operational soil moisture analysis by various weather agencies (Hess 2001, Fairbairn *et al.* 2014, de Rosnay *et al.* 2013). De Rosnay *et al.* (2013) described the operational implementation of simplified EKF (SEKF) (Draper *et al.* 2009, Mahfouf *et al.* 2009), in which EKF has been simplified by setting constant background errors in the ECMWF land surface analysis system. SEKF linearizes the observation operator with finite differences by adding individual perturbations to each element of the model state vector. Such feature significantly increases the computation cost with the increase of the dynamic model resolution. To solve this problem, one should either modify the model structure or calculate the Jacobian matrix from off-line land surface model simulation (Mahfouf *et al.* 2009). The comparison between SEKF and EnKF shows that SEKF gives similar performance to EnKF yet the latter is unable to improve on the former (Fairbairn *et al.* 2014).

Ge (1984) applied EKF on a 3-layer Storage Excess runoff model. Depending on the existence of precipitation and the fulfillment of each soil layer, they set ten different runoff generation scenarios. To avoid defining a threshold function, each scenario was assigned a state space function. By doing so, it is found that only two state space functions need to be linearized. Taylor Extension was then used with these two functions analytically. The constants from the truncation of Taylor extension are treated as the linear driven terms of model function and linear residuals of observation function respectively. Neither of them gets involved in the state error propagation explicitly, but they are the integral components of the model and observation functions and cannot be neglected.

Kitanidis and Bras (1980a) linearized the Sacramento model using statistical linearization and 'describing function' technique. Instead of linearizing the nonlinear formulas directly, a group of simple linear functions that produce similar outputs with the same inputs as the nonlinear formulas are applied. The focus is thus shifted to the determination of the coefficients of the linear functions using statistical methods. This scheme is somehow 'subjective', because the candidate linear functions are not unique depending on the selection of criteria, or the evaluation functions to assess the fitness of the linear functions. Another challenge for this scheme is it assumes that the inputs follow a constrained Gaussian distribution, and most importantly, they should fall in a very narrow range for the linear function replacement. The computation burden of statistical linearization is also a big issue.

Instead of implementing analytical linearization, Walker and Houser (2001) numerically derived the dynamic transition matrix in the propagation of the estimation error covariance. The state vector contains three intermediate soil moisture parameters and the forecast equations of the state were linearized with first order Taylor series extension. The observation is chosen as the remote sensing surface soil moisture, which is related to the state vector through a complicated nonlinear equation that is also linearized with Taylor extension. The numerical solution significantly increases the computation cost, but not as much as statistical linearization (Kitanidis and Bras 1980b). Therefore, it is important to provide a state vector of a manageable size. The problem of the numerical solution is the determination of the perturbation of the independent variables: small perturbations may lead to numerical problems, while large perturbations will cause a much greater loss of accuracy, and could also risk hitting the nonlinear threshold (Reichle *et al.* 2002b).

3.2.1.3 Assessment of EKF

The robustness issue of EKF is well discussed in hydrological DA, and much of this concerns the divergence problem (Ljung 1979). The divergence of EKF can be caused by a number of reasons (Fitzgerald 1971): The first factor concerns the inappropriate estimation of model and observation error covariance matrix, which will be discussed extensively in 4.2. Secondly, divergence can also be caused by the incorrect estimation of the initial estimation error P. Since the estimation of P at time k relies on P at time k-1, a large error in initial P could recursively propagate until there is a divergence of the EKF (Tao *et al.* 2005). The third reason concerns the truncation errors arising from the Taylor expansion: although first-order local linearization is adequate to account for differential nonlinearity in some models (Reichle *et al.* 2002b), the existence of such errors makes it harder for the model errors to meet the whiteness assumption. Lastly, the errors from the numerical calculation of the Jacobian matrix: the perturbation scale of the independent variables involves subjective trials.

Many efforts have been made to alleviate the divergence of EKF. The adaptive algorithms (Jwo and Wang 2007) are widely used to prevent the divergence caused by inappropriate error quantifications. Ljung (1979) introduces an innovation model by adding a parameterized Kalman gain term to the original EKF state space equations, he argued that the main reason for the divergence is the lack of coupling between Kalman gain K and model parameters. Tao *et al.* (2005) rescale the error covariance matrix by multiplying them by a factor; once the divergence is detected to happen, the Kalman gain is frozen until stability is restored.

Compared to other nonlinear methods, the implementation of EKF is direct and straightforward. It is arguably the de facto standard method to expand the Kalman filter to nonlinear systems for its outstanding performance in practice (Rabier 2003). However, it is not clear if EKF can outperform other DA methods in nonlinear distributed hydrological models when considering the computation burden of linearization, even though strategies such as low rank SEEK (Tuan Pham *et al.* 1998) can reduce the computation requirement of EKF.

3.2.2 EnKF

3.2.2.1 Theory of EnKF

EnKF is a Monte Carlo approach for nonlinear filtering problems (Evensen 1994a). There are two types of EnKF (Kalnay *et al.* 2007): perturbed observation EnKF (Evensen 1994b, Burgers *et al.* 1998) and square root EnKF (Anderson 2001, Bishop *et al.* 2001, Whitaker and Hamill 2002). The perturbation of the observations causes additional sample errors (Kalnay *et al.* 2007), but this type of EnKF could handle nonlinearities better than the square root EnKF (Lawson and Hansen 2004).

EnKF is based on the approximation of the conditional probability densities of the state errors covariance by a finite large number of randomly generated model trajectories. EnKF does not need any derivation of the model operator or observation operator as performed in EKF. Instead, it generates a set of realizations (the ensemble) and propagates them through the model operator independently. Then, it derives the *a priori* state error covariance through the statistical analysis of the ensemble (Evensen 1994b).

The general idea of perturbed observation EnKF is demonstrated in Fig.2. In the forecast step, the *a priori* state ensemble is created by adding random perturbations to the best estimate of the initial state. Then the *a priori* error covariance matrices are approximated from the state and model output ensemble error matrices

(Gillijns *et al.* 2006). In the analysis step, the state ensemble *x* are integrated forward in parallel, based on the original model. The observation is perturbed by adding Gaussian noises of measurement errors (Burgers *et al.* 1998). Then, the *a posteriori* error covariance matrix can be estimated from the ensemble error matrix of the *a posteriori* states. Meanwhile, the Kalman gain is calculated from the forecast error covariance matrices (Reichle *et al.* 2002b, Moradkhani *et al.* 2005b).

3.2.2.2 Application of EnKF

Although the propagation of state error covariance does not require the linearization of the model, the parallel computation of the ensemble means EnKF is an expensive method to use. Nevertheless, EnKF is a widespread approach in hydrological DA.

Weerts and El Serafy (2006) compared EnKF with two Particle filters in the flood forecast with the conceptual rainfall-runoff model HBV-96, and found that EnKF outperforms both particle filters. Dumedah and Coulibaly (2013) compared the evolutionary data assimilation (EDA) with methods based on the integration of Pareto-optimality into both an EnKF (ParetoEnKF) and a PF (ParetoPF). Rafieeinasab *et al.* (2014) compared EnKF with a maximum likelihood ensemble filter, which is an ensemble extension of variational assimilation. It is argued that the former is less sensitive to observation, model errors, and parameter uncertainties, while it does perform reasonably well with a smaller ensemble size. Borup *et al.* (2015) discussed the procedure to utilize EnKF in the assimilation of the observations that are 'out-of-range', and defined this method as 'partial EnKF'. The positive results indicate EnKF is versatile enough to take advantage of the imperfect but precious observation information, instead of simply wasting it.

Madsen *et al.* (2003) presented the procedure to assimilate observed water levels and fluxes in the MIKE 11 Flood Forecasting system using EnKF. Borup (2014) demonstrated the application of EnKF in assimilating water level and flow observations into the distributed urban drainage models. Liu *et al.* (2015) used EnKF to assimilate AMSR-E snow depth into a land surface model for stream flow predictions. Xie and Zhang (2010) implemented synthetic simulation experiments for the application of EnKF with a SWAT model. Sensitivity analysis with regards to error specification, initial realization, and ensemble size were also demonstrated in their paper. The application of EnKF in SWAT was also extensively studied in other works (Han *et al.* 2012, Lei *et al.* 2014).

