

Assimilation of ASAR Data with a Hydrologic and Semi-empirical Backscattering Coupled Model to Estimate Soil Moisture

LIU Qian, WANG Mingyu, ZHAO Yingshi

(College of Resource and Environment, Graduate University of the Chinese Academy of Sciences, Beijing 100049, China)

Abstract: The most promising approach for studying soil moisture is the assimilation of observation data and computational modeling. However, there is much uncertainty in the assimilation process, which affects the assimilation results. This research developed a one-dimensional soil moisture assimilation scheme based on the Ensemble Kalman Filter (EnKF) and Genetic Algorithm (GA). A two-dimensional hydrologic model—Distributed Hydrology-Soil-Vegetation Model (DHSVM) was coupled with a semi-empirical backscattering model (Oh). The Advanced Synthetic Aperture Radar (ASAR) data were assimilated with this coupled model and the field observation data were used to validate this scheme in the soil moisture assimilation experiment. In order to improve the assimilation results, a cost function was set up based on the distance between the simulated backscattering coefficient from the coupled model and the observed backscattering coefficient from ASAR. The EnKF and GA were used to re-initialize and re-parameterize the simulation process, respectively. The assimilation results were compared with the free-run simulations from hydrologic model and the field observation data. The results obtained indicate that this assimilation scheme is practical and it can improve the accuracy of soil moisture estimation significantly.

Keywords: Advanced Synthetic Aperture Radar (ASAR); Distributed Hydrology-Soil-Vegetation Model (DHSVM); Oh Model; couple; soil moisture; data assimilation

1 Introduction

Soil moisture is a variable that plays a leading role in surface water and energy balance. It can be estimated from point measurements, hydrologic models and remote sensing. The traditional point measurements can just provide a soil moisture distribution in a limited area. The hydrologic models can simulate the soil moisture variation in large spatial and temporal scales. However, the uncertainty in initial model state, model physical parameters, meteorological forcing variables and model equations brings much error in the soil moisture estimation. And the model errors increase gradually when the model goes forward. Remote sensing technique can be used to collect spatial data over large areas on a routine basis, which provides a capability to make frequent and spatially comprehensive measurements of the near-surface soil moisture. However, the depth of the soil moisture measurement is limited typically to the top few cen-

timeters, and the measurement can not be conducted continuously by remote sensing technique (Walker *et al.*, 2001). Many researchers suggested that the most promising approach for estimating soil moisture is the integration of remote sensing surface soil moisture data and a hydrologic model (Kostove and Jackson, 1993).

There is much uncertainty in the model initial state and model parameters in the hydrologic modeling, which brings much error to the soil moisture simulation. Many researches have been carried out on optimizing a computational model through minimizing the difference between the observation and model simulation results by the re-parameterization (Jarlan *et al.*, 2005; Yang *et al.*, 2009) or re-initialization (Heathman *et al.*, 2003; Huang *et al.*, 2008) of the computational model separately.

The re-parameterization approach can improve the model state variables by adjusting uncertain parameters. Since some parameters can hardly be obtained by field measurements, they are often estimated from the exper-

Received date: 2009-09-09; accepted date: 2010-01-21

Foundation item: Under the auspices of Major State Basic Research Development Program of China (973 Program) (No. 2007CB714400), the Program of One Hundred Talents of the Chinese Academy of Sciences (No. 99T3005WA2)

Corresponding author: WANG Mingyu. E-mail: mawang@gucas.ac.cn

© Science Press and Northeast Institute of Geography and Agroecology, CAS and Springer-Verlag Berlin Heidelberg 2010

iences or references, which brings much uncertainty to the model parameters. There are many parameter optimization methods such as Genetic Algorithm, Tabu Search, Ant Colony Optimization, Pattern Search and Scatter Search. Onwubolu and Kumalo (2001) concluded that the Genetic Algorithm is most suitable for the optimization of machining parameters.

The re-initialization approach does not attribute the simulation errors to the uncertainties of the initial conditions, model parameters or model physical equations. It can adjust the state variables directly and it adopts the updated state variables as the initial field to re-initialize the model. The initial field contains the "optimal values" of the state variables. The most effective and popular re-initialization approach is Ensemble Kalman Filter (EnKF).

