

A review on evapotranspiration data assimilation based on hydrological models

DONG Qingqing¹, *ZHAN Chesheng², WANG Huixiao¹, WANG Feiyu^{2,3}, ZHU Mingcheng¹

1. College of Water Sciences, Beijing Normal University, Beijing 100875, China;

2. Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;

3. University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: Accurate estimation of evapotranspiration (ET), especially at the regional scale, is an extensively investigated topic in the field of water science. The ability to obtain a continuous time series of highly precise ET values is necessary for improving our knowledge of fundamental hydrological processes and for addressing various problems regarding the use of water. This objective can be achieved by means of ET data assimilation based on hydrological modeling. In this paper, a comprehensive review of ET data assimilation based on hydrological modeling is provided. The difficulties and bottlenecks of using ET, being a non-state variable, to construct data assimilation relationships are elaborated upon, with a discussion and analysis of the feasibility of assimilating ET into various hydrological models. Based on this, a new easy-to-operate ET assimilation scheme that includes a water circulation physical mechanism is proposed. The scheme was developed with an improved data assimilation system that uses a distributed time-variant gain model (DTVGM), and the ET-soil humidity nonlinear time response relationship of this model. Moreover, the ET mechanism in the DTVGM was improved to perfect the ET data assimilation system. The new scheme may provide the best spatial and temporal characteristics for hydrological states, and may be referenced for accurate estimation of regional evapotranspiration.

Keywords: evapotranspiration; data assimilation; hydrological model; non-state variable

1 Introduction

Evapotranspiration (ET) is the main process controlling the water cycle and energy transport between the atmosphere, hydrosphere, and biosphere (Priestley and Taylor, 1972), and is an important subject of research on global and regional water and energy budgets (Rosenberg, 1983; Kustas and Norman, 1996; Vinukollu *et al.*, 2011). Remote sensing and hydrological

Received: 2015-06-19 **Accepted:** 2015-07-21

Foundation: National Key Basic Research Program of China (973 Program), No.2015CB452701; National Natural Science Foundation of China, No.41271003, No.41371043, No.41401042

Author: Dong Qingqing (1990–), Master Candidate, specialized in evapotranspiration data assimilation.

E-mail: dongqingqing@mail.bnu.edu.cn

***Corresponding author:** Zhan Chesheng (1975–), Associate Professor, E-mail: zhancs@igsnr.ac.cn

modeling are two key approaches to estimate evapotranspiration (Liu *et al.*, 2007). Hydrological modeling is prone to errors due to uncertainties in input and output information, the model structure, initial conditions and model parameters, which affect the accuracy of the simulation (Renard *et al.*, 2010). Using remote sensing to estimate ET has obvious advantages in accuracy and spatial resolution (Li *et al.*, 2009), but cannot provide temporally continuous values and thus cannot meet the requirements of hydrological models (Liang *et al.*, 2013). One possible approach to overcome these drawbacks involves applying multiple means or multi-source data to estimate ET, which is becoming the choice method for obtaining a continuous time series of ET with high precision (Conradt *et al.*, 2013).

Data assimilation, as an advantageous technique for combining multi-source data, has been at the frontier of research on land surface hydrology and the water cycle (Song *et al.*, 2011; Tang and Li, 2014) and has provided a new way of obtaining a highly accurate continuous time series of regional ET values. Pan *et al.* (2008) and Qin *et al.* (2008) applied the data assimilation technique and hydrological modeling to assimilate observed data, and this process improved the efficiency of the ET simulation, however, the assimilation effect was seriously affected by ET not being a state variable. Xie and Zhang (2010) assimilated flow into the Soil and Water Assessment Tool (SWAT) hydrological model, which did not significantly improve the accuracy of ET. The application of data assimilation based on hydrological models has made some progress (Xu *et al.*, 2014) when the assimilated objects are primarily state variables; data assimilation for non-state variables, such as ET, is still being attempted (Spies *et al.*, 2014). The potential of data assimilation in hydrological models should be further investigated. Therefore, we need to carry out studies of ET data assimilation based on hydrological models, and overcome the bottleneck for the use of the non-state variable ET when constructing data assimilation relationships, so as to obtain a highly precise continuous time series of ET values at the regional scale.

In this paper, a comprehensive review of evapotranspiration data assimilation based on hydrological modeling is provided. The difficulties and bottlenecks of using ET, being a non-state variable, to construct data assimilation relationships are elaborated upon, with a discussion and analysis of the feasibility of assimilating ET into various hydrological models. Based on this, a new easy-to-operate ET assimilation scheme that includes a water circulation physical mechanism is proposed. The scheme was developed with an improved data assimilation system that uses a distributed time-variant gain model (DTVGM), and the ET-soil humidity nonlinear time response relationship of this model. This scheme may provide a reference for an accurate simulation of regional evapotranspiration.

2 Evapotranspiration data assimilation

Data assimilation techniques, while only gradually being applied to hydrological simulations, have nevertheless recently become a hot research topic in the water cycle field. Data assimilation methods and their application to land surface and hydrological modeling have made several achievements in various systems, including the North American/Global Land Data Assimilation System (NLDAS/GLDAS) and European Land Data Assimilation System (ELDAS), as well as the assimilation systems in China, Canada and South Korea (Li *et al.*, 2007). However, current research is mainly focused on state variables, such as soil moisture,

leaf area index, etc., and relatively little on the non-state variable ET. Moreover, there has been little research on the emerging field of applying data assimilation in hydrology simulations for small-scale watersheds, and the potential of such data assimilation has not been fully realized (Moradkhani, 2008), especially for ET data assimilation based on hydrological models (Chen *et al.*, 2013).