Clark *et al.* (2008) attempted to apply EnKF to assimilate the stream flow observation into a distributed hydrological model; however, the researchers found it unsuccessful due to the inappropriate selection of state variables. Reichle *et al.* (2002a) assimilated remote sensing soil moisture into a land surface model with a very large state vector and a small ensemble number. They argued that a large ensemble number is needed to obtain a robust error variance estimation. Shi *et al.* (2014) found that the EnKF assimilation of multivariate observations applies strong constraints to parameter estimation in a physically-based land surface hydrologic model. Abaza *et al.* (2014a) introduced hyper-parameters perturbation factors by comparing the H-EPS spread to its mean forecast error for the perturbation of the system in EnKF. This research outlined the importance of input replicate generation in the implementation of EnKF. Panzeri *et al.* (2015) described the implementation of stochastic moment equations (MEs) based on EnKF in the groundwater flow simulation; the researchers found it more efficient than the traditional EnKF.

3.2.2.3 Assessment of EnKF

Compared to EKF, EnKF is less likely to diverge because the Taylor expansion and Jacobian matrix calculation are averted. Meanwhile, EnKF accounts for more model errors, as there is no truncation of higher order terms as in Taylor expansion (Reichle *et al.* 2002b). Considering the computation burden of the linearization, it is unrealistic to put all the representative variables from the calculation unit of the distributed model into the state vector (Gillijns *et al.* 2006). Therefore, it is difficult to consider both the horizontal correlations in the model, and the measurement errors in EKF. However, this is not an issue for EnKF (Reichle *et al.* 2002b).

Computational efficiency might be the primary concern before one applies EnKF. A major drawback of EnKF is that it assumes that the prior PDFs of the model states are Gaussian; hence, the posterior states are only determined by the first two moments of the prior density (Weerts and El Serafy 2006). Theoretically, the larger the ensemble number, the more accurate estimation of the mean and variance of the prior density (Madsen *et al.* 2003, Rasmussen *et al.* 2015). It is argued that the EnKF estimate would not converge to an optimum unless the ensemble size is large, especially in the case of high dimensional problems (Reichle *et al.* 2002a, Reichle *et al.* 2002b, Crow and Wood 2003). One of the preconditions to consider EnKF superior than EKF is that the ensemble size should reach a certain level so that the major errors are statistical noises, rather than a closure problem or unbounded error variance growth encountered by EKF (Evensen 1997). In reality, the parallel running of distributed hydrological models is a huge challenge to the computation capacity. Although the determination of the reasonable ensemble size is regarded as a case by case problem (Xie and Zhang 2010), more efforts are expected in deriving efficient algorithms to trade off between ensemble size and balanced forecast analyses, as well as representative error statistics (Mitchell *et al.* 2002).

4 Implementation of Kalman filters

4.1 Filter elements selection

An incomplete list of the recent hydrological DA works has been made to summarize their features, including the chosen DA methods, hydrological models, state, parameters (if any), and observations (see Table 1). Depending on the assimilation purpose and the available observation, the selection of which state variables to assimilate can be very flexible. A state variable can either be intermediate (e.g., soil moisture content, snow water equivalent) or prognostic (e.g., runoff/discharge). Soil moisture is by far the most popular state variable, because of its crucial role in describing the model with state space equations. Nowadays, there are many remote sensing products that can provide the temporally and spatially continuous soil moisture observations (Moradkhani 2008). On the other hand, stream flow forecast is one of the ultimate goals of hydrological modeling. With the easy availability of the stream flow observations, it is also a widely used state variable.

In most cases, the model parameters are assumed to be time invariant once they are calibrated. However, it is possible to treat some parameters as time variant, and update them together with the state variables (Samuel *et al.* 2014), even though this is still not as widespread as it should be. The selection of the parameters is less flexible than that of the state variables because it should coordinate with the state variable update. Besides that, the parameters selected should be sensitive enough to reflect the update (Xie and Zhang 2013). In the case of a model being simple enough, it is possible to include all the parameters in an assimilation; however, this would present a huge challenge for complex models. He *et al.* (2012) developed an Integrated Sensitivity and Uncertainty analysis Framework (ISURF) to screen and identify the sensitive model parameters and assess the uncertainty structure of model parameters in EnKF.

4.2 Determination of errors

Another key issue in the implementation of the Kalman filter is the determination of model and observation errors. The model errors are more difficult to describe as compared to the observation errors, because the latter can sometimes be predefined based on the measurement, sampling methods or empirical formulas (McMillan *et al.* 2010). The overestimation of model errors reduces the confidence in the model, and thus the filter would overly rely on observations; on the contrary, the underestimation of model errors exaggerates the accuracy of the model and also wastes the information from the new observations (Kitanidis and Bras 1980b). The improper selection of the model errors may lead to unacceptable results or a divergence of the Kalman filter. Puente and Bras (1987) argued that the proper error quantification of the model is even more important than the selection of the DA methods.

Clark *et al.* (2008) quantifies the errors in stream flow measurements as a fixed proportion of the discharge observations, while the model errors were generated by perturbing the precipitation forcing and model states (e.g., soil moisture and aquifer storage) with temporal varying perturbations. However, the more general method to realize the time variant estimation of model and observation error is by employing adaptive filtering. If properly used, the adaptive filtering may partly mitigate the influence of the inaccurate (if not incorrect) setup of the initial model and observation errors. The adaptive filters are divided into four categories: Bayesian, maximum likely hood, correlation and covariance matching (Mehra 1972).

Bayesian methods aim to obtain the recursive equations for the method works better for the identification of observation errors than model errors. It is also believed that covariance matching techniques give biased estimates of the true covariances (Odelson *et al.* 2006).

In general, adaptive update methods normally work well in linear systems with Gaussian errors, under stationary conditions (Reichle *et al.* 2008). However, some of them exploit the 'optimality' of the Kalman filter, especially in the assumption of the whiteness in the innovation sequences (Kitanidis and Bras 1980b). In the case of non-stationary systems or non-white errors, computation becomes extremely tedious and complicated. Crow and Reichle (2008) compared four adaptive filtering schemes in land data assimilation and subsequently found they all have the problem of low convergence. It is also argued that the current self-adjustment or adaptive methods are feasible only when system dimensionality is reduced (McLaughlin 2002). To seek a parsimonious DA approach that requires a minimum assumption to describe the unknown model and observational error characteristics, Vrugt *et al.* (2005) and Crow and Yilmaz (2014) developed the Auto-Tuned Land Data Assimilation System (ATLAS) for the integration of two remote sensing soil moisture products into a water balance model. They combined four separate adaptive filtering solutions (the Triple Collocation, Innovation, Merged, and Red Modeling Error solutions) to estimate the observation and model errors. The application of ATLAS on a simple forecast model leads to an improved surface soil moisture analysis, although its power with more complex models is yet to be investigated (Crow and Yilmaz 2014).

4.3 Parameter estimation

Due to the limited knowledge of hydrological processes, it is impossible to predetermine all the parameters in hydrological models. The most common practice first involves the initialization of the parameters, which are then adjusted with batch calibration until certain criteria are met. Once determined, the calibrated parameters are assumed to be consistent in the future simulations and are rarely updated again.

The Kalman filter is designed to update the state of the system for a better initialization of the forecast. However, it is not easy to find the state vector with a significant representativeness to update the overall estimation. Moreover, a pure state update neglects the propagation of the uncertainties in model parameters (Lü *et al.* 2011). Thus, in order to simultaneously update the model parameter with the state, it is desirable to account for both the uncertainties in parameters, and also state estimation (Liu and Gupta 2007). There are three main schemes to estimate the model parameters: state augmentation, dual state-parameters estimation and hybrid solutions.

