Response letter of hess-2019-324-RC2

Please find the responses to the comments.

Comments made by the reviewer were highly insightful. They allowed me to greatly improve the quality of the manuscript. I described the response to the comments.

Each comment made by the reviewers is written in *italic* font. I numbered each comment as (n.m) in which n is the reviewer number and m is the comment number. In the revised manuscript, changes are highlighted in yellow.

I trust that the revisions and responses are sufficient for my manuscript to be published in *Hydrology and Earth System Sciences*

Responses to the comments of Reviewer #2

The author investigated the data assimilation with a 3-D hyperresolutin land model named as ParFlow using ETKF on the various scenarios. Although I think that this manuscript is well written, I have some comments for publication.

Major comments

(2.1) 1. Ll. 316-321. Each ensemble member has different saturated hydraulic conductivity and rainfall rate using random numbers from lognormal distribution with mean = 0 and standard deviation = 0.15. Why does the author choose them? Does the author confirm their sensitivities? Please address the reason simply.

→ I chose them because this setting gives the sufficiently large error in precipitation considering the real-world applications but it does not introduce the strong non-Gaussianity to the precipitation data. This point was described in the discussion section of the original version of the paper:

> Lines 721-727: "Please note that the parameters of the lognormal distribution to model the uncertainty in rainfall were specified to make the rainfall PDF similar to the Gaussian distribution. I chose the lognormal distribution in order not to generate negative rainfall values and I intended not to introduce non-Gaussianity into the external forcing. The rainfall input which follows the Gaussian PDF was transformed into the non-Gaussian PDF of the background error by the strongly nonlinear dynamics of the topography-driven surface flow."

In the revised version of the paper, I have clarified this point in the section of experiment design when the lognormal distributed multiplicative error was introduced.

> Lines 320-325: "The two parameters of the lognormal distribution, commonly called μ and σ, were set to 0 and 0.15, respectively. These parameters were chosen to give the sufficiently large error in precipitation and saturated hydraulic conductivity. In addition, this setting makes the rainfall PDF similar to the Gaussian distribution, which is important to interpret the results of the experiments (see the discussion section)."

Although the multiplicative error used in this study was sufficiently large, the problem gets more difficult when the parameter was set to give larger uncertainty to rainfall. In addition, I could choose the parameter to give the biases and/or non-Gaussianity to the input rainfall, which may make the problem more difficult. Since these points are obvious and have been confirmed even in the conventional land surface models, I do not include the results with the different parameters. However, the specification of the prior uncertainty of rainfall and hydraulic conductivity must be important toward the real-world application, which has not been clarified in the previous version of the paper. In the revised version of the paper, I have clarified this point.

> Lines 745-748: "Although the prior uncertainty in rainfall and saturated hydraulic conductivity was arbitrary chosen in this study, the specification of the prior knowledge is not straightforward

in the real-world applications."

(2.2) 2. L. 365: RMSE is calculated by using all members, not an ensemble mean. Usually, I think the RSME is calculated by difference between an ensemble mean and truth. Although I guess the author's RMSE is better for the author's experiments, please explain why the author use all members for RMSE.

→ The NoDA has a very large spread since I gave the large uncertainty to the input rainfall and hydraulic conductivity. Despite the large spread, the ensemble members in the NoDA experiment are distributed around the truth since the model has no chaotic behavior. Therefore, I should evaluate not only if ensemble mean is consistent to the synthetic truth, but also if the ensemble spread is appropriately reduced. Therefore, here I evaluated the all members. This point was indeed unclear in the original version of the paper and I have included this point in the revised version of the paper.

> Lines 371-374: "I used all ensemble members to calculate RMSE because I should evaluate not only if the ensemble mean is consistent to the synthetic truth, but also if the extremely large ensemble spread simulated in the NoDA experiment is appropriately reduced."

I have also used the other metrics to evaluate if the ensemble spread is appropriately reduced as the response to the other reviewer comment in the previous round of the review although the results have not been included in the main manuscript.

> Lines 760-770: "The other limitation of this study is that I could not thoroughly evaluate the skill of the ensemble data assimilation to quantify the uncertainty of its prediction. Following Abbazadeh et al. (2019), I calculated the 95% exceedance ratio and found that the ensemble forecast was systematically overconfident (not shown). In the synthetic experiments of this study, the number of rainfall events was small, and the timing and magnitude of rainfall were not diversified. Due to this limited amount of data, it is difficult to deeply discuss the accuracy of the quantified uncertainty by data assimilation. While the skill of lumped hydrological models was often evaluated by the probabilistic performance measures such as the 95% exceedance ratio (e.g., Abbazadeh et al. (2019)), the uncertainty quantification of the simulation of hyper-resolution land models is in its infancy. How surface lateral flows affect the accuracy of the uncertainty quantification by data assimilation should be investigated using more realistic data."

(2.3) 3. Ll. 461, 665-666: In my understanding, the ensemble Kalman filters (EnKFs) do not assume the Gaussian PDF and linearity. The EnKFs derive an optimal value under the Gaussian PDF and linearity. This does not mean assuming the Gaussian PDF and linearity.

→ I fully agree with this comment. This point was correctly described in the method section of the original version of the paper:

> Lines 200-204: “It should be noted that the equations (10-13) give an optimal estimation only when the model and observation errors follow the Gaussian distribution. When the probabilistic distribution of the error in either model or observation has a non-Gaussian structure, results of the Kalman filter are suboptimal. This point is important to interpret the results of this study.”

However, in the results section, I described that EnKF assumed the Gaussian PDF and linearity. In the revised version of the paper, I have modified this point by describing:

> Lines 465-468: “As I mentioned in section 2.2, in the ETKF, the estimation of the state and parameter variables is optimal if and only if the model’s error has the Gaussian PDF and the relationship between observed variables and unobserved variables is linear.”
>
> Lines 672-676: “Please note that the low representativeness of the soil moisture observations in the case of the HIGH_K reference shown in section 3.1 is due to the limitation of the Kalman filter that the error PDFs need to follow the Gaussian distribution to get the optimal estimation so that the increase of the ensemble size cannot solve this issue.”

(2.4) 4. L. 616: “there are large errors in the area around 500<=x, y <=1500” I have trouble with this sentence. I cannot confirm the large errors in Fig. 9.

→ Since Figure 9 shows the improvement rate by data assimilation, the large errors cannot be directly found in Figure 9. This point was indeed unclear in the original version of the paper. I have clarified this point by simply mentioning that the large errors are not shown in the figures.

> Lines 625-627: “In the noOF configuration, there are large errors in the area around 500<=x, y <=1500 (not shown) since the increase of soil moisture in this area is caused by the topography-driven surface flow which is neglected in the noOF configuration.”

(2.5) 5. Ll. 688-699: Assimilating just one observation improves the analysis errors in whole domain despite the nonlinear equations. This means that the model has long-range spatial correlations. Therefore, I guess the ETKF works well without the localization. Also, the author mentioned that the localization scale depends on the model parameter. In order to confirm those, the author should investigate the spatial correlations.

→ I realized that this description was misleading. Because the negative impact of the non-linear and non-Gaussian on the state estimation can be found only in the edge of the area where topography-driven surface flow reaches (Figures 9b and 9c), I suggested omitting to update the state variables there as the heuristic approach. It is not identical to the localization method in the context of data assimilation, in which the spurious correlation sampled by an ensemble is eliminated by restricting the impact of assimilating observation. Therefore, what I suggested is not directly related to the spatial correlations. As the reviewer mentioned, the model has long-range spatial correlations except for the edge of flooding area. In the revised version of the paper, I have clarified this point.

> Lines 706-714: "As a possible heuristic approach to avoid the negative impact of the non-Gaussian background PDF, I can omit to update the state variables in the edge of the area where topography-driven surface flow reaches. The numerical experiments clearly indicate that the negative impact of the non-linear physics and non-Gaussian PDF is found only in the edge of flooding areas so that it is beneficial to simply omit to update the state variables in this area. It is similar but not conceptually identical to the localization method, in which the spurious correlation sampled by an ensemble is eliminated by spatially restricting the impact of assimilating observation (e.g., Rasmussen et al. 2015; Anderson 2007; Bishop and Hodyss 2009). "

(2.6) 6. Figure 8: In the OF configuration of Fig. 8 (a) and the noOF and OF configurationsof Fig. 8 (b), the DA_obs1 and DA_obs9 experiments have almost the same RMSE although the DA_obs9 experiments have 9 times observation information. Why?

→ This point was not mentioned in the paragraph where Figure 8 appears. However, I mentioned it later in the original version of the paper.

> Lines 610-615: "Figure 9b shows that the increase of the number of observations substantially improves the soil moisture simulation in the region which is affected by topography-driven surface flow (see also Figure 6c). However, the skill to simulate soil moisture is severely degraded in the lower-left corner of the domain, which causes the stalled improvement from the OF_DA_obs1 experiment to the OF_DA_obs9 experiment shown in Figure 8a."

In the revised version of the paper, I have provided the guidance for the readers to make it easier to find this statement.

> Lines 572-575: "The OF_DA_obs361 experiment has the smallest RMSE so that a denser observing network is beneficial to estimate soil moisture, although there is the stalled improvement from the OF_DA_obs1 experiment to the OF_DA_obs9 experiment (the reason for it will be explained later)."