The re-initialization and re-parameterization are the important aspects in the soil moisture assimilation, while many researches just emphasize on one aspect particularly. The variation data assimilation methods can estimate both state variables and model parameters. However, it is difficult to develop the adjoint model for the original one to evaluate the derivatives of the cost function with respect to control variables, which are composed of state variables and model parameters to be estimated. It is complicated to develop the adjoint model in practice, which posed many limitations to its application (Qin *et al.*, 2007).

In this research, the assimilation scheme combined the EnKF and GA as the assimilation algorithm and it could calibrate the model parameters and model initial state synthetically. A two-dimensional hydrologic model—Distributed Hydrology-Soil-Vegetation Model (DHSVM) was coupled with a semi-empirical backscattering model—Oh Model. The Advanced Synthetic Aperture Radar (ASAR) observation was assimilated with simulated backscattering coefficient from the coupled model. The assimilation experiment was carried out in the Heihe River Basin, Northwest China.

2 Study Area and Data

2.1 Study area and ground observation

The Heihe River Basin, located in the arid area of Northwest China, is the second largest inland river basin, with an area of 116 000 km² and an average annual precipitation of 108 mm. The middle reaches of the Heihe

River (38°45'–39°15'N, 100°00'–100°45'E) (Fig. 1) was selected as the study area. The elevation of the study area varies between 1 000 m and 2 000 m above sea level, and the average annual temperature is 6 °C. Several automatic meteorological observation stations have been established for the long-period observations (Fig. 1). The stations recorded six meteorological variables required in the DHSVM: air temperature, wind speed, relative humidity, incoming shortwave radiation, incoming longwave radiation and precipitation. And the time domain reflectometry (TDR) were used in each station to measure the soil moisture at depths of 10 cm, 20 cm, 40 cm, 80 cm, 120 cm and 160 cm. The assimilation approach proposed was examined at a typical site, Huazhaizi (38°40'N, 100°20'E), with bare soil sparsely covered by some short shrubs. The ground observation data were obtained from the: Watershed Airborne Telemetry Experiment Research of Chinese Academic of Science Action Plan for West Development Program. The vegetation type data and the soil type data were obtained from the Data Center for Resources and Environment Sciences, Chinese Academy of Sciences. Some of the model parameters were achieved from the field measurements while some of them were referenced from the literature Kang *et al.* (2008).

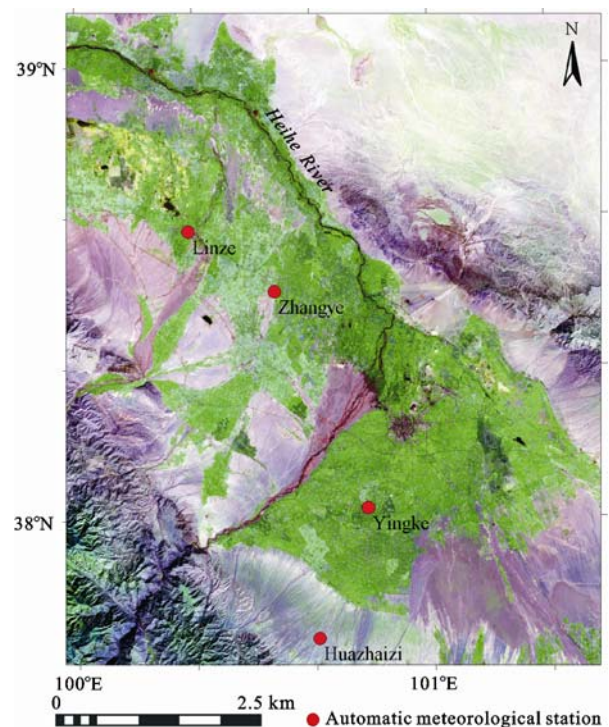


Fig. 1 Automatic meteorological stations in upper reaches of Heihe River

2.2 ASAR data

In this research, the ASAR images were used as observation data. ASAR images were acquired for several dates and different polarizations with incidence angles between 10° and 60° , operating in C band. All images were georeferenced by using a georeferenced SPOT image. The ASAR data were transferred into the backscattering coefficients by the radiometric calibrations. The spatial resolution of ASAR is about $12.5 \text{ m} \times 12.5 \text{ m}$. The spatial resolution of the DHSVM simulation was 30 m, so the spatial resolution of ASAR was resampled into 30 m by the bilinear interpolation method. The ASAR data used in this experiment were given in Table 1.