4.3.1 State augmentation

State augmentation technology is realized by adding the parameters to the original state vector; thereafter, both the original dynamic model and observation model would be modified (Hendricks Franssen and Kinzelbach 2008). It is rare that the relationship between state variables and model parameters are linear; therefore, the linear Kalman filter is normally inapplicable when the state vector is augmented. In the case of augmented EnKF, Hendricks Franssen and Kinzelbach (2008) argued that the state and parameters can be jointly updated with either an iterative or a non-iterative approach. In the case of EKF, one can normally implement the Taylor expansion and Jacobian matrix methods as usual. However, it is worth pointing out the parameters are usually independent from prognostic state variables, which means the corresponding derivatives in the Jocobian matrix are zeros. On the other hand, the prognostic state variables are not independent of the parameters.

Based on equation (8) and equation (9), supposing Z_k is the augmented new 'state vector' that contains both the original state vector X_k and parameters θ_k, the augmentation process can be demonstrated as follows (Gharamti and Hoteit 2014, Wang and Wang 1985):

$$Z_k = \begin{bmatrix} X_k \\ \theta_k \end{bmatrix} = \begin{bmatrix} M_k(X_{k-1}, \theta_{k-1}) + \eta_k \\ \theta_{k-1} \end{bmatrix} = \widetilde{M}_k(Z_k) + \tilde{\eta}_k \tag{12}$$

Where $\widetilde{M}_k$is the new model operator and $\tilde{\eta}_k$ is the new model error, assuming the parameters are constants, the new model error can be expressed as:

$$\tilde{\eta}_k = \begin{bmatrix} \eta_k \\ 0 \end{bmatrix} \sim N(0, \tilde{Q}_k) \tag{13}$$

Where

$$\tilde{Q}_k = \begin{bmatrix} Q_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{14}$$

Where Q_k is the original model error covariance and **0** is zero matrix.
Meanwhile, the observation operator is also transformed to:

$$\widetilde{H}_k = [H_k \ \ \mathbf{0}] \tag{15}$$

Where H_k is the original observation function.
Equation (13) and (14) are based on the assumption that the parameters are constants. In the case of varying parameters, the following equations should be used (Wang et al. 2009):

$$\tilde{\eta}_k = \begin{bmatrix} \eta_k \\ \varepsilon_k \end{bmatrix} \sim N(0, \tilde{Q}_k) \tag{16}$$

Where ε_k is the parameter error. Assuming ε_k is independent from η_k, then

$$\tilde{Q}_k = \begin{bmatrix} Q_k & \mathbf{0} \\ \mathbf{0} & S_k \end{bmatrix} \tag{17}$$

Where S_k is the covariance matrix of the parameter errors.

The augmentation of parameters could lead to more complex numerical solutions to the Jacobian matrix.

When numerical techniques such as forward, backward or central Euler schemes are utilized (Chang and Latif 2009), EKF might be very sensitive to the chosen step size of the parameters. EKF is also sensitive to the truncation errors and round-off errors of the parameters, which therefore makes it more difficult to choose the optimal step size (Gharamti and Hoteit 2014). Although EnKF is more capable of handling state vectors with a large degree of freedom, the increase in the degree of freedom caused by parameter augmentation, together with the spurious long-range correlations between state variables and parameters may attribute to false parameter estimation and offset the benefits of parameter updates in the EnKF assimilation (Xie and Zhang 2013, Reichle and Koster 2003). Reichle and Koster (2003) applied localization (Keppenne and Rienecker 2002) to constrain the correlations of the state vector elements beyond a certain separation distance. Xie and Zhang (2013) introduced a partitioned parameter update scheme, which is logically similar to the implementation of partially-derivative repeated updates on each parameter segment, while the rest remain fixed. This new scheme is believed superior than the regular state augmentation (Xie and Zhang 2013, Rakovec *et al.* 2015). Two major drawbacks arise for this method: first, the repetitive run of the model not only significantly increases the computation demand, but may also accumulate model errors; second, it is found the update order of the parameter segments may influence the assimilation results (Xie and Zhang 2013).

4.3.2 Dual state-parameter filter

The dual state-parameter filter method dates back to the development of Mutually Interactive State-Parameter (MISP) (Todini 1978). In this method, two Kalman Filters are run simultaneously, one for the state update and one for the model parameter update. For each time step, the two filters are run alternatively once the new observation is available: the update of the state benefits from the update of the parameters through the update of the model function, and the update of the parameters benefits from the update of the state from the update of the innovation (Bergman and Delleur 1985, Wang and Wang 1985).

The equations of a dual state-parameter EKF with a nonlinear model function and a nonlinear observation function are provided, based on the following calculation (Haykin *et al.* 2001):

First, the parameters and parameters estimate errors are propagated:

$$\theta_k^f = \theta_{k-1}^a \tag{18}$$

$$P(\theta)_k^f = \lambda P(\theta)_{k-1}^a \tag{19}$$

Where θ_k^f is the *a priori* parameters estimate, θ_{k-1}^a is the *a posteriori* parameters estimate, $P(\theta)_k^f$ is the *a priori* estimate error of the parameters, $P(\theta)_{k-1}^a$ is the *a posteriori* estimate error of the parameters, and $\lambda \in (0,1]$ is a 'forgetting factor' (Haykin *et al.* 2001, Zhong and Brown 2009).

Then, after running the state filter:

$$x_k^f = M_k(x_{k-1}^a, u_k, \theta_k^f) \tag{20}$$

$$P(x)_k^f = \dot{M}_k P(x)_{k-1}^a \dot{M}_k^T + Q_k \tag{21}$$

$$K(x)_k = P(x)_k^f \dot{H}_k^T [\dot{H}_k P(x)_k^f \dot{H}_k^T + R_k]^{-1} \tag{22}$$

$$x_k^a = x_k^f + K(x)_k \cdot (y_k^o - \dot{H}_k x_k^f) \tag{23}$$

$$P(x)_k^a = (I - K_k \dot{H}_k) P(x)_k^f \tag{24}$$

Where $P(x)_k^f$ is the *a priori* estimate error of the state, $P(x)_{k-1}^a$ is the *a posteriori* estimate error of the state, $K(x)_k$ is the Kalman gain for the state. $\dot{M}_k$ and $\dot{H}_k$ are the linearized model and observation functions calculated with equation (10) and equation (11). The rest of the variables have the same meanings as shown in equation (3) to equation (7). A major difference between the dual state-parameter filters and the state only filter is that the model function in the former is updated with the parameters.

Define:

$$v_k = (y_k^o - \dot{H}_k x_k^f) \tag{25}$$

$$C_k = \frac{\partial v_k}{\partial \theta}\Big|_{\theta=\theta_k^f} = -\frac{\partial(\dot{H}_k x_k^f)}{\partial \theta}\Big|_{\theta=\theta_k^f} = -x_k^f \cdot \frac{\partial \dot{H}_k}{\partial \theta}\Big|_{\theta=\theta_k^f} - \dot{H}_k \cdot \frac{\partial x_k^f}{\partial \theta}\Big|_{\theta=\theta_k^f} \tag{26}$$

The parameters can be updated:

$$K(\theta)_k = P(\theta)_k^f C_k^T [C_k P(\theta)_k^f C_k^T + R_k]^{-1} \tag{27}$$

$$\theta_k^a = \theta_k^f + K(\theta)_k \cdot v_k \tag{28}$$

The application of the dual state-parameter method is less common compared to state augmentation. Moradkhani *et al.* (2005b) described the application of dual state-parameter filters in EnKF. In that case, the Kalman gain of the state is calculated with the cross covariance of the states ensemble and prediction ensemble, instead of using equation (22). The Kalman gain of the parameters is calculated through the cross covariance of the parameter ensemble and prediction ensemble instead of using equation (27) (Moradkhani 2005b).

4.3.3 Hybrid solutions

The last category of the parameter update is named 'hybrid solutions', because it involves a third party scheme other than Kalman filters. Lü *et al.* (2011) coupled EKF with the optimal parameter estimation, which is realized with a particle swarm optimization to obtain the dual state-parameter estimation of root zone soil moisture within Richard's equation. This coupling was found to indeed improve the simulation results. Vrugt *et al.* (2005, 2006) presented the Simultaneous Optimization and Data Assimilation method (SODA), which uses EnKF to recursively update model states conditioned on an assumed parameter set within the inner loop, while additionally estimating time-invariant values in an outer global optimization loop using the shuffled complex evolution metropolis stochastic-ensemble optimization approach. It is proved that SODA not only improves the estimate of model parameters and state variables, but also creates reliable model prediction uncertainty bounds and a time series of valuable state and output innovations.

