Hydrol. Earth Syst. Sci. Discuss., 11, 10635–10681, 2014
www.hydrol-earth-syst-sci-discuss.net/11/10635/2014/
doi:10.5194/hessd-11-10635-2014

Hydrology and Earth System Sciences
Open Access
Discussions

Improving operational flood ensemble prediction by the assimilation of satellite soil moisture: comparison between lumped and semi-distributed schemes

C. Alvarez-Garreton[1], D. Ryu[1], A. W. Western[1], C.-H. Su[1], W. T. Crow[2], D. E. Robertson[3], and C. Leahy[4]

[1]Department of Infrastructure Engineering, The University of Melbourne, Parkville, Victoria, Australia
[2]USDA-ARS Hydrology and Remote Sensing Laboratory, Beltsville, Maryland, USA
[3]CSIRO Land and Water, Highett, Australia
[4]Bureau of Meteorology, Melbourne, Victoria, Australia

Received: 9 September 2014 – Accepted: 9 September 2014 – Published: 23 September 2014

Correspondence to: C. Alvarez-Garreton (calvarez@student.unimelb.edu.au)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Assimilation of remotely sensed soil moisture data (SM–DA) to correct soil water stores of rainfall-runoff models has shown skill in improving streamflow prediction. In the case of large and sparsely monitored catchments, SM–DA is a particularly attractive tool. Within this context, we assimilate active and passive satellite soil moisture (SSM) retrievals using an ensemble Kalman filter to improve operational flood prediction within a large semi-arid catchment in Australia (> 40 000 km^2). We assess the importance of accounting for channel routing and the spatial distribution of forcing data by applying SM–DA to a lumped and a semi-distributed scheme of the probability distributed model (PDM). Our scheme also accounts for model error representation and seasonal biases and errors in the satellite data.

Before assimilation, the semi-distributed model provided more accurate streamflow prediction (Nash–Sutcliffe efficiency, NS = 0.77) than the lumped model (NS = 0.67) at the catchment outlet. However, this did not ensure good performance at the "ungauged" inner catchments. After SM–DA, the streamflow ensemble prediction at the outlet was improved in both the lumped and the semi-distributed schemes: the root mean square error of the ensemble was reduced by 27 and 31 %, respectively; the NS of the ensemble mean increased by 7 and 38 %, respectively; the false alarm ratio was reduced by 15 and 25 %, respectively; and the ensemble prediction spread was reduced while its reliability was maintained.

Our findings imply that even when rainfall is the main driver of flooding in semi-arid catchments, adequately processed SSM can be used to reduce errors in the model soil moisture, which in turn provides better streamflow ensemble prediction. We demonstrate that SM–DA efficacy is enhanced when the spatial distribution in forcing data and routing processes are accounted for. At ungauged locations, SM–DA is effective at improving streamflow ensemble prediction, however, the updated prediction is still poor since SM–DA does not address systematic errors in the model.

HESSD

11, 10635–10681, 2014

1 Introduction

Floods have large negative impacts on society, causing destruction of infrastructure and crops, erosion, and in the worst cases, injury and loss of life (Thielen et al., 2009). To reduce flood impacts on public safety and the economy, early and accurate alert systems are needed. These systems rely on hydrologic model predictions, whose accuracy in turn is highly dependent on the quality of the data used to force and calibrate them. Therefore, in the case of sparsely monitored and ungauged catchments, flood prediction suffers from large uncertainties.

A plausible approach to reduce model uncertainties in the sparsely monitored catchments is to exploit remotely sensed hydro-meteorological observations to correct the states or parameters of the model in a data assimilation (DA) framework. Within this context, satellite soil moisture (SSM) products are appealing given the vital role of soil moisture (SM) in the runoff generation. SM influences the partitioning of energy and water (rainfall, infiltration and evapotranspiration) between the land surface and the atmosphere (Western et al., 2002). SSM observations provide global scale information and can be obtained in near real time at regular and reasonably frequent time intervals. This makes them valuable for improving the representation of catchment wetness. The accuracy of SSM has been assessed by a number of studies (Albergel et al., 2009, 2010, 2012; Draper et al., 2009; Gruhier et al., 2010; Brocca et al., 2011; Su et al., 2013). In general, they have showed promising performance, with moderate correlation between SSM and ground data, but with significant bias at some locations.

In the last decade a large number of studies have explored SSM data assimilation (SM–DA) to correct the water states of models. These studies can be categorised in two main groups; the first, and larger group, has focused their evaluation on the improvement of the SM predicted by the model (generally working with land surface models, e.g., Crow and van Loon, 2006; Crow and Reichle, 2008; Crow and Van den Berg, 2010; Reichle et al., 2008; Ryu et al., 2009). The second, and smaller group (where our study fits), has focused on the improvement of streamflow prediction

HESSD
11, 10635–10681, 2014

in rainfall-runoff models (Francois et al., 2003; Brocca et al., 2010, 2012; Alvarez-Garreton et al., 2013, 2014; Chen et al., 2014; Wanders et al., 2014).

While the studies from the first group evaluate the improvement in prediction of the same variable that is updated in the assimilation scheme (SM), studies in the second groups focus on streamflow, which involves indirect improvements due to better representation of SM. The potential improvement of streamflow predictions in the latter case is constrained by the particular runoff mechanisms operating within a catchment. Accordingly, even when a model structure and parametrisation are capable of representing the runoff mechanisms, improving streamflow prediction by reducing error in soil moisture depends on the error covariance between these two components. This error covariance (which in the model space will be defined by the representation given to the different sources of uncertainty) may become marginal when the errors in streamflow come mainly from errors in rainfall input data (Crow and Ryu, 2009). This physical constraint is case specific and determines the potential skill of SM–DA for improving streamflow prediction. To understand and assess this skill, further studies focusing on the improvement of streamflow prediction are needed with different model characteristics, such as structure, parametrisation and performance before assimilation; and with different catchment characteristics, such as climate, scale, soils, geology, land cover and density of monitoring network. Among the latter, semi-arid catchments present unique rainfall-runoff processes which have been rarely studied in SM–DA.

Here we address this gap by studying the Warrego River catchment in Australia, a large and sparsely monitored semi-arid basin. We set up the probability distributed model (PDM) within the catchment, and assimilate passive and active SSM products using an ensemble Kalman filter (Evensen, 2003), for the purpose of improving operational flood prediction. We devise an operational SM–DA scheme to answer three main questions. (1) While rainfall is presumably the main driver of flood generation in semi-arid catchments, can we effectively improve streamflow prediction by correcting the soil water state of the model? (2) What is the impact of accounting for channel

HESSD
11, 10635–10681, 2014

routing and the spatial distribution of forcing data on SM–DA performance? (3) What are the prospects for improving streamflow within ungauged inner catchments using SSM?

A series of SM–DA experiments using a lumped version of PDM have already been undertaken in this study catchment by Alvarez-Garreton et al. (2014). They found that assimilating passive microwave SSM improved flood prediction, while highlighting specific limitations in their scheme. In this paper we address those limitations by applying more robust techniques in the SM–DA framework. In particular, we improve the representation of model error by explicitly treating forcing, parameter and structural errors. We incorporate additional SSM products and apply instrument variables regression techniques for SSM seasonal rescaling and observations error estimation. Furthermore, we employ a semi-distributed scheme to evaluate the advantages of accounting for channel routing and the spatial distribution of forcing data.

In this paper, Sect. 2 presents a description of the study catchment and the data used. Section 3 presents the methodology, including a description of the rainfall-runoff model, the EnKF formulation and the specific steps for setting up the SM–DA scheme. These include the error model estimation, estimation of profile SM based on the satellite surface data, the rescaling of satellite observations and observation error estimation. Section 4 presents the results and discussion. Section 5 summarises the main conclusions of the study.

2 Study area and data

The study area is the semi-arid Warrego catchment (42 870 km^2) located in Queensland, Australia (Fig. 1). The catchment has an important flooding history, with at least three major floods within the last 15 years. The study area also features geographical and climatological conditions that enable SSM retrievals to have higher accuracy than in other areas. These conditions include the size of the catchment, the a semi-arid climate and the low vegetation cover. Moreover, the ground-monitoring

HESSD
11, 10635–10681, 2014

network within the catchment is sparse thus satellite data is likely to be more valuable than in well-instrumented catchments. The catchment has summer-dominated rainfall with mean monthly rainfall of 80 mm in January, and 20 mm in August. Mean maximum daily temperature in January is above 30 °C and in July below 20 °C. The runoff seasonality is characterised by peaks in summer months and minimum values in winter and spring. The mean annual precipitation over the catchment is 520 mm.

Daily rainfall data was collected from the gauge-interpolated dataset of the Australian Water Availability Project (AWAP), which has a spatial resolution of 0.05° (Jones et al., 2009). Hourly streamflow records were collected from the State of Queensland, Department of Natural Resources and Mines (http://watermonitoring.dnrm.qld.gov.au) for the 3 gauges within the catchment (Fig. 1). Daily discharge was calculated based on the daily AWAP time convention (09:00 a.m.–08:59 a.m. LT, UTC + 10 h). The flood classification for the study catchment (at the catchment outlet, N7) is provided by the Australian Bureau of Meteorology (BoM) as river height threshold values, corresponding to minor, moderate and major floods. These threshold values expressed as streamflow (mm day^{-1}) are 0.06, 0.55 and 2.05, respectively and relate to flood impact rather than recurrence interval. The associated annual exceedance probability for the minor, moderate and major floods at N7 are 15.7, 3.1, and 0.95 %, respectively (calculated using the complete daily streamflow record period). Potential evapotranspiration was obtained from the climatological 0.05° grid data provided by the BoM (Australian Data Archive for Meteorology database). Daily values were estimated by assuming a uniform daily distribution within a month.