The amount of research dedicated to ET data assimilation began ten years ago. Several papers have described attempts to use data assimilation techniques combined with various other models to estimate regional ET. Schuurmans *et al.* (2003) used ET estimates derived from remote sensing and the Surface Energy Balance Algorithm for Land (SEBAL) to improve distributed hydrological model (SIMGRO) simulations using a constant gain Kalman filter data assimilation algorithm, which calculated the state variable soil moisture using an empirical method after updating the ET value. In this study, the model was interpolated with ET observations relying on the calibration of empirical parameters, and hence was not statistically optimal. Pipunic *et al.* (2008) developed a data assimilation scheme with the ensemble Kalman filter (EnKF) to estimate latent heat flux and sensible heat flux based on a one-dimensional land surface model (LSM), which improved the simulation accuracy. However, assumptions in the LSM model have limited the popularity of this method. Qin *et al.* (2008) applied the extended Kalman filter to assimilate remote-sensing-derived ET estimates into a distributed hydrological model for improving the predictions of spatial water distribution over a large river basin, but the assimilation effect could not provide feedback into the model since ET is a non-state variable. The hydrological series had not been optimized as a whole, and was just equivalent to an interpolation. Irmak and Kamble (2009) proposed an assimilation methodology for the soil, water, atmosphere, and plant (SWAP) simulation model with remote sensing data from the SEBAL using genetic algorithm (GA) data assimilation. For this, uncertainty in the SEBAL model, due to its one evaporation mechanism, reduced the reliability of the methodology; besides, the assimilation method was virtually the further optimization of SWAP through genetic algorithm, no calibration between parameters within the SWAP model. Xie and Zhang (2010) tested a data assimilation system using streamflow to assimilate SWAT; this assimilation improved estimates of runoff and soil moisture, but not of ET. Dumedah and Coulibaly (2013) assimilated streamflow into the distributed hydrologic model (SWAT) using in-situ soil moisture data, and demonstrated some improvement in accuracy. However, the ET results were not effectively improved since the errors generated from the simulation were not completely corrected. Lei *et al.* (2014) assimilated synthetic surface soil moisture data into the SWAT model to evaluate their impact on other hydrological variables via the ensemble Kalman smoother (EnKS). The results showed that the assimilation of surface soil moisture can moderately improve estimates of deep layer soil moisture, surface runoff and lateral flow, while ET was still underestimated. Trudel *et al.* (2014) applied an ensemble Kalman filter into the distributed physically based hydrological model CATHY (CATchment HYdrology) to assimilate streamflow observations at different locations, as well as soil moisture at two different depths (15 and 45 cm). Assimilation of streamflow observations systematically increased the simulated soil moisture values, but no improvement in the accuracy of the ET sequence.

These studies have been useful for estimating evapotranspiration, but bottlenecks remain, and an ET assimilation system based on a hydrological model has not yet been established in

practice. Effectively combining and integrating observations and simulations is necessary. That is, by using “true values” and minimizing errors to constrain model simulations, and by effectively merging information obtained from different sources with different spatial and temporal resolutions, higher resolution as well as spatially and temporally continuous data can then be obtained, which can yield a multi-scale representation of the water and energy cycle (Li *et al.*, 2007).

3 Evapotranspiration data assimilation based on hydrological modeling

Hydrological modeling now pays more attention to the simulation of the generalized water cycle variables than previous, especially ET (Immerzeel and Droogers, 2008). A hydrological model that takes the water cycle system of a closed drainage basin as its study object can continuously simulate ET for every hydrologic unit (or space grid), and the temporal and spatial resolutions are flexible (Xu and Cheng, 2010). Moreover, it is physically based and easily applicable. If we can find a suitable hydrological model as the dynamic model of assimilation, continuous regional ET estimates may be achieved by assimilating this hydrological model. Methods used to estimate ET differ in different hydrological models, so constructing an ET assimilation system based on hydrological modeling is premised on taking full account of the relationship between ET (diagnostic variable) and other state variables, and then analyzing the feasibility of assimilating ET.

There are two groups of ET estimation methods in hydrological models: gathering methods and converting methods (Zhao *et al.*, 2013). In both methods, ET is correlated with soil moisture, which is a state variable. Physically based hydrological models usually use gathering methods such as VIC and SHE to estimate ET; others commonly apply the converting methods, for which there are different ways to estimate potential ET and different soil moisture extraction functions. These functions are constructed based on soil moisture, and describe the relationship between ET and the state variables. A comparison of the feasibilities of the different hydrological models for ET assimilation is shown in Table 1.

ET, a non-state variable, is calculated for most of the hydrological models by using soil moisture. In addition to taking this state variable as an observation to influence the simulation of ET, data assimilation for hydrological modeling based on ET observations can be carried out in two ways. First, the observation operator $AW=H(ET)$ associating soil moisture and ET can be constructed, which converts the observed ET into a state variable, so as to realize the direct assimilation. This approach can be used for various types of hydrological models, but requires extensive calculations. Moreover, rank-deficient problem may appear because the correlation between soil moisture and ET is not clear when calculating the covariance matrix. Second, we can construct a unit matrix operator $ET = H(ET)$, assimilating the “observed” and simulated ET, and then calculate the state variable after updating ET. This method is easy to compute and rank-deficient problem does not appear. But it requires explicit relationships to be made between ET and state variables in the model, which is mainly applied to the model based on converting methods. Besides, the derivation of state variables is also worthy of attention. For example, in SWAT, soil moisture (which as mentioned above is a state variable) is the combined result of the entire duration of the simulation, so a sequential assimilation system of ET based on SWAT is difficult to construct. In other hydrological models based on converting methods, soil moisture evolves over time and

Table 1 The comparison of the feasibilities of different hydrological models for evapotranspiration assimilation