4.4 Routing problem

EKF and EnKF are designed to propagate forward, with the model's forecasted state predictions being updated at the same time as the observation is obtained (Crow and Ryu 2009). In reality, the gap between the continuous hydrological states and the discrete observation usually makes this assumption hard to satisfy. For example, when the hourly rainfall-runoff model is used in DA, the discrete discharge 'observation' at the watershed outlet is not only controlled by the watershed water storage 'state' at the current hour, but also that of the last few hours, when considering the lag time that the water within the watershed took to travel to demonstrated the application of AEnKF by augmenting the state vector with past forecasted observations, and argued that the AEnKF can be considered an effective method for model state updating from the operational aspects.

It is worth to point out that the prerequisite to use the methods mentioned above is that the routing scheme used in the model is the unit hydrograph, instead of the storage routing scheme (i.e., linear reservoirs). The lag time is not an issue when the storage routing scheme is applied (Vrugt *et al.* 2006, Moradkhani *et al.* 2005a).

As a conclusion, the mainstream in the existing schemes outlines: 1) augmenting the state vector with historical observation; 2) retrospective run of EnKF. The second point may slightly violate the conception of 'filtering' , but nevertheless it seems it has been well accepted as the rule of thumb to cope with the routing issue.

5 Summary and conclusions

This paper examined and discussed the latest development of Kalman type hydrological DA. Major attention has been paid to the implementation and assessment of Kalman filter type DA, especially that of EKF and EnKF. Nevertheless, many of these issues also apply to other hydrological DA methods, while the intensive research on Kalman filter provides a better platform to discuss them.

EKF features direct linearization of the system by applying Taylor expansion and solving the Jacobian matrix. The robustness issue of EKF is well discussed in hydrological DA, and much of this concerns the divergence problem. Many reasons can cause the divergence of EKF: the erroneous determination of the model and observation error, improper initialization, truncation error of Taylor expansion, and the numerical calculation errors of Jacobian matrix, among others. EnKF is a Monte Carlo method based on the Gaussian distribution assumption of the ensemble. Despite many scholars regarding EnKF as a solution to the computation issues EKF faces, the computation efficiency of EnKF itself is a primary concern. The determination of the reasonable ensemble size is regarded as a case to case problem, although more efforts are expected for the trade-off between the ensemble size and the forecast accuracy, as well as the error representativeness.

It is critical to determine the model and observation error in Kalman filter. Adaptive filtering is one of the possible solutions to this issue. Four categories of adaptive filters are summarized: Bayesian, maximum likely hood, correlation and covariance matching. Despite the adaptive filters having already been applied to stationary linear systems with Gaussian errors, their applicability is still questioned, with regards to non-stationary systems or systems with non-white errors. It is also believed that adaptive methods only work well when the model dimensions are significantly reduced.

Joint assimilation of state and parameter has been adopted in many hydrological DA studies. There are three major schemes to realize this: state augmentation, dual filter and hybrid methods. In the foreseeable future, it is unlikely that the joint estimation of state and parameter would replace the batch calibration completely. Nevertheless, such an operation may help to extend the 'efficient period', considering that the short influence period is one of the obstacles that limits the application of Kalman filter DA in an operational hydrological forecast (Muluye 2011, Agboma and Lye 2014, Grassi and de Magistris 2014).

Despite fruitful achievements in the Kalman type hydrological DA, some challenges remain. These include, but are not limited to:

1) The quantification of model and observation errors. Even though the adaptive filters may alleviate the influence of improper initial error quantifications, it is not clear as to what extent they may alleviate. Most of the current research still tends to quantify model and observation errors using trial and error or empirical equations without further updates to adaptive filters.
2) The assimilation of model parameters through a Kalman filter is still in its infancy. The 'update' of the parameters somehow contradicts the hypothesis that the parameters are 'calibrated' constants, yet there is a lack of specific comparison studies between the batch calibration and the Kalman filter assimilation update for the determination of the parameters.
3) The application of Kalman hydrological DA in large scale watersheds. The major difference between the large and small watershed is the temporal mismatch of the outlet stream flow observation and the update of the watershed and channel water storage. The update of watershed and channel water storage may not be verified with the outlet stream flow observation right away, considering the significant travel time for the water to reach the outlet in a large watershed.
4) The unknown true state. Most of the EnKF applications utilize synthetic simulation results to represent the true state of the system (some even use them as quasi observations) (Crow and Wood 2003). In the forecast, both EKF and EnKF applications eventually compare their assimilation results with the observations, of which the historical parts have already been assimilated. This indicates that the observations are consciously or unconsciously presumed identical to the true state, and are free of errors (Ahsan and O'Connor 1994). The identification of the 'true state' is also a key issue to prompt the Kalman hydrological DA from the design and experiment stage to the operational stage.
5) The handling of nonlinearity. EKF and EnKF are two of the most recognized nonlinear variants of linear Kalman filter, yet they are both suboptimal and far from perfection. More methods are desired that could balance the assimilation accuracy, computation cost, as well as the implementation simplicity.

6) The coupling of DA with data driven models (Solomatine and Ostfeld, 2008). The traditional process based rainfall runoff models place greater emphasis on model structure and parameters; these hurdle the coupling of Kalman hydrological DA at least in two ways: the failure of creating an interface to merge multiple source observations (e.g., the remote sensing products), and the models being less sensitive to the initial condition compared to input and parameters.

References

Abaza, M.*, et al.* 2014a. Sequential streamflow assimilation for short-term hydrological ensemble forecasting. *Journal of hydrology,* 519, 2692-2706.

Abaza, M., Garneau, C. and Anctil, F. 2014b. Comparison of Sequential and Variational Streamflow Assimilation Techniques for Short-Term Hydrological Forecasting. *Journal of Hydrologic Engineering*,20(2), 04014042.

Agboma, C. and Lye, L. 2014. Hydrologic Memory Patterns Assessment over a Drought-Prone Canadian Prairies Catchment. *Journal of Hydrologic Engineering*, 20(7), 04014084.

Ahsan, M. and O'Connor, K. M. 1994. A reappraisal of the Kalman filtering technique, as applied in river flow forecasting. *Journal of hydrology,* 161(1), 197-226.

Alvarez-Garreton, C., Ryu, D., *et al*., 2013. Impact of observation error structure on satellite soil moisture assimilation into a rainfall–runoff model. In: MODSIM, 20th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, pp. 3071–3077.

Alvarez-Garreton, C.*, et al.* 2014. The impacts of assimilating satellite soil moisture into a rainfall–runoff model in a semi-arid catchment. *Journal of hydrology,* 519, 2763-2774.

Anctil, F., Perrin, C. and Andreassian, V., 2003. ANN output updating of lumped conceptual rainfall/runoff forecasting models. *Journal of the American Water Resources Association*,39(5), 1269–1279..

Anderson, J. L. 2001. An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review,* 129(12), 2884-2903.

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

Aubert, D., Loumagne, C. and Oudin, L. 2003. Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall–runoff model. *Journal of hydrology,* 280(1), 145-161.

Barnes, S. L. 1964. A technique for maximizing details in numerical weather map analysis. *Journal of Applied Meteorology,* 3(4), 396-409.

Bergman, M. and Delleur, J. 1985. Kalman filter estimation and prediction of daily stream flows : I ,Review, algorithm and simulation experiments. *Water Resources Bulletin,* 21(5), 827-832.

Bertino, L., Evensen, G. and Wackernagel, H. 2003. Sequential data assimilation techniques in oceanography. *International Statistical Review,* 71(2), 223-241.