Three SSM products are used here. The first is the Advance Microwave Scanning Radiometer – Earth Observing System (AMSR-E, AMS hereafter) version 5 VUA-NASA LPRM (Land Parameter Retrieval Model) Level 3 gridded product (Owe et al., 2008). The second product is the TU-WIEN (Vienna University of Technology) ASCAT (ASC hereafter) data produced using the change-detection algorithm (Water Retrieval Package, version 5.4) (Naeimi et al., 2009). The third SSM product is the Soil Moisture and Ocean Salinity (SMOS) satellite (SMO hereafter), version RE01 (Re-

HESSD
11, 10635–10681, 2014

processed 1 day global soil moisture product) provided by Centre Aval de Traitement des Donnees. The overpass times of the AMS, ASC and SMO over the study catchments are 01:30 a.m./p.m., 10:00 a.m./p.m. and 06:00 a.m./p.m. LT (UTC + 10 h), respectively. Figure 2 summarises the period of record of the different datasets.

For each satellite dataset, a daily averaged SSM was calculated for the complete catchment (or sub-catchment in the case of the semi-distributed scheme). The areal estimate of SSM over the catchment is given by averaging the values of ascending and descending satellite passes on days when more than 50 % of the pixels had valid data. For the case of the passive sensors (AMS and SMO), we subtracted the long-term temporal mean of the ascending and descending datasets to remove the systematic bias between them (Brocca et al., 2011; Draper et al., 2009). Then, daily SSM was calculated as the average between the mean-removed ascending and descending passes (if both were available) or directly as the mean-removed available pass. For ASC retrievals, given the unbiased ascending and descending measurements, daily SSM was calculated from the actual ascending and descending values averaged over the catchment.

3 Methods

3.1 Lumped and semi-distributed model schemes

The probability distributed model (PDM) is a conceptual rainfall-runoff model that has been widely used in hydrologic research and applications (Moore, 2007). The model treats soil moisture (S_1 in Fig. 3) as a distributed variable following a Pareto distribution function. The SM component together with the net rainfall (rainfall minus evapotranspiration and groundwater recharge) define the separation between direct runoff and subsurface runoff. Direct runoff is routed through two cascade of reservoirs (S_{21} and S_{22} in Fig. 3, with time constants k_1 and k_2, respectively). Subsurface runoff is estimated based on the drainage from S_1 and transformed into baseflow by using

HESSD
11, 10635–10681, 2014

a storage reservoir (S_3 in Fig. 3 with time constant k_b). These are then combined as total runoff, or streamflow. A detailed description of the model conceptualisation and the formulation of the different rainfall-runoff processes is presented in Moore (2007).

PDM was set up using both a lumped scheme and a semi-distributed scheme (see Fig. 1). The semi-distributed scheme was configured with 7 sub-catchments (SC1 to SC7), each using the lumped version of PDM. The area and mean annual rainfall (MAR) of each sub-catchment are summarised in Table 1. The river routing between upstream and downstream sub-catchments in the semi-distributed scheme was represented by a linear Muskingum method (Gill, 1978):

$$S = k_m(Ix + (1 - x)O), \quad (1)$$

where S is the storage within the routing reach, k_m is the storage time constant, I and O are the streamflow at the beginning and end of the reach, respectively, and x is a weighting factor parameter. The time-dependant model parameters in the semi-distributed scheme (k_1, k_2 and k_b) were scaled by the area of each sub-catchment, and k_m from Muskingum routing was scaled by the length of the river channel between corresponding nodes. The remaining model and routing parameters of the semi-distributed scheme were treated as homogeneous.

The lumped and the semi-distributed models were calibrated by using a genetic algorithm (Chipperfield and Fleming, 1995) with an objective function based on the Nash–Sutcliffe model efficiency (NS) (Nash and Sutcliffe, 1970). The calibration was done within the period 1 January 1967–31 May 2003 and the evaluation within 1 June 2003–2 March 2014. To make fair comparisons between the two model setups in a scenario where the inner catchments are ungauged, the semi-distributed scheme was calibrated using only the outlet gauge (N7 in Fig. 1). The performance of the calibrated models was evaluated based on the NS at the catchment outlet (N7, Fig. 1) and at inner nodes N1 and N3, in the case of the semi-distributed scheme.

To analyse the runoff mechanisms simulated by the lumped and the semi-distributed schemes, we calculated the lag-correlation between rainfall and streamflow, and

HESSD
11, 10635–10681, 2014

between antecedent SM and streamflow. This enables further understanding of the improvement in streamflow that can be expected by improving the simulated SM content through SM–DA.

3.2 EnKF formulation

The ensemble Kalman filter (EnKF) proposed by Evensen (2003) has been widely used in hydrologic applications given the highly nonlinear nature of runoff processes. In the EnKF, the error covariance between model and observations is calculated from Monte Carlo-based ensemble realisations. In this way, the model and observation uncertainties are propagated and the streamflow prediction is treated as an ensemble of equally likely realisations. The uncertainty of the streamflow prediction can be derived from the ensemble, which provides valuable information for operational flood alert systems.

In a state-updating assimilation approach, the state ensemble is created by perturbing forcing data, parameters and/or states of the model with unbiased errors. As we will see in Sect. 3.3, an N-member ensemble of the water content in S_1 (θ hereafter) was generated by perturbing rainfall forcing data, the model parameter k_1, and θ. Then, the error of the state member i at time t was estimated as

$$\theta_{i,t}^{-\prime} = \theta_{i,t}^{-} - \frac{1}{N}\sum_{i=1}^{N}\theta_{i,t}^{-}, \tag{2}$$

where the superscript "$^-$" denotes the state prediction prior to assimilation step. The anomaly vector for time step t was defined by $\theta^-(t)' = \{\theta_{1,t}^{-\prime}, \theta_{2,t}^{-\prime}, \ldots, \theta_{N,t}^{-\prime}\}$ and the error covariance of the model state (P^-) was estimated at each time step as:

$$P^-(t) = \frac{1}{N-1}\theta^-(t)' \cdot \left(\theta^-(t)'\right)^T. \tag{3}$$

When a daily SSM observation from AMS, ASC or SMO was available, each member of the background prediction (θ^-) was updated. Before being assimilated, each of the

HESSD
11, 10635–10681, 2014

three SSM datasets was transformed to represent a profile SM and then rescaled to remove systematic differences between the model and the transformed observation (details in Sects. 3.5 and 3.6). We sequentially assimilated a N-member ensemble of the transformed and rescaled AMS, ASC and SMO (θ^{ams}, θ^{asc} and θ^{smo}, respectively) and updated θ^- with the following 3 steps:

1. If θ^{ams} was available at time t,

$$\theta^+_{i,t} = \theta^-_{i,t} + K^1_t \left(\theta^{\text{ams}}_{i,t} - H\theta^-_{i,t} \right), \tag{4}$$

where H is an operator that transforms the model state to the measurement space. In order to correct additive and multiplicative biases between the model predictions and the microwave retrievals, remotely sensed SM is rescaled in a separate step (see Sect. 3.6), thus H reduced to a unit matrix. The Kalman gain K^1_t was calculated as

$$K^1_t = \frac{P^-_t H^T}{H P^-_t H^T + R^1_t}, \tag{5}$$

where R^1_t is the error variance of θ^{ams} estimated in the rescaling procedure (Sect. 3.6). If θ^{ams} was not available, $\theta^+_{i,t} = \theta^-_{i,t}$.

2. If θ^{asc} was available at time t, we updated the model soil moisture with

$$\theta^{++}_{i,t} = \theta^+_{i,t} + K^2_t \left(\theta^{\text{asc}}_{i,t} - H\theta^+_{i,t} \right), \tag{6}$$

where K^2_t was calculated as

$$K^2_t = \frac{P^-_t H^T}{H P^-_t H^T + R^2_t}. \tag{7}$$

HESSD
11, 10635–10681, 2014

R_t^2 is the error variance of θ^{asc} and P^- is the model error covariance re-calculated by applying Eq. (3) to the updated soil moisture $\theta_{i,t}^+$. If θ^{asc} was not available, $\theta_{i,t}^{++} = \theta_{i,t}^+$.

3. If θ^{smo} was available at time t, we updated the model soil moisture with

$$\theta_{i,t}^{+++} = \theta_{i,t}^{++} + K_t^3 \left(\theta_{i,t}^{smo} - H\theta_{i,t}^{++} \right), \quad (8)$$

where K_t^3 was calculated as

$$K_t^3 = \frac{P_t^- H^T}{H P_t^- H^T + R_t^3}. \quad (9)$$

R_t^3 is the error variance of θ^{smo} and P^- is the model error covariance re-calculated by applying Eq. (3) to the updated soil moisture $\theta_{i,t}^{++}$. If θ^{smo} was not available, $\theta_{i,t}^{+++} = \theta_{i,t}^{++}$.

3.3 Error model representation

The main sources of uncertainty in hydrologic models are the errors in the forcing data, the model structure and the incorrect specification of model parameters (Liu and Gupta, 2007). Generally, these errors are represented by adding unbiased synthetic noise to forcing variables, model state variables and/or model parameters.

The estimation of these errors is among the most crucial challenges in DA, as it determines the value of the Kalman gain. Moreover, in the case of a state updating SM–DA, the ability of the scheme to improve streamflow prediction will mainly depend on the covariance between the errors in SM states and modelled streamflow, which directly depends on the specific representation and estimation of the model errors.

To represent the forcing uncertainty, we adopted a multiplicative error model for the rainfall data (McMillan et al., 2011; Tian et al., 2013). In particular, we followed the

HESSD
11, 10635–10681, 2014

scheme used in various SM–DA studies (e.g., Chen et al., 2011; Brocca et al., 2012; Alvarez-Garreton et al., 2014) and represented a spatially homogeneous rainfall error (ϵ_p) as

$$\epsilon_p \sim \ln N\left(1, \sigma_p^2\right), \tag{10}$$

where σ_p is the standard deviation of the lognormal distribution. The above representation assumes a spatially homogeneous fraction of the error to the rainfall intensity, which could be an over simplification in a large area like the study catchment. However, it avoids the estimation of additional error parameters (e.g., spatial correlation parameter) in an already highly undetermined problem (see Sect. 3.4).

The parameter uncertainty was represented by perturbing the time constant parameter (k_1) for store S_{21}, a highly sensitive parameter of the model that directly affects the streamflow generation by influencing the water stored in both surface storages S_{21} and S_{22} (note that in the PDM formulation used, the time constant k_2 is calculated as a function of k_1). The parameter error (ϵ_k) adopted was

$$\epsilon_k \sim N\left(0, \sigma_k^2\right), \tag{11}$$

where σ_k is the standard deviation of the normal distribution.