3 Methods

In this assimilation scheme, the soil moisture was taken as the combination point of the hydrologic model and the backscattering model. The DHSVM was coupled with a semi-empirical backscattering model (Oh) to simulate backscattering coefficients. When the ASAR data were available, a cost function was set up based on the gap between the backscattering coefficients from ASAR observation and the backscattering coefficients

from the simulations. If the cost function was lower than twice the ASAR observation noise, the simulation was considered as acceptable. Then the EnKF was used to optimize the state variables. If the cost function was higher than twice the ASAR observation noises, the simulation was treated as unacceptable. Then the cost function was minimized by the GA approach. The sensitive parameters of the DHSVM were re-parameterized. The simulations were divided into acceptable and unacceptable cases. The EnKF and GA were used to re-initialize and re-parameterize the simulation process. The schematic description of the assimilation procedure is given in Fig. 2.

3.1 Models

3.1.1 Distributed Hydrology-Soil-Vegetation Model

DHSVM can explicitly represent the effects of the topography and vegetation on water fluxes. The model includes canopy interception, evaporation, transpiration, snow accumulation and melt, and runoff generation. The intensive description of DHSVM used in this study can be found in Wigmosta *et al.* (1994; 2002). DHSVM as an open source hydrological model provides a dynamics representation of watershed processes at the various sp-

Table 1 Overview of acquired ASAR data in Huazhaizi site

Acquisition data	Julian day	Orbit	Polarization mode	Spatial resolution (m)	Incidence angle ($^\circ$)
2008-06-19	171	Descending	VV/VH	12.5	38.9–42.7
2008-06-25	177	Descending	HV/HH	12.5	25.7–31.2
2008-06-28	180	Descending	HV/HH	12.5	14.2–22.3
2008-07-05	187	Descending	VV/VH	12.5	42.4–45.2
2008-07-08	190	Descending	VV/VH	12.5	38.9–42.7

Note: V means vertical polarization mode and H means horizontal polarization mode

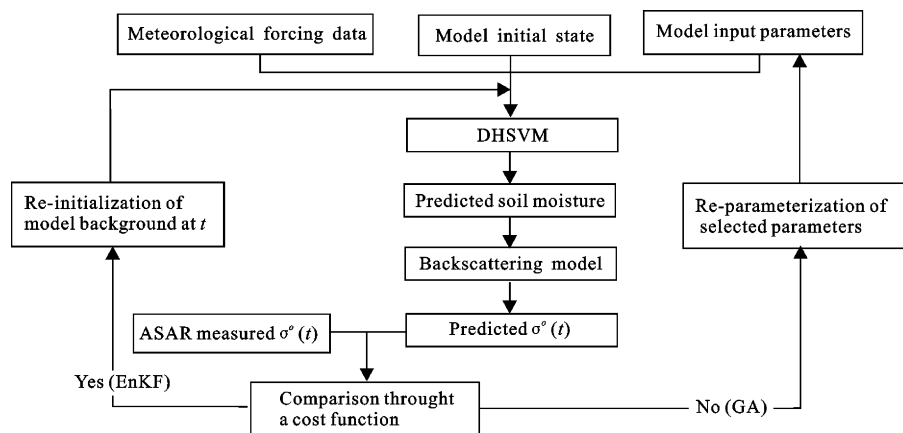


Fig. 2 Schematic description of assimilation scheme based on EnKF and GA

atial scales described by the DEM data (<http://www.Hydro.washington.edu/Lettenmaier/Models/DHSVM/index.shtml>).

In each grid cell the modeled land surface may be composed of the overstory vegetation, understory vegetation and soil. The soil is divided into three layers: the surface soil zone with depth equal to the understory vegetation root, the root zone with the extent between the overstory vegetation root and the understory vegetation root, and the saturated zone. The overstory is allowed to remove water from both the upper and lower soil zones, while the understory can only remove water from the upper zone.