Betts, A. K., *et al.* 2003. Intercomparison of water and energy budgets for five mississippi subbasins between ecmwf reanalysis (era‐40) and nasa data assimilation office fvgcm for 1990–1999. *Journal of Geophysical Research: Atmospheres (1984–2012),* 108(D16).

Bin, W. and Ying, Z. 2005. A new data assimilation approach. *Acta Meteorologica Sinica,* 63(5):694-701.

Bishop, C. H., Etherton, B. J. and Majumdar, S. J. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Monthly Weather Review,* 129(3), 420-436.

Bøgh, E., *et al.* 2004. Incorporating remote sensing data in physically based distributed agro-hydrological modelling. *Journal of hydrology,* 287(1), 279-299.

Bolzern, P., Ferrario, M. and Fronza, G. 1980. Adaptive real-time forecast of river flow-rates from rainfall data. *Journal of hydrology,* 47(3), 251-267.

Borup, M., *et al.* 2015. A partial ensemble Kalman filtering approach to enable use of range limited observations. *Stochastic Environmental Research and Risk Assessment,* 29(1), 119-129.

Borup, M., 2014. *Real Time Updating in Distributed Urban Rainfall Runoff Modelling*. Thesis (PhD). Technical University of Denmark.

Boulet, G., *et al.* 2000. A simple water and energy balance model designed for regionalization and remote sensing data utilization. *Agricultural and Forest Meteorology,* 105(1), 117-132.

Brocca, L., *et al.* 2012. Assimilation of surface-and root-zone ASCAT soil moisture products into rainfall–runoff modeling. *Geoscience and Remote Sensing, IEEE Transactions on,* 50(7), 2542-2555.

Broersen, P. M. 2007. Error correction of rainfall-runoff models with the ARMAsel program. *Instrumentation and Measurement, IEEE Transactions on,* 56(6), 2212-2219.

Burgers, G., Jan van Leeuwen, P. and Evensen, G. 1998. Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review,* 126(6), 1719-1724.

Cammalleri, C. and Ciraolo, G. 2012. State and parameter update in a coupled energy/hydrologic balance model using ensemble Kalman filtering. *Journal of hydrology,* 416, 171-181.

Chang, S.-Y. and Latif, S. M. I. 2009. Extended Kalman filtering to improve the accuracy of a subsurface contaminant transport model. *Journal of Environmental Engineering,* 136(5), 466-474.

Chen, F., Crow, W. T. and Ryu, D. 2014. Dual forcing and state correction via soil moisture assimilation for improved rainfall–runoff modeling. *Journal of hydrometeorology,* 15(5), 1832-1848.

Chen, F., *et al.* 2011. Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture. *Advances in Water Resources,* 34(4), 526-536.

Chen, H., *et al.* 2013. Hydrological data assimilation with the Ensemble Square-Root-Filter: Use of streamflow observations to update model states for real-time flash flood forecasting. *Advances in Water Resources,* 59, 209-220.

Chen, L., *et al.* 2015. Real-time error correction method combined with combination flood forecasting technique for improving the accuracy of flood forecasting. *Journal of hydrology,* 521, 157-169.

Clark, M. P., *et al.* 2006. Assimilation of snow covered area information into hydrologic and land-surface models.

Advances in Water Resources, 29(8), 1209-1221.

Clark, M. P.*, et al.* 2008. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. *Advances in Water Resources,* 31(10), 1309-1324.

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: Atmospheres (1984–2012),* 108(D22).

Coustau, M.*, et al.* 2013. Benefits and limitations of data assimilation for discharge forecasting using an event-based rainfall–runoff model. *Natural Hazards and Earth System Science,* 13(3), 583-596.

Cressman, G. P. 1959. An operational objective analysis system. *Monthly Weather Review,* 87(10), 367-374.

Crow, W. and Reichle, R. 2008. Comparison of adaptive filtering techniques for land surface data assimilation. *Water resources research,* 44, W08423, doi:10.1029/2008WR006883.

Crow, W. and Ryu, D. 2009. A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals. *Hydrology and Earth System Sciences,* 13(1), 1-16.

Crow, W. T. and Wood, E. F. 2003. The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97. *Advances in Water Resources,* 26(2), 137-149.

Crow, W. and Yilmaz, M. T. 2014. The Auto‐Tuned Land Data Assimilation System (ATLAS). *Water resources research,* 50(1), 371-385.

Das, N.*, et al.* 2008. Modeling and assimilation of root zone soil moisture using remote sensing observations in Walnut Gulch Watershed during SMEX04. *Remote Sensing of Environment,* 112(2), 415-429.

de Rosnay, P.*, et al.* 2013. A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. *Quarterly Journal of the Royal Meteorological Society,* 139(674), 1199-1213.

Dechant, C. and Moradkhani, H. 2011. Radiance data assimilation for operational snow and streamflow forecasting. *Advances in Water Resources,* 34(3), 351-364.

Dee, D. P. 1995. On-line estimation of error covariance parameters for atmospheric data assimilation. *Monthly Weather Review,* 123(4), 1128-1145.

Derber, J. and Rosati, A. 1989. A global oceanic data assimilation system. *Journal of Physical Oceanography,* 19(9), 1333-1347.

Divac, D.*, et al.* 2009. A procedure for state updating of SWAT-based distributed hydrological model for operational runoff forecasting. *Journal of the Serbian Society for Computational Mechanics,* 3(1), 298-326.

Doubková, M.*, et al.* 2012. Evaluation of the predicted error of the soil moisture retrieval from C-band SAR by comparison against modelled soil moisture estimates over Australia. *Remote Sensing of Environment,* 120, 188-196.

Draper, C.*, et al.* 2011. Assimilation of ASCAT near-surface soil moisture into the SIM hydrological model over France. *Hydrology and Earth System Sciences,* 15(12), 3829-3841.

Draper, C., Mahfouf, J. F. and Walker, J. 2009. An EKF assimilation of AMSR‐E soil moisture into the ISBA land surface scheme. *Journal of Geophysical Research: Atmospheres (1984–2012),* 114(D20).

Drécourt, J.-P. 2003. Kalman filtering in hydrological modeling, Technical Report. *Hørsholm, Denmark, DAIHM*.

Duan, Z., Han, C. and Dang, H. 2003. An adaptive Kalman Filter with dynamic rescaling of process noise. *Proceedings of the Sixth Inter-national Conference of Information Fusion*, vol.2, 1310-1315.

Dumedah, G. and Coulibaly, P. 2012. Integration of evolutionary based assimilation into Kalman-type methods for streamflow simulations in ungauged watersheds. *Journal of hydrology,* 475, 428-440.

Dunne, S. and Entekhabi, D. 2005. An ensemble ‐ based reanalysis approach to land data assimilation. *Water resources research,* 41(2).

El Serafy, G. and Mynett, A., Comparison of EKF and EnKF in Sobek River: Case study Maxau-Ijssel. ed. *Proceedings of the 6th International Conference on Hydroinformatics. World Scientific, Singapore, ISBN*, 2004, 981-238.

Ermolaev, P. and Volynsky, M., The second order extended Kalman filter and Markov nonlinear filter for data processing in interferometric systems. ed. *Journal of Physics: Conference Series*, 2014, 012015.

Evensen, G. 1994a. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *JOURNAL OF GEOPHYSICAL RESEARCH-ALL SERIES-,* 99, 10-10.

Evensen, G. 1994b. Sequential data assimilation with a nonlinear quasi ‐ geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans (1978–2012),* 99(C5), 10143-10162.

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

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

Evensen, G. and Van Leeuwen, P. J. 2000. An ensemble Kalman smoother for nonlinear dynamics. *Monthly Weather Review,* 128(6), 1852-1867.

Fairbairn, D.*, et al.* 2014. Comparing the SEKF with the DEnKF on a land surface model [online]. Available from: http://www.meteo.fr/cic/meetings/2014/CNA/presentations/3_2_Fairbairn.pdf. (Accessed March 31,3015)

Fan, Y. C. 1991. Prediction of stage-discharge in river using Kalman filter. *Journal of Hydrodynamics,* 3(1), 51-55.