Following the scheme used in most SM–DA experiments (e.g., Reichle et al., 2008; Crow and Van den Berg, 2010; Chen et al., 2011; Hain et al., 2012), the model structural error was represented by perturbing the SM prediction (θ) with a spatially homogeneous additive random error,

$$\epsilon_s \sim N\left(0, \sigma_s^2\right), \tag{12}$$

where σ_s is the standard deviation of the normal distribution.

The physical boundary conditions of SM (porosity as an upper bound and residual water content as a lower bound) are represented by the model through the storage

HESSD
11, 10635–10681, 2014

capacity of S_1. When θ approaches the limits of S_1 capacity, applying unbiased perturbation to θ can lead to truncation bias in the background prediction. The latter can result in mass balance errors and degrade the performance of the EnKF (Ryu et al., 2009). Moreover, the Kalman filter assumes unbiased state variables. This issue is of particular importance in arid regions like the study area, where the soil water content can get rapidly depleted by evapotranspiration and transmission losses, thus approaching the residual water content of the soil. To ensure that the state ensemble remained unbiased after perturbation we implemented the bias correction scheme proposed by Ryu et al. (2009).

The truncation bias correction consisted on running a single unperturbed model prediction (θ_t^{-0}) in parallel with the perturbed model prediction ($\theta_{i,t}^{-}$). A mean bias, δ_t, of the N-member ensemble prediction was calculated by subtracting θ_t^{-0} from the ensemble mean, as follows (Ryu et al., 2009):

$$\delta_t = \frac{1}{N}\sum_{i=1}^{N}\theta_{i,t}^{-} - \theta_t^{-0}. \quad (13)$$

Then, a bias corrected ensemble of state variables, $\tilde{\theta}_{i,t}^{-}$, was obtained by subtracting δ_t from each member of the perturbed ensemble, $\theta_{i,t}^{-}$.

Although the latter resulted in unbiased state ensemble, there are still some important but subtle effects that arise from the highly non-linear nature of hydrologic models that need to be guarded against in SM–DA. Representing model errors by adding unbiased perturbation to forcing, model parameters and/or model states can lead to a biased streamflow ensemble prediction (e.g., Ryu et al., 2009; Plaza et al., 2012). This biased streamflow ensemble prediction (open-loop hereafter) is degraded compared with the streamflow predicted by the unperturbed calibrated model. As a consequence, improvement of the open-loop after SM–DA may in part be due to the correction of bias introduced during the assimilation process itself.

HESSD
11, 10635–10681, 2014

3.4 Error model parameters calibration

To calibrate the error model parameters (σ_p, σ_k and σ_s), we evaluated how the open-loop ensemble prediction compared with the observed streamflow at the catchment outlet. For this we used a maximum a posteriori (MAP) scheme, a Bayesian inference procedure detailed by Wang et al. (2009) that maximises the probability of observing historical events given the model and error parameters. In other words, it maximises the probability of having the streamflow observation within the open-loop streamflow.

The N-member open-loop ($Q_{\text{sim}}^{\text{ol}}$) can be expressed as

$$Q_{\text{sim}}^{\text{ol}}(t) = Q^{\text{T}}(t) + \epsilon_m(t), \tag{14}$$

where Q^{T} is the (unknown) truth streamflow and ϵ_m is the error of the streamflow prediction and consists of forcing, parameter and states errors:

$$\epsilon_m(t) = f(\epsilon_p, \epsilon_k, \epsilon_s). \tag{15}$$

The observed streamflow at N7 (Q_{obs}) can be expressed as a function of the same (unknown) truth and the streamflow observation error (ϵ_{obs}),

$$Q_{\text{obs}}(t) = Q^{\text{T}}(t) + \epsilon_{\text{obs}}(t). \tag{16}$$

Combining Eqs. (14) and (16), the model ensemble prediction of the observed streamflow ($\hat{Q}_{\text{obs}}$) is expressed as:

$$\hat{Q}_{\text{obs}}(t) = Q_{\text{sim}}^{\text{ol}}(t) + \epsilon_m(t) + \epsilon_{\text{obs}}(t). \tag{17}$$

Following Li et al. (2014), ϵ_{obs} was assumed to be serially independent multiplicative error following a normal distribution (mean 1 and standard deviation of 0.2). Then, the likelihood function (L) defining the probability of observing the historical streamflow

HESSD
11, 10635–10681, 2014

data given the calibrated model parameters (x), and the error model parameters (σ_p, σ_k and σ_s), was expressed by

$$L(Q_{\text{obs}}|x,\sigma_p,\sigma_k,\sigma_s) = \Pi_{t=1}^{n} p(Q_{\text{obs}}(t)|\hat{Q}_{\text{obs}}(t)). \quad (18)$$

To maximise L, we applied a logarithm transformation to it and maximised the sum over time of the transformed function. The probability density function (p) at each time step was estimated by assuming that the ensemble prediction of the observed streamflow, $\hat{Q}_{\text{obs}}(t)$, follows a Gaussian distribution, with its mean and standard deviation computed using the ensemble members. The period used to calibrate the error model parameters was 1 January 1998–31 May 2003.

An important aspect to highlight about this error parameter calibration is that it is a highly-undetermined problem. Only one data set (streamflow at N7) is used to calibrate the error parameters, while there might be many combinations of error parameters to generate similar streamflow ensemble (equifinality on the error parameters).

3.5 Profile soil moisture estimation

The aim of the stochastic assimilation detailed in Sect. 3.2 is to correct θ (water content of S_1), which is a profile average SM representing a soil layer depth determined by calibration. By assuming a porosity of 0.46, (A-horizon information reported in McKenzie et al., 2000), and the model S_1 storage of 396 mm (420 mm) for the lumped (semi-distributed) scheme, this profile SM roughly represents the upper 1 m of the soil. On the other hand, the SSM observations represent only the few top centimetres of the soil column. To provide the model with information about more realistic dynamics of θ, we applied the exponential filter proposed by Wagner et al. (1999) to the SSM to estimate the soil wetness index (SWI) of the root-zone. SWI has been widely used to represent deeper layer SM based on SSM observations (e.g., Albergel et al., 2008; Brocca et al., 2009, 2010, 2012; Ford et al., 2014; Qiu et al., 2014). SWI was recursively

HESSD
11, 10635–10681, 2014

calculated as:

$$\mathrm{SWI}(t) = \mathrm{SWI}(t-1) + G_t[\mathrm{SSM}(t) - \mathrm{SWI}(t-1)], \quad (19)$$

where G_t is a gain term varying between 0 and 1 as:

$$G_t = \frac{G_{t-1}}{G_{t-1} + e^{-\left(\frac{t-(t-1)}{T}\right)}}. \quad (20)$$

T is a calibrated parameter that implicitly accounts for several physical parameters (Albergel et al., 2008). T was calibrated by maximising the correlation between SWI and θ during the first year of each SSM dataset. SWI was calculated independently for AMS, ASC and SMO datasets (named $\mathrm{SWI}_{\mathrm{AMS}}$, $\mathrm{SWI}_{\mathrm{ASC}}$ and $\mathrm{SWI}_{\mathrm{SMO}}$, respectively) and then rescaled to remove systematic differences with the model prediction (Sect. 3.6).

3.6 Rescaling and observation error estimation

The systematic differences (e.g., biases) between θ and the SWI derived from each satellite product must be removed prior to applying a bias-blind data assimilation scheme (Dee and Da Silva, 1998). We applied instrument variable (IV) regression to resolve the biases and estimate the measurement errors simultaneously (Su et al., 2014a). In three-data IV regression analysis, also known as triple collocation (TC) analysis (Stoffelen, 1998; Yilmaz and Crow, 2013), the model θ, the passive SWI and active SWI are used as data triplet. As sample size requirement for TC is stringent (Zwieback et al., 2012), a pragmatic threshold of 100 triplet sample was imposed (Scipal et al., 2008). During periods when only one satellite product was available (i.e., before ASC) or when the sample threshold for TC was not met, a two-data IV regression using lagged variables (LV) offers a practical substitute (Su et al., 2014a). Such LV analysis was performed on the model θ and a single satellite SWI, with the lagged variable coming from the model.

HESSD
11, 10635–10681, 2014

In most SM–DA experiments, the error in SSM has been treated as time-invariant (e.g., Reichle et al., 2008; Ryu et al., 2009; Crow and Van den Berg, 2010; Brocca et al., 2010, 2012; Alvarez-Garreton et al., 2014), however, studies evaluating SSM products have shown an important temporal variability in the measurement errors (Loew and Schlenz, 2011; Su et al., 2014a). Since a data assimilation scheme explicitly updates the model prediction based on the relative weights of the model and the observation errors, assuming a constant observation error may lead to over-correction of the model state if the actual error is higher, and vice versa.

Temporal characterisation of the observation error can be achieved by applying TC (or LV) to specific time windows of the observation and model prediction (for example, by grouping the triplets or doublets by month-of-the-year). There is however, a trade-off between the sampling window (which defines the temporal characterisation of the error) and the sample size (number of triplets in each subset). In an operational context this trade-off becomes more critical since only past observations are available. After analysing the temporal variability of the observation errors using the complete period of record (not shown here), we found that a 4 month sampling window can reproduce seasonality in errors while ensuring sufficient data samples for the TC and LV schemes. With this analysis we also assessed the suitability of using LV, which yielded similar results to TC although some positive bias in LV error variance estimates relative to TC was noted (not shown here).

Summarising, the procedure of rescaling and error estimation consisted of:

1. From the start of the AMS dataset, we grouped LV triplets ($SWI_{AMS}(t)$, $\theta(t)$ and $\theta(t-1)$) into three subsets: December–March, April–July and August–November.

2. We applied LV and thus, estimated the observation error and rescaling factors for a given four month subset only when a minimum of 100 samples was reached (after one year of AMS dataset). After the first year of AMS, new daily triplets were added into the corresponding 4 month data pool (retaining all earlier triplets) and LV was applied, on a daily basis, to the updated subset.

