AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2013WR014823

Special Section:
Patterns in Soil-Vegetation-Atmosphere Systems: Monitoring, Modelling and Data Assimilation

Joint assimilation of piezometric heads and groundwater temperatures for improved modeling of river-aquifer interactions

Wolfgang Kurtz[1,2], Harrie-Jan Hendricks Franssen[1,2], Hans-Peter Kaiser[3], and Harry Vereecken[1,2]

[1]Agrosphere (IBG-3), Institute of Bio- and Geosciences, Forschungszentrum Jülich GmbH, Jülich, Germany, [2]Centre for High-Performance Scientific Computing in Terrestrial Systems: HPSC TerrSys, Forschungszentrum Jülich GmbH, Jülich, Germany, [3]Water Supply of Zurich, Zurich, Switzerland

Key Points:
- Assimilation of hydraulic and thermal data into a managed river-aquifer system
- Temperature predictions could be improved with EnKF
- Temperature data gave information on spatial structure of leakage parameters

Correspondence to:
W. Kurtz,
w.kurtz@fz-juelich.de

Citation:
Kurtz, W., H.-J. Hendricks Franssen, H.-P. Kaiser, and H. Vereecken (2014), Joint assimilation of piezometric heads and groundwater temperatures for improved modeling of river-aquifer interactions, *Water Resour. Res.*, *50*, 1665–1688, doi:10.1002/2013WR014823.

Received 30 SEP 2013
Accepted 1 FEB 2014
Accepted article online 6 FEB 2014
Published online 25 FEB 2014

Abstract The ensemble Kalman filter (EnKF) is increasingly used to improve the real-time prediction of groundwater states and the estimation of uncertain hydraulic subsurface parameters through assimilation of measurement data like groundwater levels and concentration data. At the interface between surface water and groundwater, measured groundwater temperature data can provide an additional source of information for subsurface characterizations with EnKF. Additionally, an improved prediction of the temperature field itself is often desirable for groundwater management. In this work, we investigate the worth of a joint assimilation of hydraulic and thermal observation data on the state and parameter estimation with EnKF for two different model setups: (i) a simple synthetic model of a river-aquifer system where the parameters and simulation conditions were perfectly known and (ii) a model of the Limmat aquifer in Zurich (Switzerland) where an exhaustive set of real-world observations of groundwater levels (87) and temperatures (22) was available for assimilation (year 2007) and verification (year 2011). Results for the synthetic case suggest that a joint assimilation of piezometric heads and groundwater temperatures together with updating of uncertain hydraulic parameters gives the best estimation of states and hydraulic properties of the model. For the real-world case, the prediction of groundwater temperatures could also be improved through data assimilation with EnKF. For the validation period, it was found that parameter fields updated with piezometric heads reduced RMSE's of states significantly (heads −49%, temperature −15%), but an additional conditioning of parameters on groundwater temperatures only influenced the characterization of the temperature field.

1. Introduction

A special feature of river-aquifer systems is that there can be a distinct cyclic heat exchange between the river and the aquifer. Depending on the meteorological conditions, surface water temperatures are subjected to diurnal and seasonal temperature variations whereas groundwater is characterized by relatively constant temperatures. The temperature distribution around streams is therefore governed by the temperature difference and the exchange pattern between river and aquifer. For example, seepage from the river to the aquifer can result in a temperature signal that propagates from the river into the aquifer depending on the seasonal/diurnal temperature contrast between the two compartments. Vice versa, an aquifer that discharges into the river can also generate a distinct thermal profile within the river bed. This heat exchange between river and aquifer can be deployed to characterize the exchange fluxes as well as relevant material properties of the river bed and the adjacent aquifer [*Anderson*, 2005; *Constantz*, 2008]. As a tracer, groundwater temperatures are more sensitive to the connectivity patterns within an aquifer compared to hydraulic data alone and can thus provide additional information on aquifer structure. Another feature that makes the utilization of heat as a tracer very attractive is that temperature data can be measured very easily at a low cost.

Temperature measurements have, for example, been used to derive a detailed spatial picture of river-aquifer exchange fluxes under field conditions [e.g., *Conant*, 2004; *Schmidt et al*., 2006; *Hatch et al*., 2010]. A common approach for such small-scale applications is to measure vertical temperature profiles in the river bed and to apply an analytical solution to the measured temperature profiles to infer the exchange fluxes between river and aquifer [*Schornberg et al*., 2010]. In most cases, only a limited number of point

measurements of temperatures are available for the estimation of exchange fluxes. Recently, also temperature measurements with fiber-optic sensors, also called Distributed Temperature Sensing (DTS)[*Tyler et al*., 2009], have gained interest for the derivation of spatially highly resolved temperature distributions [*Vogt et al*., 2012].

Thermal data have already been used as an additional information source for different inversion techniques to constrain the estimation of subsurface parameters. For example, *Woodbury and Smith* [1988] investigated the worth of thermal data for the calibration of steady state groundwater models. They argue that thermal data can better constrain the calibration of hydraulic conductivities especially for high-permeable aquifers when heat advection is the dominant process. *Doussan et al*. [1994] calibrated a managed river-aquifer system which included bank filtration with hydraulic and thermal data and found an improved estimation of river bed parameters compared to hydraulic data alone. *Bravo et al*. [2002] also used hydraulic and thermal data for the inversion of a wetland-aquifer system to derive hydraulic conductivities and wetland inflows. They showed that the inversion gets more stable and accurate when thermal data are used compared to hydraulic data alone. *Friedel* [2005] calibrated vadose zone and transport parameters of a synthetic model of artificial aquifer recharge with different combinations of piezometric head, concentration and thermal data. In *Jiang and Woodbury* [2006], a Bayesian inversion technique was applied to an aquifer model which was conditioned on different combinations of piezometric heads, transmissivities and temperature measurements. They found that the characterization of hydraulic conductivities was improved with temperature data for different inversion scenarios.

For modeling purposes, an important implication of the temperature contrast between rivers and aquifers and the resulting heat transfer is that these temperature changes in the sediment induce a cyclic variation of water density and viscosity which also affects the hydraulic conductivity of the river bed sediments. Changes in water density can usually be neglected for the temperature range that occurs in river-aquifer systems but water viscosity could change up to a factor of 1.7 given a typical temperature range between 5 and 25°C. For example, *Constantz et al*. [1994] argue that their measured diurnal variation of river-aquifer exchange fluxes is largely attributed to the temperature dependency of hydraulic properties of the streambed. *Engeler et al*. [2011] have shown that considering the temperature dependency of water viscosity in the simulation of river-aquifer exchange can reduce the predictions errors of piezometric heads at individual measurement locations up to 30%. *Ma and Zheng* [2010] investigated the effect of regarding temperature-dependent hydraulic parameters for the modeling of heat transport in river-aquifer systems. They concluded that temperature contrasts up to 15°C lead to an average error in temperature predictions of about 3% calculated over their whole model domain. However, they did not compare the effects on individual measurement locations.

Apart from river-aquifer systems, heat transfer calculations can also be relevant to other systems of surface water-groundwater exchange. One example is artificial recharge because for such systems also a temperature contrast between surface water and groundwater is given. *Vandenbohede and Van Houtte* [2012] give an example of such an application where heat transport calculations were made for an artificial recharge system to characterize the transport behavior underneath a recharge basin. *Racz et al*. [2012] used measured temperature profiles underneath an artificial recharge basin to determine local infiltration fluxes.

An excellent example of a managed groundwater system that includes the effects of river-aquifer exchange and artificial recharge is the Limmat aquifer in Zurich (Switzerland). For the groundwater management at this site, water is pumped from several bank filtration wells close to the river Limmat and this water is then artificially recharged to the aquifer through several recharge basins and wells. This measure is taken to protect drinking water wells from a diffuse contamination that is present close to the well field. These management activities heavily influence the hydraulic and thermal situation in this aquifer. Recently, a real-time modeling system was set up for the management of this site [*Hendricks Franssen et al*., 2011]. In this framework, the ensemble Kalman filter (EnKF) [*Evensen*, 1994; *Burgers et al*., 1998] is applied to correct the piezometric head predictions and the hydraulic parameters of a 3-D groundwater model for this site with data from a dense online-monitoring network for groundwater levels on a daily basis. The updated predictions of this model can then be used to optimize the operation of the well field through a real-time control system [*Bauser et al*., 2010, 2012; *Marti et al*., 2012]. In 2005, the monitoring network was additionally equipped with several online sensors for groundwater temperature which allow a continuous monitoring of the thermal situation within the well field. As already pointed out before, groundwater temperatures are well suited

as a tracer for exchange processes between surface water and groundwater which also have a large influence on the operation of the well field in the Limmat aquifer. Thus, the additional online monitoring of groundwater temperatures can provide important information on the subsurface structure for this site which is until now not utilized in the EnKF data assimilation framework.

The aim of the following study is to extend the existing EnKF data assimilation framework, so that also temperature measurements can be used to update the predictions of groundwater states and model parameters. There is a request from groundwater managers at this site to improve the prediction of the temperature distribution in the aquifer because it is desired to avoid the pumping of drinking water that is too warm because this can negatively influence drinking water quality through bacterial contamination and increases the need for disinfection measures. EnKF offers the possibility to improve the prediction of the thermal regime of the aquifer in real time with the installed monitoring network for groundwater temperatures and groundwater managers also plan to include the results of such a real-time model in the real-time control of the well field. To the best of our knowledge, the extension of EnKF to the coupled subsurface flow-heat transport problem including parameter updating is a novel contribution. The joint assimilation of piezometric head and temperature data will first be tested with a synthetic river-aquifer model in order to assess the worth of the additional temperature assimilation on the estimation of hydraulic parameters under controlled conditions. Afterward, the extended data assimilation framework will also be applied to the real-world data of the Limmat aquifer and it will be explored how the assimilation of measured groundwater temperatures affects the prediction of heat transport in the model and also how the joint assimilation of hydraulic and thermal data affects the estimation of hydraulic parameters under real-world conditions. This dual approach allows us to monitor the effectiveness of the joint assimilation scheme at the transition from a more theoretical approach (synthetic setup) toward the more complicated situation of a real-world model. This is of special interest because the vast majority of EnKF applications in groundwater hydrology deal with synthetic cases which test data assimilation techniques under idealized conditions where unknown sources of uncertainty or model structural errors are excluded from the analysis. For the real-world case, an extensive data set of hydraulic head and temperature observations (87 and 22 daily observations respectively) is available for a period of 6 years (2006–2011). Observation data are split into an assimilation period (year 2007) and a subsequent validation period (year 2011) in which the effect of parameter updates on temperature predictions is assessed. A second important aspect of this work is therefore the verification of a calibrated coupled subsurface flow-heat transport model with both hydraulic head and groundwater temperature data, for a period which is several years separated from the assimilation period. Different updating scenarios are compared with these observation data. These scenarios study the impact of the assimilation of different data types (groundwater levels and/or groundwater temperatures) and the role of data availability and data quality on the assimilation process. An apparent difference between the joint utilization of hydraulic and thermal data in EnKF and the use of these data in other previously published inversion schemes is also that these data are not only used for the determination of uncertain model parameters but that EnKF also provides an assessment of model uncertainty and is used to for updating model states in real time. A further aspect that is investigated in this study is the effect of covariance localization on state-parameter updates with EnKF under real-world conditions. This issue has already been addressed by several studies on EnKF in subsurface characterization [e.g., *Nan and Wu*, 2011; *Devegowda et al*., 2010; *Chen and Oliver*, 2010]. However, these applications of localization were restricted to the assimilation of one measurement type in most cases (except *Chen and Oliver* [2010]) and were only done for synthetic data.

2. Methodology

In this study, we perform data assimilation experiments with a coupled flow and heat transport groundwater model. In order to illustrate the technical setup for these experiments, section 2.1 summarizes the governing equations and the utilized groundwater modeling software, sections 2.2, 2.3, and 2.4 explain the implementation of EnKF for this study and sections 3.1 and 3.2 provide an overview about the specific model setups for the synthetic and the real-world groundwater models.

2.1. Forward Model

The variably saturated subsurface flow equation can be stated as [*Bear and Chen*, 2010]:

$$\frac{\partial(n\rho_w S(\theta))}{\partial t} = \nabla\left(\frac{\rho_w k k_r(\theta)}{\mu}(\nabla p + \rho_w g \nabla z)\right) + q_f \tag{1}$$

where n is porosity [-], ρ_w is density of water [$M\,L^{-3}$], $S(\theta)$ is saturation [-], θ is water content [$L^3\,L^{-3}$], t is time [T], k is permeability [L^2], $k_r(\theta)$ is relative permeability [-], μ is dynamic viscosity [$M\,L^{-1}\,T^{-1}$], p is pressure [$M\,L^{-1}\,T^{-2}$], g is gravitational acceleration [$M\,T^{-2}$], z is height above datum [L], and q_f are additional source/sink terms [$M\,L^{-3}\,T^{-1}$].

The relative permeability can be derived with the van Genuchten parameterization [*van Genuchten*, 1980]:

$$k_r(\theta) = \hat{\theta}^{0.5}\left(1-\left(1-\hat{\theta}^{\frac{n_p}{n_p-1}}\right)^{\frac{n_p-1}{n_p}}\right)^2 \tag{2a}$$

$$\hat{\theta} = \frac{\theta-\theta_r}{\theta_s-\theta_r} \tag{2b}$$

where $\hat{\theta}$ is normalized water content, θ_r is residual water content, θ_s is saturated water content, and n_p is an empirical parameter.

