Hydrol. Earth Syst. Sci. Discuss., 12, 2267–2304, 2015
www.hydrol-earth-syst-sci-discuss.net/12/2267/2015/
doi:10.5194/hessd-12-2267-2015

Hydrology and Earth System Sciences
Open Access
Discussions

Data assimilation in integrated hydrological modeling using ensemble Kalman filtering: evaluating the effect of ensemble size and localization on filter performance

J. Rasmussen[1], H. Madsen[2], K. H. Jensen[1], and J. C. Refsgaard[3]

[1]Department of Geosciences and Natural Resource Management, University of Copenhagen, Denmark
[2]DHI, Hørsholm, Denmark
[3]Geological Survey of Denmark and Greenland, Copenhagen, Denmark

Received: 26 January 2015 – Accepted: 2 February 2015 – Published: 23 February 2015

Correspondence to: J. Rasmussen (jr@ign.ku.dk)

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

Abstract

Groundwater head and stream discharge is assimilated using the Ensemble Transform Kalman Filter in an integrated hydrological model with the aim of studying the relationship between the filter performance and the ensemble size. In an attempt to reduce the required number of ensemble members, an adaptive localization method is used. The performance of the adaptive localization method is compared to the more common local analysis localization. The relationship between filter performance in terms of hydraulic head and discharge error and the number of ensemble members is investigated for varying numbers and spatial distributions of groundwater head observations and with or without discharge assimilation and parameter estimation. The study shows that (1) more ensemble members are needed when fewer groundwater head observations are assimilated, and (2) assimilating discharge observations and estimating parameters requires a much larger ensemble size than just assimilating groundwater head observations. However, the required ensemble size can be greatly reduced with the use of adaptive localization, which by far outperforms local analysis localization.

1 Introduction

Data assimilation (DA) is frequently used in hydrological modelling for correcting errors in the models. Stemming from parameter uncertainty, model structure uncertainty, uncertainty in forcing data and boundary condition uncertainty, the errors can lead to significant bias in the model states. Data assimilation can help reduce the bias in the model sequentially, leading to improved predictive capabilities of the model. It is also commonly used for history matching, for quantifying uncertainty and for estimation of model parameters.

Application of data assimilation for state updating in hydrological modelling has been studied extensively using a number of different models with most models focusing only on a part of the hydrological cycle. These include groundwater models (e.g. Hendricks

HESSD
12, 2267–2304, 2015

Franssen et al., 2011), land-surface models (e.g. Albergel et al., 2008), rainfall–runoff models (e.g. Moradkhani et al., 2005) and others. A few studies have also used more integrated hydrological models in conjunction with assimilation of multiple types of observations, but the subject is still in its infancy as a research topic. Studies that focused on integrated hydrological modeling include Camporese et al. (2009) and Shi et al. (2014). Camporese et al. (2009) applied the Ensemble Kalman filter (EnKF) to a coupled surface–subsurface model of a synthetic tilted v-catchment and assimilated both stream discharge and groundwater hydraulic head observations with the aim of updating both groundwater and stream states. Shi et al. (2014) applied the EnKF to a coupled physically-based land surface hydrological model of a small catchment. Using observations of seven different model states ranging from discharge to transpiration, they successfully estimated six parameters pertaining to different processes in the model while simultaneously updating the model states.

Using data assimilation for parameter estimation has become common in hydrological modelling (Moradkhani et al., 2005; Vrugt et al., 2005; Hendricks Franssen and Kinzelbach, 2008; Nie et al., 2008) due to the importance of parameter uncertainty in hydrological models. Notably, Hendricks Franssen and Kinzelbach (2008) used the augmented state vector approach to update a spatially distributed groundwater hydraulic conductivity field in a groundwater model. As previously stated, Shi et al. (2014) also successfully estimated several parameters in their coupled surface–subsurface hydrological model.

The effects of observation densities and patterns on parameter estimation in hydrological modeling have been studied using a number of hydrological models and inverse modeling methods. Juston et al. (2009) studied the effect of varying the observation intervals of groundwater head and stream discharge for calibration of a hydrological model of a small catchment. They found that even relatively sparse observation subsets can provide similar restraints to the model parameters as complete (frequent) observation sets, as long as significant hydrological events are represented by the data. The effect of differing observation densities and assimilation/updating frequencies of

HESSD
12, 2267–2304, 2015

hydraulic head observations in a groundwater model was also studied by Hendricks Franssen and Kinzelbach (2008). Experimenting with 3 and 28 observation points, respectively, and updating frequencies of 1 month^{-1} and 5 month^{-1}, they found the observation point density to have the largest effect on filter performance. However, no in-depth analysis of the subject was performed in their study.

Spurious correlations in EnKF arise when the correlation cannot be properly described by the ensemble of models, and have a detrimental effect on the filter performance. Localization is a commonly used method for reducing these spurious correlations, and as such has been the subject of several studies (Anderson, 2007; Hunt et al., 2007; Sakov and Bertino, 2011). Applying localization often allows for the use of a significantly reduced ensemble size, making it particularly useful for computationally heavy models, as a means for reducing the required computational time. Several localization methods exist, with distance-based methods being the most common (Sakov and Bertino, 2011; Ott et al., 2004; Fertig et al., 2007). Distance-based methods, such as Local Analysis, specify the area of influence of an observation based on spatial distance and often removes or reduces correlation between observations and model states beyond a user-specified distance. Alternatively, several adaptive localization methods have been developed (Anderson, 2007; Bishop and Hodyss, 2009), that attempt to distinguish real correlation from spurious correlation, making them particularly useful if distance-based localization is problematic, for example due to model structure.

This study investigates the relationship between ensemble size and number of observations with filter performance in a catchment size coupled surface–subsurface model with the objective of assessing an optimal ensemble size. Furthermore, a new approach to adaptive localization is used and compared to Local Analysis localization and the possible benefits of applying adaptive localization with different ensemble sizes and groundwater head observation densities are evaluated. The study is performed using a synthetic test setup of a catchment located in Denmark, and includes the application of parameter estimation and assimilation of both groundwater head and stream

HESSD
12, 2267–2304, 2015

discharge observations. The novelty of the study lies in the extensive study in the relationship between the observation density and the required ensemble size, as well as in the application of adaptive localization, neither of which, despite potentially having big impact on the filter performance, has previously been investigated in detail for application in integrated hydrological models.

2 Methods

2.1 Model

The hydrological model used in this study is a transient, spatially distributed model based on the MIKE SHE model code (Graham and Butts, 2005). This code allows for an integrated approach to hydrological modeling, in which all the major hydrological processes are included, comprising feedback between the processes. As such, it is a good platform for investigating the assimilation of multiple observation types in hydrological models, as well as estimation of parameters related to different hydrological processes.

2.2 Study area

2.2.1 The Karup catchment

The Karup catchment, which is located in the northern part of the Jutland peninsula in Denmark, forms the basis for this study. The catchment has a size of 440 km^2 and its land use is primarily agriculture. The geology of the catchment is dominated by quaternary sand. The catchment is very flat, with a south–north slope ranging from 93 m a.s.l. in the southern part to 22 m a.s.l. in the northern part. The main drainage feature of the catchment is the Karup river, which springs at the southern edge of the catchment and runs from southeast to northwest, and is joined by seven tributaries

HESSD
12, 2267–2304, 2015

(Fig. 1). The stream is strongly groundwater dominated, meaning that the interaction between surface water and groundwater is very strong.

2.2.2 Model setup

An integrated approach to modeling of the catchment is used in this study, which includes modelling of the groundwater flow, vadose zone flow, stream flow, surface flow and evapotranspiration. Surface-, stream-, vadose zone- and groundwater flows are coupled dynamically, allowing water to be exchanged between the modules at each time step.

Modeling of the groundwater is done using a finite difference approximation of the governing 2-D Boussinesq equation, which is coupled to a one-dimensional and vertical description of unsaturated flow using the gravity flow formulation of unsaturated flow (Graham and Butts, 2005). Evapotranspiration is modeled using the Kristensen and Jensen (1975) model. Stream flow is modeled using the MIKE 11 river model with a kinematic routing description.