Mass balance equations for the surface soil layer and the root zone are summarized as follows:

$$d_1(\theta_k^{t+\Delta t} - \theta_k^t) = I_f - Q_V(\theta_k) - \sum_{j=1}^2 f_{rj1} E_{tj} - E_s + V_{ex2} - V_{ex1} \quad (1)$$

$$d_k(\theta_k^{t+\Delta t} - \theta_k^t) = Q_V(\theta_{k-1}) - Q_V(\theta_k) - \sum_{j=1}^2 f_{rjk} E_{tj} + V_{exk+1} \quad (2)$$

where d_k is the soil layer thickness in soil layer k ; θ_k^t is the average soil moisture in soil layer k at time t ; I_f is the volume of water infiltrated during the time step Δt ; Q_V is the volume of water discharged downward to the next layer; f_{rjk} is the fraction of roots from vegetation layer j in soil layer k ; V_{exk} is the volume of water supplied by a rising water table in soil layer k ; E_{tj} is the volume of evaporation during the time step Δt from vegetation layer j ; E_s is the volume of evaporated soil moisture from the surface layer. The model first calculates infiltration into the upper layer, then downward vertical moisture transfer moving from top to bottom.

DHSVM can be implemented in a wide range of spatial and temporal scales. In this study, the time step for the soil moisture simulation is 3 hours, while the spatial scale is 30 m \times 30 m, which is consistent with the resolution of resampled ASAR and DEM data.

The model inputs include the configuration file, DEM file, soil type file, soil depth file, vegetation type file, meteorological files, stream/road class files, stream/road network files, and stream/road map files. The DEM cell size of 30 m \times 30 m was used as the base for grid resampling by inverse-distance interpolation for all the variables. All spatially explicit DHSVM inputs were added to the ArcMap session as the raster layers. The raster layers were reformatted into a binary format that DHSVM could read and process. The model run at each

time step for each grid cell in the domain automatically.

The DHSVM has many parameters for describing the surface functions and the characteristics of vegetation and soil. Some parameters are rather difficult to be obtained by the instruments in the field. Different parameters influence the simulation results to different extents, some of which may considerably affect the soil moisture. The uncertainty of their values brings much error to the simulation results. Therefore, sensitivity analyses were carried out to select the parameters which primarily affect the temporal behavior of the soil moisture for the assimilation. The approach used here for the sensitivity analyses was proposed by Lenhart *et al.* (2002), in which 10% of the initial value of the respective parameters was added in the sensitivity analyses. It was concluded that the sensitivity of the parameters was ranked into four classes: low, medium, high and very high. The parameters with very high sensitivity were selected and used in the re-parameterization assimilation. The recommended range of the values for each of those selected parameters (Table 1) can be decided from the previous literature (Kang *et al.*, 2008).

Table 1 Selected parameters used in assimilation based on genetic algorithms

Symbol	Definition	Initial value	Range
k_b	Shrub layer radiation attenuation coefficient	0.4	0.1–0.5
F	Canopy fractional coverage (%)	0.8	0.8–1.0
LAI	Monthly LAI	2.8	0.5–22
K	Lateral conductivity (m/s)	10^{-5}	10^{-5} – 10^{-2}
ρ_B	Vertical conductivity (m/s)	10^{-5}	10^{-5} – 10^{-2}
θ_{wp}	Wilting point (m^3/m^3)	0.25	0.06–0.28
θ_{fc}	Field capacity (m^3/m^3)	0.40	0.18–0.41

3.1.2 Oh Model

Oh Model was used to calculate the backscattering coefficient from the soil surface, which was first introduced in 1992 (Oh *et al.*, 1992), and then improved in a later study in 1994 (Oh, 1994). The soil backscattering is given by:

$$\sigma_{vvs}^o = \frac{g \cos^3 \theta [\Gamma_v(\theta) + \Gamma_h(\theta)]}{\sqrt{p}} \quad (3)$$

$$\sigma_{hhs}^o = p \sigma_{vvs}^o \quad (4)$$

$$\sigma_{hvs}^o = q \sigma_{vvs}^o \quad (5)$$

where σ_{vvs}^o , σ_{hhs}^o , σ_{hvs}^o are the backscattering coefficients of the surface soil with different polarization modes, subscript v and h indicate the vertical and horizontal polarization modes, respectively; θ is the incidence angle in radians; $\Gamma_v(\theta)$ and $\Gamma_h(\theta)$ are the v-polarized and h-polarized reflectivity of ground surface; p , q and g are the empirical coefficients. They can be achieved from the following functions:

$$\Gamma_v(\theta) = \frac{\left| \epsilon_s \cos \theta - \sqrt{\epsilon_s - \sin^2 \theta} \right|^2}{\left| \epsilon_s \cos \theta + \sqrt{\epsilon_s - \sin^2 \theta} \right|^2} \quad (6)$$