The heat transport equation can be summarized as:

$$\frac{\partial c_v^* T}{\partial t} = -\rho_w c_w \nabla(vT) + \nabla^2(\kappa_v^* T) + \nabla(\rho_w c_w D \nabla T) + q_h \tag{3a}$$

$$c_v^* = (1-n)\rho_b c_b + n\rho_w c_w \tag{3b}$$

$$\kappa_v^* = (1-n)\kappa_b + n\kappa_w \tag{3c}$$

where T is temperature [Θ], ρ_b is bulk density [$M\,L^{-3}$], c_w and c_b are specific heat capacities of water and solid matrix [$L^2\,T^{-2}\,\Theta^{-1}$], κ_w and κ_b are thermal conductances of water and solid matrix [$M\,L\,\Theta^{-1}\,T^{-3}$], v is Darcy velocity [$L\,T^{-1}$], and q_h are additional thermal source/sink term [$M\,L^{-1}\,T^{-3}$].

For our study, the groundwater modeling software SPRING [*Delta h Ingenieurgesellschaft mbH*, 2006] was used to solve the coupled equations of 3-D variably saturated flow and heat transport (equations (1)–(3c)). River-aquifer exchange which is highly relevant in our study is implemented in SPRING as a Cauchy-type boundary condition (i.e., leakage principle):

$$Q = AL(h^{\text{river}} - h^{\text{gw}}) \tag{4}$$

where Q is leakage flux [$L^3\,T^{-1}$], A is the area assigned to a leakage node [L^2], L is leakage coefficient [T^{-1}], h^{river} is river stage [L], and h^{gw} is groundwater level underneath a river node.

2.2. Joint Assimilation of Piezometric Head and Temperature Data With EnKF

The ensemble Kalman filter [*Evensen*, 1994; *Burgers et al.*, 1998] is a data assimilation technique in which an ensemble of different model realizations (e.g., with different initial conditions, forcing terms, or parameters) is simulated forward and sequentially updated with measurement data. Updates can be made for the model states and model parameters which makes this methodology suitable for the real-time prediction of state variables as well as a model calibration tool. After its first applications in petrophysical and hydrogeological settings [e.g., *Lorentzen et al.*, 2003; *Naevdal et al.*, 2005; *Liu et al.*, 2008; *Hendricks Franssen and Kinzelbach*, 2008], this methodology recently gained increasing interest for the characterization of subsurface state variables and parameters. Different types of data have already been used for subsurface characterization, which include hydraulic head measurements [e.g., *Hendricks Franssen and Kinzelbach*, 2008; *Nowak*, 2009], concentration data [*Liu et al.*, 2008; *Camporese et al.*, 2011; *Li et al.*, 2012], or discharge measurements [*Camporese et al.*, 2009]. Temperature data have been shown to contain valuable information especially for the characterization of hydraulic parameters of river-aquifer systems and monitoring of this variable can be achieved at a low

operational cost. Improving the predictions of the thermal situation in the subsurface is also important for groundwater management close to streams in order to be able to estimate/prevent the pumping of too warm drinking water.

For the joint assimilation of piezometric heads and groundwater temperatures which is performed in this study the data assimilation scheme described in *Hendricks Franssen et al*. [2011] was extended so that measurements of hydraulic heads and groundwater temperatures can be assimilated jointly in the updating algorithm. The model states for the system under investigation are piezometric heads *h*, and groundwater temperatures *T* and the relevant uncertain parameters ϕ are hydraulic conductivities *K* and leakage coefficients *L*:

$$\psi = \begin{pmatrix} h \\ T \end{pmatrix} \tag{5}$$

$$\phi = \begin{pmatrix} \log_{10}(K) \\ \log_{10}(L) \end{pmatrix} \tag{6}$$

The state-parameter vector in the updating scheme of EnKF is then given as:

$$\Psi = \begin{pmatrix} \psi \\ \phi \end{pmatrix} = \begin{pmatrix} h \\ T \\ \log_{10}(K) \\ \log_{10}(L) \end{pmatrix} \tag{7}$$

The model states of Ψ are found by forward propagation of the uncertain parameter ensemble (*K* and *L*) with a numerical model that solves the relevant flow and transport equations from section 2.1.

For the update of Ψ, measurement data of piezometric heads y_h and groundwater temperatures y_T are used which are combined in the measurement vector *y*:

$$y = \begin{pmatrix} y_h \\ y_T \end{pmatrix} \tag{8}$$

For the assimilation of these measurement data with EnKF, *y* has to be perturbed with the expected measurement errors of *h* and *T* (ϵ_h and ϵ_T) (for details see *Burgers et al*. [1998]). For each realization *i*, a separate perturbation vector is drawn from a normal distribution $\mathcal{N}(0, \epsilon)$ with a mean of zero and a standard deviation ϵ that is derived from the expected measurement errors ϵ_h and ϵ_T. This yields a separate perturbed measurement vector $\hat{y}$ for each realization:

$$\hat{y}_i = \begin{pmatrix} y_h \\ y_T \end{pmatrix} + \begin{pmatrix} \mathcal{N}(0, \epsilon_h) \\ \mathcal{N}(0, \epsilon_T) \end{pmatrix} \tag{9}$$

The updating equation for Ψ is given as:

$$\Psi_i^a = \Psi_i^t + \alpha \mathbf{G}(\hat{y}_i - \mathbf{H}\Psi_i^t) \tag{10}$$

where Ψ_i^t and Ψ_i^a are the forecasted and updated state-parameter vectors (see equation (7)), $\hat{y}_i$ is the perturbed measurement vector of piezometric heads and groundwater temperatures (see equation (9)), **G** is the Kalman gain, and α is a damping factor for the parameter update that takes values between 0 and 1 and is used to reduce the effect of filter inbreeding [*Hendricks Franssen and Kinzelbach*, 2008]. The Kalman gain **G** is calculated with the following equation:

$$\mathbf{G}=\mathbf{C}\mathbf{H}^{\top}\left(\mathbf{H}\mathbf{C}\mathbf{H}^{\top}+\mathbf{R}\right)^{-1} \tag{11}$$

where **H** is a matrix that maps/interpolates the simulated states to the observation points, **C** is the full covariance matrix of the model states and uncertain model parameters, and **R** is the covariance matrix of measurement errors.

The following covariance matrix is used in the calculation of the Kalman gain for the joint update of *h*, *T*, and the parameters:

$$\mathbf{C}\mathbf{H}^{\top}=\begin{pmatrix}\mathbf{C}_{h\hat{h}} & \mathbf{C}_{T\hat{T}}\\ \mathbf{C}_{\log_{10}(K)\hat{h}} & \mathbf{C}_{\log_{10}(K)\hat{T}}\\ \mathbf{C}_{\log_{10}(L)\hat{h}} & \mathbf{C}_{\log_{10}(L)\hat{T}}\end{pmatrix} \tag{12}$$

where $\hat{h}$ are the simulated piezometric heads at observation points and $\hat{T}$ are the simulated groundwater temperatures at observation points. Note that the cross covariances between *h* and *T* are neglected in this assimilation scheme, i.e., neither *h* is used to update *T* nor *T* is used to update *h*.

2.3. Localization

For some of the simulations, the effect of distance-dependent localization [*Hamill et al*., 2001] was tested in the EnKF updating scheme. The rationale behind localization is to restrict the influence of observation points in space so that each observation point only has a certain area of influence around itself (which is defined by the length-scale variable λ). Model variables beyond this area are then not updated with measurements from this observation point. For this purpose, the Schur product (i.e., an element wise multiplication of two matrices) of the covariance matrix $\mathbf{C}\mathbf{H}^{\top}$ and a localization matrix Ω is calculated at each assimilation cycle and the so derived localized covariance matrix is then used in the calculation of the Kalman gain:

$$\left(\mathbf{C}\mathbf{H}^{\top}\right)_{ij}^{\text{new}}=\left(\mathbf{C}\mathbf{H}^{\top}\right)_{ij}\cdot\Omega_{ij} \tag{13}$$

where *i* and *j* are the matrix indices. The elements of Ω were calculated with the following localization function [*Hamill et al*., 2001; *Gaspari and Cohn*, 1999]:

$$\omega(\eta,\lambda)=\begin{cases}1-\frac{1}{4}\left(\frac{2\eta}{\lambda}\right)^5+\frac{1}{2}\left(\frac{2\eta}{\lambda}\right)^4+\frac{5}{8}\left(\frac{2\eta}{\lambda}\right)^3-\frac{5}{3}\left(\frac{2\eta}{\lambda}\right)^2 & 0\le\eta\le\lambda/2\\ \frac{1}{12}\left(\frac{2\eta}{\lambda}\right)^5-\frac{1}{2}\left(\frac{2\eta}{\lambda}\right)^4+\frac{5}{8}\left(\frac{2\eta}{\lambda}\right)^3+\frac{5}{3}\left(\frac{2\eta}{\lambda}\right)^2-5\left(\frac{2\eta}{\lambda}\right)+4-\frac{2}{3}\left(\frac{2\eta}{\lambda}\right)^{-1} & \lambda/2<\eta\le\lambda\\ 0 & \eta>\lambda\end{cases} \tag{14}$$

where η is the Euclidean distance between an observation point and an element of the state-parameter vector Ψ [*L*] and λ is a certain length scale which has to be defined for the problem [*L*]. This function has a shape similar to a Gaussian bell curve with the major difference that values beyond a fixed value (λ) for this function are zero whereas values for a Gaussian distribution are always greater than zero.

2.4. Parallelization of Assimilation Code

EnKF3d-SPRING is a C program that manages the forward calculations of the ensemble and calculates the EnKF updating step for each assimilation cycle. The actual solution of the flow and transport equations is calculated by the groundwater modeling software SPRING [*Delta h Ingenieurgesellschaft mbH*, 2006] which is called from within EnKF3d-SPRING as a library function for each realization and time step. EnKF3d-SPRING has already been used for the assimilation of hydraulic head data in *Hendricks Franssen et al*. [2011]. In order to perform the joint assimilation of hydraulic and thermal data, the data assimilation scheme implemented in EnKF3d-SPRING [see *Hendricks Franssen et al*., 2011] was extended with the updating equations given in section 2.2. However, as the computational burden for the calculation of a large number of realizations with a coupled flow and heat transport model is tremendously higher than the calculation of pure groundwater

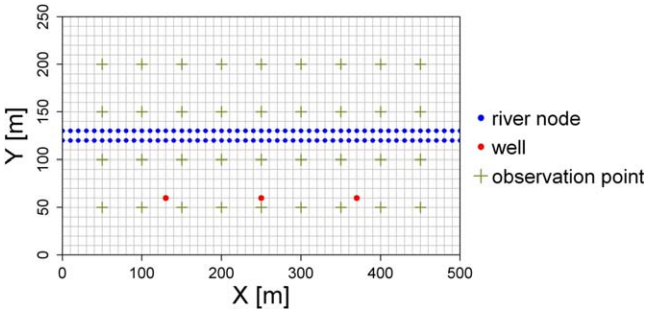

Figure 1. Model setup for synthetic experiments.

flow, it was necessary to parallelize EnKF3d-SPRING in order to retrieve results in an acceptable computation time. Principally, the parallelization of data assimilation codes can be performed at three levels:

1. Parallelizing the call of the forward runs
2. Parallelizing the forward model
3. Parallelizing the updating step

As there was no parallel version available for the forward model SPRING, only parallelization at the levels 1 and 3 could be implemented in EnKF3d-SPRING. Low-order performance tests on the supercomputing platform JUROPA at Forschungszentrum Jülich showed that the parallel version of EnKF3d-SPRING scales well up to 128 processors with an efficiency of about 70%.

3. Model Setup

3.1. Synthetic River-Aquifer Model

Assimilation experiments were first performed with a simplified synthetic model of a river-aquifer system. A sketch of the model setup is given in Figure 1. The model has a dimension of 500 × 250 × 10 m and is discretized into 50 × 25 × 10 cells. A river was placed in the middle of the model domain and is discretized into two rows of leakage nodes. Three extraction wells were placed south of the river and a regular grid of observation wells was lain out over the whole model domain. Hydraulic forcing data for the model are transient discharge of the river (expressed as river stages), transient pumping rates for the three wells, and a constant small lateral inflow/outflow at the eastern/western face of the model (assigned to the three lowest layers). Thermal forcing data are transient river temperatures.

The entire input data were based on real-world measurements for the Limmat aquifer. River stages were calculated from measured discharge data of river Sihl for the year 2006 with the modeling platform ParFlow [*Ashby and Falgout*, 1996; *Jones and Woodward,* 2001; *Kollet and Maxwell,* 2006] assuming the river geometry depicted in Figure 1. Pumping rates were taken from three bank filtration wells in the Limmat aquifer but withdrawal rates were halved in order to achieve a reasonable mass balance for the small synthetic model. River temperatures were taken from measurements of the river Sihl in 2006. A summary of the forcing data for the synthetic model is given in Figure 2.

For the synthetic experiments, a reference run with a specific K and L field (see below for details on the generation of these fields) was integrated for a 1 year period with the model described above. The h and T data sampled from this reference run (at the observation points) were then taken as input data for assimilation experiments with EnKF.

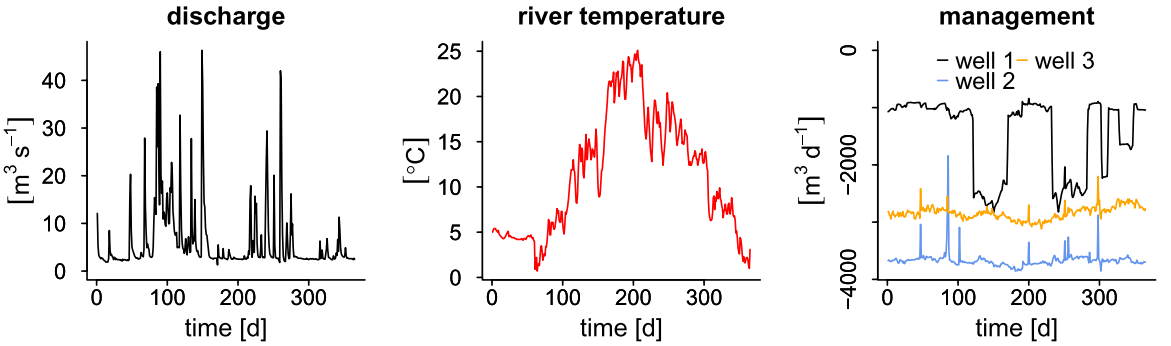

Figure 2. Forcing data for synthetic experiments.