A horizontal grid size of 1 km × 1 km is used, with the vertical discretization of the unsaturated zone gradually increasing from 0.05 m at the top to 1 m below a depth of 10 m.

2.2.3 Model parameterization

A 3-D geological model delineating six geological units forms the basis for the spatial distribution of hydrogeological parameters. Meltwater sand is the dominating geological unit, and five lenses (clay, quartz sand, mica Sand, mica clay/silt, and limestone) of varying extent make up the remaining geology. The parameter values specified in the geological model are in a preprocessing step interpolated and gridded to the horizontal 2-D computational grid to ease the computational requirements of the model. The parameters for the groundwater zone are: hydraulic conductivity (horizontal and vertical, respectively), specific yield and specific storage for the six units.

HESSD
12, 2267–2304, 2015

The parameterization of the unsaturated zone is spatially distributed and is based on texture data classified into nine soil types (Greve et al., 2007). These range from coarse sandy soil (Soil type 1) to heavy clayey soil (Soil type 8), and also includes organic soil (Soil type 11). The dominating soil type is Soil type 1, which accounts for approximately 90 % of the catchment. The parameters of the unsaturated zone are the saturated and residual moisture contents, Saturated hydraulic conductivity as well as soil matric potentials at field capacity and at wilting point.

Land use is based on data from local authorities and divided into four classes: agriculture (56 %), forest (18 %), heath (18 %) and wetlands (7 %). Forest and heath are described using constant values for the land use related parameters Leaf Area Index (LAI) and Root Depth (RD), while LAI and RD of agricultural land are seasonally dependent, divided into a growing season and a non-growing season.

Parameterization of the stream discharge model is done in a non-distributed manner, with each branch (the Karup river and each of the seven tributaries) having the same parameter values. The parameters of the stream discharge model are the Manning number, the drain level and the drain time constant describing the drainage in the wetland areas near the river, and the leakage coefficient describing the river-aquifer interaction.

2.3 Data assimilation

2.3.1 Ensemble Transform Kalman Filter

This study uses the Ensemble Transform Kalman Filter (ETKF) (Bishop et al., 2001), which is a computationally efficient implementation of the Ensemble Kalman Filter. The ETKF is also deterministic and does not require a full error covariance matrix to be generated, which makes it computationally less demanding. The practical implementation of the ETKF in this study is based on that of Harlim and Hunt (2005).

The $m \times n$ matrix, $\mathbf{X}^f$, is a forecasted ensemble of state variables (Groundwater hydraulic head, stream discharge and stream water level) composed of n numbers

HESSD
12, 2267–2304, 2015

of $1 \times m$ vectors containing the state variables of the respective ensemble members, where n is the number of ensemble members and m is the number of state variables. It is structured as

$$\mathbf{X}^f = \left[x_1^f, \ldots, x_n^f\right]. \quad (1)$$

A $k \times n$ matrix $\mathbf{Y}^f$ of model observations (k is the number of observations) is formed by applying a linear operator H that maps the state space into observation space to each column of $\mathbf{X}^f$. This matrix is averaged over the columns to form a $k \times 1$ vector of mean model observations, $\overline{\boldsymbol{y}}^f$, which is then columnwise subtracted from $\mathbf{Y}^f$ to form the $k \times n$ matrix of model observation anomalies $\mathbf{Y}^b$. Next, $\mathbf{X}^f$ is averaged columnwise to form the $m \times 1$ vector of mean model states $\overline{\boldsymbol{x}}^b$ and this vector is subtracted from each column of $\mathbf{X}^f$ to create a $m \times n$ matrix of model state anomalies $\mathbf{X}^b$.

A $n \times k$ matrix, $\mathbf{C}$, is computed as follows:

$$\mathbf{C} = (\mathbf{Y}^b) \cdot \mathbf{R}^{-1} \cdot \mathbf{P}_{\text{obs}}, \quad (2)$$

where $\mathbf{R}$ is a $k \times k$ matrix of observation covariance, and $\mathbf{P}_{\text{obs}}$ is a $k \times k$ diagonal matrix with the localization weights of each observation on the diagonal. The $k \times k$ error covariance matrix is computed by

$$\widetilde{\mathbf{P}}^a = [(k-1) \cdot \mathbf{I} + \mathbf{C}\mathbf{Y}^b]^{-1}, \quad (3)$$

where $\mathbf{I}$ is a $k \times k$ identity matrix. The $k \times k$ matrix of analysis error covariance is computed as

$$\mathbf{W}^a = [(k-1) \cdot \widetilde{\mathbf{P}}^a]^{1/2}. \quad (4)$$

The $k \times 1$ matrix of updating weights, $\boldsymbol{w}^a$, is computed as

$$\boldsymbol{w}^a = \widetilde{\mathbf{P}}^a \mathbf{C} \cdot (\boldsymbol{y} - \overline{\boldsymbol{y}}^b), \quad (5)$$

HESSD
12, 2267–2304, 2015

where $\boldsymbol{y}$ is a $k \times 1$ vector of observations and $\overline{\boldsymbol{y}}^b$ is a $k \times 1$ vector of the mean model observations. $\boldsymbol{w}^a$ is then added each column of $\mathbf{W}^a$, forming the $k \times k$ matrix of updated error covariance $\mathbf{W}$. Finally, the updated ensemble of model states is calculated:

$$\mathbf{X}^u = \mathbf{X}^b \mathbf{W}. \tag{6}$$

2.3.2 Localization

This study uses an adaptive localization method that is a combination of two separate adaptive localization methods proposed by Anderson (2001) and Bishop and Hodyss (2007) respectively. Anderson (2007) proposed to split the ensemble into a number of sub-ensembles, and for each sub-ensemble calculate the correlation coefficients between the state variables and the model observations. The correlation coefficients for each sub-ensemble are then cross-validated, and the observations are given a localization weight based on the cross-validation. That means that observations where the sub-ensembles agree on the correlation coefficient are given a high localization weight, as opposed to points in which there is disagreement between the sub-ensembles. Bishop and Hodyss (2007) proposed to calculate the sample correlation coefficient of the entire ensemble, and simply raising it to a power. The localization weight of an observation then equals the power of the correlation coefficient, giving observations with higher correlation coefficients exponentially higher localization weight than observations with low correlation. The adaptive localization method used in this study is a combination of Anderson (2001) and Bishop and Hodyss (2007), as proposed by Miyoshi (2010).

The following procedure is applied to each state variable in the state vector: the ensemble is first split into two sub-ensembles with equal number of members. For each sub-ensemble, the sample correlation coefficient between the state variable in question and each of the model observations is determined and these are then cross-validated using

$$p_{\mathrm{obs},a} = \left(1 - \frac{|c_1 - c_2|}{2}\right)^a, \tag{7}$$

HESSD
12, 2267–2304, 2015

where $p_{obs,a}$ is the localization weight, c_1 and c_2 are the correlation coefficients from the first and second sub-ensembles, and a is a constant used for tuning the localization.

Another localization weight, $p_{obs,b}$, is determined using the sample correlation coefficient for the entire ensemble, c, and another tuning constant, b:

$$p_{obs,b} = |c|^b. \quad (8)$$

The final (applied) localization weight, p_{obs} (Eq. 2), is calculated as the product of $p_{obs,a}$ and $p_{obs,b}$. Tuning the localization (i.e. determining the optimal values for the constants a and b), is in this study done by evaluating the mean Root Mean Square Error (RMSE) for the entire model domain.

For comparison, a Local Analysis (LA) localization method was used. The concept behind LA is that model grid points that are located far from each other should be expected to have little or no correlation, and so it creates a spatial window around each observation in which non-zero localization weights are applied. Model grid points located further away from the observation are given a localization weight of zero. Several methods for calculating distributions of the localization weights in the spatial window exist, but in this study the Gaussian decay is used:

$$p_{obs} = \exp\left(\frac{-d^2}{2\cdot\left(\frac{r}{2}\right)^2}\right), \quad (9)$$

where d is the physical distance between two points, and r is a user-specified localization radius. This weight is for each observation calculated for all model state variables, and results in a smooth distribution of localization weights from 1 at a distance of zero to 0 as the distance increases. At a distance of r, the localization weight is 0.135.