HESSD
11, 10635–10681, 2014

3. When ASC was available, LV triplets ($SWI_{ASC}(t)$, $\theta(t)$ and $\theta(t-1)$) subsets were formed following step 1 criteria and LV was applied after the 4 month data pools had more than 100 samples, following step 2.

4. In parallel with step 3, TC triplets were formed using the two available satellite datasets ($SWI_{AMS}(t)$, $SWI_{ASC}(t)$ and $\theta(t)$) and grouped into the four month subsets defined in step 1. TC was applied only when the 4 month data pools contained more than 100 samples (after approximately 3 years of ASC data).

5. Steps 3 and 4 were repeated when SMO was available. The triplets for TC in this case were given by $SWI_{ASC}(t)$, $SWI_{SMO}(t)$ and $\theta(t)$.

6. Once steps 1–5 were completed, a single time series of observations error and rescaling factors were constructed for each satellite-derived SWI by selecting TC results when available, and LV results if not. This criterion was adopted because LV is susceptible to bias due to auto-correlated errors in the model SM (Su et al., 2014a).

3.7 Evaluation metrics

To evaluate the SM–DA results, we used four different metrics. Firstly, the normalised root mean squared difference (NRMSE) was calculated as the ratio of the root mean square difference (RMSE) between the updated streamflow ensemble (Q_{sim}^{up}) and the observed streamflow to the RMSE between the open-loop (ensemble streamflow prediction without assimilation, Q_{sim}^{ol}) and the observed discharge:

$$\mathrm{NRMSE} = \frac{\frac{1}{N}\sum_{i=1}^{N}\sqrt{\sum_{t=1}^{T}\left(Q_{sim}^{up}(i,t) - Q_{obs}(t)\right)^2}}{\frac{1}{N}\sum_{i=1}^{N}\sqrt{\sum_{t=1}^{T}\left(Q_{sim}^{ol}(i,t) - Q_{obs}(t)\right)^2}}, \quad (21)$$

HESSD
11, 10635–10681, 2014

where $N = 1000$ is the number of ensemble members. The NRMSE provides information about both the spread of the ensemble and the performance the ensemble mean, which is considered as the best estimate of the ensemble prediction. Moreover, as it is calculated in the natural space, it gives more weight to high flows.

To further evaluate the performance of the ensemble mean, we calculated the Nash Sutcliffe efficiency (NS) for the entire evaluation period as follow (example for the open-loop case):

$$\mathrm{NS}_{\mathrm{ol}} = 1 - \frac{\sum_t \left(Q_{\mathrm{obs}}(t) - \overline{Q_{\mathrm{sim}}^{\mathrm{ol}}}(t) \right)^2}{\sum_t \left(Q_{\mathrm{obs}}(t) - \overline{Q_{\mathrm{obs}}} \right)^2}, \tag{22}$$

where $\overline{Q_{\mathrm{sim}}^{\mathrm{ol}}}$ is the open-loop ensemble mean. Similarly, $\mathrm{NS}_{\mathrm{up}}$ was calculated by applying Eq. (22) to the updated ensemble mean ($\overline{Q_{\mathrm{sim}}^{\mathrm{up}}}$).

We also estimated the probability of detection (POD) of daily flow rates (not flood events) exceeding minor, moderate and major floods, for the open-loop and the updated ensemble mean, as follows (example for the open-loop case):

$$\mathrm{POD}_{\mathrm{ol}} = \frac{\#\left(\overline{Q_{\mathrm{sim}}^{\mathrm{ol}}} \geq Q_{\mathrm{obs}}^{15.7\%} \text{ and } Q_{\mathrm{obs}} \geq Q_{\mathrm{obs}}^{15.7\%} \right)}{\#\left(Q_{\mathrm{obs}} \geq Q_{\mathrm{obs}}^{15.7\%} \right)}, \tag{23}$$

where the symbol # represents the number of times. $Q_{\mathrm{obs}}^{15.7\%}$ is the observed streamflow corresponding to a minor flood classification. This corresponds to a flow (not flood) frequency of 15.7 % (see Sect. 2). Similarly, $\mathrm{POD}_{\mathrm{up}}$ was calculated by applying Eq. (23) to the updated ensemble mean ($\overline{Q_{\mathrm{sim}}^{\mathrm{up}}}$). Finally, we estimated the false alarm ratio (FAR)

HESSD
11, 10635–10681, 2014

for daily flows as (example for the open-loop case):

$$\mathrm{FAR_{ol}} = \frac{\#\left(\overline{Q_{\mathrm{sim}}^{\mathrm{ol}}} \geq Q_{\mathrm{obs}}^{15.7\%} \text{ and } Q_{\mathrm{obs}} < Q_{\mathrm{obs}}^{15.7\%}\right)}{\#\left(Q_{\mathrm{obs}} < Q_{\mathrm{obs}}^{15.7\%}\right)}. \quad (24)$$

Similarly, $\mathrm{FAR_{up}}$ was calculated by applying Eq. (24) to the updated ensemble mean.

To evaluate the reliability of the streamflow ensemble prediction before and after SM–DA, we inspected the rank histograms of the ensemble following Anderson (1996). A reliable ensemble should have a uniform histogram while u-shape (n-shape) histogram indicates that the ensemble spread is too small (large) (De Lannoy et al., 2006).

The evaluation period of SM–DA was 1 June 2003–2 March 2014, which is independent of all scheme component calibration periods (see Sects. 3.1, 3.4 and 3.5).

4 Results and discussion

4.1 Model calibration

The streamflow at the outlet of the study catchment (N7 in Fig. 1) features long periods of zero-flow, a negligible baseflow component and sharp flow peaks after rainfall events, when the catchment has reached a threshold level of wetness (see observed streamflow in Fig. 4).

The simulated streamflow from the lumped and the semi-distributed schemes are presented in Fig. 4. To help visualisation of these time series, the calibration and evaluation periods were plotted separately. The evaluation period was further separated into two sub-periods, evaluation sub-period 1 (1 June 2003–30 April 2007), characterised by having only moderate and minor floods and evaluation sub-period 2 (1 May 2007–2 March 2014), which had three major flooding events. The plots show

HESSD
11, 10635–10681, 2014

that both the lumped and the semi-distributed models are generally able to capture the catchment hydrologic behaviour. As expected, the spatial distribution of forcing data and the channel routing accounted by the semi-distributed scheme enhanced the overall performance of the model, with lower residual values through time (Fig. 4a.2, b.2 and c.2) and consistently improved the simulation of peak flows. In the evaluation period, the NS at N7 for the lumped and the semi-distributed schemes was 0.67 and 0.77, respectively.

The good performance of the semi-distributed scheme at the catchment outlet, however, was not reflected at the inner catchments. The NS calculated at the inner gauges, N1 and N3 (not used in the calibration of the model parameters), resulted in NS of 0.28 and −1.89 in the evaluation period, respectively. Such poor performance may be attributed in part to errors in the spatially distributed rainfall dataset, but are more likely due to an inadequate model parametrisation scheme and errors in the model structure. It is however beyond of the scope of this work to address those specific factors, although they are considered in the evaluation of SM–DA experiments in the context of the expected benefits in ungauged sub-catchments.

To focus the evaluation on periods with flood events, the lag-correlation between the daily streamflow simulated at N7 and the daily rainfall (Fig. 5), and between daily streamflow and θ (Fig. 6), was calculated for daily streamflow values greater than $Q_{obs}^{15.7\%}$. The lumped scheme simulates a stronger (weaker) link between θ (rainfall) and streamflow than the semi-distributed scheme. This is represented by a higher (lower) r values in panel a compared to panels b–h in Fig. 6 (Fig. 5). These different representations of the catchment runoff response will have a direct impact on the skill of SM–DA to improve streamflow prediction. A strong relationship between θ and streamflow prediction suggests strong correlation between their errors, and therefore, greater potential improvement of streamflow resulting from an improved representation of θ.

If we assume that the semi-distributed scheme provides a better representation of runoff response within the entire catchment (based on its better model performance

HESSD
11, 10635–10681, 2014

at the outlet), Figs. 5 and 6 also suggest that daily rainfall is the main control on runoff generation and thus has a stronger impact in the streamflow prediction than soil moisture. Figure 6 shows that flood prediction strongly depends on antecedent soil moisture for up to the preceding 3 days. The strong correlation found at lag-0 suggests that the real time SM correction given by the proposed SM–DA would be a good strategy to improve flood prediction.

4.2 Error model parameters and ensemble prediction

The calibrated error parameters for the lumped and the semi-distributed schemes are $\sigma_p = 1.286$ mm and 0.977 mm; $\sigma_s = 0.099$ and 0.03 and $\sigma_k = 0.084$ and 0.018, respectively. σ_s is expressed as a percentage of the total storage capacity (396 mm in the lumped scheme and 420 mm in the semi-distributed scheme) and σ_k is expressed as a percentage of the calibrated parameter k_1.

The rank histograms of the generated ensemble prediction (open-loop) are presented in Fig. 7. The n-shaped histograms at the catchment outlet (N7), for both lumped and semi-distributed model schemes (Fig. 7a and b, respectively), suggests that the open-loop ensembles are unbiased, but feature wider spread than an ideal ensemble. The width of the spread will be critical for the evaluation of SM–DA (Sect. 4.4) since any decrease of the spread would be considered as an improvement of the ensemble prediction.

The wider spread of the open-loop ensembles at the catchment outlet can be due to factors such as an over-prediction of parameters by the MAP calibration algorithm, and the representation of the model error with time-constant error parameters. The latter becomes critical given the distinct behaviour of the intermittent streamflow response within the catchment, which could indicate a distinct behaviour in the model errors as well.

The ensemble prediction at N1 (Fig. 7c) features high bias while the ensemble prediction at N3 (Fig. 7d) presents a non-biased ensemble with adequate spread (note that observations at N1 and N3 were not used to calibrate the error parameters). The

HESSD
11, 10635–10681, 2014

large bias at N1 owes to the large errors in the calibrated model in SC1 (see Sect. 4.1). Given the highly non-linear nature of runoff mechanisms, unbiased perturbation of forcing, states or parameters can further increase the bias in streamflow prediction.