(2.7) 7. Figure S4: A green line looks like to splits into a single outlier and the others. If so, Ithink this is ensemble clustering (EC, Anderson 2010, Amezcua et al. 2012). The EC is frequently generated by ensemble square loot filters included the ETKF and may be related to the non-Gaussian PDF. Therefore, please refer to the EC in section 4.

→ Thanks for the comment. Figures S1-S4 indeed show the ensemble clustering, which strengthens the conclusion that the non-Gaussian PDF has an important role in hyperresolution land data assimilation. I have included this discussion in the revised version of the paper.

Lines 687-692: "In addition, I found ensemble clustering in which the ensemble members are split into a single outlier and the others (see Figures S1-S4). The previous studies found that this ensemble clustering is generated by the non-Gaussian PDF (Anderson 2010; Amezcua et al. 2012). Ensemble clustering shown in the analysis timeseries also implies that the non-Gaussian PDF plays an important role in the data assimilation of the hyperresolution land model."

Minor comments

(2.8) 1. Equation numbers are confused. For instance, Eq. 4 is written on the lines 141 and 148. Please correct the all equation numbers.

→ The numbering of the equations in the original version of the paper was indeed confusing. I have fixed this point following the reviewer's instruction.

References

Anderson, J. L.: A non-Gaussian ensemble filter update for data assimilation, Mon. Wea. Rev., 138, 4186-4198, 2010.

Amezcua, J., Ide, K., Bishop, C. H., and Kalnay, E.: Ensemble clustering in deterministic ensemble Kalman filters, Tellus, 64A, 1-12, 2012.

Do surface lateral flows matter for data assimilation of soil moisture observations into hyperresolution land models?

Running title: HYPERRESOLUTION LAND DATA ASSIMILATION

Yohei Sawada[1,2],

[1] Institute of Engineering Innovation, the University of Tokyo, Tokyo, Japan

[2] Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan

Corresponding author: Y. Sawada, Institute of Engineering Innovation, the University of Tokyo, Tokyo, Japan, 2-11-6, Yayoi, Bunkyo-ku, Tokyo, Japan, yohei.sawada@sogo.t.u-tokyo.ac.jp

Abstract

It is expected that hyperresolution land modeling substantially innovates the simulation of terrestrial water, energy, and carbon cycles. The major advantage of hyperresolution land models against conventional one-dimensional land surface models is that hyperresolution land models can explicitly simulate lateral water flows. Despite many efforts on data assimilation of hydrological observations into those hyperresolution land models, how surface water flows driven by local topography matter for data assimilation of soil moisture observations has not been fully clarified. Here I perform two minimalist synthetic experiments where soil moisture observations are assimilated into an integrated surface-groundwater land model by an ensemble Kalman filter. I discuss how differently the ensemble Kalman filter works when surface lateral flows are switched on and off. A horizontal background error covariance provided by overland flows is important to adjust the unobserved state variables (pressure head and soil moisture) and parameters (saturated hydraulic conductivity). However, the non-Gaussianity of the background error provided by the nonlinearity of a topography-driven surface flow harms the performance of data assimilation. It is difficult to efficiently constrain model states at the edge of the area where the topography-driven surface flow reaches by linear-Gaussian filters. It brings the

new challenge in land data assimilation for hyperresolution land models. This study highlights the importance of surface lateral flows in hydrological data assimilation.

1. Introduction

Hyperresolution land modeling is expected to improve the simulation of terrestrial water, energy, and carbon cycles, which is crucially important for meteorological, hydrological and ecological applications (see Wood et al. (2011) for a comprehensive review). While conventional land surface models (LSMs) assume that lateral water flows are negligible at the coarse resolution (>25km) and solve vertical one-dimensional Richards equation for the soil moisture simulation (e.g., Sellers et al. 1996; Lawrence et al. 2011), currently proposed hyperresolution land models, which can be applied at a finer resolution (<1km), explicitly consider surface and subsurface lateral water flows (e.g., Maxwell and Miller 2005; Tian et al. 2012; Shrestha et al. 2014; Niu et al. 2014). The fine horizontal resolution can resolve slopes, which are drivers of a lateral transport of water, and realize the fully integrated surface-groundwater modeling. Previous works indicated that a lateral transport of water strongly controls latent heat flux and the partitioning of evapotranspiration into base soil evaporation and plant transpiration (e.g., Maxwell and

Condon 2016; Ji et al. 2017; Fang et al. 2017). This effect of a lateral transport of water on land-atmosphere interactions has been recognized (e.g., Williams and Maxwell 2011; Keune et al. 2016).

Data assimilation has contributed to improving the performance of LSMs by fusing simulation and observation. The grand challenge of land data assimilation is to improve the simulation of unobservable variables using observations by propagating observations' information into model's high dimensional state and parameter space. In previous works on the conventional 1-D LSMs, many land data assimilation systems (LDASs) have been proposed to accurately estimate model's state and parameter variables, which cannot be directly observed, by assimilating satellite and in-situ observations. For example, the optimization of LSM's unknown parameters (e.g., hydraulic conductivity) has been implemented by assimilating remotely sensed microwave observations (e.g., Yang et al. 2007; Yang et al. 2009; Bandara et al. 2014; Bandara et al. 2015; Sawada and Koike 2014; Han et al. 2014). Kumar et al. (2009) focused on the correlation between surface and root-zone soil moistures to examine the potential of assimilating surface soil moisture observations to estimate root-zone soil moisture. Sawada et al. (2015) successfully improved the simulation of root-zone soil moisture by assimilating microwave brightness temperature observations which include the information of vegetation water content. Gravity Recovery and Climate Experiment total water storage observation has been intensively used to improve the simulation of groundwater and soil moisture (e.g., Li et al. 2012; Houborg et al. 2012). Improving the simulation of state variables such as soil moisture and biomass by LDASs has contributed to accurately estimating fluxes such as evapotranspiration (e.g. Martens et al. 2017) and CO_2 flux (e.g., Verbeeck et al. 2011). However, in most of the studies on the conventional 1-D LDASs, observations impacted state variables and parameters only in a single model's horizontal grid which is identical to the location of the observation. The assumption that the water flows are restricted to vertical direction in LSMs makes it difficult to propagate observation's information horizontally. It limits the potential of land data assimilation to fully use land hydrological observations.

The hyperresolution land models, which explicitly solve surface and subsurface lateral flows, provide a unique opportunity to examine the potential of land data assimilation to propagate observation's information horizontally in a model space and efficiently use land hydrological observations. Previous works successfully applied Ensemble Kalman Filters (EnKF) to 3-D Richards' equation-based integrated surface-groundwater models. For example, Camporese et al. (2009) and Camporese et al. (2010) successfully assimilated synthetic observations of surface pressure head and streamflow into the Catchment Hydrology (CATHY). Ridler et al. (2014) successfully assimilated Soil Moisture and Ocean Salinity satellite-observed surface soil moisture into the MIKE SHE distributed hydrological model (see also Zhang et al. (2015)). Kurtz et al. (2016) coupled the Parallel Data Assimilation Framework (PDAF) (Nerger and Hiller 2013) with the Terrestrial System Modelling Framework (TerrSysMP) (Shrestha et al. 2014) and successfully estimate the spatial distribution of soil moisture and saturated hydraulic conductivity in the synthetic experiment (see also Zhang et al. (2018)). In addition, Kurtz et al. (2016) indicated that their EnKF approach is computationally efficient in high-performance computers. Those studies have significantly contributed to fully assimilating the new high-resolution soil moisture observations such as Sentinel-1 (e.g., Paroscia et al. 2013)

Although the data assimilation of hydrological observations into hyperresolution land models has been successfully implemented in the synthetic experiments, it is unclear how topography-driven surface lateral water flows matter for data assimilation of soil moisture observations. Previous studies on data assimilation with high resolution models mainly focused on assimilating groundwater observations (e.g., Ait-El-Fquih et al. 2016;

Rasmussen et al. 2015; Hendricks-Franssen et al. 2008). There are some applications which focused on the observation of soil moisture and pressure head in shallow unsaturated soil layers. However, in those studies, topography-driven surface flow has not been considered in the experiment (Kurtz et al. 2016) or the role of them in assimilating observations into the hyperresolution land models has not been quantitatively discussed (Camporese et al. 2010; Camporese et al. 2009). This study aims at clarifying if surface lateral flows matter for data assimilation of soil moisture observations into hyperresolution land models by a minimalist numerical experiment.

2. Methods

2.1. Model

ParFlow is an open source platform which realizes fully integrated surface-groundwater flow modeling (Kollet and Maxwell 2006; Maxwell et al. 2015). This model can be efficiently parallelized in high performance computers and has been widely used as a core hydrological module in hyperresolution land models (e.g., Maxwell and Kollet 2008; Maxwell and Condon 2016; Fang et al. 2017; Kurtz et al. 2016; Maxwell et al. 2011; Williams and Maxwell 2011; Shrestha et al. 2014). Since I used this widely adopted solver as is and added nothing new to the model physics, I described the method of ParFlow to simulate integrated surface-subsurface water flows briefly and omitted the details of numerical methods. The complete description of ParFlow can be found in Kollet and Maxwell (2006), Maxwell et al. (2015) and references therein.