2.3.3 Parameter estimation with the ETKF

The data assimilation framework is set up as a joint state updating and parameter estimation framework, where parameter estimation is conducted using the augmented

HESSD
12, 2267–2304, 2015

state vector approach (Drécourt et al., 2005; Lui and Gupta, 2007). The state vectors (Eq. 1) are extended to also contain the parameters that are to be estimated as follows:

$$\mathbf{X}^f = \begin{bmatrix} \boldsymbol{x}_1^f, \ldots, & \boldsymbol{x}_n^f \\ \boldsymbol{\theta}_1^f, \ldots, & \boldsymbol{\theta}_n^f \end{bmatrix}, \tag{10}$$

where $\boldsymbol{\theta}_i^f$ is the set of parameters used to propagate the ith ensemble member. The mapping matrix **H** is extended accordingly, and the standard ETKF approach is applied.

2.3.4 Inflation

In order to compensate for the systematic underestimation of error variance that is common in the EnKF, covariance inflation (Anderson and Anderson, 1999) was applied to both the groundwater head states and the stream discharge states. The inflation is applied by adding a percentage to the ensemble of forecast anomalies:

$$\mathbf{X}'^b = (1 + \alpha) \cdot \mathbf{X}'^b, \tag{11}$$

where α is the inflation factor.

Covariance inflation of the ensemble of parameter values was performed by inflating the spread to a fixed spread (as described by the SD). This is done using an adaptive inflation factor that was calculated as follows:

$$\alpha = \frac{\sigma_{\text{Target}}}{\sigma_{\text{Forecast}}}, \tag{12}$$

where σ is the SD. σ_{Target} denotes the desired spread of the ensemble of parameter values and σ_{Forecast} denotes the spread of the ensemble before updating. This method is only applied if the forecast SD of the ensemble of parameters is smaller than the target SD which in this study is set to 10 % of the initial SD of the ensemble. This value has shown to produce the best results, by maintaining a sufficient spread, that does not create instabilities in any of the ensemble members.

HESSD
12, 2267–2304, 2015

2.3.5 Asynchronous assimilation

This study utilizes asynchronous assimilation, which refers to the assimilation of observations available at times different from the updating time. The AEnkF (Sakov et al., 2010) is a simple extension of the EnKF that allows for the asynchronous observations to be assimilated with little cost to the computational time or the storage requirements. The AEnKF requires the storage of model observations at the time that the asynchronous observations are available, which are then appended to the state vector, and through the covariance matrix used to update states and parameters at the time of assimilation.

The term "Assimilation window" is in the following used as the time span between two assimilation time steps. The observations collected in this assimilation window are assimilated at the time of the update, which requires that the ensemble model results stored at the observation time steps are used. So, given a set of j observations at times $t_1, \ldots, t_j$ collected in the assimilation window, the ensemble observations is formulated as follows:

$$\mathbf{HX}^f = \left[(\mathbf{H}\boldsymbol{x}^f)_1^t, \ldots, (\mathbf{H}\boldsymbol{x}^f)_j^t \right]. \tag{13}$$

Similarly, the observation vector is extended to correspond to the ensemble observations and filtering is carried out as described in Sect. 2.3.1.

2.3.6 State variables

In this study, the groundwater hydraulic head, the stream discharge, and the stream water level are updated at each assimilation step. The states are updated every two weeks, when groundwater head observations are available. Discharge observations in the assimilation window are included as asynchronous observations. This method allows all observations to be included without having to update the states at daily intervals, which would require significant computational time.

HESSD
12, 2267–2304, 2015

2.3.7 Estimated parameters

A sensitivity analysis performed using the AUTOCAL software (Madsen, 2003; MIKE by DHI, 2014) was used to choose the parameters to be estimated. Only parameters with scaled sensitivities (Table 1) of 1 % or more of the sensitivity of the most sensitive parameter were included. For practical reasons, as they tended to cause instabilities, parameters relating to the vadose zone were excluded. The exclusion of vadose zone parameters, even if they are sensitive, also means that the spread of the ensemble of parameter values is not sequentially decreased, which helps maintain a spread in the ensemble of state variables and avoid an ensemble collapse.

The hydraulic conductivities of meltwater sand and quaternary sand are included. Despite being hydrogeological parameters related to the groundwater module of the model, these are as evident from Table 1 also sensitive to the discharge observations. Also included are the drain level and drain time constant, which control the amount and dynamics of groundwater drained to the nearest stream once the groundwater table exceeds the drain level, and are as such particularly important for drain flow. The leakage coefficient, which is important with respect to base flow, is another coupling parameter, which represents the hydraulic properties of the thin layer of the sediments at the bottom of the stream.

Four of the five estimated parameters, the hydraulic conductivities of meltwater sand and quaternary sand, as well as the stream bed leakage coefficient and the drain time constant were transformed logarithmically in the filter, with drain level being the only parameter not transformed. As commonly practiced in calibration of hydrological models, the horizontal hydraulic conductivities were tied to the vertical hydraulic conductivities of the respective geological units at a fixed ratio of 10 to 1.

2.4 Twin test setup

This study is conducted as a twin test setup, in which a “true” model state is generated by perturbing the selected parameters of a model. Observations are generated by

HESSD
12, 2267–2304, 2015

adding noise to the states of the true model in the locations where observations are desired and with specified frequency. The original model, hereafter denoted base model, forms the basis for the ensemble that is used in the data assimilation. The parameter values used to generate the base model and the true model can be seen in Table 2.

2.5 Model noise

Model noise is added to the ensemble through the forcings, i.e. precipitation and reference evapotranspiration, and the parameters. Noise on forcings is added as a Gaussian noise with a SD of 20 % of the observed value, while no spatial correlation of the noise is considered.

Noise is added to a large number of model parameters relating to all model processes as seen in Table 2. In total noise is added to 66 parameters, only five of which are estimated. Adding noise to parameters that are not estimated helps maintain the spread of the ensemble even as the spread of the estimated parameters is reduced.

2.6 Data availability

The spatial distributions of hydraulic head observations studied are visualized in Fig. 2. The spatial distribution denoted “35 obs” contains observations in all the locations where actual (i.e. real world) observations are also available and presents an extensive spatial coverage of observations, with observations located in between almost all the branches of the river network and with many of the observations located in neighboring model grid cells. The spatial distribution “8 obs” is a subset of “35 obs” and represents a less extensive coverage of observations and with significantly fewer observations than “35 obs”. “2 obs” is a subset of “8 obs” and contains only two observations located approximately halfway downstream of the Karup river, and as such represents a scenario in which the spatial coverage of observations is poor. Finally, “0 obs” (not depicted in the figure) represents a scenario in which no groundwater head observations are available. Discharge observations are made available in five locations (see

HESSD
12, 2267–2304, 2015

Fig. 1), two of which are on the main river branch (one at the catchment outlet and one approximately halfway downstream) with the remaining observations located on the north-western tributary.

The frequency of groundwater head observations is 28 days^{-1}, while the frequency of discharge observations is daily. Head observations are added a normally distributed white noise with a SD of 0.05 m. Discharge observations are assigned a normally distributed white noise that is proportional to the observed value using a SD of 5 % of the observed discharge, which is a common error observed in real world observations of discharge Herschy (1999). This means that discharge measurement errors increase in peak flow situations and is larger for downstream locations, while the measurement error of head is not related to the location or the observed value.

2.7 Scenarios

This study will consist of four scenarios, with varying availability of discharge observations and with and without parameter estimation. In all four scenarios groundwater head data are assimilated.

InclParInclQ: The primary scenario in this study, in which discharge observations are assimilated and parameters are estimated, constitutes the most complex scenario. Estimating parameters makes the updates more nonlinear compared to stand-alone state updating, and assimilating discharge observations can be expected to require more ensemble members due to the complex relationship between stream discharge and groundwater head.

InclParNoQ: This scenario includes parameter estimation, but excludes discharge observations (stream discharge and water level are still included in the state vector). This means that the update of groundwater head, stream discharge and stream water level as well as the parameter estimation is based only on head observations.