Table 1. Parameters for Heat Transport Simulations for Synthetic and Real-World Experiments

Parameter	Value	Unit
Longitudinal dispersivity α_L	25	m
Transversal dispersivity α_T	2.5	m
ρ_b	2600	kg m^{-3}
ρ_w	1000	kg m^{-3}
κ_b	3.5	Js^{-1} m^{-1} K^{-1}
κ_w	0.587	Js^{-1} m^{-1} K^{-1}
c_b	800	Jkg^{-1} K^{-1}
c_w	4192	Jkg^{-1} K^{-1}

Table 2. Parameters for Data Assimilation With EnKF for Synthetic and Real-World Experiments[a]

	Value		
Parameter	Synthetic	Real World	Unit
N_{real}	128	128	-
N_{obs} (h)	36	87	-
N_{obs} (T)	36	22	-
α	0.1	0.1	-
Update frequency	10	10	d
ε_h	0.05	0.05	m
ϵ_T	0.1	0.1	K

[a]N_{real} is number of ensemble members and N_{obs} is the number of observation points.

The reference $\log_{10}$ (K) field was generated with Sequential Gaussian Simulation [*Gómez-Hernández and Journel*, 1993] with a mean of -2 $\log_{10}$ (ms^{-1}), a range of 100 m, and a sill of 0.5 $\log_{10}$ (m^{2}s^{-2}) using a spherical variogram. The initial ensemble of K fields was created in a similar way but with a mean value of -1 instead of -2 $\log_{10}$ (ms^{-1}) (range and sill were the same as for the creation of the reference field). The reference field for leakage coefficients L was created by reducing the K values underneath the river by one log unit and dividing this value by the discharge area of the respective leakage node. The initial ensemble of L was created in a similar way by taking the initial K values underneath the river, multiplying them by 0.5 and also dividing them by the discharge area of the corresponding leakage node. Note that K and L fields used in the data assimilation experiments are therefore slightly biased with respect to the reference fields, for this setup. Tables 1 and 2 give an overview of the settings used for heat transport simulations and for the data assimilation with EnKF. The measurement errors given in Table 2 correspond to the estimated monitoring errors of observation wells in the Limmat aquifer. The measurement error of hydraulic heads includes errors in the determination of groundwater levels as well as uncertainty related to determining the exact vertical height of the piezometer (e.g., through subsidence of the piezometer pipe). The measurement error of groundwater temperatures is an estimate of the precision of the online temperature sensors. With respect to the heat transport simulations it is also noteworthy that only temperature effects on fluid viscosity were considered in the calculations and thermal changes of fluid density were assumed to be of minor importance for the considered temperature range of the simulations. Table 3 additionally summarizes the different updating scenarios for the synthetic case.

3.2. Model and Input Data for Real-World Case

Simulations for the real-world case were performed with a 3-D-model of the Limmat aquifer in Zurich. The model domain has a spatial extent of approximately 6 $\times$ 2 km and is discretized into 92,015 model nodes, 173,599 finite elements, and 25 layers (Figure 3). The horizontal discretization varies between 1 and 50 m and the vertical discretization is set to 1.6 m. The rivers Limmat and Sihl are chosen as the northern and eastern boundaries of the model domain which is in correspondence with the geological and hydrological conditions for this site. Within the model there is an area of intensive management activities (Hardhof area) where water is pumped from several bank filtration and drinking water wells. Additionally, artificial groundwater recharge is performed through several infiltration wells and three recharge basins.

Transient boundary conditions of the flow model include recharge flux at the top of the model domain, small lateral inflows from surrounding hills on the south and north face of the model, fixed head boundary

Table 3. Simulation Scenarios for Synthetic Case[a]

	Update of:						
Scenario Name	h	T	K	L	N_{obs} (h)	N_{obs} (T)	λ (m)
SY_{uc}					0	0	-
SY_{hT}	√	√			36	36	-
SY_{hKL}	√		√	√	36	36	-
SY_{TKL}		√	√	√	36	36	-
SY_{hTKL}	√	√	√	√	36	36	-/100/200/350/500

[a]Multiple column entries indicate that the scenario was simulated with varying values for this variable.

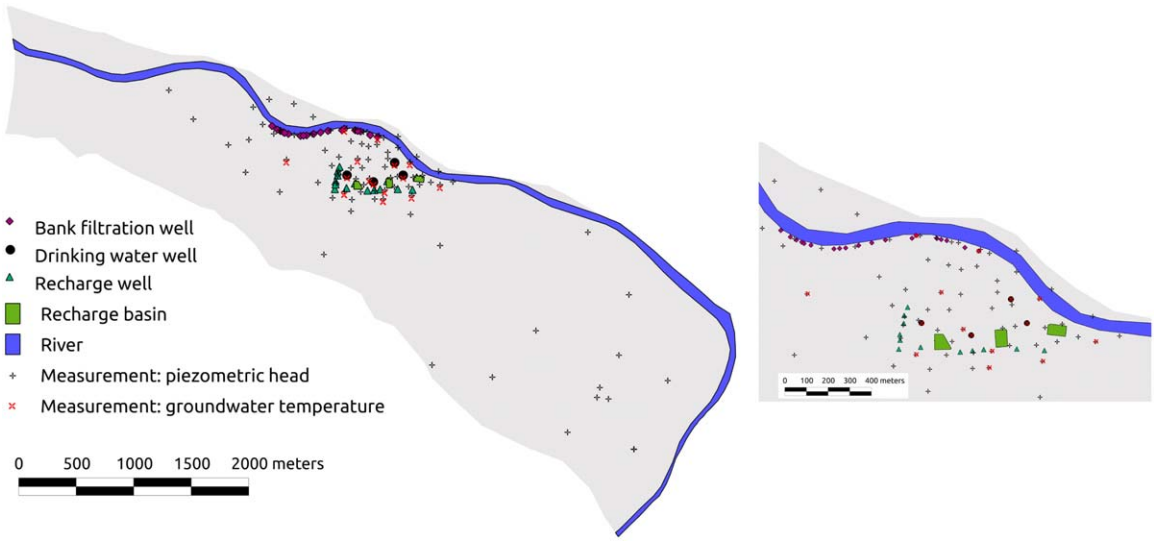

Figure 3. Position of observation points for piezometric heads and groundwater temperatures: (left) distribution of observation points over the whole Limmat aquifer model and (right) distribution of observation points in the Hardhof area.

conditions on the western face, river stages of the rivers Limmat and Sihl as well as infiltration and pumping rates of the management wells. Details on the calculation of transient boundary conditions for the flow model can be found in *Hendricks Franssen et al*. [2011]. Transient boundary conditions for the heat transport simulations include river temperatures, the temperature of injected water in infiltration wells/recharge basins, and the temperature of the first model layer which is set equal to the 20 cm soil temperature measured at the meteorological station Zurich-Reckenholz (MeteoSwiss). The temperatures for the small lateral inflows were set to a value of 13°C which is the average groundwater temperature further away from the river. The thermal regime in the model domain is mainly influenced by river-aquifer exchange and management activities and the contribution of lateral inflows only plays a subordinate role.

Simulations were performed with 128 different realizations of hydraulic conductivity fields and leakage coefficients. These model parameters are characterized by a considerable variability in natural settings and are supposed to be the most influencing factors on the model dynamics and on the propagation of uncertainty within the model. Of course, other parameters like thermal conductivities or porosities could also contribute to the uncertainty in the model but their variability is usually less pronounced than the one of permeability data. Therefore, we restricted our analysis to hydraulic parameters although this poses certain limitations on the overall assessment of model performance. The initial ensemble of hydraulic conductivities was generated on the basis of a precalibration of the flow model with 87 piezometric head data for the time periods June 2004 and July 2005 with the pilot point method including a regularization term [*Alcolea et al*., 2006]. This calibrated $\log_{10}$ (K) field served as the mean value for the $\log_{10}$ (K) ensemble. The individual ensemble members were then derived by perturbing this calibrated $\log_{10}$ (K) field with perturbation fields that were generated by Sequential Gaussian Simulation [*Gómez-Hernández and Journel*, 1993]. This approach was chosen to provide a reasonable estimate for the average $\log_{10}$ (K) within the model area in order to achieve numerically stable forward simulations. The geostatistical parameters for the creation of the perturbation fields (nugget: 0 $\log_{10}$ (ms^{-1}), sill: 0.584 $\log_{10}$ (m^2s^{-2}), range in horizontal direction: 99 m, range in vertical direction: 3.2 m, spherical variogram) were estimated from about 857 small-scale $\log_{10}$ (K)-measurements that were conducted for this area [see *Hendricks Franssen et al*., 2011]. The perturbation fields were generated on a very fine grid (1 m $\times$ 1 m $\times$ 0.01 m) and then upscaled to the simulation grid through simplified renormalization [*Renard et al*., 2000] and added to the calibrated $\log_{10}$ (K) field.

The ensemble of leakage coefficients $\log_{10}$ (L) was also generated by Sequential Gaussian Simulation. As there were no measurements of river bed conductivities available for this site the geostatistical parameters for the generation of $\log_{10}$ (L) fields were sampled from a wide range of values. A spherical variogram was used for the generation of $\log_{10}$ (L) fields with a range sampled from a uniform distribution between 50 and 5000 m, a sill sampled from a uniform distribution between 0.1 and 2 $\log_{10}$ (m^2s^{-2}) and a nugget of 0 $\log_{10}$

Table 4. Simulation Scenarios for Real-World Case[a]

	Update of:								
Scenario	h	T	K	L	N_{obs} (h)	N_{obs} (T)	λ (m)	ϵ_h(m)	ϵ_T(°C)
RW_{uc}					0	0	-	0.05	0.1
RW_{hT}	√	√			87/40	22/11	-	0.05	0.1
RW_{hKL}	√		√	√	87	22	-	0.05	0.1
RW_{hTKL}	√	√	√	√	87/40	22/11	-/2000/3000	0.05/0.25/0.5	0.1/0.5/1.0

[a]Multiple column entries indicate that the scenario was simulated with varying values for this variable.

(ms^{-1}). Note that for these geostatistical simulations we used a modified formulation of the leakage coefficient (L^*) which is L multiplied with an estimation of the river width for the particular river node. This is done because L^* is the direct input parameter required by the model. Throughout the rest of this work we will use L and L^* interchangeably because the river width is relatively uniform along the river reach and we will not discuss on the absolute value of this model parameter.

Within the model area 87 observation points for piezometric heads and 22 observation points for groundwater temperatures are available. The spatial position of these observation points is shown in Figure 3. Hydraulic head observations are distributed over the whole model domain with the highest density in the Hardhof area where most of the management activities take place. The observation points for groundwater temperatures are only clusters in the Hardhof area mostly between the recharge basins and the river in different depths.

Data assimilation experiments were conducted for different settings of EnKF which include the update of states only, the joint update of states and parameters and the use of localization. Tables 1, 2, and 4 give an overview of the standard settings for heat transport calculations, EnKF, and the performed simulation scenarios.

Online sensors for measuring groundwater temperature are available from autumn 2005. Therefore, heat transport simulations were performed with data ranging from 2006 to 2012. Four simulation periods were distinguished. The year 2006 was used as a warm-up period for the model in order to have a more realistic estimate of the temperature distribution within the aquifer. This spin-up was conducted with the initial parameter ensembles of K and L without data assimilation. The initial h and T fields for the spin-up period were generated with steady state calculations using averaged parameter fields as input. The starting value of groundwater temperature for this steady state calculation was set to 13°C for the entire aquifer which is roughly the mean groundwater temperature measured further away from the rivers. The final h and T fields from the spin-up period for each ensemble member were then used as initial conditions for the different assimilation experiments that were conducted for the year 2007. The updated parameter ensembles for this assimilation period were then used in validation runs to additionally access the performance of the different assimilation strategies for parameter estimation. From January 2008 to October 2010 some major reconstructions were performed in the Hardhof area. These activities included the utilization of additional pumping wells and nonstandard management activities for which only incomplete information was available. Therefore, this time period could not be used for the validation of the different assimilation scenarios and the validation simulations were done for the hydrological year 2011 (November 2010 to October 2011). The period from January 2008 to October 2010 was used as a spin-up for the validation period (November 2010 to October 2011) which started with the final results (state ensembles and updated parameter ensembles) from the assimilation period.

Table 5 summarizes the different time periods for the heat transport calculations and Figure 4 gives an overview on the forcing data that were used for the initial spin-up, the assimilation period and the verification period.

3.3. Performance Measures

The performance of the different assimilation runs was evaluated with the root mean square error (RMSE) of model states (h and T) or model parameters (K and L). In the case of the synthetic study, RMSE is calculated over all model nodes or model elements because the values for the reference are known exactly. For the real-world application, RMSE is calculated only for states at measurement locations because these are the

Table 5. Time Periods for Heat Transport Simulations (Real-World Case)

Period		Purpose	Initial h/T	Initial K/L
(I)	1/1/2006–31/12/2006	Initial spin-up	Steady state	Initial ensemble
(II)	1/1/2007–31/12/2007	Assimilation period	From unconditional simulation (I)	Initial ensemble
(III)	1/1/2008–31/10/2010	Spin-up verification period	Final ensembles from (II)	Final ensembles from (II)
(IV)	1/11/2010–31/10/2011	Verification period	Final ensembles from (III)	Final ensembles from (II)

only available data sources for evaluating the performance of the simulations. RMSE is usually calculated per time step and with respect to the mean value of the specific variable. A general formula for the calculation of RMSE for variable x for time step t is:

$$\mathrm{RMSE}\,(x,t)=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\bar{x}_i(t)-x_i^{\mathrm{ref}}(t)\right)^2} \tag{15}$$

where N is the number of model nodes/model elements for synthetic experiments and the number of observation points for real-world applications. x^{ref} are values from reference runs in synthetic experiments and measurement values for the real-world case.