Hydrological model	Method used to calculate ET_a	Feasibility of assimilating ET_a	References
SWAT	Converting method including canopy interception, soil moisture evaporation and plant transpiration	$ET_a = ET_{water} + f(sd, wc) + f(ET_p, LAI)$. ET_a consists of water surface evaporation, soil evaporation and vegetation transpiration. With no time-recursive conditions, it is not suitable for constructing explicit assimilation relations.	Arnold <i>et al.</i> , 1998; Xu <i>et al.</i> , 2009
VIC	Gathering method including canopy interception, vegetation transpiration and bare soil evaporation	$ET_a = E_c + ET_{veg} + ET_{soil}$. ET_a is determined by different vegetation types, and calculated by leaf area index, vegetation impedance and proportion of vegetation roots in the soil. With no time-recursive conditions, it is not suitable for constructing explicit assimilation relations.	Liang <i>et al.</i> , 1994; Yu <i>et al.</i> , 2014
SHE	Gathering method using Penman-Monteith formula	$E_a = R_a \Delta + \frac{\rho c_p \delta_e}{r_a} / \lambda \left[\Delta + \gamma \left(1 + \frac{r_s}{r_a} \right) \right]$ ET_a is determined by the climatic conditions, canopy water content, soil moisture, etc. With no time-recursive conditions, it is not suitable for constructing explicit assimilation relations.	Xu <i>et al.</i> , 2009; Freeze <i>et al.</i> , 1969
VIP	Gathering method using a dual-source energy balance model for canopy and surface-based Penman-Monteith equation	$ET_a = E_c + E_s$. ET_a is calculated by the canopy transpiration and surface evaporation, determined by net radiation, canopy resistance, vapor pressure, etc. With no time-recursive conditions, it is not suitable for constructing explicit assimilation relations.	Wang <i>et al.</i> , 2008; Mo and Liu, 2001; Wang <i>et al.</i> , 2010
MIKE SHE	Gathering method using Penman-Monteith equation and Kristensen-Jensen model	$ET_a = E_c + E_{soil} + E_{water} + E_{veg}$. ET_a is calculated by interception, evaporation from soil and water surfaces, and plant transpiration. With no time-recursive conditions, it is not suitable for constructing explicit assimilation relations.	Xu <i>et al.</i> , 2009; Refshaard <i>et al.</i> , 1995; Vázquez, 2003; Huang <i>et al.</i> , 2010
TOPMODEL	Converting method using calculation based on the evaporation capacity E_p	$E_{a,i} = E_p \left(1 - \frac{S_{r,i}}{S_{max,i}} \right)$ ET_a is related to the water shortage and maximum storage capacity of vegetation root zone, and evaporation capacity. With no time-recursive conditions, it is not suitable for constructing explicit assimilation relations.	Xu <i>et al.</i> , 2009; Beven and Kirkby, 1979; Li <i>et al.</i> , 2006; Beven <i>et al.</i> , 1984
IHDM	Converting method using the EVAP routine	$E_a/E_p = f(\psi_w, \psi_s)$. ET_a is related to the capillary potential at the wilting point and anaerobiosis point, temperature, solar radiation, etc. With no time-recursive conditions, it is not suitable for constructing explicit assimilation relations.	Beven <i>et al.</i> , 1987; Feddes <i>et al.</i> , 1976; Feddes <i>et al.</i> , 1976
Xinanjia	Converting method using calculation by three layers of evapotranspiration	$EU = K \times EM, EL = (K \times EM - EU) \times WL/LM$ $ED = C \times (K \times EM - EU), ET_a = EU + EL + ED$ ET_a is primarily related to climate and underlying conditions, the upper and lower water storage capacities, coverage area of deep-rooted plants, etc. With no time-recursive conditions, it is not suitable for constructing explicit assimilation relations.	Hao <i>et al.</i> , 2010; Cao <i>et al.</i> , 2005
HIMS	Converting method using a conceptual model, and related to soil water storage and potential evaporation	$ET_a(t) = ET_p(t) \cdot \left[1 - \left(1 - \frac{W_s(t)}{W_{sm}} \right)^c \right]$ ET_a is calculated by soil water storage, solar radiation, latent heat of vaporization, and temperature. With no time-recursive conditions, it is not suitable for constructing explicit assimilation relations.	Liu <i>et al.</i> , 2006; Wu <i>et al.</i> , 2012
DTVGM	Converting method using improved Bagrov model that considers soil moisture	$AW_{t+1} = \left\{ P_t - KS_t - ET_{t+1} + \left[1 - \frac{K_s}{2} \right] AW_t \right\} / \left(1 + \frac{K_s}{2} \right)$ $ET_{t+1} = \begin{cases} ET_{p,t+1} \frac{AW_{t+1}}{WM} & (P_{t+1} + AW_{t+1} > W_{mm}) \\ (P_{t+1} + AW_{t+1}) \cdot KAW & (P_{t+1} + AW_{t+1} \leq W_{mm}) \end{cases}$ A relationship between ET_a and precipitation, soil moisture and potential evapotranspiration is established. The explicit assimilation relation can be constructed through soil moisture.	Xia <i>et al.</i> , 2003

can provide feedback of ET assimilation results, so we can first assimilate evapotranspiration, and then adjust soil moisture to calibrate the hydrological model. But it is hard to construct the sequential assimilation system without explicit relationships between ET and state variables. In addition, it is most convenient to directly assimilate simulated ET with observed ET under the conditions of few observations, since the assimilation efficiency is high when using a unit matrix.