NoParInclQ: This scenario includes the assimilation of discharge observations but excludes parameter estimation. This way, the influence of differing parameter sets is removed, allowing the direct results of updating the states to be seen.

HESSD
12, 2267–2304, 2015

NoParNoQ: This scenario excludes both the assimilation of discharge observations and parameter estimation. The simplest of the included scenarios, this scenario, when compared to scenario NoParInclQ illustrates the value of discharge observations on state updating.

2.8 Performance indicators

The performance of the filter will be evaluated based on three indicators:

- Mean root mean square error of hydraulic head for the entire domain (every model grid), calculated based on the mean of the ensemble at each time step (12 h time steps) and the true model state. Hereafter denoted head RMSE.
- Mean root mean square error of discharge in all grid point in the river network model, calculated based on the mean of the ensemble at each time step (2 h time steps) and the true model state. Hereafter denoted discharge RMSE. Note that this indicator inadvertently is dominated by downstream grid points with higher flow.
- The convergence of estimated parameters to the true value, including the spread and mean of the ensemble of parameters.

3 Results and discussion

3.1 Localization tuning

Tuning of the localization algorithm is carried out using a scenario in which two hydraulic head observations and all five discharge observations are available. An ensemble size of 50 is used, as experience had shown that this ensemble size resulted in significant spurious correlation with this number of observations.

HESSD
12, 2267–2304, 2015

The head RMSE as a function of the two localization constants can be seen in Fig. 4. Based on these results, constant values of *a* and *b* of 2 are used in the remainder of this study. Due to the computational time required, only integer values of the constants are tested, although it may have been possible to fine-tune the localization algorithm by using fractions.

Localization using LA was analyzed with varying localization distances and compared to using adaptive localization and no localization, as seen in Fig. 4. Overall localization distances of 20 and 10 km that apply to both the groundwater domain and the stream domain were tested. Compared to using no localization, a small reduction in head RMSE is obtained in the case of 20 km localization distance, whereas a significant increase in head RMSE is seen when a localization distance of 10 km is used. Compared to adaptive localization worse results are obtained, and it is clear that simple LA localization with localization distances that apply to both groundwater variables and stream variables is not sufficient. It should be noted that the LA method applied do not distinguish between model processes and that the localization distance also applies to the cross-correlation between the different state variables (i.e. groundwater observations are localized with the described distance with regards to stream variables and vice versa).

As a result, using lower localization distances for correlation across model processes was tested. This means that head observations are localized with a smaller radius with regards to stream discharge and water level, and vice versa. Two scenarios were analysed in which the localization distance within the same model process is Infinite (i.e. no localization) and with localization distances across processes of 5 and 0 km, respectively. The latter scenario means that there is no update across model processes and the two model states (groundwater and stream) are therefore updated independently from each other. As Fig. 3 shows, both scenarios led to a reduction in head RMSE compared to not using localization, yet head RMSE is still significantly higher than for the scenario with adaptive localization.

HESSD
12, 2267–2304, 2015

The effect of localization becomes clear when studying the time series of head RMSE (Fig. 4). Using no localization, the spurious correlations become dominant, as evident from the regular spikes visible in the dark blue line in Fig. 4. Using the LA localization method with 20 km localization radius does little to remove the spikes (and by extension, the spurious correlation), and a localization radius of 10 km only exacerbates them. Using differentiated localization radii for intra- and cross-process correlation removes a significant part of the spurious correlation with 0 km radius significantly outperforming 5 km radius. However, a few spikes do persist, and compared to the adaptive localization the decrease in head RMSE during some updates is not as big, suggesting that real correlation is removed.

The lower graph of Fig. 4, which shows the discharge RMSE as a function of time, illustrates why spurious correlation is a particular problem for discharge modelling. While the filter can reduce the discharge RMSE to almost zero at the updating time, peaks in the RMSE often appear in the time step immediately after the update. These peaks are the results of spurious correlation in the groundwater manifesting itself in the discharge and are due to the nature of the groundwater-stream flow interaction. Spurious correlation in groundwater appears where little real correlation with the observation points is present, which makes the grid cells that exchange water with the stream more sensitive than others. The dynamics of these cells are significantly different from the slow changing dynamics of most groundwater model cells, and any change in the interaction cells are reflected in the stream flow. Put simply, a change in the groundwater head of a few centimeters is barely noticeable in most grid cells, but may result in a significant change in the stream flow if the change is found in the grid cells that controls the interaction with the stream flow.

Figure 5 gives an insight into why the LA localization methods perform worse than the adaptive localization. As the figure shows, the localization weight derived from the adaptive localization algorithm is not a simple function of distance. In the case of groundwater localization weights, they seem to be highly correlated with the proximity to the stream network, with model grid points located next to the stream network

HESSD
12, 2267–2304, 2015

(and therefore exchanging water with the stream network) generally assigned very low weights. This may be explained by the dynamics of these groundwater model grid points displaying significantly different dynamics as previously discussed. As for stream model grid points, the distribution of localization weights appears to primarily depend on the branch on which the observation is located. The most downstream observation (at the model outlet) has the highest localization weights on the main branch of the network while the observation located on the tributary has the highest weights on that tributary (and on its tributaries). For both observations there is an alternation between high and medium to low localization weights visible in the figure, which is a result of the alternation between discharge and water level calculation used in the model. The adaptive localization algorithm assigns a lower weight to water level variables than to discharge variables due to the fact that discharge is what is being observed. One could think that this could lead to some discrepancies in the stream network states after updates, due to improper discharge-water level relations, but no effects (i.e. post update instability or state fluctuations) of this were seen, leading to the conclusion that the model is able to adjust any discrepancies there might be quickly after the states have been updated.

3.2 InclParInclQ

Head RMSE as a function of ensemble size for the scenario InclParInclQ can be seen in Fig. 6. Using 35 observations and no localization, a small improvement in head RMSE is seen when increasing the ensemble size from 25 to 50, but no change is seen when increasing from 50 to 100, and it appears that an ensemble size of 50 is sufficient in this case. When applying localization, almost no improvement is seen when increasing the ensemble size, suggesting that an ensemble size of 25 (or even less) is sufficient. Reducing the number of observations to 8, an improvement is seen when increasing the ensemble size from 25 to 50, 100 and even 200, with the largest improvement achieved when increasing from 25 to 50. The results suggest that some improvement may still be possible with ensemble sizes larger than 200. However, an

HESSD
12, 2267–2304, 2015

ensemble size of 200 is already very time consuming, despite the relatively small area and large spatial discretization of the model, and larger ensemble sizes are not realistic for practical applications. Using localization, the improvement seen when increasing the ensemble size is very small, and a small increase in head RMSE is even seen when increasing the ensemble size from 50 to 100, suggesting that 50 ensemble members is sufficient in this case. Using two or no head observations also results in improvement when increasing the ensemble size to 200. Using localization results in improvement in both scenarios, but even with localization applied, an ensemble size of 200 may not be sufficient.

The results show that the ensemble size needed depends strongly on the number of observations available. With fewer observations, the information from the observations needs to be spread further spatially, and the small correlation that exists relatively far from an observation needs to be estimated well, which requires more ensemble members. With a more extensive spatial coverage of observations, the small correlation of an observation and a distant model grid point becomes less important, as there is more likely to be another observation closer with a higher correlation which is better estimated with fewer ensemble members to estimate. As a result, a larger ensemble size is required when the spatial coverage of observations is poor. Localization removes the spurious correlation and thus improves the performance at lower ensemble sizes or lower observation numbers, but does not affect the sampling error itself. As a result, very large ensemble sizes are necessary when very few observations are available, even if localization is applied.

In terms of discharge RMSE (the lower graph in Fig. 1) the relationship between ensemble size and RMSE is a bit more unclear, due to spurious correlation being significant in most cases except “35 obs”. However, a general improvement is observed when using localization in most cases, with a significant improvement observed in the “0 obs” case, where spurious correlations are also most apparent.

As Fig. 7 shows, the performance of parameter estimation is related to the ensemble size and the number of observations as well as to the application of localization.