4.3 SWI estimation and rescaling

The SSM derived from AMS, ASC and SMO are presented in Fig. 8a, for the lumped model. The SSM datasets feature significantly higher noise than the modelled θ. This can be explained by factors such as random errors in the satellite retrievals (Su et al., 2014b), and the rapid variation of water content in the surface layer of soil due to infiltration and evapotranspiration losses. Figure 8b presents the SWI derived from the SSM products, after seasonal rescaling. This plot shows better agreement between model and observations due to SWI filtering/transformation, even when the higher noise in the rescaled SWI time series is still present.

Table 2 summarises the results of the SWI calibration and seasonal rescaling for the lumped model, showing the T parameter for each SWI and the correlation coefficient (r) between θ and SWI before and after rescaling (SWI^r). These results confirm the visual assessment of plots in Fig. 8 by showing an important increase in the linear correlation coefficient with θ when SSM is transformed into SWI. The correlation is further increased after rescaling, which illustrates that there is clear benefit from performing seasonal bias correction. Note that applying a constant rescaling factor would have no impact on on the correlation between θ and SWI^r.

The optimal T values (Table 2) are difficult to validate since there is no ground data to compare with and, it has been shown that they strongly depend on physical processes of the study site (Ceballos et al., 2005), thus direct comparison with other studies cannot be made reliably. Indeed, previous studies have shown a wide range of optimal T values for soil depths ranging between 10 and 100 cm. As an example, in Fig. 9 we summarised the optimal T found in 5 different studies.

There are some key theoretical issues that should be considered when using SWI as a profile SM estimator. Firstly, the parameter T in Eq. (20) was estimated by

HESSD
11, 10635–10681, 2014

maximising the correlation between SWI and θ, which could introduce cross-correlated errors between them. This would violate the IV regression assumption of no correlation between the errors among the triplets (Sect. 3.6). A way to overcome this issue, if data requirements are meet, would be to estimate a profile SM independently of the rainfall-runoff model prediction, for example by using a physically-based model to transfer surface SM into deeper layers (e.g., Richards, 1931; Beven and Germann, 1982; Manfreda et al., 2014).

Secondly, the SWI formulation explicitly incorporates autocorrelation terms, which would result in autocorrelated errors in the observation, which violates an EnKF assumption: independence between observation and prediction errors. The autocorrelation in the observation error can be transferred to the updated θ^+ during the SM–DA updating step. In that case, the θ^- background prediction error covariance at time $t+1$ would be correlated to the error of the new available SWI at time $t+1$. In contrast with the first issue listed above, the violation of the EnKF assumption can not be avoided by replacing SWI with a physically-based model, since the latter would result in profile SM strongly correlated with previous states as well. Indeed, given the physical mechanisms of water flux in the unsaturated soil, this problem will be present whenever a profile SM estimated from SSM is used as an observation in an EnKF-based DA framework. A way to overcome this could be to work with models that explicitly account for the water in the top few centimetres of soil and therefore can directly assimilate a (rescaled) SSM retrieval. However, the errors in SSM retrievals are probably already autocorrelated (Crow and Van den Berg, 2010).

Breaching some of the EnKF-based DA and the IV-based rescaling assumptions could in theory degrade the performance of the SM–DA scheme, when the variable analysed is directly soil moisture (Crow and Van den Berg, 2010; Reichle et al., 2008; Ryu et al., 2009). The performance of the SM–DA with respect to the improvement in streamflow however, seems to be less sensitive to these violated assumptions, as shown by Alvarez-Garreton et al. (2013, 2014). This highlights the need for further

HESSD
11, 10635–10681, 2014

studies focusing on SM–DA for the purposes of improving streamflow prediction from rainfall-runoff models.

4.4 Satellite soil moisture data assimilation

The ensemble predictions of streamflow and θ, before and after SM–DA, for the lumped and the semi-distributed schemes at N7, are presented in Fig. 10. It can be seen from this figure that the truncation bias correction (Sect. 3.3) was successful in creating a non-biased θ ensemble when the unperturbed model approached the soil water storage bounds (Fig. 10a.2 and b.2).

The rank histograms at N7, N1 and N3 are presented in Fig. 7. The ensemble predictions at the catchment outlet are more reliable for both the lumped (Fig. 7a) and the semi-distributed schemes after SM–DA (Fig. 7b). However, there is a slight tendency towards overestimation of the observed streamflow. The ensemble predictions at the "ungauged" inner catchments after SM–DA are reliable for SC1 (Fig. 7c) and biased for SC3 (Fig. 7d). These results are similar to the ones found in the ensemble prediction before SM–DA (Fig. 7c and d).

The evaluation statistics of the SM–DA are summarised in Table 3. The streamflow data of the inner catchments (N1 and N3) were used only for evaluation purposes in the semi-distributed scheme, therefore they are representative of "ungauged" inner catchments.

The NRMSE in Table 3 (all values below 1) demonstrates that the SM–DA was effective at reducing the streamflow prediction uncertainty (RMSE) across all gauged and ungauged catchments. The reductions in the RMSE ranged from 24 to 31 % for the different evaluation nodes. The NRMSE combines precision improvement (i.e., reduction of ensemble spread) with prediction accuracy improvement (i.e., enhancement of ensemble mean performance) resulting from the SM–DA. Given that the ensemble open-loop spread was larger than an ideal ensemble (based on the n-shaped rank histograms in Fig. 7), the reduction of the ensemble spread may be in part artificial.

HESSD
11, 10635–10681, 2014

The performance of the ensemble mean was assessed by computing the NS_{ol} and NS_{up} (Table 3). A first thing to note is that NS_{ol} presents a slight degradation of the unperturbed models (NS of 0.67 and 0.77 for the unperturbed lumped and semi-distributed schemes at N7, respectively, as presented in Sect. 4.1). This degradation was introduced in the ensemble generation process and is attributed to the highly non-linear runoff mechanisms represented by the model. The latter can be drawn since in the ensemble generation process (Sect. 3.3), the forcing data, the model states and the model parameters were perturbed with un-biased errors. The perturbation of forcing was un-biased and did not result in biased θ ensemble. The perturbation of parameter K_1 was unbiased. The θ perturbation was un-biased and, by applying the bias correction scheme (Ryu et al., 2009), we ensured that the generated θ ensemble was un-biased as well. However, the state perturbation bias correction by Ryu et al. (2009) was not able to completely remove biases in fluxes originating from non-linear model physics.

At the catchment outlet, the NS of the ensemble mean after SM–DA improved by 7 % and 38 %, for the lumped and semi-distributed models, respectively. At the ungauged catchments, SM–DA was also effective at improving the performance of the ensemble mean, compared to the open-loop. However, the performance of the model in those catchments was still poor (NS_{up} below 0.18). This can be explained by the systematic errors of the model on those catchments before assimilation, which were not addressed by the SM–DA.

The POD values at the catchment outlet (N7) show that before and after SM–DA, the model is consistently capable of detecting minor floods. Although this does not demonstrate an advantage of the SM–DA scheme proposed here, it does reflect the adequacy of the model ensemble prediction for simulating minor (and larger) floods. Consistently with previous results, the prediction of the semi-distributed model at the inner catchments is poorer in terms of detecting minor floods. The lower FAR values after SM–DA demonstrates the efficacy of the scheme in reducing the number of times

HESSD
11, 10635–10681, 2014

the model predicted an unobserved minor flood, at both the gauged and the ungauged catchments.

In summary, SM–DA was effective at improving streamflow ensemble predictions in the gauged and the ungauged catchments. The improvement was reflected by a reduction in the RMSE, an increase in the NS efficiency, and a decrease in the FAR. By accounting for rainfall spatial distribution and routing process within the large study catchment, we improved the model performance at the outlet compared to a lumped homogeneous scheme, which in turn improved the performance of the SM–DA. The latter was achieved even though the relation between θ and the streamflow prediction was weaker in the semi-distributed scheme (Fig. 6). The proposed SM–DA scheme therefore, has the merits of improving streamflow ensemble predictions by correcting the SM state of the model, even when rainfall appears to be the main driver of the runoff mechanism (see Sect. 4.1).

5 Conclusions

This paper presents an evaluation of the assimilation of passive and active satellite soil moisture observations (SM–DA) into a conceptual rainfall-runoff model (PDM) for the purpose of reducing flood prediction uncertainty in a sparsely monitored catchment. We set up the experiments in the large semi-arid Warrego River Basin ($> 40\,000\,km^2$) in south central Queensland, Australia. Within this context, we explore the advantages of accounting for the forcing data spatial distribution and the routing processes within the catchment.

The framework proposed here rigorously addressed the two main stages of a SM–DA scheme: model error representation and satellite data processing. We applied the different methods in the context of a sparsely monitored large catchment (i.e., limited data), under operational streamflow and flood forecasting scenarios (i.e., not future information is used in any of the presented methods).

HESSD
11, 10635–10681, 2014

The model error representation was the most critical step in the SM–DA scheme, since it determined the error covariance between observations and model state, and thus the potential efficacy of SM–DA. Moreover, the SM–DA evaluation was done against the open-loop ensemble prediction. The open-loop ensembles at the catchment outlet were un-biased, however, they featured sub-optimal (too large) spread. The latter was highlighted as the main limitation of the proposed scheme (Sect. 4.2), which had direct implications for the evaluation on the SM–DA results. Further exploration of model error representation (sources of error and the structure of those errors) and error parameter estimation should be explored to improve the characteristics of the open-loop ensemble prediction.

In the satellite data processing, we highlighted that the use of an exponential filter to transfer surface information into deeper layers may potentially lead to violation of some of TC and EnKF assumptions (Sect. 4.3). Possible solutions to overcome this would be to use more physically-based methods to transfer SSM into deeper layers or to use a rainfall-runoff model that explicitly accounts for the surface soil layer that can directly assimilate a (rescaled) SSM product. Both solutions however, are constrained by the ancillary data available for satisfactory implementation of a physically-based model. In the rescaling and error estimation procedure, we applied TC and LV to avoid error-in-variable biases. Applying these to correct biases in the SWI, showed improved agreement between observed and modelled SM.