To apply the assimilation procedures described above into hydrological models and determine ET levels with high precision, we should consider the requirements of the assimilation algorithm itself for hydrological models: first, the simulation variables must be state variables that can describe the time-domain behavior of a dynamic system; second, the simulation variable must evolve over time. In addition, ET is a prognostic variable to construct a data assimilation system, so we need to construct the time response relationship between ET and state variables by assimilating feedback and then optimizing the simulation, so as to obtain accurate estimates of ET. Thus, the key to construct an ET assimilation system is to select the hydrological model that meets the requirements of the assimilation algorithm, which means developing the time response relationship between ET and the state variable being used. By effectively transforming ET into a state variable, the hydrological model was modified by the assimilation conditions, and could be combined with sequential assimilation to establish a feasible ET data assimilation system.

4 A new evapotranspiration assimilation scheme based on hydrological modeling

To construct the data assimilation system based on evapotranspiration (ET), which is a non-state variable, we need to choose the hydrological model with a time response relationship between ET and the state variable, as the model operator of the assimilation system (Table 1). Based on the nonlinear theory of rain-derived flooding, and considering the preceding precipitation, Xia *et al.* put forward the distributed time-variant gain model (DTVGM) for watershed hydrological simulation (Xia *et al.*, 1997; Xia, 2002; Xia *et al.*, 2004). The model established the time response relationship between ET and soil moisture, and could be used to construct the data assimilation system based on ET. This section of the paper will use the DTVGM as an example to elaborate the new ET assimilation scheme.

4.1 Distributed time-variant gain model (DTVGM)

Xia *et al.* (2003) proposed the distributed time variant gain model (DTVGM) with a combination of a hydrological nonlinear system approach and distributed hydrological simulation technology, constructing the model structure from the perspective of feasibility and practicability. The DTVGM can simulate the nonlinear relationship between rainfall, runoff and other hydrological variables under the influence of human activities (Li *et al.*, 2010), with the ability, using a small amount of computation, to provide a real-time response of underlying surfaces, especially in arid and semi-arid basins such as the Heihe and Yellow rivers (Xia *et al.*, 2005). In this model, the soil moisture at a given point in time is calculated by the precipitation, surface runoff, soil moisture and actual ET at a previous point in time using the water balance equation, combined with a simple nonlinear relationship between ac-

tual ET and soil moisture, to effectively transform ET into a state variable. That is, the time response relationship of ET and soil moisture is constructed (Eqs. 1 and 2), and this construction is suitable for ET data assimilation.

$$AW_{t+1} = \left\{ P_t - RS_t - ETa_t + \left[1 - \frac{K_r}{2} \right] \cdot AW_t \right\} / \left(1 + \frac{K_r}{2} \right) \quad (1)$$

$$ETa_{t+1} = \begin{cases} ETp_{t+1} \frac{AW_{t+1}}{WM} & (P_{t+1} + AW_{t+1} > W_{\min}) \\ (P_{t+1} + AW_{t+1}) \cdot KAW & (P_{t+1} + AW_{t+1} \leq W_{\min}) \end{cases} \quad (2)$$

where ETa_t is the actual evapotranspiration at time t ; ETa_{t+1} is the actual evapotranspiration at time $t+1$; K_r is soil water discharge coefficient; AW_t is the soil moisture at time t ; AW_{t+1} is the soil moisture at time $t+1$; W_{\min} is the minimum soil moisture; ETp_{t+1} is the potential evapotranspiration at time $t+1$; WM is the saturated soil moisture; P_t is the precipitation at time t ; RS_t is the surface runoff at time t ; and KAW is the coefficient of evapotranspiration conversion.

4.2 Improvement of DTVGM evapotranspiration estimation

The integrated converting method was used to calculate the actual ET in the DTVGM, and Zhao Lingling (2013) improved the ET formula by constructing a nonlinear soil water availability function based on the logistic distribution. A good estimation of actual ET was obtained based on the logistic function, not only in the winter and spring, when there is an adequate water supply, but also in summer, when the water content is low. The results indicate that the conversion model based on the logistic function has a wide range of applications, and can simulate the actual ET under various water supply conditions. The soil water availability function based on the logistic function is used to improve the ET formula in the DTVGM; then the improved time response relationship between ET and soil moisture can be described as:

$$\begin{cases} AW_{t+1} = \left\{ P_t - RS_t - ETa_t + \left[1 - \frac{K_r}{2} \right] \cdot AW_t \right\} / \left(1 + \frac{K_r}{2} \right) \\ ETa_{t+1} = \frac{1}{1 + e^{(-0.5(AW_{t+1} - 0.5))}} ETp_{t+1} \end{cases} \quad (3)$$

where the meanings of the parameters are as described above. Eq. (3) completes the transition from ET to the state variable, i.e., soil moisture, yielding a forecast of soil moisture and updates of actual ET.

In order to improve the DTVGM, the time response relationship between ET and soil moisture will be further improved by considering the impact of vegetation (Zhao *et al.*, 2014). Andersen believes that actual ET is related to potential ET, soil moisture, leaf area index and root depth (Andersen *et al.*, 2002). Therefore, we use the leaf area index (or root depth) to improve the relationship between actual and potential ET in the DTVGM, thereby affecting the time response relationship of ET and soil moisture (Eq. (4)). Improving this relationship perfects the empirical ET formula and improves simulation accuracy, with better feedback and updates of the assimilation performance.

$$\begin{cases} AW_{t+1} = \left\{ P_t - RS_t - ETa_t + \left[1 - \frac{K_r}{2} \right] \cdot AW_t \right\} / \left(1 + \frac{K_r}{2} \right) \\ ETa_{t+1} = \frac{1}{1 + e^{(-0.5(AW_{t+1} - 0.5))}} \cdot f(\text{LAI}) \cdot f(\text{RD}) \cdot ETp_{t+1} \end{cases} \quad (4)$$

where $f(\text{LAI})$ is the leaf area index function, and $f(\text{RD})$ is the root depth function.