HESSD
12, 2267–2304, 2015

When assimilating eight observations, a slight improvement in the estimation of the drain level and drain constant is observed, while little or no improvement is observed when estimating the remaining parameters. The improvement with localization is more pronounced when only two or no head observations are assimilated, where an improvement can also be observed when the ensemble size is increased.

3.3 InclParNoQ

The head RMSE as a function of ensemble size for the scenario InclParNoQ can be seen in the left-most graphs in Fig. 8. Whether using eight observations or two observations, the use of localization and the increase in ensemble size has a much smaller effect on the performance in terms of head RMSE compared to the scenario in which discharge observations are assimilated. This suggests that the issue of spurious correlation is most dominant in the cross-process correlation, as is also suggested by the localization weights seen in Fig. 5. Generally, an improvement in terms of head RMSE is seen compared to the scenario in which discharge observations are assimilated (InclParInclQ), but the convergence of the two scenarios with increasing ensemble size suggests that this is due to spurious correlation in the InclParInclQ scenario. This means that if one is only interested in optimizing the filter for groundwater head updating, discharge observations could be left out, as they result in little or no improvement in the groundwater domain and requires a larger ensemble size.

The discharge RMSE in Fig. 7 shows that clear improvements in the discharge RMSE is achieved with the assimilation of discharge observations. Both the scenario with 8 observations and with 2 observations show increasing trends with respect to discharge RMSE vs. ensemble size, which seems to be related to the estimation of parameters, particularly the leakage coefficient which was estimated worse with increasing ensemble size (i.e. the mean of the ensemble of parameter values was offset from the true value). This is presumably done by the filter to optimize the groundwater state, but led to significant biases in the estimated parameter values.

HESSD
12, 2267–2304, 2015

The estimated parameter values can be seen in Fig. 7, which shows that little or no improvement in the estimation of parameters is obtained by increasing the ensemble size or by adding localization. When comparing to the parameter estimation of the InclParInclQ scenario, the estimation of all parameters is clearly worse in InclParNoQ, both in terms of the mean and the spread of the ensemble, underlining the necessity of assimilating discharge observations in integrated hydrological models, if the aim is to estimate parameters.

3.4 NoParInclQ

For the scenario NoParInlQ the head and discharge RMSE as a function of ensemble size can be seen in the two middle graphs in Fig. 8. In the case of two head observations, a significant reduction in head RMSE is observed when increasing the ensemble size from 25 to 50, followed by an increase in head RMSE with increasing ensemble size. The decrease in head RMSE is followed by a corresponding increase in discharge RMSE, indicating that the trade-off between groundwater and stream flow has shifted. Due to bias in the parameters that control the groundwater-stream flow interaction which is not being sequentially reduced due to the exclusion of parameter estimation, a more correct determination of the groundwater head will inevitably result in larger errors in the discharge. A visual study of the head RMSE as a function of time shows that while the update is approximately equally effective when using 25 or 50 ensemble members (in terms of discharge RMSE at the updating time), the increase in discharge RMSE between updates is larger for the 50 ensemble scenario, suggesting that the error in the interaction with the groundwater is more pronounced. It seems that the combination of parameter noise and stream discharge present in the 50 ensemble member case favors the correct description of groundwater head over discharge, even if all other factors (observation noise, uncertainty, and inflation) are the same in all scenarios. The increase in discharge and head RMSE observed from 100 to 200 ensemble members is due to an increase in spurious correlation. It seems counterintuitive that an increase in ensemble size would increase spurious correlation, but an increase in en-

HESSD
12, 2267–2304, 2015

semble size also increases the parameter space spanned by the ensemble, which may lead to spurious correlation appearing. This increase is not observed in InclParInclQ, as the parameter spread there is sequentially reduced.

When using eight observations, a general decrease in head and discharge RMSE is observed with increasing ensemble size. There is a significant difference when using localization, with localization increasing the head RMSE and decreasing the discharge RMSE substantially. The decrease in discharge RMSE is explained by the removal of spurious correlation, which can cause significant problems for the discharge in particular (see Sect. 3.1). The general increase in head RMSE may be explained by the trade-off effect shifting between the groundwater and the discharge observations.

In both cases when using either eight or two observations, the effect on the head RMSE is relatively small. This is most likely due to the slow changing dynamics of groundwater, which means that the groundwater head is well constrained, and does not deviate significantly from the "true" groundwater head in between updates. The discharge on the other hand, changes very rapidly, and the effect of updating the discharge at a specific time will quickly disappear after the model is started again. Adding to this is the problem with spurious correlation and its relevance to discharge modelling (see Sect. 3.1) which often results in very high discharge RMSE and makes direct comparison of discharge RMSE between scenarios difficult.

3.5 NoParNoQ

The head and discharge RMSE when parameter estimation and discharge observations are omitted can be seen in the two rightmost graphs of Fig. 8. Both when using two observations and eight observations, the resulting changes in head and discharge RMSE with increasing ensemble size are very small. Likewise, the benefit of using localization is negligible. This may be explained by the updating being much more linear than in any of the other scenarios, thus reducing the need for a large ensemble size.

When comparing the NoParNoQ results with the NoParInclQ results, it becomes clear that the impact of assimilating discharge (without estimating parameters) is small

HESSD
12, 2267–2304, 2015

with respect to both head RMSE and discharge RMSE both in the case of using eight observations and two observations. However, the trade-off issue does not exist in the NoParNoQ scenario, causing this scenario to perform better with respect to head RMSE when two observations are used.

4 Conclusion

This study investigated the impact of localization and ensemble size when applying data assimilation to a coupled surface–subsurface model, considering different types and varying amount of observation data and parameter estimation.

The adaptive localization method used in this study was in many cases able to reduce the required ensemble size significantly. The method resulted in a complex distribution of localization weights in both domains of the model (groundwater and stream flow) that depended heavily on the geology and the position of the observation relative to the stream network. This distribution could not be obtained using the common local analysis (LA) methods, and direct comparison of the adaptive localization and LA localization also showed that adaptive localization outperformed LA localization with respect to head RMSE. Adaptive localization is not only easily implemented in the ETKF, it also automatically ensures that the cross-process correlation is localized differently than the intra-process correlation, making it particularly suitable for data assimilation in coupled surface–subsurface models. Others have encountered the problem with cross-process correlation, notably Zupanski (2013), although no definitive solution to the problem has been presented. Adaptive localization, such as the method applied in this study, may be one possible solution.

When assimilating both groundwater head observations and estimating parameters, localization and large ensemble sizes are important due to the nonlinearity of the state and parameter updates. This tendency is increasingly pronounced with decreasing number of observations assimilated due to the small correlations between observations and model states being more important when the spatial distribution of obser-

HESSD
12, 2267–2304, 2015

vations is poor. Excluding discharge observations reduces the benefits of localization and increasing ensemble size, as does the exclusion of parameter estimation. When excluding both discharge observation assimilation and parameter estimation applying localization or increasing the ensemble size from 25 to 50, 100 or 200 has almost no effect on the filter performance. The effects of increasing ensemble size in hydrological modelling has previously been studied (Chen et al., 2013; Xie and Xhang, 2010), with findings similar to the ones of this study. Both studies found that increases in ensemble size improved filter performance, with Xie and Xhang (2010) increasing the ensemble size to 1000, and still observed improvements. However, neither of the studies related the ensemble size to the amount of observations assimilated or to the estimation of parameters.

Like with state updating, estimation of parameters was primarily improved by an increasing ensemble size when discharge observations were assimilated. With discharge observations assimilated, clear improvements in parameter estimation were observed when applying localization, and to some extent when increasing the ensemble size (depending on the number of assimilated head observations). However, no improvement was observed when applying localization or increasing the ensemble size when discharge observations were not assimilated.

In conclusion, the required ensemble size depends heavily on the assimilation of discharge observations and estimation of parameters, as well as on the available number of observations. A large ensemble size is necessary when discharge observations are assimilated, parameters are estimated and few observations are available, while a significantly smaller ensemble size is sufficient when only groundwater head is assimilated and updated. However, the best overall filter performance (i.e. a combination of groundwater head and stream flow modelling) is found when discharge observations are assimilated and parameters are estimated. While the findings of this study could to a certain extent be derived intuitively, this is to our knowledge the first time that they have been quantified and documented in integrated hydrological modeling.