Fitzgerald, R. J. 1971. Divergence of the Kalman filter. *Automatic Control, IEEE Transactions on,* 16(6), 736-747.

Francois, C., Quesney, A. and Ottlé, C. 2003. Sequential assimilation of ERS-1 SAR data into a coupled land surface-hydrological model using an extended Kalman filter. *Journal of hydrometeorology,* 4(2), 473-487.

Galantowicz, J. F., Entekhabi, D. and Njoku, E. G. 1999. Tests of sequential data assimilation for retrieving profile soil moisture and temperature from observed L-band radiobrightness. *Geoscience and Remote Sensing, IEEE Transactions on,* 37(4), 1860-1870.

Gandin, L. S. 1965. Objective analysis of meteorological fields (Vol. 242). R. Hardin (Ed.). Jerusalem: Israel program for scientific translations.

Ge, G. 1984. The Kalman Filtering Algorithm for the Model of Runoff Formation at the Natural Storage Point.

Journal of Chendu Univeristy of Technology, 1(4), 69-78.

Georgakakos, K. P. and Bras, R. L. 1982. Real‐time, statistically linearized, adaptive flood routing. *Water resources research,* 18(3), 513-524.

Gharamti, M. and Hoteit, I. 2014. Complex step-based low-rank extended Kalman filtering for state-parameter estimation in subsurface transport models. *Journal of hydrology,* 509, 588-600.

Gift Dumedah and Paulin Coulibaly.2013. Integration of an evolutionary algorithm into the ensemble Kalman filter and the particle filter for hydrologic data assimilation. *Journal of Hydroinformatics*, Vol 16 No 1 pp 74–94. doi:10.2166/hydro.2013.088.

Gillijns, S., *et al.* 2006. "What is the ensemble Kalman filter and how well does it work?", Proc. *Amer. Control Conf.,* pp.4448 -4453.

Grassi, S. and de Magistris, P. S. 2014. When long memory meets the Kalman filter: A comparative study. *Computational Statistics & Data Analysis,* 76, 301-319.

Hamilton, J. D., 1994. *Time series analysis.* Princeton, NJ: Princeton University press.

Han, E., Merwade, V. and Heathman, G. C. 2012. Implementation of surface soil moisture data assimilation with watershed scale distributed hydrological model. *Journal of hydrology,* 416, 98-117.

Haykin, S. S.2001. *Kalman filtering and neural networks.* New York, NY: John Wiley & Sons Ltd.

He, M., *et al.* 2012. An integrated uncertainty and ensemble-based data assimilation approach for improved operational streamflow predictions. *Hydrology and Earth System Sciences,* 16(3), 815-831.

Hendricks Franssen, H. and Kinzelbach, W. 2008. Real‐time groundwater flow modeling with the ensemble Kalman filter: Joint estimation of states and parameters and the filter inbreeding problem. *Water resources research,* 44(9).

Hess, R. 2001. Assimilation of screen-level observations by variational soil moisture analysis. *Meteorology and Atmospheric Physics,* 77(1-4), 145-154.

Houser, P. R., De Lannoy, G. J. and Walker, J. P. 2005. Hydrologic Data Assimilation. In Advances in Water Science Methodologies, Aswathanarayana, A. (ed.), AA Balkema, The Netherlands, pp. 230.

Houser, P. R., *et al.* 1998. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water resources research,* 34(12), 3405-3420.

Houtekamer, P. L. and Mitchell, H. L. 1998. Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review,* 126(3), 796-811.

Huang, C., Li, X. and Lu, L. 2008. Retrieving soil temperature profile by assimilating MODIS LST products with ensemble Kalman filter. *Remote Sensing of Environment,* 112(4), 1320-1336.

Huang, W.-C. 1999. Kalman Filter Effective to Hydrologic Routing? *Journal of Marine Science and Technology,* 7(1), 65-71.

Ines, A. V. and Mohanty, B. P. 2008. Near‐surface soil moisture assimilation for quantifying effective soil hydraulic properties using genetic algorithm: 1. Conceptual modeling. *Water resources research,* 44(6).

Jonsdottir, H., Madsen, H. and Palsson, O. P. 2006. Parameter estimation in stochastic rainfall-runoff models.

Journal of hydrology, 326(1), 379-393.

Jwo, D.-J. and Wang, S.-H. 2007. Adaptive fuzzy strong tracking extended Kalman filtering for GPS navigation. *Sensors Journal, IEEE,* 7(5), 778-789.

Kaheil, Y. H.*, et al.* 2008. Downscaling and assimilation of surface soil moisture using ground truth measurements. *Geoscience and Remote Sensing, IEEE Transactions on,* 46(5), 1375-1384.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering,* 82(1), 35-45.

Kalnay, E.*, et al.* 2007. 4 - D - Var or ensemble Kalman filter? *Tellus A,* 59(5), 758-773.

Keppenne, C. L. and Rienecker, M. M. 2002. Initial testing of a massively parallel ensemble Kalman filter with the Poseidon isopycnal ocean general circulation model. *Monthly Weather Review,* 130(12), 2951-2965.

Kim, S., Tachikawa, Y. and Takara, K. 2005. Real-time prediction algorithm with a distributed hydrological model using Kalman filter. *Annual journal of hydraulic engineering, JSCE,* (49), 163-168.

Kitanidis, P. K. and Bras, R. L. 1980a. Real-time forecasting with a conceptual hydrologic model 1. Analysis of Uncertainty. *WATER RESOURCES RES.,* 16(6), 1034-1044.

Kitanidis, P. K. and Bras, R. L. 1980b. Real - time forecasting with a conceptual hydrologic model: 1. Analysis of Uncertainty. *Water resources research,* 16(6), 1025-1033.

Knight, D. and Shamseldin, A., 2006. *River basin modelling for flood risk mitigation.* D. W. Knight and A. Y. Shamseldin, eds., A. A. Balkema, Brookfield, VT, 301–334..

Kolmogorov, A. N., Doyle, W. and Selin, I. 1941. Interpolation and extrapolation of stationary random sequences. *Izv. Akad. Nauk. SSSR Ser. Mat., 1941,5, 3–14. (in Russian) Translation in A. N. Shiryagev (Ed.), Selected works of A. N. Kolmogorov (Vol. 23, pp. 272–280). Norwell, MA: Kluwer*.

Kumar, P. and Kaleita, A. L. 2003. Assimilation of near-surface temperature using extended Kalman filter. *Advances in Water Resources,* 26(1), 79-93.

Kumar, S. V.*, et al.* 2008. A land surface data assimilation framework using the land information system: Description and applications. *Advances in Water Resources,* 31(11), 1419-1432.

Lawson, W. G. and Hansen, J. A. 2004. Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Monthly Weather Review,* 132(8), 1966-1981.

Lee, Y. and Singh, V. 1999. Tank model using Kalman filter. *Journal of Hydrologic Engineering,* 4(4), 344-349.

Lei, F.*, et al.* 2014. Improving the estimation of hydrological states in the SWAT model via the ensemble Kalman smoother: Synthetic experiments for the Heihe River Basin in northwest China. *Advances in Water Resources,* 67, 32-45.

Lewis, F. L., Xie, L. and Popa, D., 2008. *Optimal and robust estimation: with an introduction to stochastic control theory.* CRC.

Li, Y.*, et al.* 2013. Assimilation of stream discharge for flood forecasting: The benefits of accounting for routing time lags. *Water resources research,* 49(4), 1887-1900.

Li, Y.*, et al.* 2014. An integrated error parameter estimation and lag-aware data assimilation scheme for real-time flood forecasting. *Journal of hydrology,* 519, 2722-2736.

Liu, J., DOAN, C. D. and Liong, S.-Y. 2011. Conceptual Rainfall-Runoff Model with Kalman Filter for Parameter and Outflow Updating. *Advances in Geosciences (a 6-volume Set)-Volume 23: Hydrological Science (hs),* 23, 133.

Liu, S.*, et al.* 2012a. Improved regional hydrologic modelling by assimilation of streamflow data into a regional hydrologic model. *Environmental Modelling & Software,* 31, 141-149.