4. Results

4.1. Synthetic Experiments

The different updating scenarios are first compared with respect to the deviation of state variables from the reference values. Figure 5 shows the RMSE of piezometric heads and groundwater temperatures for the different scenarios. If only states are updated, there is only a minor improvement compared to unconditional simulations. This is related to the bias in the parameter values of the initial K and L ensembles. When piezometric heads are jointly updated with parameters (scenarios SY_{hKL} and SY_{hTKL}) errors are significantly reduced. For these scenarios, the joint update of piezometric heads and groundwater temperatures leads to lower RMSE(h) values in the first phase of the simulation period but is outperformed by SY_{hKL} in the final stage of the simulations. RMSE(T) shows a seasonal variation for all scenarios as opposed to RMSE(h). The highest values for RMSE(T) occur during the summer months when higher river temperatures propagate into the aquifer. These higher deviations between reference and ensemble predictions stem from a mismatch concerning the location of the heat plume between the reference and the ensemble simulations (due to the different parameterization of K and L) and from the higher absolute temperature difference between river and aquifer. The lowest RMSE(T) is found for scenario SY_{hTKL} where all states and uncertain parameters are jointly updated. For the scenario without an update of T (scenario SY_{hKL}), temperature predictions are similar to unconditional simulation and partly show a higher RMSE(T). Thus, a parameter update which is only based on hydraulic data might have little impact on prediction of transport processes.

The effect of h and T assimilation on parameter updates can be seen in Figure 6 which shows RMSE of K and L for scenarios SY_{hKL}, SY_{hTKL}, and SY_{TKL}. When piezometric head data are used to update parameters

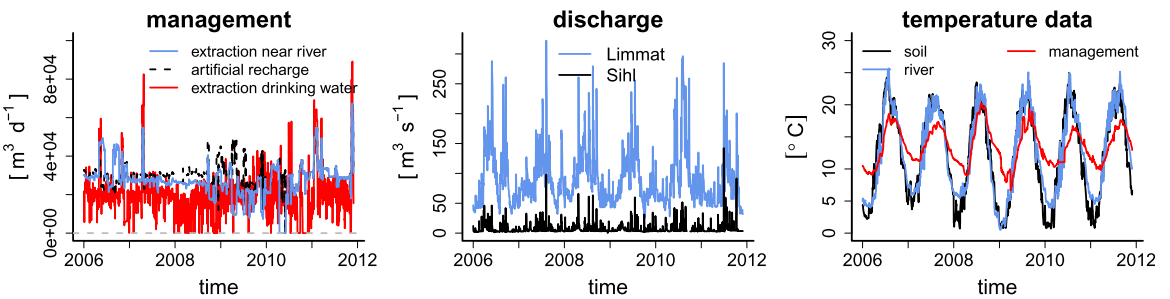

Figure 4. Forcing data of the real-world study for the years 2006–2011.

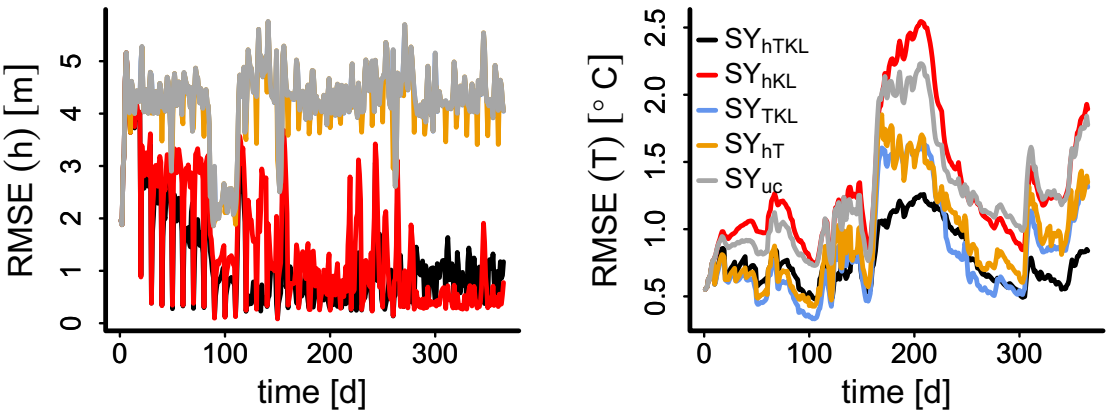

Figure 5. Root mean square error of (left) updated piezometric heads and (right) groundwater temperatures for the synthetic test case and different updating scenarios (Table 3).

there is a fast decrease of RMSE within the first assimilation cycles. This decrease is mainly caused by a correction of the bias between the reference fields and the initial ensembles. When only temperature data are used to condition the parameter fields (scenario SY_{TKL}) there is almost no correction of parameter values which explains the similarity between scenario SY_{hT} and SY_{TKL} in terms of temperature distribution and RMSE(*T*). From Figure 6, it can also be seen that the joint assimilation of *h* and *T* gives slightly better parameter estimates for *K* and *L* than the scenario where only *h* data are used. The conditioning of *K* and *L* on *h* leads to a RMSE(*K*) reduction of 30.0% and a RMSE(*L*) reduction of 73.1% whereas a joint conditioning on *h* and *T* gives a RMSE(*K*) reduction of 36.7% and a RMSE(*L*) reduction of 77.3%.

Figure 7 compares the final ensemble means of *K* for scenarios SY_{hKL}, SY_{hTKL}, and SY_{TKL} with the reference field. For scenario SY_{TKL}, there are only minor updates compared to the initial ensemble which do not correspond very well with the distribution of *K* in the reference field. Therefore, temperature data alone seem not to be very sensitive on the distribution of *K* values for this synthetic model setup. When piezometric head data are assimilated with EnKF (scenarios SY_{hKL} and SY_{hTKL}) the ensemble mean of *K* is much closer to the reference field. *K* values are mainly updated along the river reach because the highest sensitivity of the model is related to river stage fluctuations. For scenario SY_{hKL}, *K* values along the river reach are lower than for scenario SY_{hTKL} which causes slightly higher RMSE values for scenario SY_{hKL}. The different magnitude of *K* values along the river reach for scenarios SY_{hKL} and SY_{hTKL} is probably related to their different updating behavior with respect to *L* (Figure 8). Here it can be seen that the uncertainty in the final *L* ensemble is higher for SY_{hKL} than for SY_{hTKL}, i.e., with assimilation of *h* data alone the distribution of *L* could not be constrained as well as with a joint assimilation of *h* and *T*. As a result, also the residuals for *K* were higher for

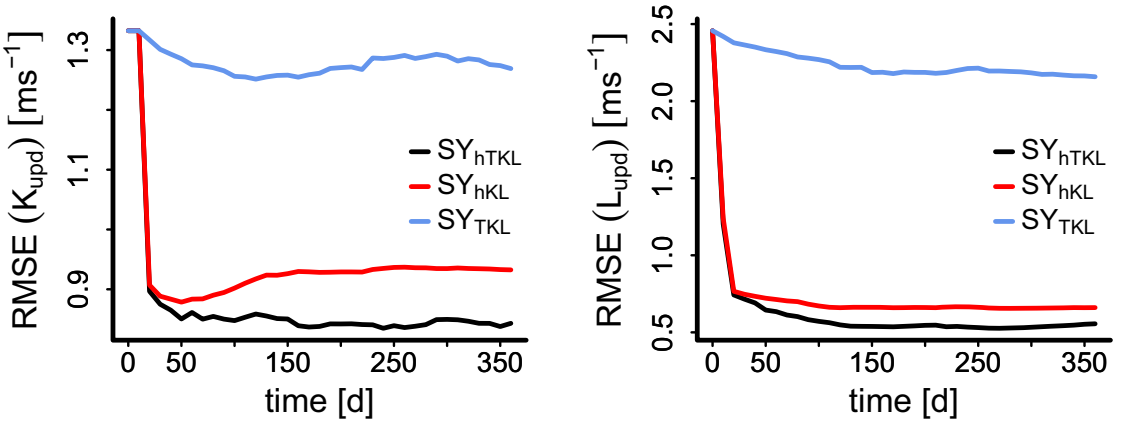

Figure 6. Root mean square error for (left) updated hydraulic conductivities and (right) updated leakage coefficients for scenarios SY_{hKL}, SY_{hTKL}, and SY_{TKL} (Table 3).

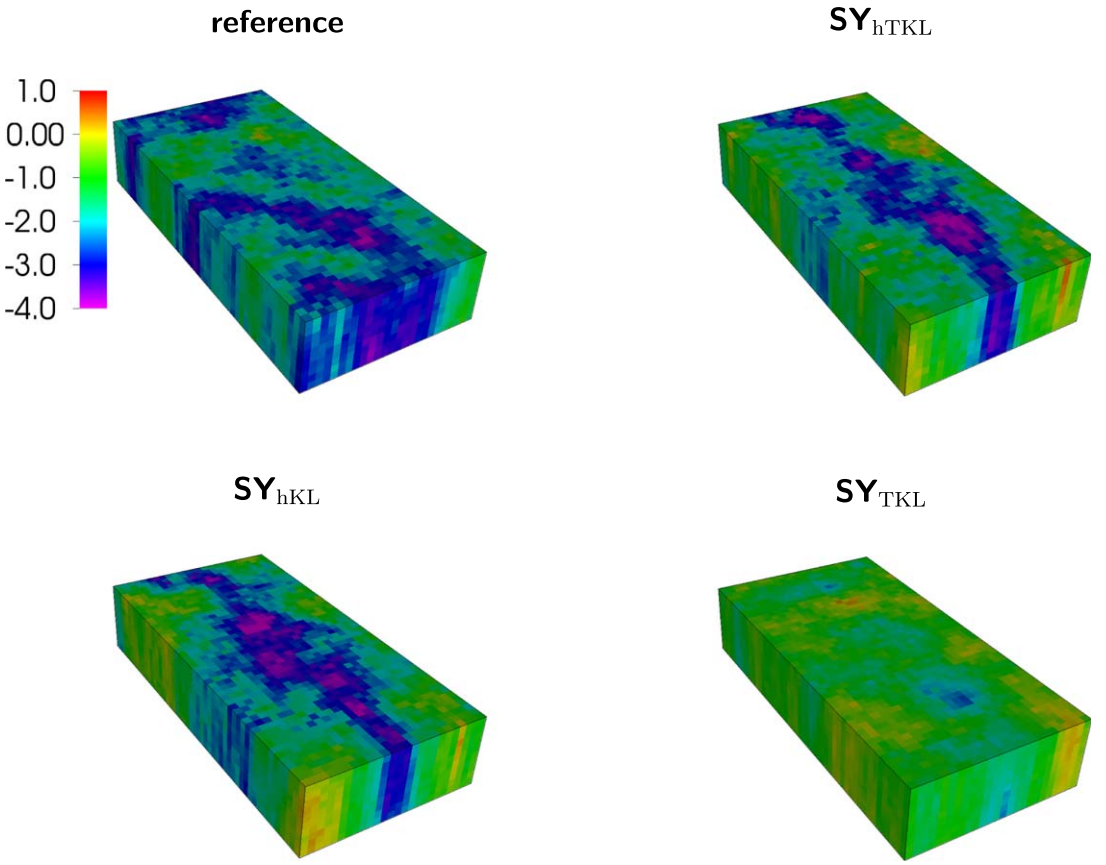

Figure 7. Updated fields of hydraulic conductivities (end of simulation period) for scenarios SY_{hKL}, SY_{hTKL}, and SY_{TKL} (Table 3).

scenario SY_{hKL} compared to SY_{hTKL}. From Figure 8, it also becomes obvious that the assimilation of *h* mostly led to a correction of the bias of the initial *L* ensemble whereas the assimilation of temperature data (scenarios SY_{TKL} and SY_{hTKL}) led to a more precise determination of the spatial structure of leakage coefficients. The main advantage of using a joint assimilation of *h* and *T* for the synthetic experiments is that both state variables have a different information content which allows a better estimation of leakage parameters. *h* data provide information on the net exchange between river and groundwater and thus are sensitive to the bias between initial ensemble and the reference values. *T* data give additional information on the spatial location of exchange fluxes and therefore constrain the spatial structure of leakage parameters.