4.3 A new evapotranspiration assimilation scheme based on DTVGM

The new evapotranspiration assimilation scheme based on the DTVGM consists of three components: a dynamic model (model operator), a set of observations (observation operator) and assimilation algorithm. The improved DTVGM model can be used as a model operator, and the daily values of ET derived from remote sensing can be used as observed data. The implicit relationship between observed data and the state variable is contained in the assimilation process, and the ensemble Kalman filter algorithm can be chosen as the assimilation algorithm.

Using the ensemble Kalman filter (EnKF) algorithm, the ET data assimilation scheme takes ET values from remote sensing as observations, to correct the state variables and improve the performance of the DTVGM, which is the driving model. The specific processes of this assimilation scheme are as follows: 1) addition of perturbation model errors to generate the initial background field, and drive model simulation combined with the basic input data; 2) addition of perturbation errors to, when available, the observed data of the day, i.e., ET values derived from daily remote sensing, and then assembly of these data to obtain the observation field; 3) at the same time, assembly of the ET values from the model running in parallel to obtain the forecast field, and then assimilation of the forecast field and observation field by the ensemble Kalman filter algorithm to obtain the analysis field; 4) feedback provided by the assimilation effect to the model operator to replace and update the background field. Then the simulation of next day is driven, and the assimilation cycle begins again when there are observations in the next day (Figure 1).

The ensemble Kalman filter (EnKF) can be used for the assimilation algorithm in this scheme. The main calculation steps are as follows:

(1) Forecast. The background field is initialized with N Gaussian random variables X_i ($i=1, \dots, N$), namely watershed actual ET, and the forecast $X_{i,k+1}^f$ is calculated for each random variable at time $k+1$ as follows:

$$X_{i,k+1}^f = M_{k,k+1} \left(X_{i,k}^a \right) + w_{i,k} \quad w_{i,k} \sim N(0, Q_k) \quad (5)$$

where $X_{i,k}^a$ is the analysis of the i th ensemble member at time k , namely the ET_a of the DTVGM simulation at time k ; $X_{i,k+1}^f$ is the forecast of the i th ensemble member at time $k+1$, which uses the transition from ET to the soil moisture to obtain the ET_a of the DTVGM simulation at time $k+1$; $M_{k,k+1}$ is the change of the state (which is generally a nonlinear model operator) from time k to $k+1$, and this change is the DTVGM in this scheme is the change of the relationship of state from time k to $k+1$, generally nonlinear model operator, and it is the DTVGM in this scheme; $w_{i,k}$ is the model error, drawn from normal distribution

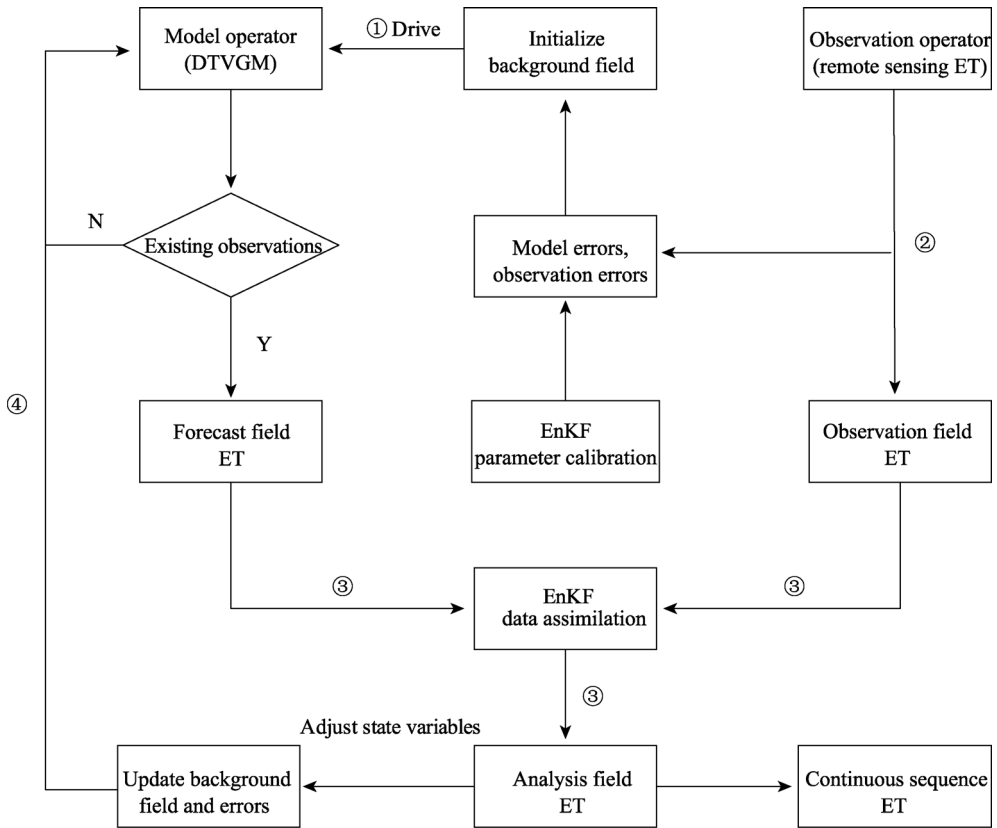


Figure 1 The flow chart of ET assimilation based on distributed time-variant gain model

with zero mean and covariance matrix R .