In the subsurface, ParFlow solves the variably saturated Richards equation in three dimensions.

$$S_S S_W(h)\frac{\partial h}{\partial t} + \phi S_W(h)\frac{\partial S_W(h)}{\partial t} = \nabla \cdot \mathbf{q} + q_r \quad (1)$$

$$\mathbf{q} = -\boldsymbol{K_s}(\boldsymbol{x})k_r(h)[\nabla(h+z)cos\theta_x + sin\theta_x] \quad (2)$$

In equation (1), h is the pressure head [L]; z is the elevation with the z axis specified as upward [L]; S_S is the specific storage [L^{-1}]; S_W is the relative saturation; ϕ is the porosity [-]; q_r is a source/sink term. Equation (2) describes the flux $\mathbf{q}$ [LT^{-1}] by Darcy's law, and $\boldsymbol{K_s}$ is the saturated hydraulic conductivity tensor [LT^{-1}]; k_r is the relative permeability [-]; θ is the local angle of topographic slope (see Maxwell et al. 2015). In this paper, the saturated hydraulic conductivity is assumed to be isotropic and a function of z:

$$\boldsymbol{K_s} = K_s(z) = K_{s,surface}\exp\left(-f\left(z_{surface} - z\right)\right) \quad (3)$$

where $K_{s,surface}$ is the saturated hydraulic conductivity at the surface soil, and $z_{surface}$ is the elevation of the soil surface. The saturated hydraulic conductivity decreases exponentially as the soil depth increases (Beven 1982). A van Genuchten relationship (van Genuchten 1980) is used for the relative saturation and permeability functions.

$$S_W(h) = \frac{S_{sat} - S_{res}}{(1+(\alpha h)^n)^{(1-\frac{1}{n})}} + S_{res} \quad (4)$$

$$k_r(h) = \frac{\left(1 - \frac{(\alpha h)^{n-1}}{(1+(\alpha h)^n)^{\left(1-\frac{1}{n}\right)}}\right)^2}{(1+(\alpha h)^n)^{\frac{\left(1-\frac{1}{n}\right)}{2}}} \quad (5)$$

where α [L-1] and n [-] are soil parameters, S_{sat} is the relative saturated water content and S_{res} is the relative residual saturation.

Overland flow is solved by the two-dimensional kinematic wave equation. The dynamics of the surface ponding depth, h [L], can be described by:

$$\mathbf{k} \cdot [-K_s(z) k_r(h) \cdot \nabla(h+z)] = \frac{\partial \|h,0\|}{\partial t} - \nabla \cdot \|h,0\| \boldsymbol{v}_{sw} + q_r \quad (6)$$

In equation (6), $\mathbf{k}$ is the unit vector in the vertical and $\|h,0\|$ indicates the greater value of the two quantities following the notation of Maxwell et al. (2015). This formulation results in the overland flow equation being represented as a boundary condition to the variably saturated Richards equation (Kollet and Maxwell 2006). If h < 0, equation (6) describes that vertical fluxes across the land surface is equal to the source/sink term q_r (i.e., rainfall and evapotranspiration). If h > 0, the terms on the right-hand side of equation

(6), which indicate water fluxes routed according to surface topography, are active. $\boldsymbol{v}_{sw}$ is the two-dimensional depth-averaged water flow velocity [LT^{-1}] and estimated by the Manning's law:

$$\boldsymbol{v}_{sw,x} = \left(\frac{\sqrt{S_{f,x}}}{n_M} h^{\frac{2}{3}}\right), \boldsymbol{v}_{sw,y} = \left(\frac{\sqrt{S_{f,y}}}{n_M} h^{\frac{2}{3}}\right) \quad (7)$$

where $S_{f,x}$ and $S_{f,y}$ are the friction slopes [-] for the x- and y-direction, respectively; n_M is the Manning's coefficient [$TL^{-1/3}$]. In the kinematic wave approximation, the friction slopes are set to the bed slopes. The methodology of discretization and numerical method to solve equations (1-7) can be found in Kollet and Maxwell (2006).

2.2. Data Assimilation

In this paper, the ensemble Kalman filter (EnKF) was applied to assimilate soil moisture observations into ParFlow. The EnKF has widely been applied to hyper-resolution land models (e.g., Camporese et al. (2009); Camporese et al. (2010); Ridler et al. (2014); Zhang et al. (2015); Kurtz et al. (2016); Zhang et al. (2018)). I examined if surface lateral flows matter for data assimilation of soil moisture observations into hyperresolution land models using this widely adopted data assimilation method.

The Parflow model can be formulated as a discrete state-space dynamic system:

$$\boldsymbol{x}(t+1) = f\big(\boldsymbol{x}(t), \boldsymbol{\theta}, \boldsymbol{u}(t)\big) + \boldsymbol{q}(t) \qquad (8)$$

where $\boldsymbol{x}(t)$ is the state variables (i.e. pressure head), $\boldsymbol{\theta}$ is the time-invariant model parameters (i.e. saturated hydraulic conductivity), $\boldsymbol{u}(t)$ is the external forcing (i.e., rainfall and evapotranspiration), and $\boldsymbol{q}(t)$ is the noise process which represents the model error. In data assimilation, it is useful to formulate an observation process as follows:

$$\boldsymbol{y}^f(t) = \mathcal{H}\big(\boldsymbol{x}(t)\big) + \boldsymbol{r}(t) \qquad (9)$$

where $\boldsymbol{y}^f(t)$ is the simulated observation, $\mathcal{H}$ is the observation operator which maps the model's state variables into the observable variables, and $\boldsymbol{r}(t)$ is the noise process which represents the observation error. The purpose of EnKF (and any other data assimilation methods) is to find the optimal state variables $\boldsymbol{x}(t)$ based on the simulation $\boldsymbol{y}^f(t)$ and observation (defined as $\boldsymbol{y}^o$) considering their errors ($\boldsymbol{q}(t)$ and $\boldsymbol{r}(t)$)

The general description of the Kalman filter is the following:

$$\boldsymbol{x}^f(t) = \mathcal{M}[\boldsymbol{x}^a(t-1)] \quad (10)$$

$$\boldsymbol{x}^a(t) = \boldsymbol{x}^f(t) + \boldsymbol{K}[\boldsymbol{y}^o - \mathcal{H}(\boldsymbol{x}^f(t))] \quad (11)$$

$$\mathbf{K} = \boldsymbol{P}^f \boldsymbol{\mathcal{H}}^{\boldsymbol{T}} (\boldsymbol{\mathcal{H}} \boldsymbol{P}^{\boldsymbol{f}} \boldsymbol{\mathcal{H}}^{\boldsymbol{T}} + \boldsymbol{R})^{-\mathbf{1}} \quad (12)$$

$$\boldsymbol{P^a} = (\boldsymbol{I} - \boldsymbol{K\mathcal{H}})\boldsymbol{P^f} \quad (13)$$

I follow the notation of Houtekamer and Zhang (2016). Superscripts f and a are forecast and analysis, respectively. In equation (10), a forecast model $\mathcal{M}$ (ParFlow in this study) is used to obtain a prior estimate at time t, $\boldsymbol{x}^f(t)$, from the estimation at the previous time $\boldsymbol{x}^a(t-1)$. In equation (11), a prior estimate $\boldsymbol{x}^f(t)$ is updated to the analysis state, $\boldsymbol{x}^a(t)$, using new observations y^o. The Kalman gain matrix **K,** calculated by equation (12), gives an appropriate weight for the observations with an error covariance matrix $\boldsymbol{R}$, and the prior with an error covariance matrix $\boldsymbol{P}^f$. $\boldsymbol{P^a}$ is an updated analysis error covariance. To calculate **K**, the observation operator $\boldsymbol{\mathcal{H}}$ is needed to map from model space to observation space. It should be noted that the equations (10-13) give an optimal estimation only when the model and observation errors follow the Gaussian distribution. When the probabilistic distribution of the error in either model or observation has a non-Gaussian structure, results of the Kalman filter are suboptimal. This point is important to interpret the results of this study.

EnKF is the Monte Carlo implementation of equations (10-13). To compute the Kalman gain matrix, **K,** ensemble approximations of $\boldsymbol{P}^f\boldsymbol{\mathcal{H}^T}$ and $\boldsymbol{\mathcal{H}P^f\mathcal{H}^T}$ can be given by:

$$\boldsymbol{P^f\mathcal{H}^T} \equiv \frac{1}{k-1}\sum_{i=1}^{k}\left(\boldsymbol{x}_i^f - \overline{\boldsymbol{x}^f}\right)\left(\mathcal{H}\boldsymbol{x}_i^f - \overline{\mathcal{H}\boldsymbol{x}^f}\right)^T \quad (14)$$

$$\boldsymbol{\mathcal{H}P^f\mathcal{H}^T} \equiv \frac{1}{k-1}\sum_{i=1}^{k}\left(\mathcal{H}\boldsymbol{x}_i^f - \overline{\mathcal{H}\boldsymbol{x}^f}\right)\left(\mathcal{H}\boldsymbol{x}_i^f - \overline{\mathcal{H}\boldsymbol{x}^f}\right)^T \quad (15)$$

where $\boldsymbol{x}_i^f$ is the ith member of a k-member ensemble prior and $\overline{\boldsymbol{x}^f} = \frac{1}{k}\sum_{i=1}^{k}\boldsymbol{x}_i^f$ and $\overline{\mathcal{H}\boldsymbol{x}^f} = \frac{1}{k}\sum_{i=1}^{k}\mathcal{H}\boldsymbol{x}_i^f$.