In a next step, the effect of localization on parameter updates was tested. For this purpose, scenario SY_{hTKL} was rerun with different λ values of the localization function (equation (14)) ranging from 100 to 500 m. The lower bound of λ of 100 m corresponds to the geostatistical range parameter of the generated *K* fields (reference and ensemble) and the upper bound of λ of 500 m corresponds to the size of the model domain. RMSE of temperature predictions and *K* values for these simulation runs are shown in Figure 9. The simulation with a localization limit of 100 m consistently performed worse compared to the simulation without any localization. In this case, the region of influence for the observations is too restricted and only the model cells adjacent to the grid cell of observation points are updated. When the localization limit is increased there is no observable effect on simulated groundwater temperatures but the estimation of *K* values is improved compared to the simulation without localization. It also can be seen from Figure 9 that the estimation of *K* gets slightly worse again when the localization limit is increased from 350 to 500 m. For the simulation without localization there are some parts where *K* values were highly increased during the update (e.g., at the most right corner of the model domain, see Figure 7). However, these zones with high *K* values are not present in the reference field and are possibly caused by spurious correlations related to the limited ensemble size of 128 members. When localization is used, these zones of elevated *K* values are eliminated or at least reduced during the updating procedure whereby the general structure of updated *K*

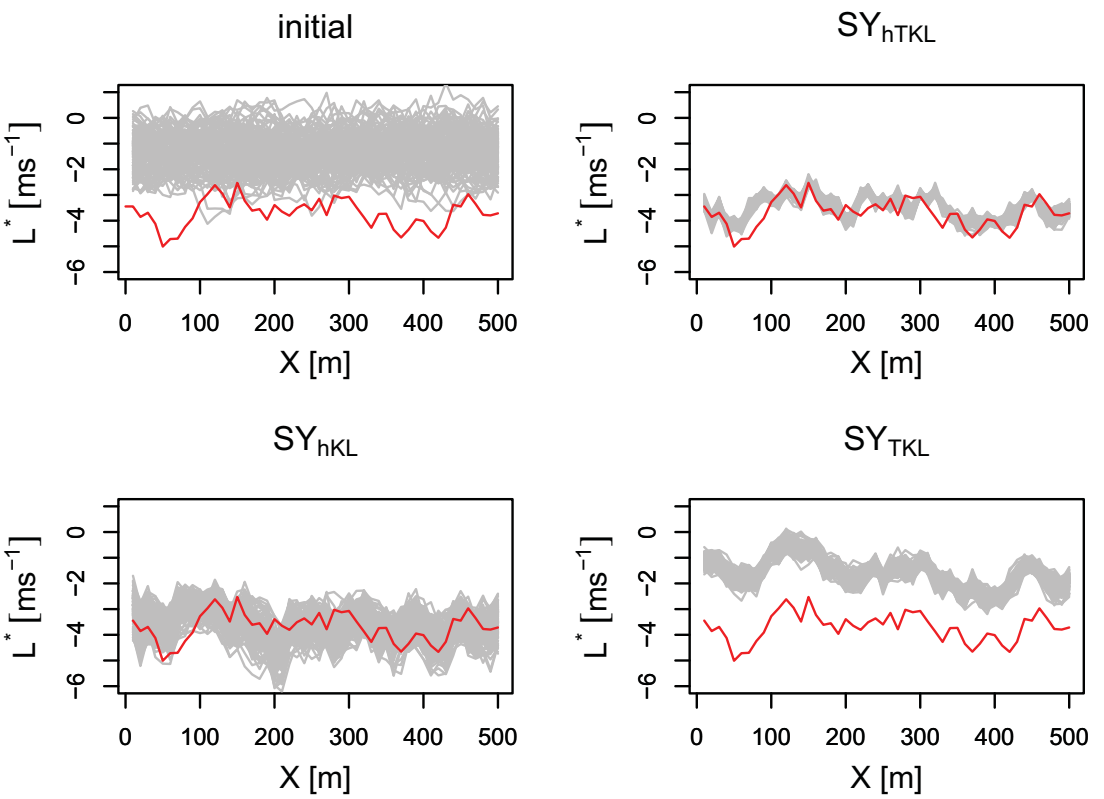

Figure 8. Updated fields of leakage coefficients (end of simulation period) for scenarios SY_{hKL}, SY_{hTKL}, and SY_{TKL} (Table 3).

values (i.e., reduced values along the river reach) is preserved. This finally leads to lower RMSE(*K*) values for simulations with localization.

4.2. Assimilation of Groundwater Temperature Data for Real-World Case

The results for the assimilation of temperature data for the real-world case cannot be evaluated for the whole model domain because only data for the measurement locations are available. Additionally, for these simulations the true parameter distribution for *K* and *L* is unknown and the evaluation of updated parameter values is restricted to visual inspection of the updated parameter fields. Therefore, the errors of *h* and *T*

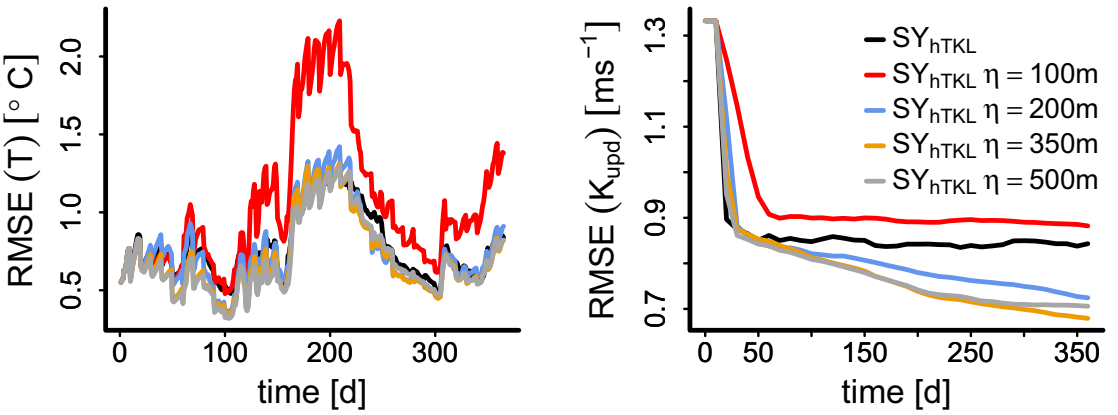

Figure 9. Root mean square error of (left) groundwater temperatures and (right) hydraulic conductivities for different localization distances (scenario SY_{hTKL} in Table 3).

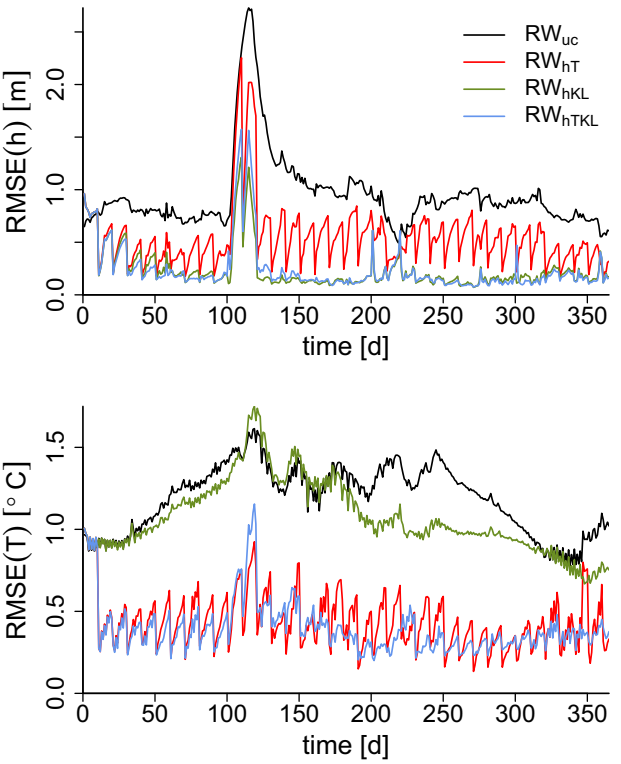

Figure 10. Temporal evolution of RMSE of *h* and *T* for different updating scenarios (see Table 4).

are additionally monitored for a validation period in which the quality of parameter updates for the different updating scenarios is compared on the basis of *h* and *T* predictions.

4.2.1. Assimilation Period: Effect of Updating Scheme

Figure 10 compares the basic updating scenarios (see Table 4) in terms of the temporal evolution of RMSE values for *h* and *T* and Figure 11 summarizes the statistics of RMSE values at observation points (calculated for the whole assimilation period). All RMSE(*h*) time series are characterized by some day to day fluctuations due to the transient boundary conditions and by a sharp increase around day 110 which is caused by a period of increased pumping rates. For RW_{uc}, the overall errors are relatively high due to the highly uncertain *K* and *L* ensembles. For all updating scenarios, the assimilation with EnKF improves the prediction of *h* and *T* compared to unconditional simulations. Piezometric heads at measurement locations show high temporal fluctuations when only state variables are updated. An additional parameter update leads to relatively constant values of RMSE(*h*) and there is no observable effect of temperature assimilation on RMSE(*h*). The average improvement of RMSE(*h*) with parameter update is about 74% (compared to RW_{uc}) leading to average RMSE(*h*) values of about 0.33 m. For the real-world case, the seasonal cycle of RMSE(*T*) is less pronounced than for the synthetic case which is probably attributed to the more complex model dynamics in the real-world case. In addition, for RMSE(*T*), the effect of parameter update is smaller than for RMSE(*h*) because the mean values for scenario RW_{hT} are very similar to the mean values of scenario RW_{hTKL}. RW_{hT} shows an average RMSE(*T*) improvement of 62.5% (compared to RW_{uc}) whereas the average reduction for RW_{hTKL} is 65.8%. The main difference between scenarios RW_{hT}

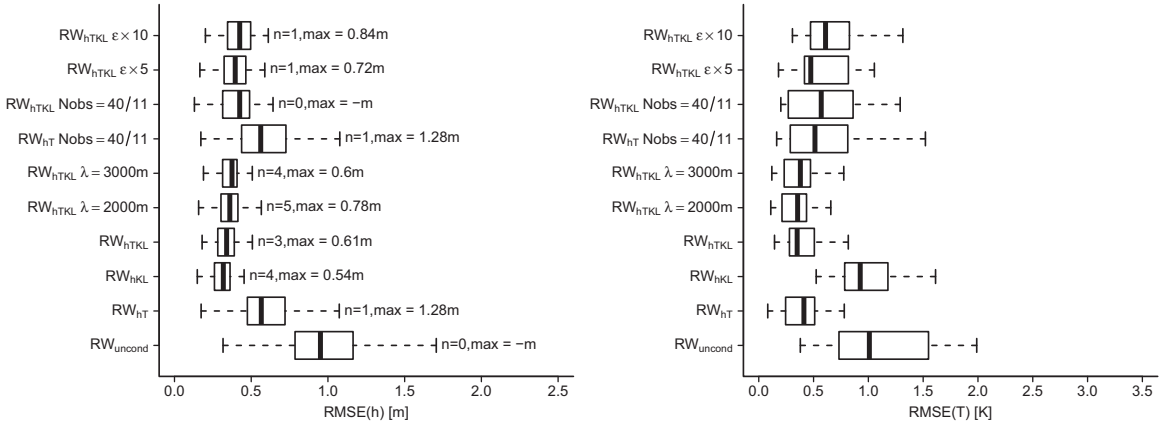

Figure 11. Statistics of root mean square error of (left) mean hydraulic heads and (right) mean groundwater temperatures for all observation points (calculated for the whole assimilation period). Numbers on the right-hand side of boxplots for hydraulic heads refer to the number of outliers (n) and the value of the highest outlier (max).

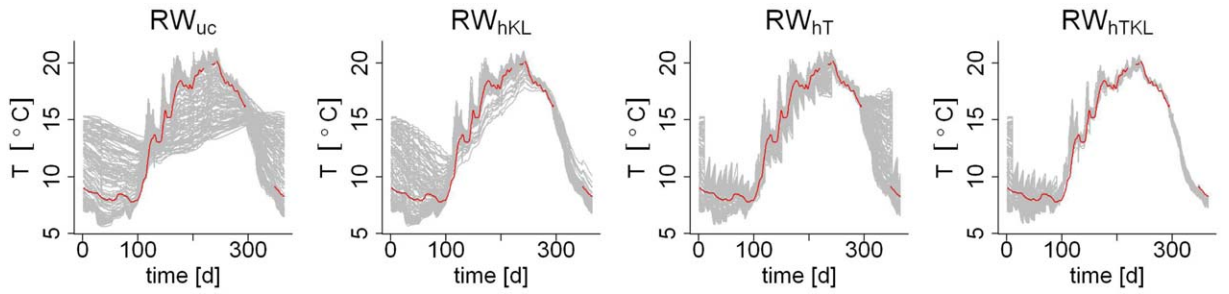

Figure 12. Evolution of groundwater temperatures during the assimilation period for one observation point with partly missing data. Red lines display measured groundwater temperatures and gray lines display different realizations of the ensemble.

and RW_{hTKL} lies in the temporal fluctuations of RMSE(T) which are higher for scenario RW_{hT}. The different effect of parameter updates for h and T is also apparent from the statistics of the residuals. For piezometric heads, there is a significant discrepancy between the median of residuals for scenarios RW_{hT} and RW_{hKL}/RW_{hTKL}. Also the variability of errors is higher when only states are updated. In contrast, for groundwater temperatures there is not so much difference in the error distribution between scenarios RW_{hT} and RW_{hTKL}. However, the parameter update becomes important also for T predictions when there are gaps in the time series of temperature measurement data. This is exemplified in Figure 12 which shows the temperature evolution for an observation point with partly missing measurement data. Here it can be seen that the missing reduction of parameter uncertainty in scenario RW_{hT} can lead to much larger ensemble spread when no measurement data are available.

Final hydraulic conductivity fields differ only marginally between RW_{hKL} and RW_{hTKL} (data not shown) but an effect is visible for the final parameter distribution for leakage coefficients (Figure 13). Here an additional assimilation of groundwater temperatures led to an overall reduction of uncertainty of the L ensemble. However, the basic structure of updated L fields is still very similar to the scenarios without T assimilation.