(2) Update. The obtained state forecast is updated by the observations at time $k+1$ when available, and the updated state analysis and its error covariance are obtained. The process used to update ET can be expressed by Eq. (6).

$$X_{i,k+1}^a = X_{i,k+1}^f + K_{k+1} \left[(y_{k+1} + v_{i,k+1}) - H(x_{i,k+1}^f) \right] v_{i,k+1} \sim N(0, R_k) \quad (6)$$

where $X_{i,k+1}^f$ is the forecast of the i th ensemble member at time $k+1$, namely the ET_a of the DTVGM simulation at time $k+1$; $X_{i,k+1}^a$ is the analysis of the i th ensemble member at time $k+1$, that is, the assimilated ET after the update that combines the model forecast and observations; K_{k+1} is the Kalman gain matrix at time $k+1$ (Evensen, 2003), which weights the relative uncertainty of the simulated estimation and observation; y_{k+1} is the observations at time $k+1$, namely the ET measured by remote sensing; $H(\cdot)$ is the observation operator, being the unit matrix in the scheme; $v_{i,k+1}$ is white Gaussian noise with zero mean and specified covariance R_k ; and R_k is the observation error covariance matrix.

(3) Update background field. Initialize the model using the state estimation at time $k+1$. That is, update the background field by calculating the soil moisture according to assimilated ET, with the assimilating feedback to the model operator. When observations are available, the assimilation at the next point in time is executed, and the above steps are repeated until the forecast and update of the entire ET process is completed.

In summary, a new easy-to-operate evapotranspiration assimilation scheme with a water circulation physical mechanism has been proposed in this paper. Using ensemble Kalman filter, a data assimilation system that combined a remote sensing ET model and a hydrological model was constructed in the scheme depending on the time response relationship between ET and soil moisture in the DTVGM. This scheme can be expected to yield a continuous time series of regional ET estimates with high accuracy.

5 Discussion and prospect

In this paper, a new scheme to estimate evapotranspiration on a regional scale has been proposed. However, since the applications of this scheme are still at the experimental stage (Yin *et al.*, 2014), more tests of these applications should be carried out at the regional level to validate the scheme. Research in ET data assimilation is still in the exploratory stage, and there are many bottlenecks that need to be overcome in this work, in contrast to assimilation research of other variables (e.g., land surface temperature, soil temperature, etc.). The biggest challenge in the assimilation of regional ET, a non-state variable in hydrological models, is how to improve the ET mechanism in the models as well as establish a rational isomorphic relationship of ET assimilation. Moreover, it is necessary to carry out further research on several aspects of this work, including on the accuracy of observational models, the stability of model operators, the capacity of data assimilation, the validation of regional ET processes, the optimization of the data assimilation method, the mechanism of the ET process and so on.

Accuracy of observational models. At present, most observational models adopt remote sensing (RS) ET models, which involve a number of parameters that describe the physical characteristics of the land surface. Due to external factors such as clouds, the atmosphere, the solar angle and observation angle, the accuracy of remote sensing data is limited to some extent. Besides, the cumulative effect of errors from retrieval of land surface parameters also affects the accuracy of observational models (Zhang *et al.*, 2012). Therefore, it would be important to further explore the relationships between land surface physical characteristics and remote sensing information, so that we can reveal the basic laws underlying these relationships, and then realize cooperative retrieval of input data for the ET model using multi-source remote sensing information. This would improve the accuracy and spatio-temporal continuity of land surface parameters, so as to improve the accuracy of ET estimation at the regional scale.

Stability of model operators. The stability of model operators in data assimilation system, usually hydrological models, is of vital importance for obtaining accurate ET simulation results. Hydrological models synthetically consider the interactions between precipitation-runoff and landform, physiognomy, climate, and human activities. Since these models provide reliable mechanistic explanations for the hydrological cycle and use many parameters, uncertainties in these models have a great effect on ET simulation. Due to the high spatio-temporal heterogeneity of land surface variables, it is difficult to avoid large deviations between the hydrological model simulation and the observed data. Development of more mature, simple and effective model operators is necessary to achieve a reliable assimilation system. Therefore, further research is needed on the optimization and improvement of the hydrological model, as well as research on assessing the universality of the model. The convergence, rate of convergence and stability of the assimilation algorithm should also be

comprehensively assessed.

Capacity of data assimilation. The capacity of data assimilation is closely related to the input frequency of “observations,” which means that an increase in the number of observed samples leads to better assimilation of ET, which promotes the influence of hydrological model calibration. When high-quality remote sensing data are lacking, it can be helpful to consider other sources of data, such as obtaining an observation field via spatial interpolation based on highly accurate flux data from regionally representative observational stations, moving/ fixed-point observations by radar remote sensing techniques, or calculation of energy flux by eliminating cloud interference for remote sensing data with large cloud coverage. These approaches will be the focus of future research.

Validation of the regional ET process. Direct comparison between assimilated ET and observed ET is the most efficient validation approach. Of these approaches, large-aperture scintillometry (LAS) is the most effective at a regional level, but a large-scale use of this method is often impeded by the high price of the equipment. The validation of actual ET still depends on the simulation data from a third party; thus, research focusing on finding an effective means of validating ET at the regional level is being carried out.

Development of data assimilation techniques. To further improve estimation of ET using hydrological models, it can be helpful in future research to use data fusion methods, such as the wavelet transform, to extract contact information between the grids of observational products and fuse them into the assimilation. Meanwhile, to enhance the accuracy of the estimation of hydrological variables and at the same time ensure high calculation efficiency, a more appropriate assimilation mechanism should be established. Research should be focused on the following: development of techniques to combine multiple and complementary remote sensing sources for hydrologic modeling (Xu *et al.*, 2014); joint assimilation of multi-scale, multi-sensor products; and development of new techniques such as integration of new potential sensor products with hydrological models.