Once $\overline{\boldsymbol{x}^a} = \sum_{i=1}^{k}\boldsymbol{x}_i^a$ ($\boldsymbol{x}_i^a$ is the ith member of a k-member ensemble analysis) and $\boldsymbol{P^a} = \frac{\boldsymbol{1}}{\boldsymbol{k-1}}\sum_{i=1}^{k}(\boldsymbol{x}_i^a - \overline{\boldsymbol{x}^a})(\boldsymbol{x}_i^a - \overline{\boldsymbol{x}^a})^T$ are computed by equations (10-15), there are many choices of an analysis ensemble. Although equations (10-15) can calculate the mean and variance of the ensemble members, they do not tell how to adjust the state of the ensemble members in order to realize the estimated mean and variance. There are many proposed flavors of EnKF and one of the differences among them is the method to choose the analysis $\boldsymbol{x}_i^a$. In this paper, the Ensemble Transform Kalman Filter (ETKF; Bishop et al. 2001; Hunt et al. 2007) was used to transport forecast ensembles to analysis ensembles. ETKF has been used for hyperresolution land data assimilation (e.g., Kurtz et al. 2016). Please refer to Hunt et al. (2007) for the complete description of the ETKF and its localized version, the Local Ensemble Transform Kalman Filter (LETKF). The open source available at https://github.com/takemasa-miyoshi/letkf was used in this study as the ETKF code library.

In many ensemble Kalman filter systems, the ensemble spread, $\boldsymbol{P^a}$, tends to become too underdispersive to stably perform data assimilation cycles without any ensemble inflation methods (Houtekamer and Zhang, 2016). To overcome this limitation, $\boldsymbol{P^a}$ is arbitrarily inflated after data assimilation. In this paper, the relaxation to prior perturbation method (RTPP) of Zhang et al. (2004) was used to maintain an appropriate ensemble spread. In the RTPP, the computed analysis perturbations are relaxed back to the forecast perturbations:

$$\boldsymbol{x}^a_{i,new} - \overline{\boldsymbol{x}^a} = (1 - \alpha)(\boldsymbol{x}^a_i - \overline{\boldsymbol{x}^a}) + \alpha\left(\boldsymbol{x}^f_i - \overline{\boldsymbol{x}^f}\right), \ 0 \leq \alpha \leq 1 \quad (16)$$

where α was set to 0.975 in this study. If $\alpha = 1$, the analysis spread is identical to the background spread. Many studies show that the ensemble inflation works well when α remains fairly close to 1 (see also the comprehensive review by Houtekamer and Zhang 2016).

In the data assimilation experiments, I adjusted pressure head by data assimilation so that $\boldsymbol{x}^f$ is pressure head. Since the surface saturated hydraulic conductivity was also adjusted, $\boldsymbol{x}^f$ includes log-transformed $K_{s,surface}$. I assimilated volumetric soil moisture observations so that $\boldsymbol{y}^f$ and $\boldsymbol{y}^o$ are simulated and observed volumetric soil moisture, respectively. The van Genuchten relationship converts the adjusted state variables $\boldsymbol{x}^f$ to the observable variables $\boldsymbol{y}^f$ and can be recognized as an observation operator $\mathcal{H}$. However, since volumetric soil moisture $\boldsymbol{y}^f$ has already been calculated by Parflow, I did not need the van Genuchten relationship in data assimilation.

2.3. Kullback-Leibler divergence

To evaluate the non-Gaussianity of the background error sampled by an ensemble, I used the Kullback-Leibler divergence (KLD) (Kullback and Leibler 1951):

$$D_{KL}(p,q) = \sum_i p(i) log \frac{p(i)}{q(i)} \quad (17)$$

where $D_{KL}(p,q)$ is the KLD between two probabilistic distribution functions (PDFs), p and q. If two PDFs are equal for all i, $D_{KL}(p,q) = 0$. A large value for $D_{KL}(p,q)$ indicates that the two PDFs, p and q, substantially differ from each other. Therefore, the KLD can be used as an index to evaluate the closeness of two PDFs. In this study, I compared the PDF of the ensemble simulation (p in equation (17)) with the Gaussian PDF which has the mean and variance of the ensembles (q in equation (17)). A large value for $D_{KL}(p,q)$ indicates the state variables simulated by ensembles do not follow the Gaussian PDF. It should be noted that the KLD is not symmetric ($D_{KL}(p,q) \neq D_{KL}(q,p)$). The KLD has been used to quantitatively evaluate the Gaussianity of the sampled background error in the studies on data assimilation (e.g., Kondo and Miyoshi 2019; Duc and Saito 2018).

3. Synthetic experiments

In this study, I performed two synthetic experiments. In the synthetic experiments, I generated the synthetic truth of the state variables by driving ParFlow with the specified parameters and input data. Then the synthetic observations were generated by adding the Gaussian white noise to this synthetic truth. The performance of data assimilation was evaluated by comparing the estimated state and parameter values by ETKF with the synthetic truth. This synthetic experiment has been recognized as an important research method to analyze how data assimilation works (e.g., Moradkhani et al. 2005; Camporese et al. 2009; Vrugt et al. 2013; Kurtz et al. (2016); Sawada et al. 2018)

3.1. Simple 2-D slope with homogeneous hydraulic conductivity

3.1.1. Experiment Design

The synthetic experiment was implemented to examine how topography-driven surface lateral flows contribute to efficiently propagating observation's information horizontally in the data assimilation of soil moisture observation. Two synthetic reference runs were created by Parflow. The 2-D domain has a horizontal extension of 4000m and a vertical extension of 5m. The domain of the virtual slope was horizontally discretized into 40 grid cells with a size of 100m and vertically discretized into 50 grid cells with a size of 0.10m. The domain has a 25% slope. In two synthetic reference runs, it heavily rains only in the upper half of the slope (2000m<x<4000m). Although this rainfall distribution is unrealistic, the effect of surface lateral flows on data assimilation can clearly be discussed in this simplified problem setting. More realistic rainfall distribution will be used in the next synthetic experiment (see section 3.2). A constant rainfall rate of 50mm/h was applied for 3 hours and then the period with no rainfall and evaporation of 0.075mm/h lasted for 117 hours. This 120-hour rain/no rain cycle was repeatedly applied to the domain. There is no rainfall in the lower half of the slope (0m<x<2000m). The configurations described above were schematically shown in Figure 1a. The parameters of the van Genuchten relationship, alpha and n, were set to 1.5 [m^{-1}] and 1.75, respectively. Those values are in the reasonable range estimated by the published literature (e.g., Ghanbarian-Alavijeh et al. 2010). The porosity, ϕ in equation (1), was set to 0.40. The

Manning's coefficient, n_M in equation (5), was set to 5.52×10^{-6} [$m^{-1/3}h$]. These clayey soil properties described above are applied to the whole domain. The groundwater table was located at z=3m and the hydrostatic pressure gradient was assumed for the initial pressure heads in the unsaturated soil layers.

The difference between two synthetic reference runs is the value of saturated hydraulic conductivity. The surface saturated hydraulic conductivity, $K_{s,surface}$ in equation (3), was set to 0.005 [m/h] in one reference, and 0.02 [m/h] in the other. These surface saturated hydraulic conductivities described above are applied to the whole domain. Figure 1 shows the difference of the response to heavy rainfall between the two synthetic reference runs. In the case of the low saturated hydraulic conductivity (hereafter called the LOW_K reference), larger surface lateral flows are generated than the case of the high saturated hydraulic conductivity (hereafter called the HIGH_K reference). In the LOW_K reference, the topography-driven surface lateral flows reach the left edge of the domain (Figure 1b). In the HIGH_K reference, supplied water moves vertically rather than horizontally and the topography-driven surface flow reaches around x = 1000~1500m (Figure 1d).

For the data assimilation experiment, an ensemble of 50 realizations was generated. Each ensemble member has different saturated hydraulic conductivity and rainfall rate. Lognormal multiplicative noise was added to surface saturated hydraulic conductivity and rainfall rate of the synthetic reference runs. This specification of uncertainty in rainfall was also adopted in Crow et al. (2011). The two parameters of the lognormal distribution, commonly called μ and σ, were set to 0 and 0.15, respectively. These parameters were chosen to give the sufficiently large error in precipitation and saturated hydraulic conductivity. In addition, this setting makes the rainfall PDF similar to the Gaussian distribution, which is important to interpret the results of the experiments (see the discussion section). The initial groundwater depth of each ensemble member was drawn from the uniform distribution from 2.0m to 3.5m. The hydrostatic pressure gradient was assumed for the initial pressure heads in the unsaturated soil layers.