The evaluation of the SM–DA results led to several insights. (1) The SM–DA was successful at improving the open-loop ensemble prediction at the catchment outlet, for both the lumped and the semi-distributed case. (2) Accounting for spatial distribution in the model forcing data and for the routing processes within the large study catchment improved the skill of the SM–DA at the catchment outlet. (3) The SM–DA was effective at improving streamflow prediction at the ungauged locations, compared to the open-loop. However, the updated prediction in those catchments was still poor, because the systematic errors before assimilation are not addressed by a SM–DA scheme.

HESSD
11, 10635–10681, 2014

This work provides new evidence of the efficacy of SM–DA to improve streamflow ensemble prediction in sparsely instrumented catchments. We demonstrate that SM–DA skill can be enhanced if the spatial distribution of forcing data and routing processes within the catchment are accounted for in large catchments. We show that SM–DA performance is directly related to the model quality before assimilation, therefore we recommend that efforts should be focused on ensuring adequate models, while evaluating the trade-offs between more complex models and data availability.

Acknowledgements. This research was conducted with financial support from the Australian Research Council (ARC Linkage Project No. LP110200520) and the Australian Bureau of Meteorology. C. Alvarez-Garreton was supported by Becas Chile scholarship.

References

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., and Martin, E.: From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations, Hydrol. Earth Syst. Sci., 12, 1323–1337, doi:10.5194/hess-12-1323-2008, 2008. 10649, 10650

Albergel, C., Rüdiger, C., Carrer, D., Calvet, J.-C., Fritz, N., Naeimi, V., Bartalis, Z., and Hasenauer, S.: An evaluation of ASCAT surface soil moisture products with in-situ observations in Southwestern France, Hydrol. Earth Syst. Sci., 13, 115–124, doi:10.5194/hess-13-115-2009, 2009. 10637

Albergel, C., Calvet, J.-C., de Rosnay, P., Balsamo, G., Wagner, W., Hasenauer, S., Naeimi, V., Martin, E., Bazile, E., Bouyssel, F., and Mahfouf, J.-F.: Cross-evaluation of modelled and remotely sensed surface soil moisture with in situ data in southwestern France, Hydrol. Earth Syst. Sci., 14, 2177–2191, doi:10.5194/hess-14-2177-2010, 2010. 10637

Albergel, C., de Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y., and Wagner, W.: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, Remote Sens. Environ., 118, 215–226, 2012. 10637

HESSD
11, 10635–10681, 2014

Alvarez-Garreton, C., Ryu, D., Western, A. W., Crow, W. T., and Robertson, D. E.: Impact of observation error structure on satellite soil moisture assimilation into a rainfall-runoff model, in: MODSIM2013, 20th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, 1–6 December 2013, Adelaide, Australia, edited by: Piantadosi, J., Anderssen, R., and Boland, J., 3071–3077, 2013. 10638, 10658

Alvarez-Garreton, C., Ryu, D., Western, A. W., Crow, W. T., and Robertson, D. E.: The impacts of assimilating satellite soil moisture into a rainfall-runoff model in a semi-arid catchment, J. Hydrol., doi:10.1016/j.jhydrol.2014.07.041, online first, 2014. 10638, 10639, 10646, 10651, 10658

Anderson, J. L.: A method for producing and evaluating probabilistic forecasts from ensemble model integrations, J. Climate, 9, 1518–1530, 1996. 10654

Beven, K. and Germann, P.: Macropores and water flow in soils, Water Resour. Res., 18, 1311–1325, 1982. 10658

Brocca, L., Melone, F., Moramarco, T., and Morbidelli, R.: Antecedent wetness conditions based on ERS scatterometer data, J. Hydrol., 364, 73–87, 2009. 10649

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., and Hasenauer, S.: Improving runoff prediction through the assimilation of the ASCAT soil moisture product, Hydrol. Earth Syst. Sci., 14, 1881–1893, doi:10.5194/hess-14-1881-2010, 2010. 10638, 10649, 10651

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., Dorigo, W., Matgen, P., Martínez-Fernández, J., Llorens, P., Latron, J., Martin, C., and Bitelli, M.: Soil moisture estimation through ASCAT and AMSR-E sensors: an intercomparison and validation study across Europe, Remote Sens. Environ., 115, 3390–3408, 2011. 10637, 10641

Brocca, L., Moramarco, T., Melone, F., Wagner, W., Hasenauer, S., and Hahn, S.: Assimilation of surface-and root-zone ASCAT soil moisture products into rainfall–runoff modeling, IEEE T. Geosci. Remote, 50, 2542–2555, 2012. 10638, 10646, 10649, 10651

Ceballos, A., Scipal, K., Wagner, W., and Martínez-Fernández, J.: Validation of ERS scatterometer-derived soil moisture data in the central part of the Duero Basin, Spain, Hydrol. Process., 19, 1549–1566, 2005. 10657

Chen, F., Crow, W. T., Starks, P. J., and Moriasi, D. N.: Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture, Adv. Water Resour., 34, 526–536, 2011. 10646

HESSD
11, 10635–10681, 2014

Chen, F., Crow, W., and Ryu, D.: Dual forcing and state correction via soil moisture assimilation for improved rainfall-runoff modeling, J. Hydrometeorol., doi:10.1175/JHM-D-14-0002.1, online first, 2014. 10638
Chipperfield, A. and Fleming, P.: The MATLAB genetic algorithm toolbox, in: Applied Control Techniques Using MATLAB, IEE Colloquium on, 26 January 1995, London, UK, 10/1–10/4, doi:10.1049/ic:19950061, 1995. 10642
Crow, W. T. and Reichle, R. H.: Comparison of adaptive filtering techniques for land surface data assimilation, Water Resour. Res., 44, W08423, doi:10.1029/2008WR006883, 2008. 10637
Crow, W. T. and Ryu, D.: A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals, Hydrol. Earth Syst. Sci., 13, 1–16, doi:10.5194/hess-13-1-2009, 2009. 10638
Crow, W. T. and Van den Berg, M. J.: An improved approach for estimating observation and model error parameters in soil moisture data assimilation, Water Resour. Res., 46, W12519, doi:10.1029/2010WR009402, 2010. 10637, 10646, 10651, 10658
Crow, W. T. and van Loon, E.: Impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture, J. Hydrometeorol., 7, 421–432, 2006. 10637
De Lannoy, G. J., Houser, P. R., Pauwels, V., and Verhoest, N. E.: Assessment of model uncertainty for soil moisture through ensemble verification, J. Geophys. Res.-Atmos., 111, D10101, doi:10.1029/2005JD006367, 2006. 10654
Dee, D. P. and Da Silva, A. M.: Data assimilation in the presence of forecast bias, Q. J. Roy. Meteor. Soc., 124, 269–295, 1998. 10650
Draper, C. S., Walker, J. P., Steinle, P. J., de Jeu, R. A., and Holmes, T. R.: An evaluation of AMSR–E derived soil moisture over Australia, Remote Sens. Environ., 113, 703–710, 2009. 10637, 10641
Evensen, G.: The ensemble Kalman filter: theoretical formulation and practical implementation, Ocean Dynam., 53, 343–367, 2003. 10638, 10643
Ford, T. W., Harris, E., and Quiring, S. M.: Estimating root zone soil moisture using near-surface observations from SMOS, Hydrol. Earth Syst. Sci., 18, 139–154, doi:10.5194/hess-18-139-2014, 2014. 10649
Francois, C., Quesney, A., and Ottlé, C.: Sequential assimilation of ERS-1 SAR data into a coupled land surface-hydrological model using an extended Kalman filter, J. Hydrometeorol., 4, 473–487, 2003. 10638

HESSD
11, 10635–10681, 2014

Gill, M. A.: Flood routing by the Muskingum method, J. Hydrol., 36, 353–363, 1978. 10642
Gruhier, C., de Rosnay, P., Hasenauer, S., Holmes, T., de Jeu, R., Kerr, Y., Mougin, E., Njoku, E., Timouk, F., Wagner, W., and Zribi, M.: Soil moisture active and passive microwave products: intercomparison and evaluation over a Sahelian site, Hydrol. Earth Syst. Sci., 14, 141–156, doi:10.5194/hess-14-141-2010, 2010. 10637
Hain, C. R., Crow, W. T., Anderson, M. C., and Mecikalski, J. R.: An ensemble Kalman filter dual assimilation of thermal infrared and microwave satellite observations of soil moisture into the Noah land surface model, Water Resour. Res., 48, W11517, doi:10.1029/2011WR011268, 2012. 10646
Jones, D. A., Wang, W., and Fawcett, R.: High-quality spatial climate data-sets for Australia, Australian Meteorological and Oceanographic Journal, 58, 233–248, 2009. 10640
Li, Y., Ryu, D., Western, A. W., Wang, Q., Robertson, D. E., and Crow, W. T.: An integrated error parameter estimation and lag-aware data assimilation scheme for real-time flood forecasting, J. Hydrol., doi:10.1016/j.jhydrol.2014.08.009, online first, 2014. 10648
Liu, Y. Q. and Gupta, H. V.: Uncertainty in hydrologic modeling: toward an integrated data assimilation framework, Water Resour. Res., 43, W07401, doi:10.1029/2006WR005756, 2007. 10645
Loew, A. and Schlenz, F.: A dynamic approach for evaluating coarse scale satellite soil moisture products, Hydrol. Earth Syst. Sci., 15, 75–90, doi:10.5194/hess-15-75-2011, 2011. 10651
Manfreda, S., Brocca, L., Moramarco, T., Melone, F., and Sheffield, J.: A physically based approach for the estimation of root-zone soil moisture from surface measurements, Hydrol. Earth Syst. Sci., 18, 1199–1212, doi:10.5194/hess-18-1199-2014, 2014. 10658
McKenzie, N. J., Jacquier, D., Ashton, L., and Cresswell, H.: Estimation of Soil Properties Using the Atlas of Australian Soils, CSIRO Land and Water, Canberra, 2000. 10649
McMillan, H., Jackson, B., Clark, M., Kavetski, D., and Woods, R.: Rainfall uncertainty in hydrological modelling: an evaluation of multiplicative error models, J. Hydrol., 400, 83–94, 2011. 10645
Moore, R. J.: The PDM rainfall-runoff model, Hydrol. Earth Syst. Sci., 11, 483–499, doi:10.5194/hess-11-483-2007, 2007. 10641, 10642
Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S., and Wagner, W.: An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations, IEEE T. Geosci. Remote, 47, 1999–2013, 2009. 10640