References

- Andersen J, Dybkjaer G, Jensen K H *et al.*, 2002. Use of remotely sensed precipitation and leaf area index in a distributed hydrological model. *Journal of Hydrology*, 264(1): 34–50.
- Arnold J G, Williams J R, Srinivasan R *et al.*, 1998. Large area hydrologic modeling and assessment part I: Model development. *Journal of the American Water Resources Association*, 34(1): 73–89.
- Beven K, Calver A, Morris E M, 1987. The Institute of Hydrology Distributed Model, 1–33. Wallingford: Institute of Hydrology, 1-33. (IH Report No.98)
- Beven K J, Kirkby M J, 1979. A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Journal*, 24(1): 43–69.
- Beven K J, Kirkby M J, Schofield Net *et al.*, 1984. Testing a physically based flood-forecasting model (TOPMODEL) for three UK catchments. *Journal of Hydrology*, 69(1–4): 119–143.
- Cao Lijuan, Liu Jingmiao, Ren Liliang, 2005. Improving on the evapotranspiration calculation of Xinanjinag model. *Hydrology*, 25(3): 5–9, 19. (in Chinese)
- Chen H, Yang D, Hong Y *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.
- Conradt T, Wechsung F, Bronstert A, 2013. Three perceptions of the evapotranspiration landscape: Comparing spatial patterns from a distributed hydrological model, remotely sensed surface temperatures, and sub-basin water balances. *Hydrology and Earth System Sciences Discussions*, 10(1): 1127–1183.

- Dumedah G, Coulibaly P, 2013. Evolutionary assimilation of streamflow in distributed hydrologic modeling using in-situ soil moisture data. *Advances in Water Resources*, 53: 231–241.
- Evensen G, 2003. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4): 343–367.
- Feddes R A, Kowalik P, Kolinska-Malinka K *et al.*, 1976. Simulation of field water uptake by plants using a soil water dependent root extraction function. *Journal of Hydrology*, 31(1): 13–26.
- Feddes R A, Kowalik P, Neuman S P *et al.*, 1976. Finite difference and finite element simulation of field water uptake by plants. *Hydrological Sciences Journal*, 21(1): 81–98.
- Freeze R A, Harlan R L, 1969. Blueprint for a physically-based, digitally-simulated hydrologic response model. *Journal of Hydrology*, 9(3): 237–258.
- Hao Zhenchun, Li Li, Wang Jiahu *et al.*, 2010. Theory and Method of Distributed Hydrological Model. Beijing: Science Press. (in Chinese)
- Huang Yue, Chen Xi, Bao Anming *et al.*, 2010. Distributed hydrological modeling in Kaidu Basin: MIKE-SHE model calibration and uncertainty estimation. *Journal of Glaciology and Geocryology*, 32(3): 567–572. (in Chinese)
- Immerzeel W, Droogers P, 2008. Calibration of a distributed hydrological model based on satellite evapotranspiration. *Journal of Hydrology*, 349(3/4): 411–424.
- Irmak A, Kamble B, 2009. Evapotranspiration data assimilation with genetic algorithms and SWAP model for on-demand irrigation. *Irrigation Science*, 28(1): 101–112.
- Kustas W P, Norman J M, 1996. Use of remote sensing for evapotranspiration monitoring over land surfaces. *Hydrological Sciences Journal*, 41(4): 495–516.
- Lei F, Huang C, Shen H *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.
- Li L, Xia J, Xu C *et al.*, 2010. Evaluation of the subjective factors of the GLUE method and comparison with the formal Bayesian method in uncertainty assessment of hydrological models. *Journal of Hydrology*, 390(3): 210–221.
- Li Xin, Huang Chunlin, Che Tao *et al.*, 2007. Progress and prospect of research on land data assimilation system in China. *Progress in Natural Science*, 17(2): 163–173. (in Chinese)
- Li Z, Tang R, Wan Z *et al.*, 2009. A review of current methodologies for regional evapotranspiration estimation from remotely sensed data. *Sensors*, 9(5): 3801–3853.
- Li Zhijia, Zhang Ke, Yao Cheng, 2006. Comparison of distributed geological models based on GIS technology and DEM. *Journal of Hydraulic Engineering*, 37(8): 1022–1028. (in Chinese)
- Liang Shunlin, Li Xin, Xie Xianhong *et al.*, 2013. Land Surface Observations, Modeling and Data Assimilation. Beijing: Higher Education Press, 97–114. (in Chinese)
- Liang X, Lettenmaier D P, Wood E F *et al.*, 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, 99(D7): 14415–14428.
- Liu Sanchao, Zhang Wanchang, Gao Maofang *et al.*, 2007. Simulation of land surface evapotranspiration using distributed hydrological model, remote sensing and GIS technology. *Scientia Geographica Sinica*, 27(3): 354–358. (in Chinese)
- Mo X, Liu S, 2001. Simulating evapotranspiration and photosynthesis of winter wheat over the growing season. *Agricultural and Forest Meteorology*, 109(3): 203–222.
- Moradkhani H, 2008. Hydrologic remote sensing and land surface data assimilation. *Sensors*, 8(5): 2986–3004.
- Pan M, Wood E F, Wójcik R *et al.*, 2008. Estimation of regional terrestrial water cycle using multi-sensor remote sensing observations and data assimilation. *Remote Sensing of Environment*, 112(4): 1282–1294.
- Pipunic C, Walker P, Western A, 2008. Assimilation of remotely sensed data for improved latent and sensible heat flux prediction: A comparative synthetic study. *Remote Sensing of Environment*, 112(4): 1295–1305.
- Priestley C H B, Taylor R J, 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, 100(2): 81–92.
- Qin C, Jia Y, Su Z *et al.*, 2008. Integrating remote sensing information into a distributed hydrological model for improving water budget predictions in large-scale basins through data assimilation. *Sensors*, 8(7): 4441–4465.
- Refshaard J C, Storm B, Singh V P, 1995. MIKE SHE. In: Computer Models of Watershed Hydrology. Water Resources Publications, 809–846.