HESSD
12, 2267–2304, 2015

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.

Anderson, J. L., Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter, Physica D, 230, 99–111, doi:10.1016/j.physd.2006.02.011, 2007.

Bishop, C. H. and Hodyss, D.: Ensemble covariances adaptively localized with ECO-RAP – Part 1: Tests on simple error models, Tellus A, 61, 84–96, 2009a.

Bishop, C. H. and Toth, Z.: Ensemble transformation and adaptive observations, J. Atmos. Sci., 56, 1748–1765, 1999.

Bishop, C. H., Etherton, B. J., and Majumdar, S. J.: Adaptive sampling with the Ensemble Transform Kalman Filter – Part I: Theoretical aspects, Mon. Weather Rev., 129, 420–436, 2001.

Camporese, M., Paniconi, C., Putti, M., and Salandin, P.: Ensemble Kalman filter data assimilation for a process-based catchment scale model of surface and subsurface flow, Water Resour. Res., 45, W10421, doi:10.1029/2008WR007031, 2009.

Drécourt, J.-P., Madsen, H., Rosbjerg, D.: Bias aware Kalman filters: comparison and improvements, Adv. Water Resour., 29, 707–718, doi:10.1016/j.advwatres.2005.07.006, 2005.

Fertig, E. J., Hunt, B. R., Ott, E., and Szunyogh, I.: Assimilating non-local observations with a local ensemble Kalman filter, Tellus A, 59, 719–730, doi:10.1111/j.1600-0870.2007.00260.x, 2007.

Graham, D. N. and Butts, M. B.: Flexible, integrated watershed modelling with MIKE SHE, in: Watershed Models, edited by: Singh, V. P. and Frevert, D. K., CRC Press, Boca Raton, Florida, USA, 245–272, 2005.

Greve, M. H., Greve, M. B., Bøcher, P. K., Balstrøm, T., Breuning-Madsen, H., and Krogh, L.: Generating a Danish raster-based topsoil property map combining choropleth maps and point information, Geogr. Tidsskr., 107, 1–12, 2007.

Harlim, J. and Hunt, B. R.: Local Ensemble Transform Kalman Filter: an efficient scheme for assimilating atmospheric data, Department of Atmospheric & Oceanic Science, University of Maryland, available at: http://www.atmos.umd.edu/~ekalnay/pubs/harlim_hunt05.pdf (last access: 18 February 2015), 2005.

HESSD
12, 2267–2304, 2015

Hendricks Franssen, H. J. and Kinzelbach, W.: Real-time groundwater flow modeling with the Ensemble Kalman Filter: joint estimation of states and parameters and the filter inbreeding problem, Water Resour. Res., 44, W09408, doi:10.1029/2007WR006505, 2008.

Hendricks Franssen, H. J., Kaiser, H. P., Kuhlmann, U., Bauser, G., Stauffer, F., Muller, R., and Kinzelbach, W.: Operational real-time modeling with ensemble Kalman filter of variably saturated subsurface flow including stream-aquifer interaction and parameter updating, Water Resour. Res., 47, W02532, doi:10.1029/2010WR009480, 2011.

Herschy, R. W.: Hydrometry – Principles and Practices, 2nd edn., Wiley & Sons Ltd., Hoboken, New Jersey, USA, 1999.

Hunt, B., Kostelich, E., and Syzunogh, I.: Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter, Physica D, 230, 112–126, 2007.

Juston, J., Seibert, J., and Johansson, P.-O.: Temporal sampling strategies and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment, Hydrol. Process., 23, 3093–3109, doi:10.1002/hyp.7421, 2009.

Kristensen, K. J. and Jensen, S. E.: A model of estimating actual evapotranspiration from potential evapotranspiration, Nord. Hydrol., 6, 170–188, 1975.

Liu, Y. and Gupta, H. V.: Uncertainty in hydrologic modeling: toward an integrated data assimilation framework, Water Resour. Res., 43, W07401, doi:10.1029/2006WR005756, 2007.

Madsen, H.: Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, Adv. Water Resour., 26, 205–216, doi:10.1016/S0309-1708(02)00092-1, 2003.

Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P.: Dual state-parameter estimation of hydrological models using Ensemble Kalman Filter, Adv. Water Resour., 28, 2135–2147, 2005.

Miyoshi, T.: An adaptive covariance localization method with the LETKF, 14th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS), Atlanta, GA, USA, 21 January, 2010.

Nie, S., Zhu, J., and Luo, Y.: Simultaneous estimation of land surface scheme states and parameters using the ensemble Kalman filter: identical twin experiments, Hydrol. Earth Syst. Sci., 15, 2437–2457, doi:10.5194/hess-15-2437-2011, 2011.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. A.: A local ensemble Kalman filter for atmospheric data assimilation, Tellus A, 56, 415–428, doi:10.1111/j.1600-0870.2004.00076.x, 2004.

HESSD
12, 2267–2304, 2015

Refsgaard, J. C.: Parametrisation, calibration and validation of distributed hydrological models, J. Hydrol., 198, 69–97, 1997.

Sakov, P., Evensen, G., and Bertino, L.: Asynchronous data assimilation with the EnKF, Tellus A, 62, 24–29, doi:10.1111/j.1600-0870.2009.00417.x, 2010.

Shi, Y., Davis, K. J., Zhang, F., Duffy, C. J., and Yu, Z.: Parameter estimation of a physically-based land surface hydrologic model using the Ensemble Kalman Filter: a synthetic experiment, Water Resour. Res., 50, 706–724, doi:10.1002/2013WR014070, 2014.

Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W., and Verstraten, J. M.: Improved treatment of uncertainty in hydrologic modeling: combining the strengths of global optimization and data assimilation, Water Resour. Res., 41, W01017, doi:10.1029/2004WR003059, 2005.

Xie, X. and Zhang, D.: Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter, Adv. Water Resour., 33, 678–690, doi:10.1016/j.advwatres.2010.03.012, 2010.

Zupanski, M.: All-sky satellite radiance data assimilation: methodology and challenges, in: Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications, vol. 2, edited by: Park, S. K. and Xu, L., Springer-Verlag, Berlin, Heidelberg, 2, 1–25, doi:10.1007/978-3-642-35088-7_1, 2013.

HESSD
12, 2267–2304, 2015

Table 1. List of parameters included for estimation, including their normalized sensitivity coefficients to head and discharge observations respectively.

Parameter name	Parameter description	Normalized sensitivity coefficient (Head)	Normalized sensitivity coefficient (Discharge)
HK_mws	Horizontal hydraulic conductivity of meltwater sand	1.00	1.00
Leakage	Stream bed leakage coefficient	0.22	0.29
HK_qs	Horizontal hydraulic conductivity of quaternary sand	0.06	0.11
Drain level	Drain level	0.03	0.07
Drain constant	Drain constant	0.01	0.04

HESSD
12, 2267–2304, 2015

Table 2. List of parameters perturbed to create the true model and to add noise to the filter ensemble. Parameters in bold are included in the parameter estimation. Parameters with very low sensitivity were omitted. Parameters are perturbed using Gaussian noise with SD shown in the table.