HESSD
11, 10635–10681, 2014

Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I: A discussion of principles, J. Hydrol., 10, 282–290, 1970. 10642

Owe, M., de Jeu, R., and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, J. Geophys. Res.-Earth, 113, F01002, doi:10.1029/2007JF000769, 2008. 10640

Plaza, D. A., De Keyser, R., De Lannoy, G. J. M., Giustarini, L., Matgen, P., and Pauwels, V. R. N.: The importance of parameter resampling for soil moisture data assimilation into hydrologic models using the particle filter, Hydrol. Earth Syst. Sci., 16, 375–390, doi:10.5194/hess-16-375-2012, 2012. 10647

Qiu, J., Crow, W. T., Nearing, G. S., Mo, X., and Liu, S.: The impact of vertical measurement depth on the information content of soil moisture times series data, Geophys. Res. Lett., 41, 4997–5004, 2014. 10649

Reichle, R. H., Crow, W. T., and Keppenne, C. L.: An adaptive ensemble Kalman filter for soil moisture data assimilation, Water Resour. Res., 44, W03423, doi:10.1029/2007WR006357, 2008. 10637, 10646, 10651, 10658

Richards, L. A.: Capillary conduction of liquids through porous mediums, Physics, 1, 318–333, 1931. 10658

Ryu, D., Crow, W. T., Zhan, X., and Jackson, T. J.: Correcting unintended perturbation biases in hydrologic data assimilation, J. Hydrometeorol., 10, 734–750, 2009. 10637, 10647, 10651, 10658, 10660

Scipal, K., Holmes, T., De Jeu, R., Naeimi, V., and Wagner, W.: A possible solution for the problem of estimating the error structure of global soil moisture data sets, Geophys. Res. Lett., 35, L24403, doi:10.1029/2008GL035599, 2008. 10650

Stoffelen, A.: Toward the true near-surface wind speed: error modeling and calibration using triple collocation, J. Geophys. Res.-Oceans, 103, 7755–7766, 1998. 10650

Su, C.-H., Ryu, D., Young, R. I., Western, A. W., and Wagner, W.: Inter-comparison of microwave satellite soil moisture retrievals over the Murrumbidgee Basin, southeast Australia, Remote Sens. Environ., 134, 1–11, 2013. 10637

Su, C., Ryu, D., Crow, W. T., and Western, A. W.: Beyond triple collocation: applications to soil moisture monitoring, J. Geophys. Res.-Atmos., 119, 6416–6439, 2014a. 10650, 10651, 10652

HESSD

11, 10635–10681, 2014

Su, C.-H., Ryu, D., Crow, W. T., and Western, A. W.: Stand-alone error characterisation of microwave satellite soil moisture using a Fourier method, Remote Sens. Environ., 154, 115–126, 2014b. 10657

Thielen, J., Bartholmes, J., Ramos, M.-H., and de Roo, A.: The European Flood Alert System – Part 1: Concept and development, Hydrol. Earth Syst. Sci., 13, 125–140, doi:10.5194/hess-13-125-2009, 2009. 10637

Tian, Y., Huffman, G. J., Adler, R. F., Tang, L., Sapiano, M., Maggioni, V., and Wu, H.: Modeling errors in daily precipitation measurements: additive or multiplicative?, Geophys. Res. Lett., 40, 2060–2065, 2013. 10645

Wagner, W., Lemoine, G., and Rott, H.: A method for estimating soil moisture from ERS scatterometer and soil data, Remote Sens. Environ., 70, 191–207, 1999. 10649

Wanders, N., Karssenberg, D., de Roo, A., de Jong, S. M., and Bierkens, M. F. P.: The suitability of remotely sensed soil moisture for improving operational flood forecasting, Hydrol. Earth Syst. Sci., 18, 2343–2357, doi:10.5194/hess-18-2343-2014, 2014. 10638

Wang, Q., Robertson, D., and Chiew, F.: A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites, Water Resour. Res., 45, W05407, doi:10.1029/2008WR007355, 2009. 10648

Western, A. W., Grayson, R. B., and Blöschl, G.: Scaling of soil moisture: a hydrologic perspective, Annu. Rev. Earth Pl. Sc., 30, 149–180, 2002. 10637

Yilmaz, M. T. and Crow, W. T.: The optimality of potential rescaling approaches in land data assimilation, J. Hydrometeorol., 14, 650–660, 2013. 10650

Zwieback, S., Scipal, K., Dorigo, W., and Wagner, W.: Structural and statistical properties of the collocation technique for error characterization, Nonlin. Processes Geophys., 19, 69–80, doi:10.5194/npg-19-69-2012, 2012. 10650

HESSD

11, 10635–10681, 2014

Table 1. Area and mean annual rainfall (MAR) of the catchments used in the lumped and semi-distributed schemes.

Catchment	Area (km^2)	MAR (mm)
SC1	14 670	492
SC2	4453	532
SC3	8070	596
SC4	5431	524
SC5	4067	503
SC6	2130	467
SC7	4049	418
Total	42 870	512

HESSD
11, 10635–10681, 2014

Table 2. Summary of SSM and SWI for the entire catchment.

Dataset	T (days)	r between θ and SSM	SWI	SWI^r
AMS	3	0.65	0.74	0.94
ASC	11	0.77	0.92	0.97
SMO	40	0.46	0.79	0.93

HESSD
11, 10635–10681, 2014

Table 3. SM–DA evaluation statistics calculated at the catchment outlet (N7) and at the inner catchments (N1 and N3).

Statistic	Lumped scheme (N7)	Semi-distributed scheme (N7)	(N1)	(N3)
NRMSE	0.73	0.69	0.76	0.75
NS_{ol}	0.61	0.53	−0.02	−5.36
NS_{up}	0.65	0.73	0.18	−2.47
POD_{ol}	1.00	0.99	0.61	0.74
POD_{up}	0.99	0.98	0.58	0.71
FAR_{ol}	0.20	0.20	0.13	0.17
FAR_{up}	0.17	0.15	0.10	0.14

HESSD
11, 10635–10681, 2014

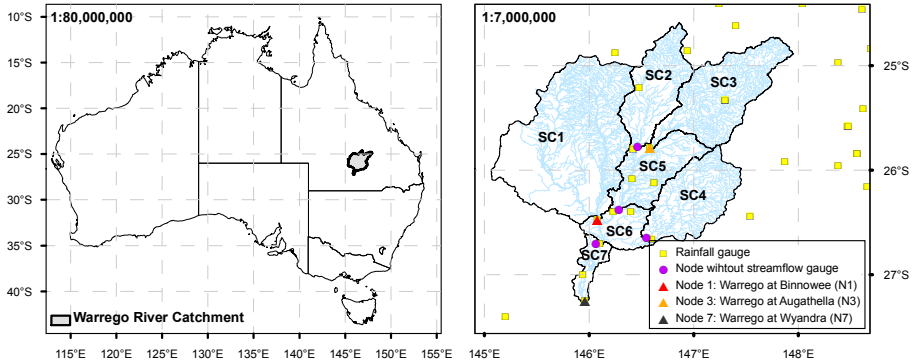

Figure 1. The Warrego river basin located in Queensland, Australia (left panel). A close-up of the area is presented on the right panel. The lumped PDM scheme is set up over the entire catchment, while the semi-distributed scheme divides the total catchment in 7 sub-catchments (SC1–SC7).

HESSD
11, 10635–10681, 2014

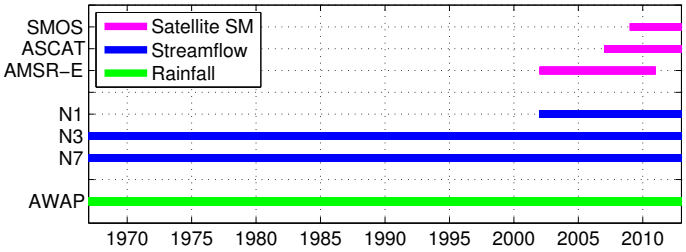

Figure 2. Period of record of the different datasets. The initial date of the plot was set as the beginning of the streamflow data record.

HESSD
11, 10635–10681, 2014

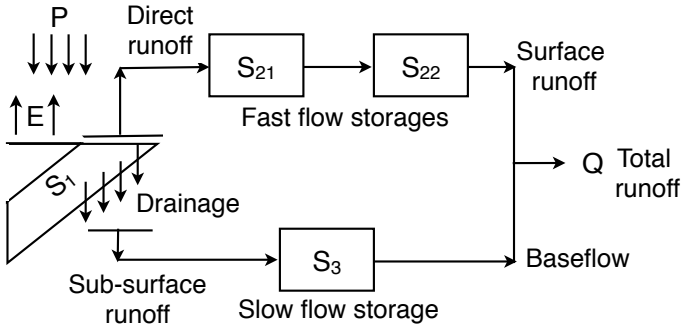

Figure 3. The PDM scheme.

HESSD
11, 10635–10681, 2014

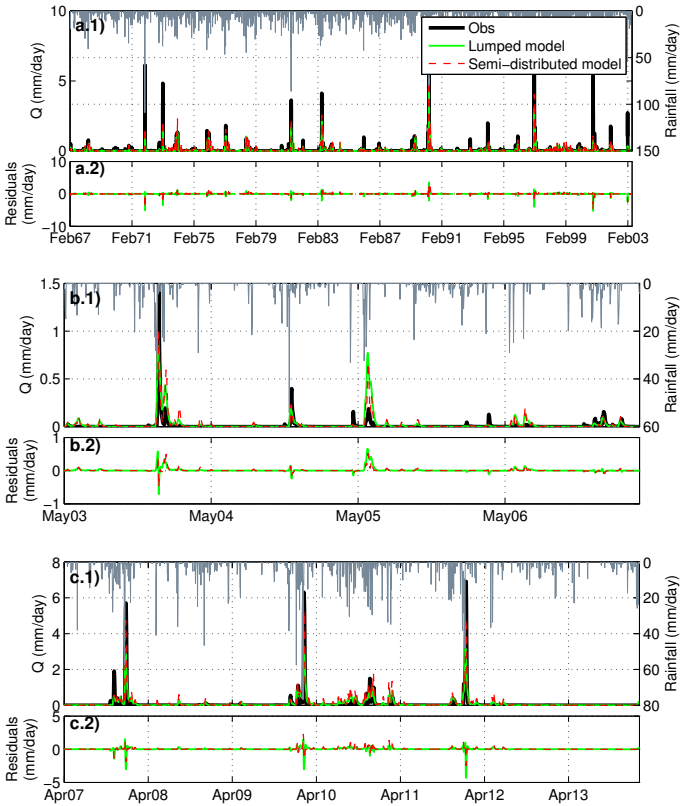

Figure 4. Simulated, observed daily streamflow (Q) and model streamflow prediction residuals (simulated minus observed) at the catchment outlet (N7). **(a.1)** and **(a.2)** present the calibration period. **(b.1)** and **(b.2)** present the evaluation sub-period 1, which has only moderate and minor flood events. **(c.1)** and **(c.2)** present the evaluation sub-period 2, which has 3 major flood events. The daily rainfall plotted in the right axis corresponds to the averaged rainfall over the entire catchment.