$$\Gamma_h(\theta) = \frac{\left| \cos \theta - \sqrt{\epsilon_s - \sin^2 \theta} \right|^2}{\left| \cos \theta + \sqrt{\epsilon_s - \sin^2 \theta} \right|^2} \quad (7)$$

$$g = 0.7[1 - e^{-0.65(ks)^{1.8}}] \quad (8)$$

$$p = \frac{\sigma_{hhs}^o}{\sigma_{vvs}^o} = [1 - (\frac{2\theta}{\pi})^{(0.314/\Gamma_0)} \cdot \exp(-ks)]^2 \quad (9)$$

$$q = \frac{\sigma_{hvs}^o}{\sigma_{vvs}^o} = 0.25\sqrt{\Gamma_0(0.1 + \sin^{0.9} \theta)} \cdot [1 - e^{-(1.4-1.6\Gamma_0)ks}] \quad (10)$$

$$\Gamma_0 = \frac{\left| \sqrt{\epsilon_s} - 1 \right|^2}{\left| \sqrt{\epsilon_s} + 1 \right|^2} \quad (11)$$

where k equals to $2\pi/\lambda$ and λ is the wavelength; s is the root mean square (RMS) height of ground surface; Γ_0 is the Fresnel surface reflectivity; ϵ_s is the relative complex dielectric constant of the soil, which depends mainly on the soil moisture (m_v) (Dobson *et al.*, 1985). Moreover, the following can be given:

$$\epsilon_s^\alpha \cong 1 + \frac{\rho_b}{\rho_{ss}}(\epsilon_{ss}^\alpha - 1) + m_v^\beta \epsilon_{fw}^\alpha - m_v \quad (12)$$

$$\beta = 1.09 - 0.11S + 0.18C \quad (13)$$

where $\alpha = 0.65$; S and C are percents of sand and clay, respectively; ρ_b is the soil bulk density; ρ_{ss} is the soil specific density; ϵ_{ss} is the dielectric constant of soil having the extremely low moisture contents; ϵ_{fw} is the dielectric constant of free water.

The soil moisture simulated by the DHSVM was introduced into the Oh model to simulate backscattering coefficient.

3.2 Cost function

The purpose of the data assimilation is to minimize the

distance between ASAR observations and the simulated backscattering coefficients from the DHSVM-Oh coupled model using optimization algorithms. The discrepancy between the ASAR observations at time t_i (σ_{o,t_i}^o) and the corresponding model simulation (σ_{s,t_i}^o) is computed using the following cost function (J):

$$J = \left[\frac{1}{n} \sum_{i=1}^n (\sigma_{o,t_i}^o - \sigma_{s,t_i}^o)^2 \right]^{\frac{1}{2}} \quad (14)$$

The simulations were divided into acceptable and unacceptable cases. If the cost function was lower than twice the ASAR observation noise, the simulation was considered as acceptable. The EnKF was then used to optimize the state variables. If the cost function was higher than twice the ASAR observation noise, the simulation was treated as unacceptable. Accordingly, the cost function was minimized using GA while the selected parameters were updated. The ASAR observation noise level was referenced from Baup *et al.* (2007).

3.3 Assimilation algorithms

3.3.1 Ensemble Kalman Filter

EnKF was first introduced by Evensen (1994) which is based on the stochastic dynamic prediction theory. Burgers *et al.* (1998) modified the analysis scheme proposed by Evensen, in which the random perturbations were added with the correct statistics to the observations to generate the ensembles of observations that then were used in updating the ensembles of model states. The EnKF based on the Monte-Carlo approach was generally used. The method is formulated with nonlinear dynamics and can properly handle the error covariance evolution in nonlinear models.

Use of EnKF implicitly assumes that the observations are related to the true state (x^t) through

$$d = Hx^t + \varepsilon \quad (15)$$

where d is the observation vector; H is the operator that maps the model variable space to the observation space; x is the state variable, the superscript t represents true state; ε is the measurement error, a Gaussian random error vector with a mean of zero and measurement error covariance R . Furthermore, the forecast of x^t values follow the Gaussian distribution with the mean x^f and error covariance P^f . By using these assumptions, the estimated state of the profile soil moisture and error covariance is

updated as follows:

$$x^a = x^f + P^f H^T (H P^f H^T + R)^{-1} (d - H x^f) \quad (16)$$

$$P^a = P^f - P^f H^T (H P^f H^T + R)^{-1} H P^f \quad (17)$$

where the superscripts f and a represent the respective prior (forecast) and posterior (analysis) estimates, respectively, T represents the matrix transpose, while the script t (time) is omitted in these equations, while the Kalman gain matrix K is defined as:

$$K = P^f H^T (H P^f H^T + R)^{-1} \quad (18)$$

It is essential that the observations are treated as random variables having a distribution with a mean equal to the first-guess observations and covariance equal to R . The ensemble of observations is defined as:

$$d_j = d + \varepsilon_j \quad (19)$$

where j counts from 1 to the number of model state ensemble members N . The ensemble is generated by perturbing a first guess value so that ensemble mean is equal to the first guess value. The variance is specified based on the uncertainty in the first guess. Then EnKF forecast and analysis error covariance come directly from an ensemble of model simulation as:

$$P^f H^T = (N-1)^{-1} \sum_{j=1}^N (x_j^f - \bar{x}^f) (H x_j^f - H \bar{x}^f)^T \quad (20)$$

where N is the number of ensemble members, and the subscript j represents j th individual ensemble member. The overbar represents the ensemble mean. Ensemble members are integrated independently and updated in accordance with the Kalman filter methods when the new observations become available. In this specific application, when the simulated variables and observed variables conclude only one variable (the backscattering coefficient), the error covariance may transform into the error variance of the single variable.

3.3.2 Genetic Algorithm

If the cost function was considered to be unacceptable, the sensitive parameters were re-parameterized. The selected parameters were optimized by minimizing the cost function using GA. GA is an advanced search and optimization technique. GA was first proposed by John (1973) at the University of Michigan in the United States. It simulates the principle of encouraging good structure "survival of the fittest" from Darwinian. GA adopts the genetic variation theory from Mendelian,

which maintains some of the structure and at the same time finds a better structure of an adaptive search algorithm for a global optimization.

To solve the identification problem, each of the seven selected parameters is codified as a gene, and the seven genes are grouped in a chromosome. In the GA, the seven parameters constitute the chromosome of an individual. Equation (21) represents the fitness function (f):

$$\begin{aligned} f &= \frac{1}{1+J} = \frac{1}{1 + \left[\frac{1}{n} \sum_{i=1}^n (\sigma_{o,t_i}^o - \sigma_{s,t_i}^o)^2 \right]^{\frac{1}{2}}} \\ &= \frac{1}{1 + \left[\frac{1}{n} \sum_{i=1}^n (F(\Omega, X_i) - \sigma_{s,t_i}^o)^2 \right]^{\frac{1}{2}}} \\ a_i &\leq \alpha_i \leq b_i, \quad (i = 1, 2, \dots, n) \end{aligned} \quad (21)$$

where J is the cost function (Equation (14)), $\Omega = \{\alpha_i\}$ is the 7-dimensional vector of the model selected parameters; $[a_i, b_i]$ is the initial interval of α_i , which is the search space; X_i is the model input; F is the DHSVM-Oh coupled model.

The GA is carried out according to the following steps: 1) Creating a population with 100 individuals; 2) evaluating the fitness for each individual; 3) selecting individuals according to their fitness values and building a temporary population; 4) implementing single point real crossover with the probability of 0.6 on the temporary population; 5) mutating the current population with the probability of 0.05; 6) repeating the steps 2, 3, 4, 5 and 6 until the number of the generations is met. It should be pointed out that the specific parameters in GA are referenced from Tutkun (2009).

4 Results

To evaluate the impact of the ASAR observations on the soil moisture simulations and examine the assimilation scheme, we conducted the data assimilation experiments in the point scale. The experiments were performed from 27 May 2008 to 10 July 2008 at the Huzhaizi site in the Heihe watershed. In this study, a hydrologic and semi-empirical backscattering coupled model was applied to estimate the simulated backscattering coefficients. The difference between simulated and observed backscattering coefficient was evaluated using a cost function. Table 2 provides the calculation results in

terms of the surface soil moisture from DHSVM simulations (m_v), backscattering coefficients from the coupled model (σ_{s,t_i}^o), the backscattering coefficients from the ASAR observations (σ_{o,t_i}^o) and the results of the cost function (J) during the assimilation process when the ASAR data appear.

As shown in Table 2, it can be found that the result of the cost function J was quite large on Julian day 171 and the GA was carried out to re-parameterize the sensitive parameters. The backscattering coefficients from the simulations were less than those from the observations. Additionally, J decreased when this data assimilation was performed while J increased on Julian day 187 when rain occurred and it decreased on Julian day 190. It is indicated that the difference between the observed and simulated backscattering coefficients can be decreased significantly during data assimilation process.