The virtual hourly observations were generated by adding the Gaussian white noise whose mean is zero to the volumetric soil moisture simulated by the synthetic reference runs. The observation error (the standard deviation of the added Gaussian white noise) was set to 0.05 m^3/m^3. It was assumed that the volumetric soil moistures can be observed in every model's soil layer from surface to the depth of 1m at the specific location. These soil moisture observations can be obtained in the in-situ observation sites (e.g., Dorigo et al., 2017). In the section 3.2, I will assume that only surface soil moisture observation can be accessed, which is more realistic since satellite sensors can observe only surface soil moisture. I assumed that the small part of the domain can be observed. The two scenarios of the observation's location are provided. In the first scenario (hereafter called the UP_O scenario), the volumetric soil moisture at the upper part of the slope (x = 2500m) was observed. In the UP_O scenario, I could observe the volumetric soil moisture in the upper part of the slope where it heavily rains and tried to infer the soil moisture in the lower part of the slope where it does not rain by propagating the observation's information downhill. In the second scenario (hereafter called the DOWN_O scenario), the volumetric soil moisture at the lower part of the slope (x = 1500m) was observed. In the DOWN_O scenario, I could observe the volumetric soil moisture in the lower part of the slope where it does not rain and tried to infer the soil moisture in the upper part of the slope where it heavily rains by propagating the observation's information uphill.

Since I had the two synthetic reference runs (the HIGH_K and LOW_K references) and the two observation scenarios (the UP_O and DOWN_O scenarios), I implemented totally four data assimilation experiments. Table 1 summarizes the data assimilation experiments implemented in this study. For instance, in the HIGH_K-UP_O experiment, I chose the HIGH_K reference and generated an ensemble of 50 realizations from the HIGH_K reference. The soil moisture observations were generated from the HIGH_K reference at the location of x = 2500m and assimilated into the model every hour. The simulated volumetric soil moisture of the data assimilation experiment was compared with that of the HIGH_K reference.

In addition to the data assimilation (DA) experiments, I implemented the NoDA experiment (also called the open-loop experiment in the literature of the LDAS study) in which the ensemble was used but no observation data were assimilated. Please note that in the NoDA experiment, the true rainfall rate and saturated hydraulic conductivity were unknown so that I could not accurately estimate the synthetic true state variables. I will evaluate how this negative impact of uncertainties in rainfall and saturated hydraulic conductivity can be mitigated by data assimilation in the DA experiment.

As evaluation metrics, root-mean-square-error (RMSE) was used:

$$\mathrm{RMSE} = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(F_i - T)^2} \quad (18)$$

where k is the ensemble number, F_i is the volumetric soil moisture simulated by the i-th member in the DA or NoDA experiment, T is the volumetric soil moisture simulated by the synthetic reference run. I used all ensemble members to calculate RMSE because I should evaluate not only if the ensemble mean is consistent to the synthetic truth, but also if the extremely large ensemble spread simulated in the NoDA experiment is appropriately reduced.

To evaluate the impact of data assimilation, the improvement rate (IR) was defined and calculated by the following equation:

$$\mathrm{IR} = \frac{\overline{RMSE_{DA}} - \overline{RMSE_{NoDA}}}{\overline{RMSE_{NoDA}}} \quad (19)$$

where $\overline{RMSE_{DA}}$ and $\overline{RMSE_{NoDA}}$ are time-mean RMSE of the DA and NoDA experiments, respectively. The negative IR indicates that data assimilation positively impacts the simulation of soil moisture. The metrics described above was calculated in the whole domain. In the DA experiment, soil moisture values before the update by ETKF (i.e. initial guess) were used to calculate the metrics.

Four of 120-hour rain/no rain cycles were applied so that the computation period was 480 hours. The spin-up results in the first 120 hours were not used to calculate the evaluation metrics. Since the steady state of groundwater level is not the scope of this paper, the long spin-up is not absolutely necessary.

3.1.2. Results

Figure 2a shows the IR of the LOW_K-UP_O experiment. The time series of the DA and NoDA experiment and the synthetic reference run in the LOW_K-UP_O experiment can be found in Figure S1. The data assimilation efficiently propagates the information of the observations located in the upper part of the slope (see the black arrow in Figure 2a) both horizontally and vertically. Despite the uncertainty in rainfall and hydraulic conductivity, RMSE is reduced by data assimilation not only directly under the observation but also the lower part of the slope where it does not rain. The estimated $K_{s,surface} \approx 0.00508$ [m/h] by ETKF is mostly identical to the synthetic truth. However, the increase of RMSE by data assimilation can be found at the left edge of the domain, which is far from the location of the observation. The impact of data assimilation on the surface soil moisture simulation is small because the volumetric soil moisture's RMSE of the NoDA experiment in this surface soil layer is already small ($\leq 0.01 m^3/m^3$) in the case of the LOW_K reference so that any improvements do not make sense.

Figure 2b shows the IR of the LOW_K-DOWN_O experiment (see also Figure S2 for time series). The IR's spatial pattern of the LOW_K-DOWN_O experiment is similar to that of the LOW_K-UP_O experiment except for the left edge of the domain. It is promising that I can accurately infer soil moisture in the region where it heavily rains from the shallow soil moisture observations in the region where it does not rain. The estimated $K_{s,surface} \approx 0.00512$ [m/h] by ETKF is mostly identical to the synthetic truth.

Figure 3a shows the difference of time-mean RMSEs ($\overline{RMSE_{DA}}$ in equation (18)) between the LOW_K-UP_O and LOW_K-DOWN_O experiments. Although observing the lower part of the slope slightly improves the soil moisture simulation at the left edge of the domain compared with observing the upper part of the slope (the reason for it will be explained later), there are few differences between the UP_O and DOWN_O scenarios in the case of the LOW_K reference. The soil moisture observations have large representativeness and I can efficiently infer soil moisture in the soil columns which are horizontally and vertically far from the observations.

Figure 2c shows the IR of the HIGH_K-UP_O experiment (see also Figure S3 for time series). The data assimilation significantly reduces RMSE of the soil moisture simulation directly under the observations (see the black arrow in Figure 2c), which indicates that the data assimilation efficiently propagates the information of the observations vertically. The saturated hydraulic conductivity estimated by ETKF is mostly identical to the synthetic truth ($K_{s,surface} \approx 0.0204$ [m/h]). However, the impact of the data assimilation on the soil moisture simulation in the lower part of the slope around x=1500m is marginal although there are large RMSE in the NoDA experiment (>0.05m^3/m^3) at the edge of the area where topography-driven surface flow reaches in the HIGH_K reference (see Figure 1d).

Figure 2d shows the IR of the HIGH_K-DOWN_O experiment (see also Figure S4 for time series). Although the observations in the lower part of the slope (see the black arrow in Figure 2d) significantly contribute to improving the soil moisture simulation in the downstream area of the observation and accurately estimating $K_{s,surface} \approx 0.0208$ [m/h], the impact of the data assimilation on the shallow soil moisture simulation around x=500~1000m is marginal. As I found in the LOW_K-DOWN_O experiment, the shallow soil moisture observations in the region where it does not rain can improve the soil moisture simulation in the region where it heavily rains. However, the IR of the HIGH_K-DOWN_O experiment in the upper part of the slope is smaller than that of the LOW_K-DOWN_O experiment (see Figure 2b and 2d).

The high representativeness of the observations which I found in the case of the LOW_K reference (i.e. the small difference of RMSEs between two observation scenarios) cannot be found in the case of the HIGH_K reference. Figure 3b shows the difference of time-mean RMSEs ($\overline{RMSE_{DA}}$ in equation (18)) between the HIGH_K-UP_O and HIGH_K-DOWN_O experiments. Compared with the LOW_K reference case (Figure 3a), there are significant differences between the UP_O and DOWN_O scenarios in the case of higher saturated hydraulic conductivity. In this case, the vertical propagation of the observations' information is more efficient than the horizontal propagation.

The relatively low efficiency of the data assimilation and the low representativeness of the soil moisture observations in the case of the HIGH_K reference are caused by the non-Gaussian background error distribution. I calculated KLD by comparing the PDF of the NoDA ensemble (p in equation (17)) with the Gaussian PDF which has the mean and variance of the NoDA ensemble (q in equation (17)). Figure 4 shows that the NoDA ensemble in the case of the HIGH_K reference has stronger non-Gaussianity than the case of the LOW_K reference especially in the shallow soil layers. The strong non-Gaussianity of the NoDA ensemble generated from the HIGH_K reference can be found at the edge of the area where the topography-driven surface flow reaches (Figure 1d). Figure 5 shows that there is the bifurcation of the ensemble in this region when the ensemble is generated from the HIGH_K reference. The process of topography-driven surface flows is switched on if and only if the surface soil is saturated (see equation (6)) so that the ensemble tends to be bifurcated into the members with surface flows and without surface flows. As I mentioned in section 2.2, in the ETKF, the estimation of the state and parameter variables is optimal if and only if the model's error has the Gaussian PDF and the relationship between observed variables and unobserved variables is linear. Therefore, the non-Gaussianity of the prior ensemble induced by the strong non-linear dynamics of surface lateral flows makes the ETKF inefficient. It is more difficult to reconstruct 3-D fields of soil moisture in high conductivity soils since the 1-D vertical water movement is more dominant. The absolute RMSE of the NoDA experiment in the HIGH_K reference is larger than the LOW_K reference in many places (not shown). Please note that the non-Gaussianity can also be found in the LOW_K reference at the edge of the domain (x=500m) due to the non-linear dynamics of surface lateral flows, which causes the degradation of the soil moisture simulation in the LOW_K-UP_O experiment (see Figure 2a).