4.2.2. Assimilation Period: Effect of Observation Density and Data Quality

In a next step, the effects of observation density and measurement errors on the update of groundwater temperatures and piezometric heads were explored in further detail. The assigned values of measurement errors for h and T that were used for the previous simulations ($\epsilon_h = 0.05\text{m}, \epsilon_T = 0.1°\text{C}$) can be seen as optimal values for the measurement devices for the site. However, the accuracy of online sensors may also decrease over time due to alteration of the measurement device or drifts in the calibration function. Also the flow conditions within the bore hole and the support volume of the measurement device may have an effect on the accuracy of the measurements which are hard to quantify and were not regarded in the previous used values of measurement error. Therefore, the measurement errors for h and T were multiplied with a factor of 5 and 10 and results for the different measurement errors were compared for scenario RW_{hTKL} (Figure 14). Generally, the increase of measurement errors led to higher errors for h and T. This is visible from the temporal evolution of RMSE and also from the higher spread of the residuals. This occurs because increasing measurement uncertainty leads to lower weights in the Kalman gain matrix and thus to a decreasing

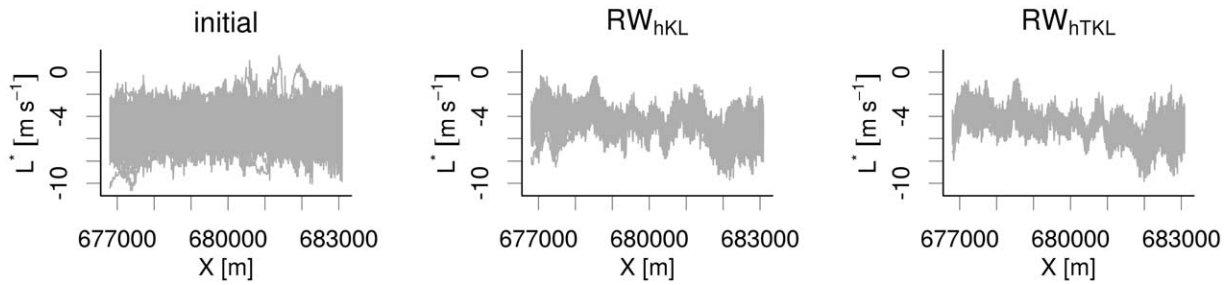

Figure 13. Initial and updated fields of leakage coefficients after the last assimilation cycle. Updated leakage coefficients are shown for a scenario (middle) without and (right) with assimilation of groundwater temperatures.

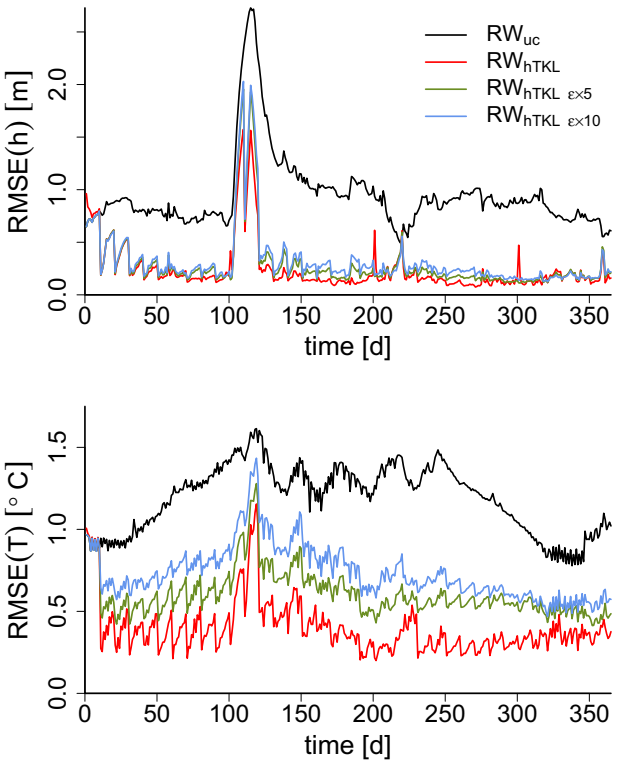

Figure 14. Temporal evolution of RMSE of *h* and *T* for different measurement errors. $\epsilon\times5$ and $\epsilon\times10$ means a 5 or 10 times higher measurement error compared to the settings give in Table 2.

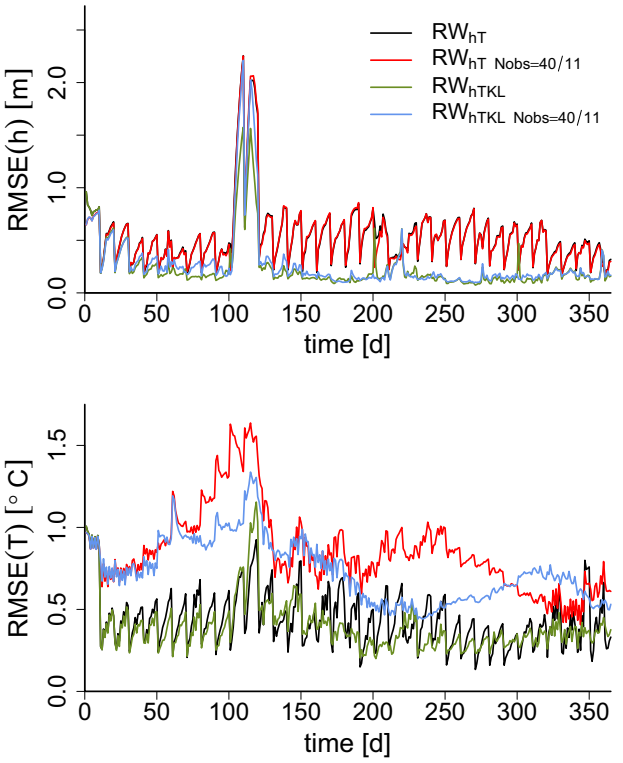

Figure 15. Temporal evolution of RMSE of *h* and *T* for different observation densities with and without parameter update.

adaptation toward the measurements. However, the effect of measurement errors is slightly different for *h* and *T*. For piezometric heads, the RMSE does not differ so much between the scenarios with different measurement errors (ΔRMSE(h) $\sim$ 0.08m). For groundwater temperatures, measurement errors have a higher impact on the assimilation (ΔRMSE(T) up to 0.33°C). This discrepancy is related to the different prediction uncertainty for *h* and *T*. After the first few assimilation cycles the ensemble spread for groundwater temperatures is very low for most of the observation points and EnKF therefore assigns a high confidence to temperature predictions. Forecasts for piezometric heads show more variability and therefore the relative increase of measurement error has less effect on the update compared to groundwater temperatures.

The effect of observation density was investigated for scenarios RW$_{hT}$ and RW$_{hTKL}$ by reducing the number of observation points for *h* and *T* to about half of the values that are available for the site. The results for these scenarios are compared in Figure 15. For piezometric heads there is not much difference for the scenarios with low and high observation density. Only when parameters are included in the assimilation with EnKF there is some tendency for a higher spread in the error statistics when a lower amount of measurements is used. For groundwater temperatures, the situation is very different compared to piezometric heads. In this case the errors rise significantly when only half of the observation points are available for conditioning with EnKF. This relationship can be found for both updating scenarios (states only and state-

parameter update) with a slight tendency toward higher errors for the state-only scenario. The relatively low sensitivity toward observation density for piezometric heads is probably an effect of the higher abundance of piezometric head measurements for this site (87) compared to groundwater temperature data (22). Using only 40 piezometric head measurements instead of 87 thus still gives enough information of system dynamics and there is a certain amount of redundant information for the 87 piezometers. An additional reason for the different error statistics of h and T with respect to observation density are the different correlation lengths of both variables. Groundwater levels for the site are relatively smooth due to the high permeability of the unconfined aquifer whereas groundwater temperature is mainly driven by river-aquifer exchange due to pumping and by the artificial recharge which leads to small-scale variability of the temperature distribution within the model domain. As a consequence, a single measurement of piezometric head gives more information on the surrounding area than a measurement of groundwater temperature. This affects the critical amount of data that is necessary to constrain the system states with EnKF.

4.2.3. Assimilation Period: Effect of Localization

In a next step, it was tried to identify the effect of localization on the state-parameter updates with EnKF. Localization with two different λ values (2000 and 3000 m) was compared with the standard assimilation scheme for scenario RW_{hTKL}. RMSE values for the different localization scenarios indicate that the errors of piezometric heads slightly increase when localization is used. In contrast, the observed errors for groundwater temperature are not so sensitive to localization. For a localization length scale λ of 2000 m there is a slight decrease in error variance but the median value is approximately the same as for the simulation without localization.

A reason for the different behavior of h and T with respect to localization could be the different correlation length of both variables. A comparison of the correlations length of h and T for various observation points and time steps shows that the covariances for h have a much longer range than for T ($\sim$1000 m versus $\sim$400 m). When such a long range covariance function is then tapered with a localization function that has a smaller spatial extent, information may be lost in the assimilation with EnKF. For groundwater temperatures, the spatial correlation is much lower and therefore fits better within the range of the utilized localization function. The effect of localization is also visible for the parameter updates. For example, without localization the variability in the L ensemble decreased for the whole river reach and a distinct spatial pattern is visible in the final ensemble. When localization is used, only the part of the river that is close to the management activities shows a significant reduction in ensemble variance because the density of observation points is highest here. In addition, the observation points are also much closer to the river in this area than in other parts of the model. The spatial distribution of updated mean hydraulic conductivities at the end of the assimilation period is visualized in Figure 16. It is visible that the parameter update for K with unlocalized EnKF leads to a relatively patchy structure and a high degree of spatial variability where mean K values between neighboring cells can exhibit considerable contrasts. On the one hand, this is related to the large number of degrees of freedom for the parameter update with EnKF, i.e., observation data are used to update the parameters of each grid cell separately to reduce the misfit between observations and forecasted state variables. On the other hand, spurious correlations due to the limited ensemble size may also play a role in this context and could emphasize a certain randomness in the assimilation procedure. Localization generally leads to a smoothing of the updated K fields. It can be seen that the scenarios with and without localization yield approximately the same spatial structure of K fields at the end of the assimilation period. However, localization leads to less extreme K values and the transition between neighboring cells is much more gentle than for the update with unlocalized EnKF. Note however that the ensemble variance at individual cells is tendentially higher for the simulations with localization (especially further away from observation points).

4.2.4. Validation Period

In this section, the quality of the updated parameter ensembles from the assimilation period is assessed with validation runs for the hydrological year 2011. Figure 17 shows the errors for h and T for different updating scenarios (unconditional, with updated parameters on the basis of h only and with updated parameters also conditioned to T) and a variation of observation density (87/22 versus 40/11 observation points for h and T in scenario RW_{hTKL}) over time. Generally, the overall RMSE(h) values for the validation period are relatively high compared to the assimilation period. Figure 18 additionally shows the distribution of RMSE(h) at observation points for the different updating scenarios. It can be seen that the relatively high

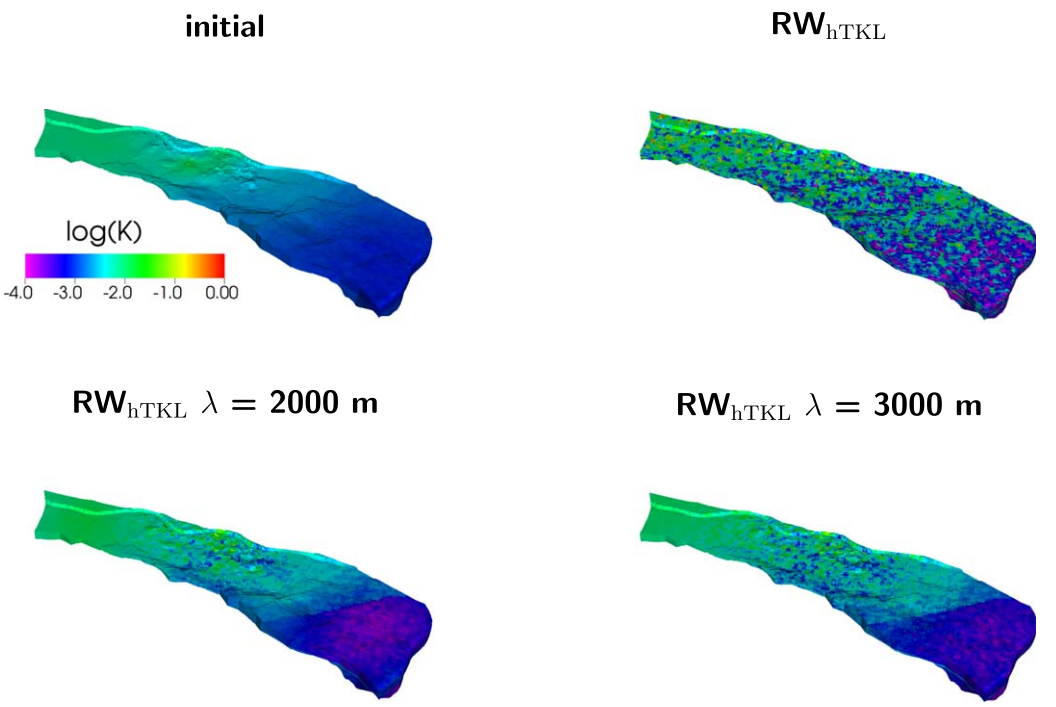

Figure 16. Mean fields of hydraulic conductivity after the last assimilation cycle for parameter update without localization and with localization for two localization length scales λ (2000 and 3000 m). As a comparison the average initial field of hydraulic conductivities is shown.

overall errors are caused by a certain percentage of outliers and that the median RMSE(*h*) values are much lower than the RMSE(*h*) that is calculated for all observation points (Figure 17). These outliers are mainly located in the eastern part of the model domain where the observation density is low and the groundwater dynamics are not much influenced by the management activities. These relatively high RMSE(*h*) values point to a systematic model error in this region that is probably caused by the unsaturated conditions that are present in the first model layers which lead to some nonlinearities in this model region and by an insufficient parameterization of leakage parameters. However, when the scenarios with updated parameters are compared with the unconditional simulation they show a significant reduction in the overall RMSE(*h*) (on average 0.67 m or 49%) and in the median values of RMSE(*h*) (approximately 0.40 m or 59%). For groundwater temperatures the updated parameter ensembles do not show such a large RMSE reduction as piezometric heads. An improvement is clearly visible but it is lower in magnitude compared to piezometric

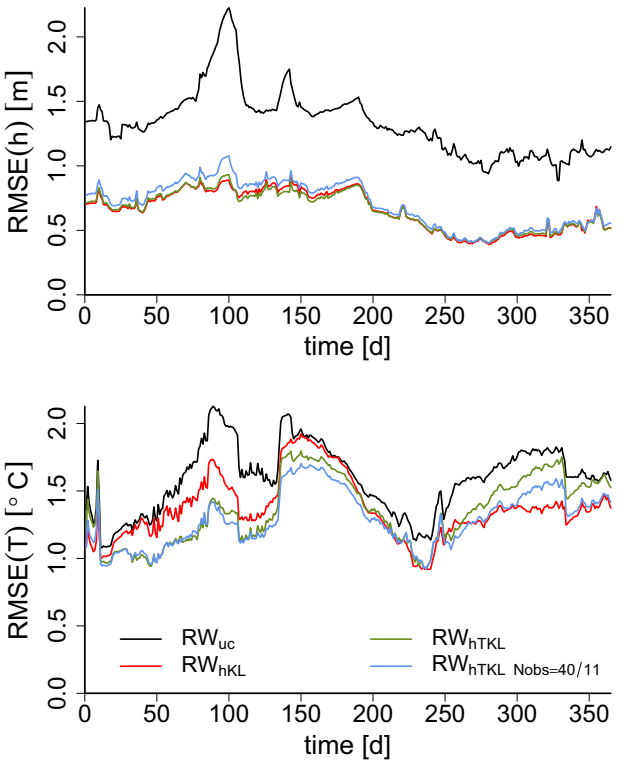

Figure 17. Temporal evolution of RMSE of *h* and *T* for different updating scenarios (validation period).