Parameter name	Distribution	True value	Initial value	SD	Log transformed
Hor. Hyd. conductivity	**Meltwater sand**	**–8.52**	**–7.60**	**0.818**	**x**
Hor. Hyd. conductivity	**Quarternary sand**	**–6.21**	**–7.01**	**1.151**	**x**
Hor. Hyd. conductivity	Clay	−15.42	−15.42	0.194	x
Hor. Hyd. conductivity	Mica Sand	−11.74	−11.74	0.201	x
Hor. Hyd. conductivity	Mica clay/sand	−16.34	−16.34	0.213	x
Specific yield	Meltwater sand	0.25	0.25	0.025	
Specific yield	Clay	0.05	0.05	0.004	
Specific yield	Quarternary sand	0.25	0.25	0.023	
Specific yield	Mica Sand	0.20	0.20	0.022	
Specific yield	Mica clay/sand	0.05	0.05	0.005	
Specific storage	Meltwater sand	−11.62	−9.90	0.308	x
Specific storage	Clay	−9.57	−9.90	0.335	x
Specific storage	Quarternary sand	−11.74	−9.90	0.318	x
Specific storage	Mica Sand	−9.21	−9.90	0.320	x
Specific storage	Mica clay/sand	−6.21	−9.21	0.367	x
Drain level	**Global**	**–1.00**	**–0.90**	**0.215**	
Drain time constant	**Global**	**–14.33**	**–15.02**	**0.381**	**x**
Leakage coefficient	**Global**	**–15.48**	**–14.79**	**0.885**	**x**
Overland manning no.	Global	4.00	5	0.213	
Overland detention	Global	0.01	0.02	0.001	
Leaf area index	Forest	5.00	6	0.431	
Leaf area index	Heath	2.50	2	0.209	
Leaf area index	Agriculture	4.00	5	0.413	
Root depth	Agriculture	900	1000	43.87	
Leaf area index	Wetland	5.00	6	0.435	
Root depth	Wetland	710	700	98.91	
Stream manning number	Global	20	25	1.964	
Sat. moisture content	Soil type 1	0.40	0.40	0.020	
Soil matric pot. (field cap)	Soil type 1	2.00	2.00	0.088	
Soil matric pot. (wilting point)	Soil type 1	4.20	4.20	0.168	
Sat. Hyd. conductivity	Soil type 1	−11.18	−11.18	0.345	x

HESSD
12, 2267–2304, 2015

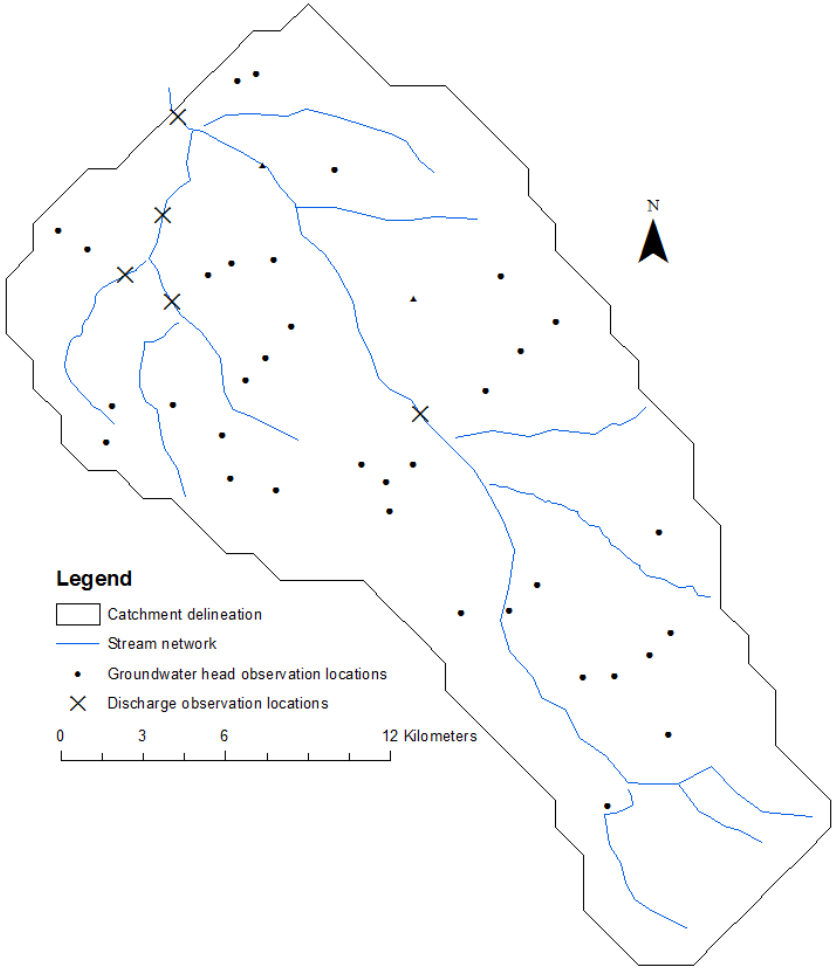

Figure 1. Karup catchment with locations of discharge and hydraulic head observations.

HESSD
12, 2267–2304, 2015

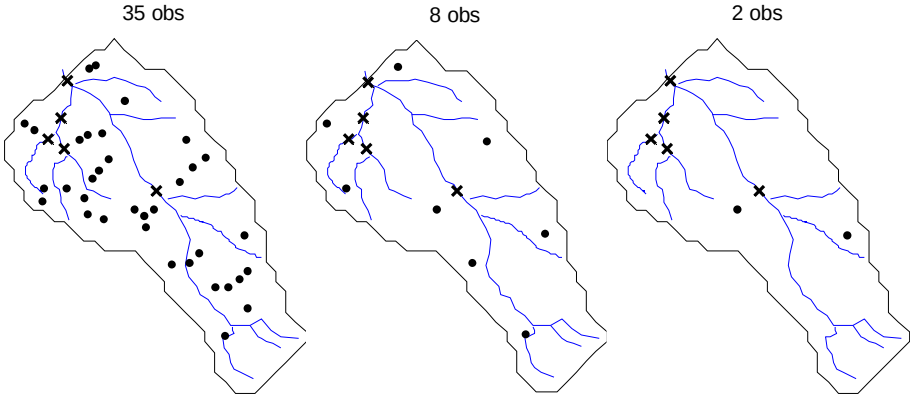

Figure 2. Spatial distributions of observations. Dots and crosses respectively denote groundwater hydraulic head and stream discharge observation locations, respectively.

HESSD
12, 2267–2304, 2015

Const. B \ Const. A	1	2	3
3	0.100	0.106	0.106
2	0.098	0.098	0.104
1	0.111	0.110	0.109

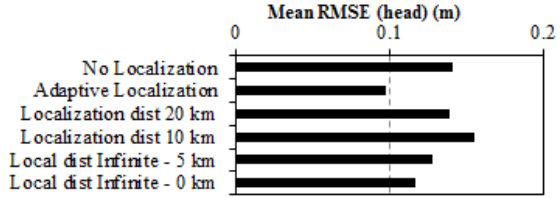

Figure 3. Head RMSE as a function of the localization constants *a* and *b* (left) and head RMSE using different localization methods (right).

HESSD
12, 2267–2304, 2015

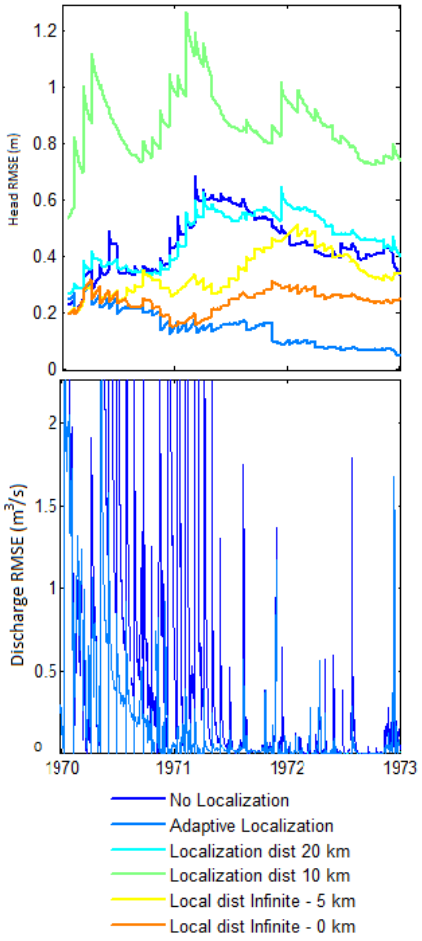

Figure 4. Head (top) and discharge (bottom) RMSE for the entire model domain as a function of time with different types of localization applied.