One of the major simplifications in this experiment is spatially homogeneous surface saturated hydraulic conductivity. The optimization of it can efficiently improve the soil moisture simulation in the whole domain. However, the optimization of this homogeneous surface saturated hydraulic conductivity has a limited impact on the soil moisture simulation. Figure S5 shows the IR of the HIGH_K-DOWN_O experiment where the parameter optimization by ETKF is switched off. Even if I do not optimize the surface saturated hydraulic conductivity, I could obtain the similar IR to the original experiment and the shallow soil moisture observations in the region where it does not rain can improve the soil moisture simulation in the region where it heavily rains. The horizontal propagation of the observations' information shown in this experiment was brought out not only by the estimation of spatially homogeneous saturated hydraulic conductivity but also by the adjustment of state variables (i.e., pressure head and volumetric soil moisture).

Please note that the improvement of the soil moisture simulation cannot be found if the topography-driven surface flow is neglected. Figure S6 shows the IR of the LOW-K_DOWN-O experiment where the topography-driven surface flow is neglected in the ParFlow simulation. Please note that although many conventional land surface models neglected or parameterized lateral flows, this assumption can be applied only in the coarse spatial resolution (>25km), which is not the case of this experimental setting. The imperfect model physics of ParFlow substantially degrades the skill to simulate soil moisture and data assimilation cannot compensate this degradation. This point will also be discussed in the section 3.2 more deeply.

3.2. Simple 3-D slope with heterogeneous hydraulic conductivity

3.2.1. Experiment design

To further demonstrate how land data assimilation works with topography-driven surface lateral flows, I implemented another synthetic experiment which is more realistic than that shown in section 3.1. The 3-D domain has a horizontal extension of 4000 m$\times$4000m and a vertical extension of 3m. The domain was horizontally discretized into 40$\times$40 grid cells with a size of 100m$\times$100m and vertically discretized into 30 grid cells with a size of 0.1m. The domain has a 10% slope in both x and y directions (see Figure 6a). The parameters of the van Genuchten relationship, porosity and Manning's coefficient were set to the same variables for the synthetic experiment in section 3.1.

The spatially heterogeneous surface saturated hydraulic conductivity was generated following Kurtz et al. (2016). The field of $log_{10}(K_{s,surface})$ was generated by two-dimensional unconditioned sequential Gaussian simulation. A Gaussian variogram with nugget, sill, and range values of 0.0 $log_{10}(\mathrm{m/h})$, 0.1 $log_{10}(m^2h^2)$, and 12 model grids (1200m), respectively was used to simulate the spatial distribution of $log_{10}(K_{s,surface})$. A constant value of -2.30 $log_{10}(\mathrm{m/h})$ (i.e. 0.005 (m/h)) was added to the generated field so that the mean of the logarithm of surface saturated hydraulic conductivity was set to -2.30 (i.e. 0.005(m/h)). This method to generate the field of the saturated hydraulic conductivity has been used previously (e.g., Kurtz et al. 2016). Subsurface saturated hydraulic conductivity was calculated by equation (3). An ensemble of 51 realizations of $log_{10}(K_{s,surface})$ was generated and one of them was chosen as a synthetic reference (Figure 6a). The remaining 50 members were used for data assimilation experiments.

A rainfall rate $\mathrm{R(x, y)}$ (mm/h) was modelled by a logistic function:

$$R(x,y) = \frac{R_{max}}{1+100\exp\left(-0.2\times\frac{x+y}{2}\right)} \quad (20)$$

where x and y are horizontal grid numbers ($1 \le x \le 40, 1 \le y \le 40$). In the synthetic reference, the maximum rainfall rate in the domain, R_{max}, was set to 50 (mm/h) (Figure 6b). This rainfall rate was applied for 3 hours and then the period with no rainfall and evaporation of 0.075mm/h lasted for 117 hours. For data assimilation experiment, an ensemble of 50 realization of $R(x,y)$ was generated by adding a lognormal multiplicative noise to R_{max} of the synthetic reference. The two parameters of the lognormal distribution, commonly called μ and σ, were set to 0 and 0.15, respectively.

Figure 6c shows the distribution of surface soil moisture in the synthetic reference run. Strong rainfall rate applied in the upper part of the slope generates the topography-driven surface lateral flows. The virtual hourly observations were generated by adding the Gaussian white noise, whose mean is zero and standard deviation is 0.05 m^3/m^3, to the volumetric surface soil moisture simulated by the synthetic reference run. Unlike the experiment in section 3.1, only surface soil moisture can be observed in this synthetic experiment, which makes this experiment more realistic since satellite sensors can observe only surface soil moisture. Three different observing networks with different observation densities were used (Figure 7). The observing networks shown in Figure 7a, 7b, and 7c have totally 1, 9, and 361 observations and are called obs1, obs9, and obs361, respectively.

In the DA experiments, those virtual observations of surface soil moisture were assimilated every hour to adjust pressure head and saturated hydraulic conductivity. As I did in the section 3.1, the NoDA experiments were also implemented. The two different configurations of ParFlow were used for both DA and NoDA experiments. In the first configuration, called OF (Overland Flow), Parflow explicitly solves overland flows. In the second configuration, called noOF, Parflow assumes the flat terrain for surface flows so that no overland flows are generated. Since the synthetic reference run explicitly considers the topography-driven surface flow, the configuration of noOF assumes that the model physics is imperfect. I implemented 8 numerical experiments which are summarized in Table 2. For example, the OF_DA_obs9 experiment is the data assimilation experiment with the observing network shown in Figure 7b, in which Parflow explicitly solves the topography-driven surface flow. The noOF_NoDA is the model run without assimilating observations, in which Parflow does not consider the topography-driven surface flow.

3.2.2. Results

Figure 8a shows the RMSE of soil moisture simulation of a second soil layer (i.e. 10-20cm soil depth) in all 8 experiments (the same conclusion described below can be obtained by analyzing all of shallow soil layers). When Parflow explicitly solves the topography-driven surface flow, data assimilation substantially reduces RMSE of the soil moisture simulation (green bars in Figure 8a). The OF_DA_obs361 experiment has the smallest RMSE so that a denser observing network is beneficial to estimate soil moisture, although there is the stalled improvement from the OF_DA_obs1 experiment to the OF_DA_obs9 experiment (the reason for it will be explained later). Figure 8b shows the RMSE of the estimation of saturated surface hydraulic conductivity in all 8 experiments. Data assimilation also reduces the uncertainty in model's parameters (green bars in Figure 8b). However, the OF_DA_obs361 experiment has larger RMSE than the other DA experiments. This is because the adjustment of hydraulic conductivity in the OF_DA_obs361 experiment greatly mitigates not only the errors induced by uncertainty in hydraulic conductivity but those induced by uncertainty in rainfall rate. In the OF configuration, there are two sources of errors, rainfall rate and hydraulic conductivity. However, data assimilation can adjust only hydraulic conductivity in this study. Although it is expected that the adjustment of hydraulic conductivity mainly mitigates the errors of simulated volumetric soil moisture induced by uncertainty in hydraulic conductivity, it also greatly mitigates those induced by uncertainty in rainfall rate by adjusting the parameter in the incorrect direction when the number of observations is large. Therefore, the assimilation of a large number of observations degrades the estimation of saturated hydraulic conductivity despite the improvement of the soil moisture simulation.

The noOF_NoDA experiment has larger RMSE than the OF_NoDA experiment due to the negligence of the topography-driven surface flow. In the noOF configuration, data assimilation also improves the soil moisture simulation (red bars in Figure 8a). The noOF_DA_obs361 experiment outperforms the OF_NoDA experiment so that data assimilation with a dense observing network can compensate the negative impact of neglecting the topography-driven surface flow. Although data assimilation positively impacts the parameter estimation, the denser observing network cannot reduce RMSE of hydraulic conductivity estimation (red bars in Figure 8b). The negative impact of the dense observations in the noOF_DA_obs361 experiment on the parameter estimation is larger than in the OF_DA_obs361 experiment. In addition to rainfall rate and hydraulic conductivity, the imperfect model physics (i.e., no topography-driven surface flow) is the source of error in the noOF configuration. The assimilation of a large number of observations degrades the estimation of saturated hydraulic conductivity because it greatly mitigates the impact of all systematic errors which comes from three different sources only by adjusting hydraulic conductivity.