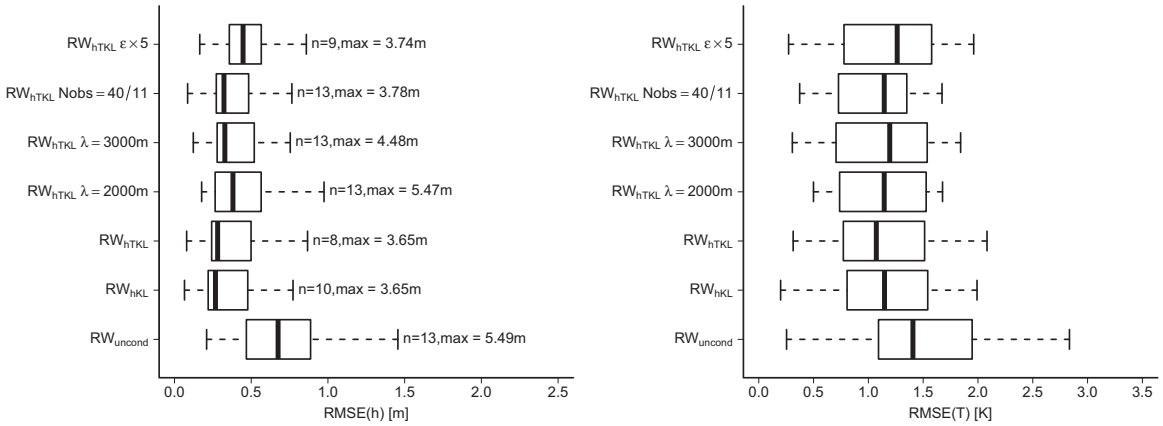

Figure 18. Statistics of root mean square error of (left) mean hydraulic heads and (right) mean groundwater temperatures for all observation points (calculated over the whole validation period). Numbers on the right-hand side of boxplots for hydraulic heads refer to the number of outliers (n) and the value of the highest outlier (max).

head data (average RMSE(*T*) reduction: 0.2°C or 15%). From the temporal evolution of RMSE(*T*) it can not clearly be distinguished which of the updating scenarios performs better throughout the validation period. In the first phase the scenarios with a *T* update in the assimilation period have lower errors than the ensemble without *T* update. However, in the last phase of the validation period this relationship reverses and the ensemble for RW_{hKL} gives better results than the one for RW_{hTKL}.

When the measurement error is increased during the assimilation period, RMSE(*h*) increases for an event of increased groundwater withdrawal around time step 100 but is very similar to the ensemble with lower measurement error for the rest of the validation period. For scenario $RW_{hTKL,}$ the reduction of observation density lead to a slight increase in RMSE(*h*) (see Figure 18). However, RMSE(*T*) values were equal or lower for the simulations with reduced observation points compared to the simulation with all available observation points.

Localization has almost no effect in terms of RMSE(*T*), only the spread of residuals is a bit higher for the scenarios with localization. A stronger effect of localization was found for RMSE(*h*) where the ensembles with localization show significantly higher values. The relatively constant bias between the scenarios with and without localization suggests that there is a systematic misinterpretation in the groundwater dynamics for the localized ensembles. This systematic bias could be related to the effect of localization on the estimation of *L* fields. When localization is used, the parameter fields for *L* are not constrained as well as for the simulations without localization. This leads to a higher uncertainty in simulated exchange fluxes which directly affects the mass balance in the aquifer and biases the predicted groundwater levels. In order to investigate this in more detail we repeated the scenario with a localization length of 2000 m with variable localization length scales for each state/parameter couple. We set the localization length scale between *L* and the states variables to a value of 10,000 m in order to reduce the effect of localization on the leakage parameters. The correlation length scale between *h* and *h*/*K* were increased to 5000 m/3000 m and the correlation length scale between *T* and *T*/*K* remained at 2000 m. Results showed that the updated parameter fields with this setup lead to similar RMSE(*h*) values as the simulation without localization (for the validation period). At the same time, RMSE(*T*) is comparable to the simulation with a uniform localization length scale of 2000 m.

5. Discussion

Results for the synthetic experiments suggest that a joint assimilation of piezometric heads and groundwater temperatures with EnKF principally can lead to an improvement in the estimation of subsurface parameters compared to an assimilation of piezometric head data alone. Assimilated temperatures mainly gave additional information on the spatial distribution of river-aquifer exchange and the corresponding leakage parameters whereas assimilated piezometric head data gave information on the overall magnitude of water exchange. This relationship was not so clear for the real-world data of the Limmat aquifer. In this case, the additional assimilation of groundwater temperatures also led to a certain reduction of the uncertainty for the spatial distribution of leakage coefficients compared to an assimilation of piezometric head

data alone. However, this additional restriction of parameter values did obviously not have a persistent effect on the state predictions. This could be seen from the validation runs with updated parameter fields with and without assimilation of groundwater temperatures which gave very similar results in terms of temperature and piezometric head predictions. Generally, the model for the real-world case seems to already capture the most important system dynamics with respect to heat transport. This becomes obvious from the relatively low errors for the unconditional simulations (average RMSE(T) of about 1.5°C) and also from the simulated values at the observation points which already capture the seasonal temperature dynamics that were measured.

However, from the assimilation experiments for the real-world case it also became obvious that the model dynamics for some parts of the model could not be corrected well with EnKF. This observation especially refers to the temperature dynamics monitored at two multilevel piezometers south of the recharge basins and recharge wells. At these locations, the model indicated a small seasonal temperature variation during the assimilation and validation period but the measured groundwater temperatures are nearly constant for both periods. In the simulations these piezometers seem to be influenced by the artificial recharge which causes the seasonal temperature variation but this does not reflect the true conditions in that area. EnKF was obviously not able to adapt model parameters to adjust model dynamics for this area toward the true conditions. From geological surveys for the Hardhof area it is known that there are some channel structures within the Hardhof area which facilitate preferential flow conditions for a part of the managed site. These geological structures are not included in the initial ensemble of K values which was assumed to be multi-Gaussian and hence does not include a possible layering or non-Gaussian structures that could be present at this site. Our results emphasize that such errors in the model structure can only partly be compensated by data assimilation (i.e., only during the assimilation period). Furthermore, from the assimilation of head data alone this specific model structural error (possible wrong parameterization of K values south of the recharge wells) was not detected which is possibly attributed to the fact that heat transport is more sensitive to the 3-D aquifer structure than groundwater flow. Thus, these results show that the additional assimilation of temperature data can aid in a better understanding of the model dynamics and possible structural errors.

The results also emphasize that the application of EnKF for real-world data is accompanied by a higher degree of uncertainty than for synthetic studies which could reduce the effectiveness of the method. The large majority of published studies on subsurface hydrological inverse modeling and data assimilation is synthetic and the outcome of those studies is therefore overoptimistic concerning the impact of conditioning to measurement data on improved model predictions. The differences in the worth of conditioning data for the transition from synthetic to real-world cases becomes apparent in this study because the synthetic setup indicated a potential improvement in parameter estimation whereas for the real-world setup there is no direct indication that the additional assimilation of groundwater temperatures improves the estimation of unknown hydraulic properties.

Another issue that became apparent from the joint assimilation of heads and temperatures is that the preservation of ensemble variance (or the effect of filter inbreeding) is different for these two variables. For hydraulic heads, a certain degree of variability is maintained throughout the assimilation experiments due to the highly uncertain parameter fields and the fast hydraulic reaction of the aquifer toward changes in the boundary conditions. For groundwater temperatures, a relatively fast decrease of ensemble variance is visible from the experiments. This is caused by the fact that heat propagation is a slower process and therefore the buildup of variance in temperature between the updating cycles is retarded compared to hydraulic heads. The problem of filter inbreeding for temperatures could principally be compensated by different measures, e.g., by increasing the ensemble size, by introducing more variability in the boundary conditions for heat transport or by inflating the ensemble before each updating step. Increasing of the ensemble size is problematic in this setup because the computational demand for the coupled flow and heat transport simulations with 128 ensemble members is already very high. Introducing more uncertainty in the boundary conditions would be possible but this additional source of uncertainty would still propagate slowly through the aquifer. Therefore, we decided to repeat the simulation of scenario RW_{hTKL} including covariance inflation with a constant inflation factor [*Hamill et al.*, 2001] before each updating step. In this simulation each ensemble member of both state variables was inflated by a constant factor of 1.1 around its mean value before the updating step with EnKF. Figure 19 gives an impression on the development of RMSE(h) and RMSE(T) for scenario RW_{hTKL} with and without covariance inflation. The errors for hydraulic heads are

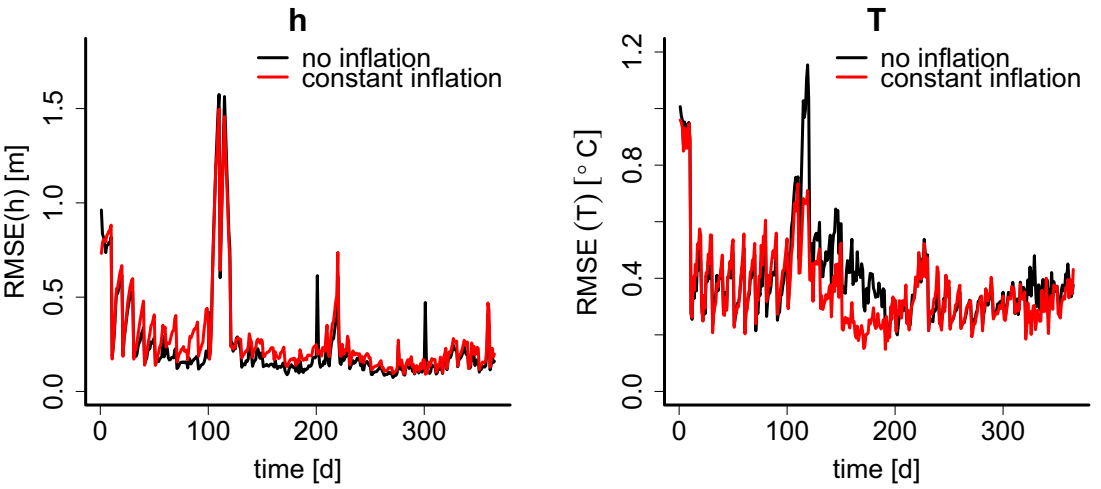

Figure 19. Root mean square error of (left) mean hydraulic heads and (right) mean groundwater temperatures for scenario RW_{hTKL} with and without constant covariance inflation.

slightly higher when covariance inflation is used whereas there is some improvement at certain time periods for groundwater temperatures. This underpins that both state variables show a different evolution of ensemble variance where the variability in hydraulic heads is still adequate for updating with EnKF because an increase of variability by inflating the ensemble lead to slightly higher errors whereas an increase of ensemble variance for groundwater temperatures had a positive effect on the prediction capability of the model. The highest improvements for groundwater temperatures were achieved during the period of rising groundwater temperatures which indicates that there is a deficiency in ensemble variance during spring time when river and groundwater temperatures are nearly equal.

Localization had a relatively clear effect on the estimation of states and parameters for the synthetic experiments. Here, localization could reduce the effect of spurious correlation during the assimilation with EnKF. For the real-world case this relationship is not so clear as for the synthetic case. In the real-world application the model dynamics are more complicated and also the spatial distribution of observation points is not as regular as for the synthetic model setup. The comparison of covariance structures for piezometric heads and groundwater temperatures showed that there are large differences concerning the spatial correlation for the two variables. It was also found that localization can greatly affect the prediction of exchange fluxes between river and aquifer when it is applied for the estimation of leakage parameters. For the model setup of the real-world case, the initial ensemble of leakage coefficients was highly variable and the observation points were not distributed evenly along the river reach. When localization is used for such conditions it is not possible to constrain the parameters that are very distant from the observations which could then lead to a significant misinterpretation of exchange fluxes. Therefore, a localization scheme that uses only one length scale for all variables is probably too generalized for the application in the real-world case. For the different state variables, it seems desirable to introduce a separate localization length scale λ for piezometric heads and for groundwater temperatures because their covariance structures vary considerably. Additionally, the relationship between the ensemble variance of model parameters and the localization scheme should be considered carefully before localization is applied in data assimilation with EnKF. For example, the results for applying localization to river bed properties show that when the initial ensemble spread is very high this can lead to a systematic bias between model predictions and measurement data because the highly variable parameters in some parts of the model cannot be updated anymore when localization is used. Results for the application of variable-dependent localization length scales suggest that introducing such a localization scheme makes the assimilation more flexible with respect to the specific model conditions and variables. The application of localization on the update of hydraulic conductivities seems to have a more positive effect than for leakage parameters because the tendency toward relatively extreme values for hydraulic conductivities with high spatial variations was reduced during the updating procedure. For hydraulic conductivity, the observation points were distributed more evenly over the model domain and also the initial ensemble was better constrained compared to the initial ensemble of leakage coefficients. Thus, it is concluded that the application of distance-dependent localization requires a careful tuning and

adaption toward the specific conditions of a model. This adaption should consider the estimation of the variability of model parameters, the specific covariance structures of the model states and parameters as well as the availability and spatial distribution of observations.