HESSD
12, 2267–2304, 2015

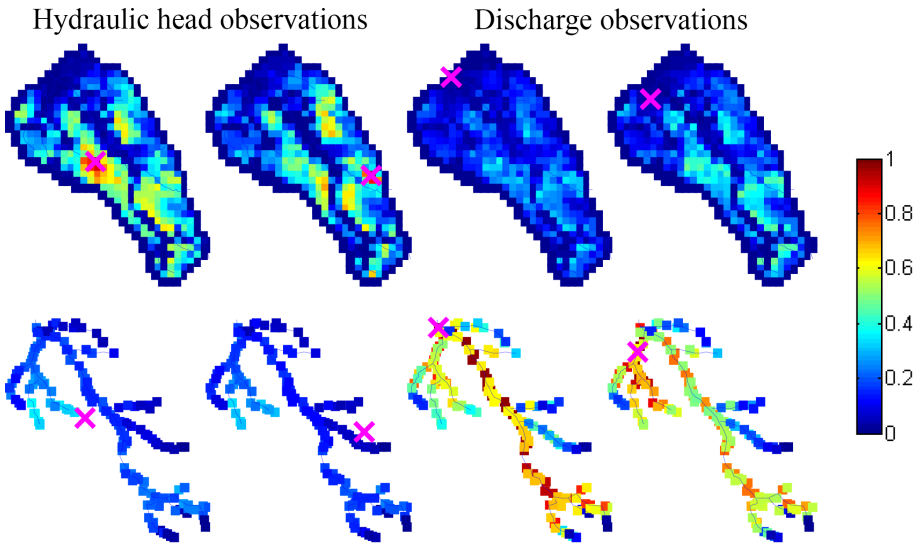

Figure 5. Mean localization weight derived from the localization algorithm. Top row shows the localization weight in groundwater head and the bottom row shows the localization weights in stream variables (discharge and water level). Magenta crosses indicate the observation locations. Only two selected discharge observations are shown.

HESSD
12, 2267–2304, 2015

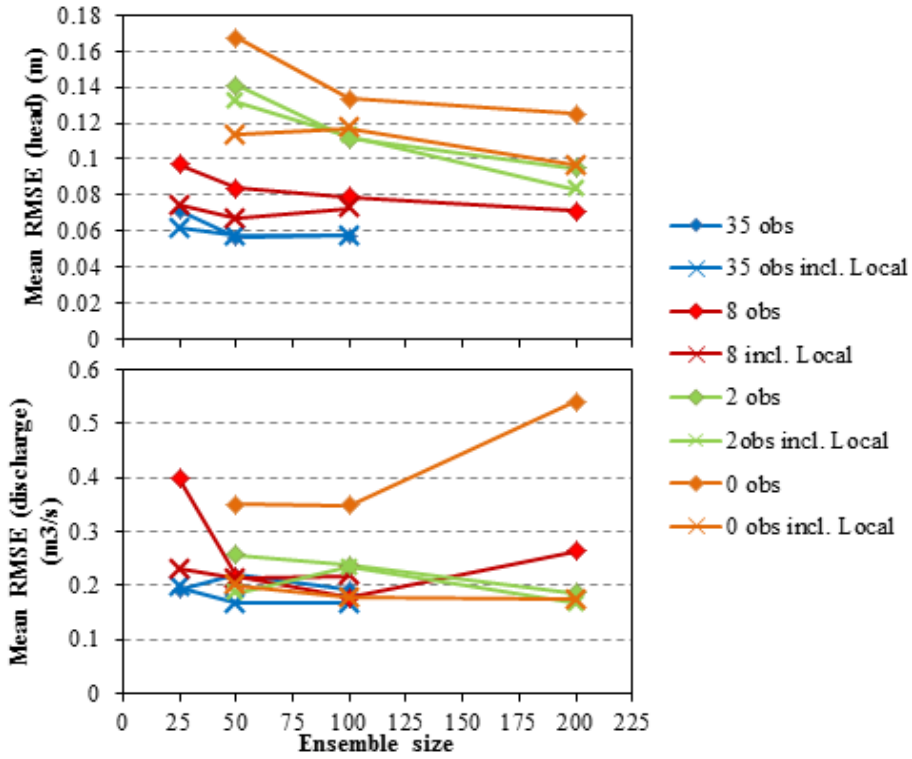

Figure 6. Head RMSE (top) and discharge RMSE (bottom) as a function of ensemble size in the InclParInclQ scenario for different numbers of groundwater head observations.

HESSD
12, 2267–2304, 2015

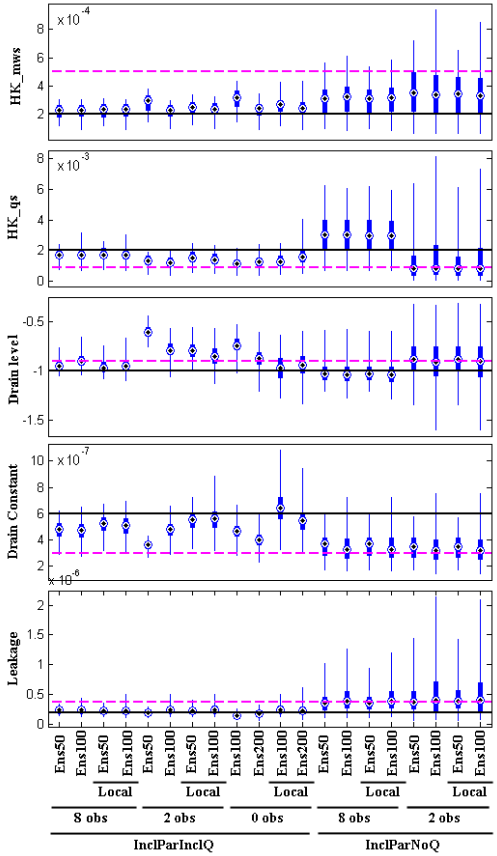

Figure 7. Spread of estimated parameters at the final update. Thin blue lines show the total spread of the ensemble and thick blue lines show the 25th and 75th percentile. Dots show the mean of the ensemble. The horizontal black and magenta lines show the true and the base parameter values, respectively.

HESSD
12, 2267–2304, 2015

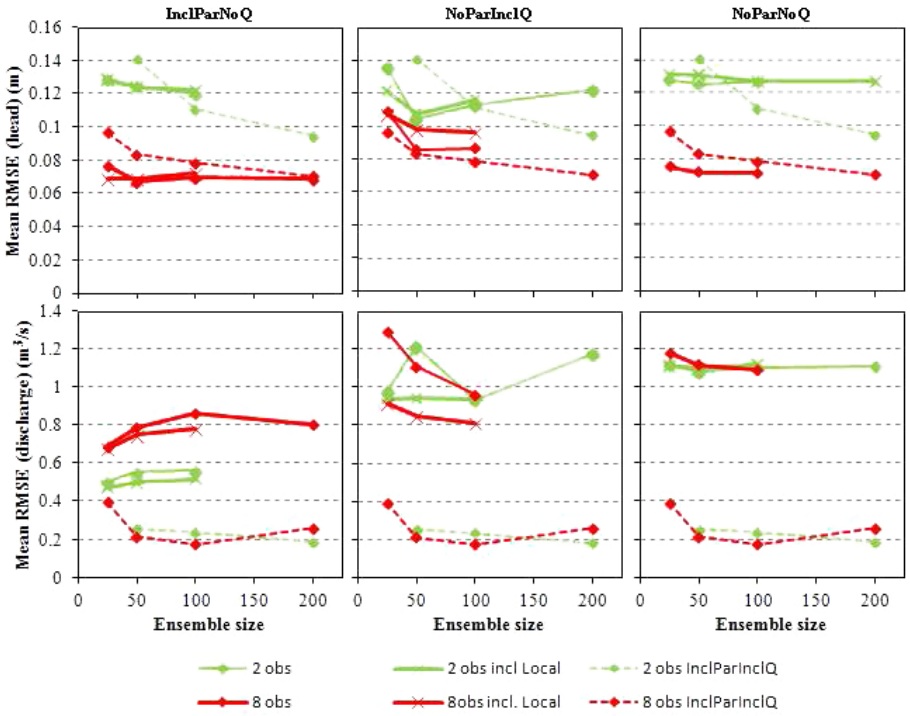

Figure 8. Head RMSE (top) and discharge RMSE (bottom) as a function of ensemble size for three of the scenarios. For comparison, the dashed lines indicate the head and discharge RMSE of the InclParInclQ scenario (without localization).

HESSD
12, 2267–2304, 2015