Figure 9 shows the difference of RMSE of the soil moisture simulation between the DA experiments and the OF_NoDA experiment. In the DA configuration, the improvement of the soil moisture estimation can be found in a large area even if there is a single observation in the center of the domain (Figure 9a). Figure 9b shows that the increase of the number of observations substantially improves the soil moisture simulation in the region which is affected by topography-driven surface flow (see also Figure 6c). However, the skill to simulate soil moisture is severely degraded in the lower-left corner of the domain, which causes the stalled improvement from the OF_DA_obs1 experiment to the OF_DA_obs9 experiment shown in Figure 8a. Figure 9c shows that although the far denser observing network can slightly mitigate this degradation, increasing the number of observations cannot efficiently solve this issue. This degradation is caused by the bifurcation of ensemble members at the edge of the area where topography-driven surface flow reaches (Figure S7). Figure 10 shows KLD in the OF_NoDA and noOF_NoDA experiments. Figure 10a clearly shows that the ensemble simulation of volumetric soil moisture generates the strong non-Gaussianity at the edge of the area where topography-driven surface flow reaches, which harms the efficiency of the ETKF. This finding is consistent to what I found in the previous experiment in section 3.1.

In the noOF configuration, there are large errors in the area around 500<=x, y <=1500 (not shown) since the increase of soil moisture in this area is caused by the topography-driven surface flow which is neglected in the noOF configuration. Figures 9d and 9e show that the sparse observations cannot completely remove this degradation caused by imperfect model physics. Figure 9f shows that the noOF_DA_obs361 can outperform the OF_NoDA experiment in exchange for the degradation of the parameter estimation as I found in Figure 8. The unstable behavior of the ETKF found in the OF configuration does not occur when the topography-driven surface flow is neglected since the ensemble simulation does not generate the non-Gaussian background distribution (Figure 10b). Although ETKF can significantly improve the simulation skill of the hyperresolution land model in many cases, I found its limitation when it is applied to the problems with the topography-driven surface lateral flows. Figure 10 clearly indicates that this limitation appears only if lateral water flows are explicitly considered.

4. Discussion

In this study, I revealed that the hyperresolution integrated surface-subsurface hydrological model gives the unique opportunity to effectively use soil moisture observations to improve the soil moisture simulation in terms of a horizontal propagation of observation's information in a model space. I found that the explicit calculation of the topography-driven surface flow has an important role in propagating the information of soil moisture observation horizontally by data assimilation even if there is considerable heterogeneity of meteorological forcing. It is possible that the soil moisture observations in the area where it does not heavily rain can improve the soil moisture simulation in the severe rainfall area.

This potential cannot be brought out in the conventional 1-D LSM where sub-grid scale surface runoff is parameterized and the surface flows in one grid do not move to the adjacent grids. I found that neglecting the topography-driven surface flow causes significant bias in the soil moisture simulation and this bias cannot be completely mitigated by data assimilation especially in the case of a sparse observing network. However, I found that even if the model uses imperfect physics which neglects the interaction between topography-driven surface lateral flows and subsurface soil moisture, assimilating soil moisture observations into the model's three-dimensional state and parameter space can improve the skill to estimate soil moisture and hydraulic conductivity. This finding implies that the conventional 1-D LSM with full 3-D data assimilation may be a computationally cheap and reasonable choice in some cases although many land data assimilation systems with the conventional 1-D LSM currently update state variables only in a single model's horizontal grid which is identical to the location of the observation.

The conventional ensemble data assimilation (i.e. ETKF) severely suffers from the non-Gaussian background error PDFs caused by the strongly nonlinear dynamics of the topography-driven surface flow although it has been widely used by previous studies (e.g., Camporese et al. (2009); Camporese et al. (2010); Ridler et al. (2014); Zhang et al. (2015); Kurtz et al. (2016); Zhang et al. (2018)). The efficiency of ETKF to propagate the information of observations horizontally in the model space is limited in the edge of the area where the topography-driven surface flow reaches. Please note that the low representativeness of the soil moisture observations in the case of the HIGH_K reference shown in section 3.1 is due to the limitation of the Kalman filter that the error PDFs need to follow the Gaussian distribution to get the optimal estimation so that the increase of the ensemble size cannot solve this issue. I implemented the data assimilation experiment in the case of the HIGH_K reference with an ensemble size of 500, which is 10 times larger than the original experiments shown in section 3.1, and found no significant improvement of the soil moisture simulation (not shown). Some studies revealed that volumetric soil moisture distributions follow the Gaussian distribution better than pressure head so that they recommend updating soil moisture as a state variable (e.g., Zhang et al. (2018)). However, in this study, I found that volumetric soil moisture distributions have bimodal structure and do not follow the Gaussian distribution. The limitation of ensemble Kalman filters found in this study does not depend on the updated state variables.

In addition, I found ensemble clustering in which the ensemble members are split into a single outlier and the others (see Figures S1-S4). The previous studies found that this ensemble clustering is generated by the non-Gaussian PDF (Anderson 2010; Amezcua et al. 2012). Ensemble clustering shown in the analysis timeseries also implies that the non-

Gaussian PDF plays an important role in the data assimilation of the hyperresolution land model.

The spatially dense soil moisture observations are needed to efficiently constrain state variables at the edge of surface flows. High resolution soil moisture remote sensing based on satellite active and passive combined microwave observations at the 1 km spatial resolution (e.g., He et al. 2018) and the assimilation of those data (Lievens et al. 2017) may be important in the era of the hyperresolution land modeling. High resolution observations of surface inundated water from satellite imagery with a spatial resolution finer than 100 m (e.g., Sakamoto et al. 2007; Arnesen et al. 2013) may also be useful. However, the numerical experiment in section 3.2 implies that the dense observing network of surface soil moisture cannot completely remove the negative impact of the non-Gaussian background PDF.

As a possible heuristic approach to avoid the negative impact of the non-Gaussian background PDF, I can omit to update the state variables in the edge of the area where topography-driven surface flow reaches. The numerical experiments clearly indicate that the negative impact of the non-linear physics and non-Gaussian PDF is found only in the edge of flooding areas so that it is beneficial to simply omit to update the state variables in this area. It is similar but not conceptually identical to the localization method, in which the spurious correlation sampled by an ensemble is eliminated by spatially restricting the impact of assimilating observation (e.g., Rasmussen et al. 2015; Anderson 2007; Bishop and Hodyss 2009).

Reducing the uncertainty in rainfall positively impacts the efficiency of data assimilation since the bifurcation of simulated soil moisture found in Figure 5c is originally induced by the uncertainty in rainfall. Although assimilating land hydrological observations to improve the rainfall input has been intensively investigated (e.g., Sawada et al. 2018; Herrnegger et al. 2015; Crow et al. 2011; Vrugt et al. 2008), it has yet to be applied to hyperresolution land models. Please note that the parameters of the lognormal distribution to model the uncertainty in rainfall were specified to make the rainfall PDF similar to the Gaussian distribution. I chose the lognormal distribution in order not to generate negative rainfall values and I intended not to introduce non-Gaussianity into the external forcing. The rainfall input which follows the Gaussian PDF was transformed into the non-

Gaussian PDF of the background error by the strongly nonlinear dynamics of the topography-driven surface flow.

To explicitly consider non-Gaussianity and non-linear relationship between observed and unobserved variables induced by the topography-driven surface flow, the particle filters may be useful. The particle filter can represent a probability distribution (including non-Gaussian distributions) directly by an ensemble. Particle filters have been intensively applied to conventional 1-D LSMs (e.g., Sawada et al. 2015; Qin et al. 2009) and lumped hydrological models (e.g., Yan and Moradkhani 2016; Vrugt et al. 2013). Although particle filtering in a high dimensional system suffers from the "curse of dimensionality" (e.g., Snyder et al. 2008), some studies developed the methodology to improve the efficiency of particle filtering (e.g., van Leeuwen 2009; Poterjoy et al. 2019). The applicability of particle filtering to 3-D hyperresolution land models should be assessed in the future.

Since the synthetic numerical experiments in this paper adopted the simple and minimalistic setting, the findings of this paper may be exaggerated. There are no river channels in the synthetic experiment so that the skill to simulate river water level and discharge cannot be discussed, which is the major limitation of this study. The simple representation of soil properties is also a limitation of this study. Although the prior uncertainty in rainfall and saturated hydraulic conductivity was arbitrary chosen in this study, the specification of the prior knowledge is not straightforward in the real-world applications. In future work, the contributions of the topography-driven surface runoff process to the data assimilation of hydrological observations should be quantified in real-world applications. In addition, in the virtual experiment of this paper, I neglected some of the important land processes such as transpiration, canopy interception, snow, and frozen soil. These processes affect the source term of equation (1) in hyper-resolution land models (e.g., Shrestha et al. 2014). Since the inclusion of the neglected processes do not change the structure of the original ParFlow, the findings of this study can be robust to the models which include these processes. Although they are generally not primary factors in the propagation of overland flows generated by extreme rainfall, which has a shorter timescale than the neglected processes, those processes should be considered in the future.

The other limitation of this study is that I could not thoroughly evaluate the skill of the ensemble data assimilation to quantify the uncertainty of its prediction. Following

Abbazadeh et al. (2019), I calculated the 95% exceedance ratio and found that the ensemble forecast was systematically overconfident (not shown). In the synthetic experiments of this study, the number of rainfall events was small, and the timing and magnitude of rainfall were not diversified. Due to this limited amount of data, it is difficult to deeply discuss the accuracy of the quantified uncertainty by data assimilation. While the skill of lumped hydrological models was often evaluated by the probabilistic performance measures such as the 95% exceedance ratio (e.g., Abbazadeh et al. (2019)), the uncertainty quantification of the simulation of hyper-resolution land models is in its infancy. How surface lateral flows affect the accuracy of the uncertainty quantification by data assimilation should be investigated using more realistic data.