The results of the data assimilations, DHSVM model free-run simulation and the field measurement were compared (Fig. 3) and then evaluated by the root mean square error (RMSE) and mean bias errors (MBE).

The root mean square errors (RMSEs) were 0.0314 and 0.0112 without assimilation and with assimilation, respectively, in the surface soil layer, while the RMSEs were 0.0068 and 0.0025 without assimilation and with assimilation, respectively, in the root zone. For the two

different scenarios the mean bias errors (MBEs) for the surface soil layer simulation were -0.0126 and -0.0060 without and with assimilations respectively, while for the root zone the MBEs were -0.0063 and -0.0021 correspondingly. The result shows that the model estimate of surface soil moisture was improved by being assimilated with the ASAR data. Additionally, the data assimilation improves the DHSVM model outputs which are more close to the actual values. It is observed that the field measured soil moisture fluctuated during one day time. However, the model simulation can hardly show this variation. Nevertheless, the trend of the model simulations was consistent with the field measurements at a longer time scale. It is quite apparent that, the data assimilation can actually improve the trend of simulation compared with the free-runs. Accordingly, the assimilation method is effective.

5 Discussion and Conclusions

This paper proposes a one-dimensional soil moisture assimilation scheme combining the EnKF and GA, which has been illuminated by the assimilation experiment carried out in the Heihe River Basin in Northwest China in 2008. A cost function was set up based on the gap between backscattering coefficients from ASAR

Table 2 The calculation results during the assimilation process when the ASAR data appear

Julian day	171	177	180	187	190
m_v	0.034	0.072	0.054	0.062	0.055
σ_{s,t_i}^o (dB)	-10.341	-8.170	-6.752	-9.361	-8.755
σ_{o,t_i}^o (dB)	-8.238	-7.954	-6.264	-8.855	-8.439
J	2.103	1.295	0.954	1.114	1.044

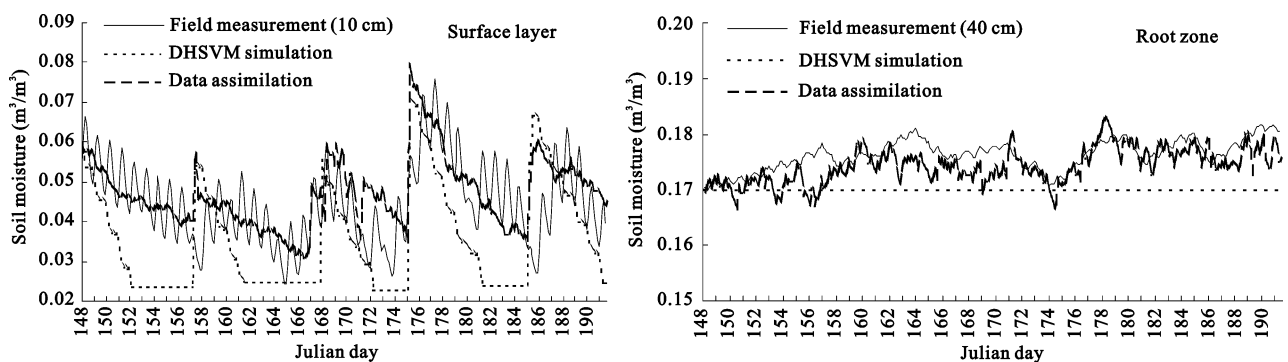


Fig. 3 Comparisons among field measurement, DHSVM simulation and assimilation results at Huazhaizi site from 27 May (Julian day 148) to 10 July (Julian day 190), 2008

observation and the backscattering coefficients from the hydrologic and radiation transfer coupled model. In order to improve the simulation precision of the soil moisture, the EnKF and GA were used to re-initialize and re-parameterize the simulation processes, respectively. The following conclusions can be drawn:

(1) Assimilation of the ASAR observations with the DHSVM-Oh coupled model improves the simulation results of soil moisture significantly compared to the free-run DHSVM model.

(2) In the assimilation scheme proposed in this paper, both the initial model conditions and the model sensitive parameters can be optimized using a combined assimilation approach based on EnKF and GA, which can improve the simulation precision further.