6. Conclusions

In this study, we applied the ensemble Kalman filter to jointly assimilate piezometric heads and groundwater temperature data into a managed river-aquifer system. A synthetic model of a river-aquifer system as well as a real-world model of the Limmat aquifer in Zurich were used to identify the usefulness of this data assimilation approach for the real-time prediction of aquifer states and the identification of hydraulic subsurface parameters. For the real-world case, an extensive data set of head and temperature measurements was available for verification. Results for the synthetic river-aquifer model which mimics a losing stream problem through groundwater withdrawal showed that the joint assimilation of piezometric heads and groundwater temperatures resulted in the best estimate of hydraulic properties (i.e., hydraulic conductivities and leakage coefficients). Conditioning of parameters on hydraulic head data lead to a RMSE reduction of 30% for hydraulic conductivities and 73% for leakage coefficients compared to the initial parameter ensembles. An additional assimilation of temperature data led to a total RMSE reduction of 36% and 77%, respectively. Assimilated piezometric head data gave information on the magnitude of river-aquifer exchange fluxes and corrected for a bias in the leakage parameters. Assimilation of temperature data mainly led to a better characterization of the spatial distribution of leakage parameters. Results for the Limmat aquifer indicate that the prediction of hydraulic heads and groundwater temperatures can be significantly improved through data assimilation with EnKF under real-world conditions. The best results were obtained with a simultaneous update of model states and parameters (average RMSE reduction of 74% and 64% for heads and temperatures respectively). However, validation experiments revealed that for the real-world case most of the improvement of model parameters is due to the assimilation of hydraulic head data with no significant additional improvement by assimilating temperature data. Nevertheless, the conditioning of parameters on hydraulic heads and temperatures lead to an average RMSE reduction of 49% for heads and 15% for temperatures in the validation period (compared to unconditioned parameters). The effect of distance-dependent localization was relatively straightforward for the synthetic river-aquifer system because it could easily be identified from the updated parameter fields that nonphysical parameter updates were suppressed through the tapering with the localization function. For the real-world case, localization also seems to have a significant impact on the updated parameter fields. This was mainly observed for the fields of aquifer hydraulic conductivities where localization lead to a relatively strong smoothing effect compared to the updating scheme without localization. This smoothing also prevented the buildup of extremer parameter values which was observed for the standard EnKF. However, it also became clear that the localization limit should be carefully chosen in order to honor the different physical relationships of model variables. Additionally, the effective range of the localization function and the variance of the initial parameter ensemble have to be balanced very carefully. Further research is required to find an optimal strategy for the simultaneous localization of different model variables under complex system dynamics.

Acknowledgments

We thank Water Supply of Zurich for providing the measurement data for the real-world model, Uli Kuhlmann for technical support with SPRING and Beatrice Marti and Wolfgang Kinzelbach for fruitful discussions about the model. We also would like to acknowledge Jülich Supercomputing Center for granting computation time on JUROPA and three anonymous reviewers for their helpful comments and suggestions. This work was partially financed by the German Research Foundation DFG (Transregional Collaborative Research Centre 32—Patterns in Soil–Vegetation–Atmosphere Systems: Monitoring, modeling and data assimilation).

References

Alcolea, A., J. Carrera, and A. Medina (2006), Pilot points method incorporating prior information for solving the groundwater flow inverse problem, *Adv. Water Resour.*, *29*(11), 1678–1689, doi:10.1016/j.advwatres.2005.12.009.

Anderson, M. (2005), Heat as a ground water tracer, *Ground Water*, *43*(6), 951–968, doi:10.1111/j.1745–6584.2005.00052.x.

Ashby, S. F., and R. D. Falgout (1996), A parallel multigrid preconditioned conjugate gradient algorithm for ground water flow simulations, *Nucl. Sci. Eng.*, *124*(1), 145–159.

Bauser, G., H. J. Hendricks Franssen, H. P. Kaiser, U. Kuhlmann, F. Stauffer, and W. Kinzelbach (2010), Real-time management of an urban groundwater well field threatened by pollution, *Environ. Sci. Technol.*, *44*(17), 6802–6807, doi:10.1021/es100648j.

Bauser, G., H. J. Hendricks Franssen, F. Stauffer, H.-P. Kaiser, U. Kuhlman, and W. Kinzelbach (2012), A comparison study of two different control criteria for the real-time management of urban groundwater works, *J. Environ. Manage.*, *105*, 21–29, doi:10.1016/j.jenvman.2011.12.024.

Bear, J., and A. H.-D. Chen (2010), *Modeling Groundwater Flow and Contaminant Transport*, 834 pp., Springer, Dordrecht, Netherlands.

Bravo, H. R., F. Jiang, and R. J. Hunt (2002), Using groundwater temperature data to constrain parameter estimation in a groundwater flow model of a wetland system, *Water Resour. Res.*, *38*(8), doi:10.1029/2000WR000172.

Burgers, G., P. J. van Leeuwen, and G. Evensen (1998), Analysis scheme in the ensemble Kalman filter, *Mon. Weather Rev.*, *126*(6), 1719–1724, doi:10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Camporese, M., C. Paniconi, M. Putti, and P. Salandin (2009), 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.

Camporese, M., G. Cassiani, R. Deiana, and P. Salandin (2011), Assessment of local hydraulic properties from electrical resistivity tomography monitoring of a three-dimensional synthetic tracer test experiment, *Water Resour. Res.*, *47*, W12508, doi:10.1029/2011WR010528.

Chen, Y., and D. S. Oliver (2010), Cross-covariances and localization for EnKF in multiphase flow data assimilation, *Comput. Geosci.*, *14*(4), 579–601, doi:10.1007/s10596-009-9174-6.
Conant, B. (2004), Delineating and quantifying ground water discharge zones using streambed temperatures, *Ground Water*, *42*(2), 243–257, doi:10.1111/j.1745-6584.2004.tb02671.x.
Constantz, J. (2008), Heat as a tracer to determine streambed water exchanges, *Water Resour. Res.*, *44*, W00D10, doi:10.1029/2008WR006996.
Constantz, J., C. L. Thomas, and G. Zellweger (1994), Influence of diurnal variations in stream temperature on streamflow loss and groundwater recharge, *Water Resour. Res.*, *30*(12), 3253–3264, doi:10.1029/94WR01968.
Delta h Ingenieurgesellschaft mbH (2006), *Spring 3.3, Software*, Witten, Germany.
Devegowda, D., E. Arroyo-Negrete, and A. Datta-Gupta (2010), Flow relevant covariance localization during dynamic data assimilation using EnKF, *Adv. Water Resour.*, *33*(2), 129–145, doi:10.1016/j.advwatres.2009.10.001.
Doussan, C., A. Toma, B. Paris, G. Poitevin, E. Ledoux, and M. Detay (1994), Coupled use of thermal and hydraulic head data to characterize river-groundwater exchanges, *J. Hydrol.*, *153*(1-4), 215–229, doi:10.1016/0022-1694(94)90192-9.
Engeler, I., H. J. Hendricks Franssen, R. Müller, and F. Stauffer (2011), The importance of coupled modelling of variably saturated groundwater flow-heat transport for assessing river-aquifer interactions, *J. Hydrol.*, *397*(3-4), 295–305, doi:10.1016/j.jhydrol.2010.12.007.
Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using monte-carlo methods to forecast error statistics, *J. Geophys. Res.*, *99*(C5), 10,143–10,162, doi:10.1029/94JC00572.
Friedel, M. J. (2005), Coupled inverse modeling of vadose zone water, heat, and solute transport: Calibration constraints, parameter non-uniqueness, and predictive uncertainty, *J. Hydrol.*, *312*, 148–175, doi:10.1016/j.jhydrol.2005.02.013.
Gaspari, G., and S. E. Cohn (1999), Construction of correlation functions in two and three dimensions, *Q. J. R. Meteorol. Soc.*, *125*(554), 723–757, doi:10.1002/qj.49712555417.
Gómez-Hernández, J., and A. Journel (1993), Joint sequential simulation of multi-Gaussian fields, in *Geostatistics Tróia '92*, vol. 1, edited by A. Soares, pp. 85–94, Kluwer Acad., New York.
Hamill, T., J. Whitaker, and C. Snyder (2001), Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, *Mon. Weather Rev.*, *129*(11), 2776–2790, doi:10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.
Hatch, C. E., A. T. Fisher, C. R. Ruehl, and G. Stemler (2010), Spatial and temporal variations in streambed hydraulic conductivity quantified with time-series thermal methods, *J. Hydrol.*, *389*(3-4), 276–288, doi:10.1016/j.jhydrol.2010.05.046.
Hendricks Franssen, H. J., and W. Kinzelbach (2008), 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.
Hendricks Franssen, H. J., H. P. Kaiser, U. Kuhlmann, G. Bauser, F. Stauffer, R. Mueller, and W. Kinzelbach (2011), 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.
Jiang, Y., and A. D. Woodbury (2006), A full-Bayesian approach to the inverse problem for steady-state groundwater flow and heat transport, *Geophys. J. Int.*, *167*(3), 1501–1512, doi:10.1111/j.1365-246X.2006.03145.x.
Jones, J. E., and C. S. Woodward (2001), Newton-Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems, *Adv. Water Resour.*, *24*(7), 763–774.
Kollet, S. J., and R. M. Maxwell (2006), Integrated surface–groundwater flow modeling: a free-surface overland flow boundary condition in a parallel groundwater flow model. *Adv. Water Resour., 29*, 945–958.
Li, L., H. Zhou, J. J. Gómez-Hernández, and H. J. Hendricks Franssen (2012), Jointly mapping hydraulic conductivity and porosity by assimilating concentration data via ensemble Kalman filter, *J. Hydrol.*, *428–429*, 152–169, doi:10.1016/j.jhydrol.2012.01.037.
Liu, G., Y. Chen, and D. Zhang (2008), Investigation of flow and transport processes at the MADE site using ensemble Kalman filter, *Adv. Water Resour.*, *31*(7), 975–986, doi:10.1016/j.advwatres.2008.03.006.
Lorentzen, R., G. Naevdal, and A. Lage (2003), Tuning of parameters in a two-phase flow model using an ensemble Kalman filter, *Int. J. Multiphase Flow*, *29*(8), 1283–1309, doi:10.1016/S0301-9322(03)00088-0.
Ma, R., and C. Zheng (2010), Effects of density and viscosity in modeling heat as a groundwater tracer, *Ground Water*, *48*(3), 380–389.
Marti, B. S., G. Bauser, F. Stauffer, U. Kuhlman, H.-P. Kaiser, and W. Kinzelbach (2012), An expert system for real-time well field management, *Water Sci. Technol.*, *12*(5), 699–706, doi:10.2166/ws.2012.021.
Naevdal, G., L. Johnsen, S. Aanonsen, and E. Vefring (2005), Reservoir monitoring and continuous model updating using ensemble Kalman filter, in 2003 SPE Annual Technical Conference and Exhibition, *SPE J.*, *10*(1), 66–74.
Nan, T., and J. Wu (2011), Groundwater parameter estimation using the ensemble Kalman filter with localization, *Hydrogeol. J.*, *19*(3), 547–561, doi:10.1007/s10040-010-0679-9.
Nowak, W. (2009), Best unbiased ensemble linearization and the quasi-linear Kalman ensemble generator, *Water Resour. Res.*, *45*, W04431, doi:10.1029/2008WR007328.
Racz, A. J., A. T. Fisher, C. M. Schmidt, B. S. Lockwood, and M. Los Huertos (2012), Spatial and temporal infiltration dynamics during managed aquifer recharge, *Ground Water*, *50*(4), 562–570, doi:10.1111/j.1745-6584.2011.00875.x.
Renard, P., G. Le Loc'h, E. Ledoux, G. de Marsily, and R. Mackay (2000), A fast algorithm for the estimation of the equivalent hydraulic conductivity of heterogeneous media, *Water Resour. Res.*, *36*(12), 3567–3580, doi:10.1029/2000WR900203.
Schmidt, C., M. Bayer-Raich, and M. Schirmer (2006), Characterization of spatial heterogeneity of groundwater-stream water interactions using multiple depth streambed temperature measurements at the reach scale, *Hydrol. Earth Syst. Sci.*, *10*(6), 849–859.
Schornberg, C., C. Schmidt, E. Kalbus, and J. H. Fleckenstein (2010), Simulating the effects of geologic heterogeneity and transient boundary conditions on streambed temperatures—Implications for temperature-based water flux calculations, *Adv. Water Resour.*, *33*(11), 1309–1319, doi:10.1016/j.advwatres.2010.04.007.
Tyler, S. W., J. S. Selker, M. B. Hausner, C. E. Hatch, T. Torgersen, C. E. Thodal, and S. G. Schladow (2009), Environmental temperature sensing using raman spectra dts fiber-optic methods, *Water Resour. Res.*, *45*, W00D23, doi:10.1029/2008WR007052.
van Genuchten, M. T. (1980), A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.*, *44*(5), 892–898.
Vandenbohede, A., and E. Van Houtte (2012), Heat transport and temperature distribution during managed artificial recharge with surface ponds, *J. Hydrol.*, *472–473*, 77–89.
Vogt, T., M. Schirmer, and O. A. Cirpka (2012), Investigating riparian groundwater flow close to a losing river using diurnal temperature oscillations at high vertical resolution, *Hydrol. Earth Syst. Sci.*, *16*, 473–487.
Woodbury, A. D., and L. Smith (1988), Simultaneous inversion of hydrogeologic and thermal data: 2. Incorporation of thermal data, *Water Resour. Res.*, *24*(3), 356–372.