HESSD
11, 10635–10681, 2014

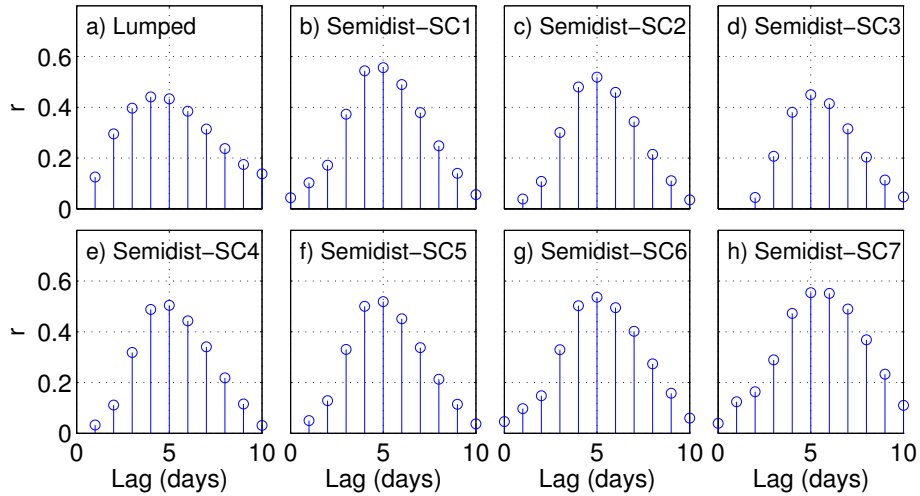

Figure 5. Lag-correlation coefficient (r) between the simulated streamflow at N7 (mm day^{-1}), and the daily rainfall (mm) of the entire catchment **(a)** and the 7 sub-catchments **(b–h)**.

HESSD
11, 10635–10681, 2014

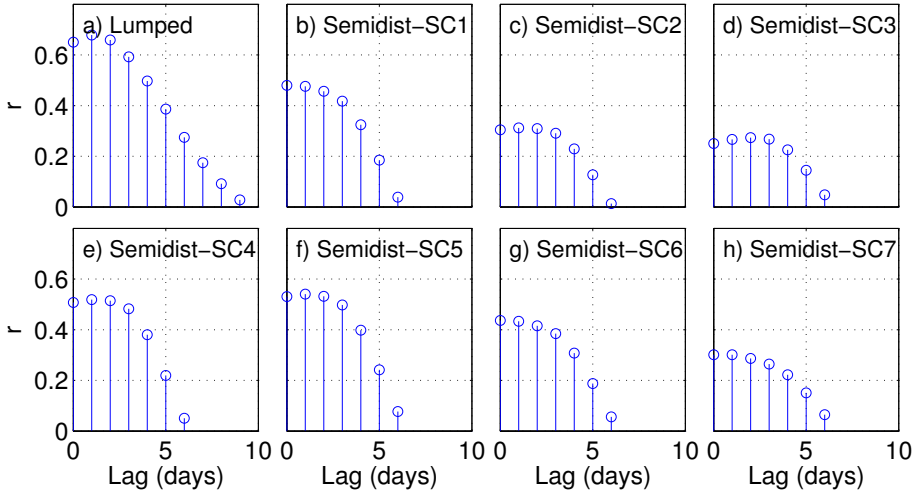

Figure 6. Lag-correlation coefficient (r) between the simulated streamflow at N7 (mm day^{-1}), and θ from the lumped **(a)** and the semi-distributed **(b–h)** model schemes.

HESSD
11, 10635–10681, 2014

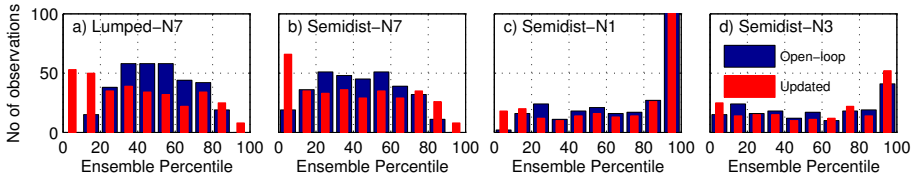

Figure 7. Rank histograms of the open-loop and updated streamflow ensemble predictions. **(a)** presents the results from the lumped scheme at node N7. **(b–d)** present the results from the semi-distributed (semidist) scheme at nodes N7, N1 and N3.

HESSD
11, 10635–10681, 2014

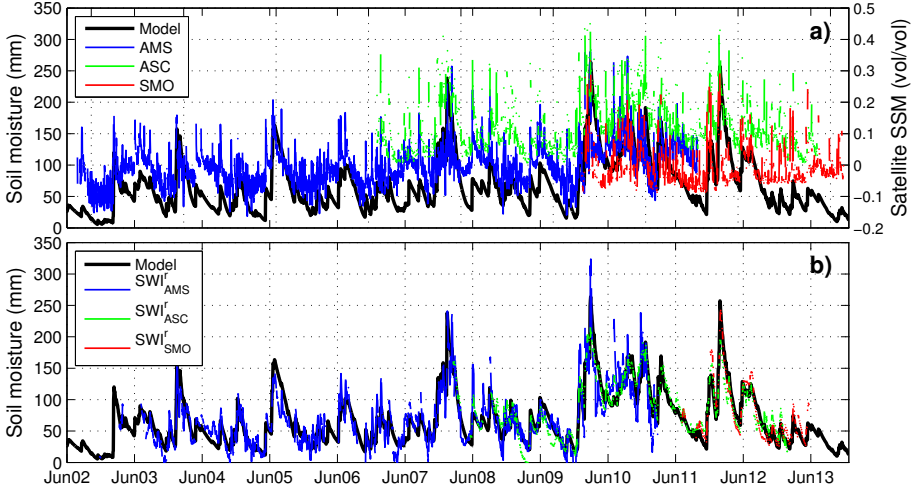

Figure 8. (a) shows the model soil water content in the left axis and the satellite soil moisture (SSM) observations in the right axis. **(b)** shows the soil moisture in the model space, after the three SSM datasets were transformed into a soil wetness index (SWI) and then rescaled by using TC or LV (SWI^r_{AMS}, SWI^r_{ASC} and SWI^r_{SMO}).

HESSD
11, 10635–10681, 2014

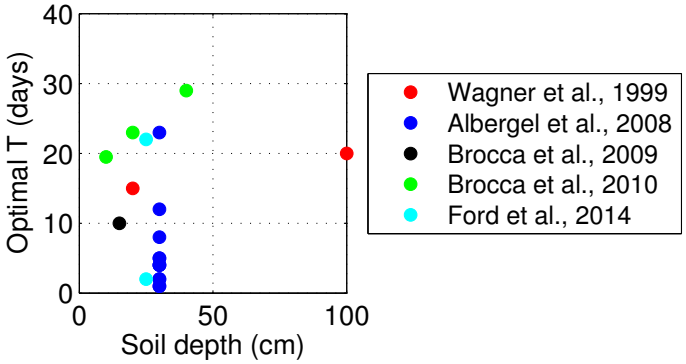

Figure 9. Optimal T parameter against soil depth found in previous studies.

HESSD
11, 10635–10681, 2014

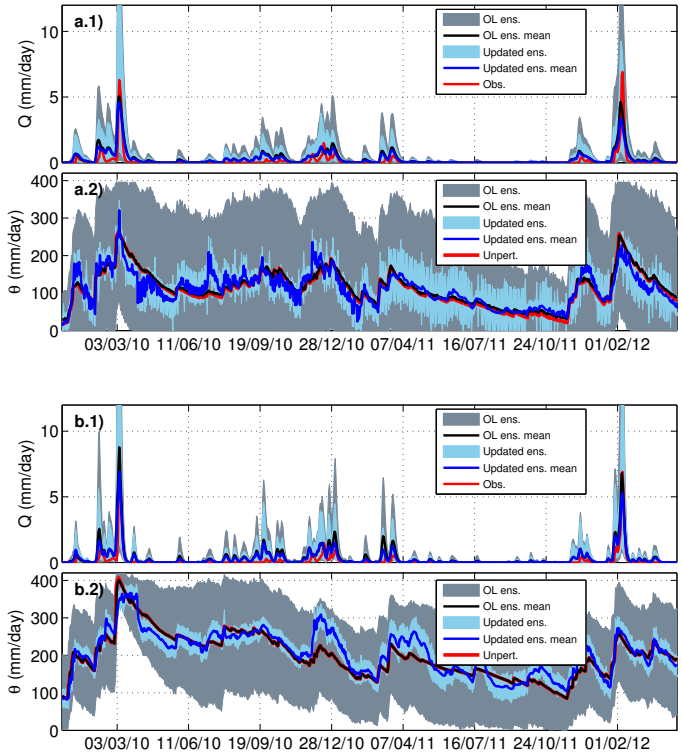

Figure 10. Streamflow (Q) and soil moisture (θ) ensemble prediction at the catchment outlet, before and after SM–DA. **(a.1)** and **(a.2)** present the results for the lumped model. **(b.1)** and **(b.2)** present the results for the semi-distributed model.

HESSD
11, 10635–10681, 2014