Liu, Y. and Gupta, H. V. 2007. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water resources research,* 43(7). DOI:10·1029/2006WR005756.

Liu, Y.*, et al.* 2015. Blending satellite ‐ based snow depth products with in situ observations for streamflow predictions in the Upper Colorado River Basin. *Water resources research*. 51(2), 1182-1202.

Liu, Y.*, et al.* 2012b. Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. *Hydrology and Earth System Sciences,* 16(10), 3863-3887.

Ljung, L. 1979. Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems. *Automatic Control, IEEE Transactions on,* 24(1), 36-50.

Lorenc, A. 1981. A global three-dimensional multivariate statistical interpolation scheme. *Monthly Weather Review,* 109(4), 701-721.

Lorenc, A. C. 2003. The potential of the ensemble Kalman filter for NWP—a comparison with 4D ‐ Var. *Quarterly Journal of the Royal Meteorological Society,* 129(595), 3183-3203.

Lü, H.*, et al.* 2010. Using a H∞ filter assimilation procedure to estimate root zone soil water content. *Hydrological processes,* 24(25), 3648-3660.

Lü, H.*, et al.* 2011. Dual state-parameter estimation of root zone soil moisture by optimal parameter estimation and extended Kalman filter data assimilation. *Advances in Water Resources,* 34(3), 395-406.

Madsen, H.*, et al.* 2003. Data assimilation in the MIKE 11 Flood Forecasting system using Kalman filtering. *International Association of Hydrological Sciences, Publication,* (281), 75-81.

Mahfouf, J. F.*, et al.* 2009. A comparison of two off ‐ line soil analysis schemes for assimilation of screen level observations. *Journal of Geophysical Research: Atmospheres (1984 ‐ 2012),* 114(D8).

Massari, C.*, et al.* 2014. Potential of soil moisture observations in flood modelling: Estimating initial conditions and correcting rainfall. *Advances in Water Resources,* 74, 44-53.

McLaughlin, D. 1995. Recent developments in hydrologic data assimilation. *Reviews of Geophysics,* 33(S2), 977-984.

McLaughlin, D. 2002. An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering. *Advances in Water Resources,* 25(8), 1275-1286.

McMillan, H.*, et al.* 2010. Impacts of uncertain river flow data on rainfall ‐ runoff model calibration and discharge predictions. *Hydrological processes,* 24(10), 1270-1284.

McMillan, H.*, et al.* 2013. Operational hydrological data assimilation with the recursive ensemble Kalman filter. *Hydrology and Earth System Sciences,* 17(1), 21-38.

Mehra, R. K. 1970. On the identification of variances and adaptive Kalman filtering. *Automatic Control, IEEE*

comparison of approaches to mapping observational information to state space. *Water resources research,* 45(8).

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

Pham, D. T. 2001. Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Monthly Weather Review,* 129(5), 1194-1207.

Piani, C., Haerter, J. and Coppola, E. 2010. Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology,* 99(1-2), 187-192.

Puente, C. E. and Bras, R. L. 1987. Application of nonlinear filtering in the real time forecasting of river flows. *Water resources research,* 23(4), 675-682.

Rabier, F., Liu,Z., 2003. *Variational assimilation:theory and overview: theory and overview. Proc. Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean. Reading, United Kingdom, ECMWF, 2003,29–43.*.

Randrianasolo, A.*, et al.* 2014. Impact of streamflow data assimilation and length of the verification period on the quality of short-term ensemble hydrologic forecasts. *Journal of hydrology,* 519, 2676-2691.

Rafieeinasab, A.*, et al.* 2014. Comparative evaluation of maximum likelihood ensemble filter and ensemble Kalman filter for real-time assimilation of streamflow data into operational hydrologic models. *Journal of hydrology,* 519, 2663-2675.

Rakovec, O.*, et al.* 2015. Operational aspects of asynchronous filtering for hydrological forecasting. *Hydrology and Earth System Sciences Discussions,* 12(3), 3169-3203.

Rasmussen, J.*, et al.* 2015. Data assimilation in integrated hydrological modeling using ensemble Kalman filtering: evaluating the effect of ensemble size and localization on filter performance. *Hydrology and Earth System Sciences Discussions,* 12(2), 2267-2304.

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

Reichle, R. H., Crow, W. T. and Keppenne, C. L. 2008. An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water resources research,* 44(3).

Reichle, R. H. and Koster, R. D. 2003. Assessing the impact of horizontal error correlations in background fields on soil moisture estimation. *Journal of hydrometeorology,* 4(6), 1229-1242.

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

Reichle, R. H., McLaughlin, D. B. and Entekhabi, D. 2002a. Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review,* 130(1), 103-114.

Reichle, R. H.*, et al.* 2002b. Extended versus ensemble Kalman filtering for land data assimilation. *Journal of hydrometeorology,* 3(6), 728-740.

Hydroinformatics, 10(1), 3–22. doi:10.2166/hydro.2008.015.

Sun, L., *et al.*, 2013. Prediction of Daily Discharge at Bakel (Senegal) using Multiple Linear Regression, Kalman Filer and Artificial Neural Networks. *CSCE 2013, 3rd Specialty Conference on Disaster Prevention and Mitigation.* Montreal,Canada.

Szöllősi-Nagy, A. and Mekis, E. 1988. Comparative analysis of three recursive real-time river flow forecasting models: deterministic, stochastic, and coupled deterministic-stochastic. *Stochastic Hydrology and Hydraulics,* 2(1), 17-33.

Tanizaki, H., 1996. *Nonlinear filters: estimation and applications.* New York: Springer Science & Business Media.

Tao, W., *et al.* 2005. A Method of Preventing Divergence of EKF. *JOURNAL OF PROJECTILES, ROCKETS, MISSILES AND GUIDANCE,* 25(2), 484-486.

Todini, 1978. E. Todini Mutually interactive state-parameter (MISP) estimation, C.L. Chiu (Ed.), Proc. AGU Chapman Conf. on Application of Kalman Filter to Hydrology, Hydraulics and Water Resources, Department of Civil Engineering, University of Pittsburgh, Pittsburgh (1978), pp. 135–151

Tran, A. P., *et al.* 2014. Joint estimation of soil moisture profile and hydraulic parameters by ground - penes(5(p)11.0)7.3( )-92

Walker, J. P., Houser, P. R. and Reichle, R. H. 2003. New technologies require advances in hydrologic data assimilation. *EOS, Transactions American Geophysical Union,* 84(49), 545-551.

Walker, J. P. and Houser, P. R. 2005. Hydrologic data assimilation. *Advances in water science methodologies. Londres: Taylor & Francis, ed,* 1, 25-48.

Wan, E. A. and Van Der Merwe, R., The unscented Kalman filter for nonlinear estimation. ed. *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, 2000, 153-158.

Wanders, N.*, et al.* 2014a. The benefits of using remotely sensed soil moisture in parameter identification of large‐scale hydrological models. *Water resources research,* 50(8), 6874-6891.

Wanders, N.*, et al.* 2014b. The suitability of remotely sensed soil moisture for improving operational flood forecasting. *Hydrology and Earth System Sciences,* 18(6), 2343-2357.

Wang, D. and Cai, X. 2008. Robust data assimilation in hydrological modeling–A comparison of Kalman and H-infinity filters. *Advances in Water Resources,* 31(3), 455-472.

Wang, D., Chen, Y. and Cai, X. 2009. State and parameter estimation of hydrologic models using the constrained ensemble Kalman filter. *Water resources research,* 45(11).

Wang, W. and Kou, X. 2009. Methods for hydrological data assimilation and advances of assimilating remotely sensed data into rainfall-runoff models. *Journal of Hohai University: Natural Sciences,* 37(5), 556-562.

Wang, Y. and Wang, S. 1985. The application of adaptive data assimilation in hydrological forecast. *Hydrology,* 1, 56-58.

Weerts, A. H. and El Serafy, G. Y. 2006. Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall‐runoff models. *Water resources research,* 42(9).

Whitaker, J. S. and Hamill, T. M. 2002. Ensemble data assimilation without perturbed observations. *Monthly Weather Review,* 130(7), 1913-1924.