- Renard B, Kavetski D, Kuczera G *et al.*, 2010. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, 46(5): W05521.
- Rosenberg N J, 1983. *Microclimate: The Biological Environment*. New York: John Wiley & Sons.
- Schuurmans M, Troch A, Veldhuizen A *et al.*, 2003. Assimilation of remotely sensed latent heat flux in a distributed hydrological model. *Advances in Water Resources*, 26(2): 151–159.
- Song Xiaomeng, Zhan Chesheng, Kong Fanzhe *et al.*, 2011. A review on uncertainty analysis of large-scale hydrological cycle modeling system. *Acta Geographica Sinica*, 66(3): 396–406. (in Chinese)
- Spies R R, Franz K J, Hogue T S *et al.*, 2014. Distributed hydrologic modeling using satellite-derived potential evapotranspiration. *Journal of Hydrometeorology*, 16(1): 129–146.
- Tang H, Li Z L, 2014. Estimation and validation of evapotranspiration from thermal infrared remote sensing data. In: *Quantitative Remote Sensing in Thermal Infrared*. Berlin and Heidelberg: Springer, 145–201.
- Trudel M, Leconte R, Paniconi C, 2014. Analysis of the hydrological response of a distributed physically-based model using post-assimilation (EnKF) diagnostics of streamflow and in situ soil moisture observations. *Journal of Hydrology*, 514: 192–201.
- Vázquez R F, 2003. Effect of potential evapotranspiration estimates on effective parameters and performance of the MIKE SHE-code applied to a medium-size catchment. *Journal of Hydrology*, 270(3): 309–327.
- Vinukollu R K, Wood E F, Ferguson C R *et al.*, 2011. Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: Evaluation of three process-based approaches. *Remote Sensing of Environment*, 115(3): 801–823.
- Wang Kun, Mo Xingguo, Lin Zhonghui *et al.*, 2010. Improvement and validation of vegetation interface process model. *Chinese Journal of Ecology*, 29(2): 387–394. (in Chinese)
- Wang Yongfen, Mo Xingguo, Hao Yanbin *et al.*, 2008. Simulation seasonal and interannual variations of ecosystem evapotranspiration and its components in Inner Mongolia steppe with VIP model. *Journal of Plant Ecology*, 32(5): 1052–1060. (in Chinese)
- Wu Mengying, Wang Zhonggen, Dang Suzhen, 2012. Simulation and analysis of runoff in the upper reaches of the Heihe River basin. *Resources Science*, 34(10): 1913–1921. (in Chinese)
- Xia J, O'Connor K M, Kachroo R K *et al.*, 1997. A non-linear perturbation model considering catchment wetness and its application in river flow forecasting. *Journal of Hydrology*, 200(1): 164–178.
- Xia Jun, 2002. *Theory and Method of Hydrologic Nonlinearity*. Wuhan: Wuhan University Press. (in Chinese)
- Xia Jun, Wang Gangsheng, Lv Aifeng *et al.*, 2003. A research on distributed time variant gain modeling. *Acta Geographica Sinica*, 58(5): 789–796. (in Chinese)
- Xia Jun, Wang Gangsheng, Tan Ge *et al.*, 2004. Hydrology nonlinear systems and distributed time-variant gain model. *Science in China: Series D*, 34(11): 1062–1071. (in Chinese)
- Xia Jun, Ye Aizhong, Wang Gangsheng, 2005. A distributed time-variant gain model applied to Yellow River (I): Model theories and structures. *Engineering Journal of Wuhan University*, 38(6): 10–15. (in Chinese)
- Xie X, Zhang D, 2010. Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. *Advances in Water Resources*, 33(6): 678–690.
- Xu X, Li J, Tolson B A, 2014. Progress in integrating remote sensing data and hydrologic modeling. *Progress in Physical Geography*. doi: 10.1177/0309133314536583.
- Xu Zongxue *et al.*, 2008. *Hydrological Model*. Beijing: Science Press. (in Chinese)
- Xu Zongxue, Cheng Lei, 2010. Progress on studies and applications of the distributed hydrological models. *Journal of Hydraulic Engineering*, 1(3): 5–6. (in Chinese)
- Yin Jian, Zhan Chesheng, Gu Hongliang *et al.*, 2014. A case study of evapotranspiration data assimilation based on hydrological model. *Advances in Earth Science*, 29(9): 1075–1084. (in Chinese)
- Yu Z, Deng J, Liu C, 2014. Application of VIC model to hydrological response caused by urbanization in Dongjiang Basin. *Journal of Water Resources Research*, 3(1): 78–83. (in Chinese)
- Zhang Ronghua, Du Junping, Sun Rui, 2012. Review of estimation and validation of regional evapotranspiration based on remote sensing. *Advances in Earth Science*, 27(12): 1295–1307. (in Chinese)
- Zhao Lingling, 2013. *The evapotranspiration estimation methods study in hydrological cycle simulation [D]*. Beijing: University of Chinese Academy of Sciences. (in Chinese)
- Zhao Lingling, Xia Jun, Xu Chongyu *et al.*, 2013. A review of evapotranspiration estimation methods in hydrological models. *Acta Geographica Sinica*, 68(1): 127–136. (in Chinese)
- Zhao L W, Zhao W Z, 2014. Evapotranspiration of an oasis-desert transition zone in the middle stream of Heihe River, Northwest China. *Journal of Arid Land*, 6(5): 529–539.