(3) ASAR images were used as observation data, that were operated in the band C (their wavelength were 5.6 cm). It should be mentioned that the penetration ability of the band C is relatively weak. Thus, the radar signals were primarily affected by the scattering from the surface soil. The semi-empirical model—Oh was used, which was established on the relationship between the surface soil moisture and backscattering coefficients. Therefore, only the surface soil moisture on the contribution of backscattering coefficients was considered and the soil scattering effects from the root zone were ignored.

References

- Baup F, Mougin E, Rosnay P *et al.*, 2007. Surface soil moisture estimation over the AMMA Sahelian site in Mali using ENVISAT/ASAR data. *Remote Sensing of Environment*, 109(4): 473–481. DOI: 10.1016/j.rse.2007.01.015
- Burgers G, Leeuwen P J, Evensen G, 1998. Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, 126(6): 2884–2903.
- Dobson M C, Ulaby F T, Hallikainen M *et al.*, 1985. Microwave dielectric behavior of wet soil Part II: Four-component dielectric mixing models. *IEEE Transactions on Geoscience and Remote Sensing*, 23(1): 35–46.
- Evensen G, 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5): 10143–10162.
- Heathman G C, Starks P J, Ahuja L R *et al.*, 2003. Assimilation of surface soil moisture to estimate profile soil water content. *Journal of Hydrology*, 279(1–4): 1–17.
- Huang C L, Li X, Lu L *et al.*, 2008. Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter. *Remote Sensing of Environment*, 112(3): 888–900. DOI: 10.1016/j.rse.2007.06.026
- Jarlan L, Mougin E, Mazzega P *et al.*, 2005. Using coarse remote sensing radar observations to control the trajectory of a simple Sahelian land surface model. *Remote Sensing of Environment*, 94(2): 269–285. DOI: 10.1016/j.rse.2004.10.005
- John H H, 1973. Genetic algorithms and the optimal allocation of trials. *Siam Journal on Computing*, 2(2): 88–105.
- Kang L L, Wang S R, Gu J Q, 2008. The simulation test of the distributed hydrological model DHSVM on the runoff change of Lanjiang River Basin. *Journal of Tropical Meteorology*, 24(2): 176–183. (in Chinese)
- Kostove K G, Jackson T J, 1993. Estimating profile soil moisture from surface layer measurements: A review. *International Society for Optical Engineering*, 1941: 125–136. DOI: 10.1117/12.154681
- Lenhart T, Eckhardt K, Fohrer N *et al.*, 2002. Comparison of two different approaches of sensitivity analysis. *Physics and Chemistry of the Earth*, 27(9–10): 645–654.
- Oh Y, 1994. An inversion algorithm for retrieving soil moisture and surface roughness from polarimetric radar observation. *IEEE International Geoscience and Remote Sensing Symposium*, Pasadena, USA, 1582–1584.
- Oh Y, Sarabandi K, Ulaby F T, 1992. An empirical model and an inversion technique for radar scattering from bare soil surfaces. *IEEE Geoscience and Remote Sensing*, Florida, USA, 370–381.
- Onwubolu G C, Kumalo T, 2001. Multi-pass turning operations optimization based on genetic algorithms. *Journal of Engineering Manufacture*, 215(4): 117–124.
- Qin J, Liang S, Liu R *et al.*, 2007. A weak-constraint based data assimilation scheme for estimating surface turbulent fluxes. *IEEE Geoscience and Remote Sensing Letters*, 4(4): 649–653.
- Tutkun N, 2009. Parameter estimation in mathematical models using the real coded genetic algorithms. *Expert Systems with Applications*, 36(2): 3342–3345.
- Walker J P, Willgoose G R, Kalma J D, 2001. One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: A comparison of retrieval algorithms. *Advances in Water Resources*, 24(6): 631–650.
- Wigmosta M S, Nijssen B, Storck P *et al.*, 2002. The distributed hydrology soil vegetation model. In: *Mathematical Models of Small Watershed Hydrology and Applications*. Colorado: Water Resource Publications, LLC, 7–42.
- Wigmosta M S, Vail L, Lettenmaier D P, 1994. A distributed hydrology-vegetation model for complex terrain. *Water Resource Research*, 30(6): 1665–1679.
- Yang S B, Shen S H, Li B B *et al.*, 2009. Mapping rice yield based on assimilation of ASAR data with rice growth model. *Journal of Remote Sensing*, 13(2): 282–290.