Wiener, N., 1949. *Extrapolation, interpolation, and smoothing of stationary time series.* MIT press Cambridge, MA.

Wu, X.-l.*, et al.* 2008. Kalman filtering correction in real-time forecasting with hydrodynamic model. *Journal of Hydrodynamics, Ser. B,* 20(3), 391-397.

Xie, X. and Zhang, D. 2010. Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. *Advances in Water Resources,* 33(6), 678-690.

Xie, X. and Zhang, D. 2013. A partitioned update scheme for state‐parameter estimation of distributed hydrologic models based on the ensemble Kalman filter. *Water resources research,* 49(11), 7350-7365.

Xu, X., Li, J. and Tolson, B. A. 2014. Progress in integrating remote sensing data and hydrologic modeling. *Progress in Physical Geography*, 0309133314536583.

Yu, P.-S. and Chen, S.-T. 2005. Updating Real-Time Flood Forecasting Using a Fuzzy Rule-Based Model/Mise à Jour de Prévision de Crue en Temps Réel Grâce à un Modèle à Base de Règles Floues. *Hydrological Sciences Journal,* 50(2), 265-278.

Zhong, J. and Brown, M. 2009. Joint State and Parameter Estimation for Biochemical Dynamic Pathways with Iterative Extended Kalman Filter: Comparison with Dual State and Parameter Estimation. *The Open*

Automation and Control Systems Journal, 2(1), 69-77.

List of Figures

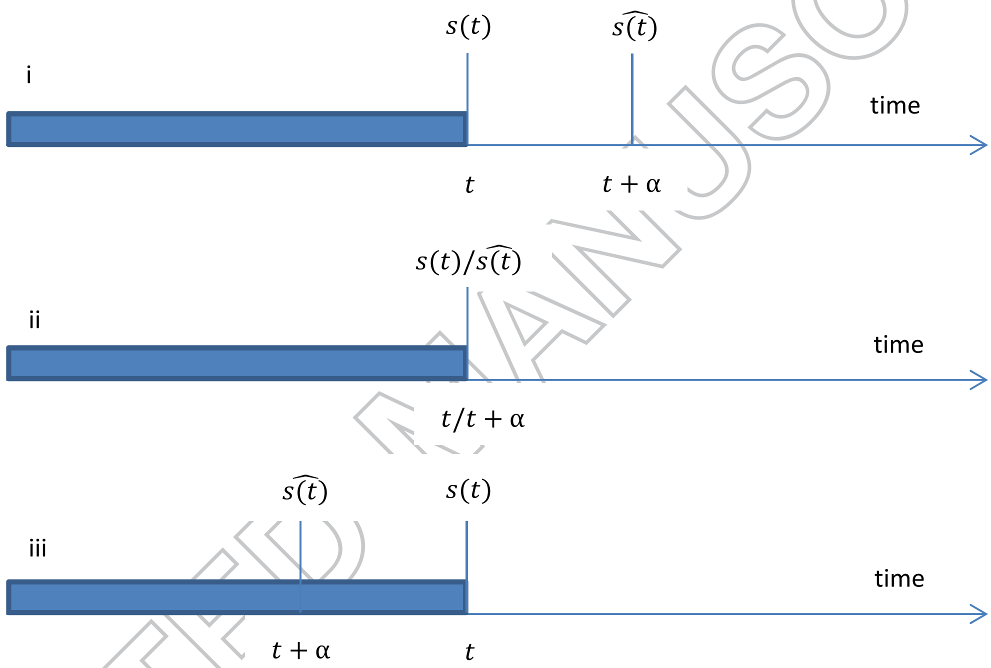

Fig. 1 Types of estimation problem: i) prediction; ii) filtering; iii) smoothing

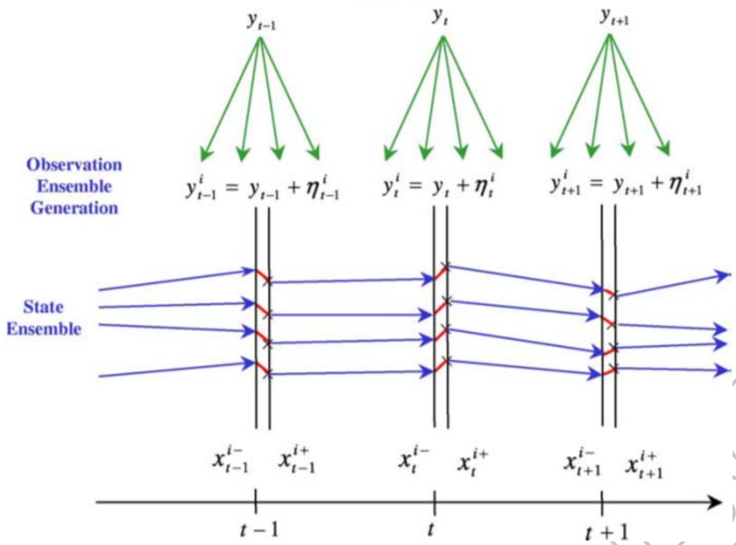

Fig.2 Schematic of EnKF with perturbed observation (Moradkhani *et al.* 2005b)

Table

Table 1 Literature summary of state and parameter(s) selection

Literature	Method	Model	State	Parameter(s)	Observations
(Reichle *et al.* 2002a)	EnKF	(Reichle *et al.* 2001)	soil moisture and soil temperature	-	L-band microwave radio brightness
(Madsen *et al.* 2003)	EnKF	MIKE 11	water level and flux	-	water level and flux
(Moradkhani *et al.* 2005b)	EnKF	HyMOD	water storage	Rq, Rs, α, bexp, Cmax	stream flow
(Vrugt *et al.* 2006)	EnKF	SAC-SMA	Six soil zone reservoir variable, three routing variables	14 time invariant parameters	stream flow
(Weerts and El Serafy 2006)	EnKF, Particle filter	HBV96	soil moisture, snow water, interception storage, lower and upper zone, etc	-	discharge
(Clark *et al.* 2008)	EnKF	Topnet	soil storage, aquifer storage, surface storage, stream flow prediction	-	stream flow
(Xie and Zhang 2010)	EnKF	SWAT	runoff, water storage. evapotranspiration, etc.	CN2	stream flow, soil water content, evapotranspiration
(Dechant and Moradkhani 2011)	EnKF, Particle filter	Snow 17, SAC-SMA	snow water equivalent (SWE)	All model parameters	microwave radiance
(Chen *et al.* 2011)	EnKF	SWAT	soil moisture	-	surface Soil moisture
(Cammalleri and Ciraolo 2012)	EnKF	SVAT	soil moisture	All model parameters	surface temperature, latent heat flux, soil water content
(Han *et al.* 2012)	EnKF	SWAT	soil moisture	-	surface soil moisture
(Trudel *et al.* 2014)	EnKF	CATHY	pressure head, incoming and outgoing discharge	-	stream flow
(Samuel *et al.* 2014)	EnKF	SAC-SMA	soil moisture content, SWE, etc	10 time variant parameters	stream flow, soil moisture
(Kitanidis and Bras 1980b)	EKF	NWS	soil water content, additional	-	stream flow

			impervious storage, channel storage, routing coefficient		
(Puente and Bras 1987)	EKF	A conceptual lumped model	water equivalent mass in cloud, soil water content, channel reservoir storage, etc.	-	discharge, precipitation
(Francois *et al.* 2003)	EKF	GRKAL	soil moisture	-	surface soil moisture
(Aubert *et al.* 2003)	EKF	GR4J	water level in soil reservoir and water level in routing reservoir	-	soil moisture, stream flow
(Kumar and Kaleita 2003)	EKF	LSM	soil temperature	-	near surface soil temperature
(Jonsdottir *et al.* 2006)	EKF, maximum likelihood	A stochastic model similar to HBV and NAM	temperature, snow, upper and lower water content	All model parameters	discharge
(Lü *et al.* 2011)	EKF, PSO	1-D Richards' Equation	soil moisture content	K_s,θ_r, θ_s,α,n	surface soil moisture