5. Conclusions

The simplified synthetic experiments of this study indicate that topography-driven lateral surface flows induced by heavy rainfalls do matter for data assimilation of hydrological observations into hyperresolution land models. Even if there is extreme heterogeneity of rainfall, the information of soil moisture observations can be propagated horizontally in the model space and the soil moisture simulation can be improved by the ensemble

Kalman filter. However, the nonlinear dynamics of the topography-driven surface flow induces the non-Gaussianity of the model error, which harms the efficiency of data assimilation of soil moisture observations. It is difficult to efficiently constrain model states at the edge of the area where the topography-driven surface flow reaches by linear-Gaussian filters, which brings the new challenge in land data assimilation for hyperresolution land models. Future work will focus on the real-world applications using intense in-situ soil moisture observation networks and/or high-resolution satellite soil moisture observations.

Acknowledgement

This study was supported by the JSPS KAKENHI grant JP17K18352 and JP18H03800. I thank two anonymous reviewers for their constructive comments.

Code/Data Availability

All data used in this paper are stored in the repository of the University of Tokyo for 5 years and available upon request to the author. The ETKF code used in this study can be found at https://github.com/takemasa-miyoshi/letkf.

Author Contribution

YS designed the study, executed numerical experiments, analyzed the results, and wrote the paper.

Competing interests

The author declares no competing interests.

References

Abbaszadeh, P., Moradkhani, H., and Daescu, D., N.: The Quest for Model Uncertainty Quantification: A Hybrid Ensemble and Variational Data Assimilation Framework. *Water Resources Research, 55*, 2407– 2431. https://doi.org/10.1029/2018WR023629, 2019.

Ait-El-Fquih, B., El Gharamti, M., and Hoteit, I.: A Bayesian consistent dual ensemble Kalman filter for state-parameter estimation in subsurface hydrology. *Hydrology and Earth System Sciences*, *20*(8), 3289–3307. https://doi.org/10.5194/hess-20-3289-2016, 2016

Amesen et al.: Monitoring flood extent in the lower Amazon River floodplain using ALOS/PALSAR ScanSAR images. *Remote Sensing of Environment, 130,* 51-61. https://doi.org/10.1016/j.rse.2012.10.035, 2013

Amezcua, J., Ide, K., Bishop, C. H., and Kalnay, E.: Ensemble clustering in deterministic ensemble Kalman filters, *Tellus, 64A,* 1-12, https://doi.org/10.3402/tellusa.v64i0.18039, 2012.

Anderson, J. L.: Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D: Nonlinear Phenomena*, *230*(1–2), 99–111. https://doi.org/10.1016/j.physd.2006.02.011, 2007

Anderson, J. L.: A non-Gaussian ensemble filter update for data assimilation*, Mon. Wea. Rev., 138,* 4186-4198, https://doi.org/10.1175/2010MWR3253.1, 2010.

Bandara, R., Walker, J. P., and Rüdiger, C.: Towards soil property retrieval from space: Proof of concept using in situ observations. *Journal of Hydrology, 5*12, 27–38. https://doi.org/10.1016/j.jhydrol.2014.02.031, 2014

Bandara, R., Walker, J. P., Rüdiger, C., and Merlin, O.: Towards soil property retrieval from space: An application with disaggregated satellite observations. *Journal of Hydrology, 522,* 582-593, https://doi.org/10.1016/j.jhydrol.2015.01.018, 2015

Beven, K.: On subsurface stormflow: an analysis of response times, *Hydrological Science Journal, 27,* 505-521, doi:10.1080/02626668209491129, 1982

Bishop, C.H., Etherton, B., J., and Majumdar, S., J.: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Monthly. Weather. Review.,* ***129***, 420–436, https://doi.org/10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2, 2001.

Bishop, C. H., and Hodyss, D.: Ensemble covariances adaptively localized with ECO-RAP. Part 1: Tests on simple error models. *Tellus, 61A,* 84–96, https://doi.org/10.1111/j.1600-0870.2008.00371.x, 2009.

Camporese, M., Paniconi, C., Putti, M., and Salandin, P.: Ensemble Kalman filter data assimilation for a process-based catchment scale model of surface and subsurface flow. *Water Resources Research*, *45*(10), 1–14. https://doi.org/10.1029/2008WR007031, 2009.

Camporese, M., Paniconi, C., Putti, M., and Orlandini, S.: Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data. *Water Resources Research*, *46*(2). https://doi.org/10.1029/2008WR007536, 2010.

Crow, W. T., Van Den Berg, M. J., Huffman, G. J., and Pellarin, T.: Correcting rainfall using satellite-based surface soil moisture retrievals: The Soil Moisture Analysis Rainfall Tool (SMART). *Water Resources Research*, *47*(8), 1–15. https://doi.org/10.1029/2011WR010576, 2011.

Dorigo et al.: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions, *Remote Sensing of Environment, 203,* 185-215, https://doi.org/10.1016/j.rse.2017.07.001, 2017

Duc, L., and Saito, K.: Verification in the presence of observation errors: Bayesian point of view. *Q. J. R. Meteorol. Soc., 144,* 1063– 1090. https://doi.org/10.1002/qj.3275, 2018.

Fang, Y., L. R. Leung, Z. Duan, M. S. Wigmosta, R. M. Maxwell, J. Q. Chambers, and J. Tomasella: Influence of landscape heterogeneity on water available to tropical forests in an Amazonian catchment and implications for modeling drought response, *J. Geophys. Res. Atmos., 122,* https://doi.org/10.1002/2017JD027066, 2017.

Ghanbarian-Alavijeh, B., Liaghat, A., Huang, G., H., and van Genuchten, M. Th.,: Estimation of the van Genuchten soil water retention properties from soil textural data. *Pedosphere. 20(4)*: 456–465., 2010

Han, X., Franssen, H.-J. H., Montzka, C. and Vereecken, H.: Soil moisture and soil properties estimation in the Community Land Model with synthetic brightness temperature observations, *Water Resources Research., 50,* 6081 - 6105, https://doi.org/10.1002/2013WR014586, 2014

Herrnegger, M., Nachtnebel, H. P., and Schulz, K.: From runoff to rainfall : inverse rainfall – runoff modelling in a high temporal resolution, 4619–4639. https://doi.org/10.5194/hess-19-4619-2015, 2015

He, L., Hong, Y., Wu, X., Ye, N., Walker, J. P. and Chen, X.: Investigation of SMAP Active–Passive Downscaling Algorithms Using Combined Sentinel-1 SAR and SMAP Radiometer Data. *IEEE Transactions on Geosciences and Remote Sensing, 56,* 4906-4918, https://doi.org/10.1109/TGRS.2018.2842153, 2018

Hendricks Franssen, H. J., and Kinzelbach, W.: Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem. *Water Resources Research*, *44*(9), 1–21. https://doi.org/10.1029/2007WR006505, 2008

Houborg, R., Rodell, M., Li, B., Reichle, R., and Zaitchik, B. F.: Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. *Water Resources Research*, *48*(7). https://doi.org/10.1029/2011WR011291, 2012

Houtekamer, P. L., and Zhang, F.: Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*, MWR-D-15-0440.1. https://doi.org/10.1175/MWR-D-15-0440.1 , 2016

Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, *230*(1–2), 112–126. https://doi.org/10.1016/j.physd.2006.11.008, 2007

Ji, P., Yuan, X., and Liang, X. Z.: Do Lateral Flows Matter for the Hyperresolution Land Surface Modeling? *Journal of Geophysical Research: Atmospheres*, 1–16. https://doi.org/10.1002/2017JD027366, 2017

Keune, J., F. Gasper, K. Goergen, A. Hense, P. Shrestha, M. Sulis, and S. Kollet: Studying the influence of groundwater representations on land surface-atmosphere feedbacks during the European heat wave in 2003, *Journal of Geophysical Research. Atmospheres*, 121, 13,301–13,325, https://doi.org/10.1002/2016JD025426, 2016.

Kollet, S. J., and Maxwell, R. M.: Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. *Advances in Water Resources*, *29*(7), 945–958. https://doi.org/10.1016/j.advwatres.2005.08.006, 2006.

Kondo, K., and Miyoshi, T.: Non-Gaussian statistics in global atmospheric dynamics: a study with a 10 240-member ensemble Kalman filter using an intermediate atmospheric general circulation model. *Nonlinear Processes in Geophysics, 26,* 211-225. https://doi.org/10.5194/npg-26-211-2019, 2019

Kumar, S. V., Reichle, R. H., Koster, R. D., Crow, W. T., and Peters-Lidard, C. D.: Role of Subsurface Physics in the Assimilation of Surface Soil Moisture Observations. *Journal of Hydrometeorology*, *10*(6), 1534–1547. https://doi.org/10.1175/2009JHM1134.1, 2009

Kurtz, W., He, G., Kollet, S. J., Maxwell, R. M., Vereecken, H., and Franssen, H. J. H.: TerrSysMP-PDAF (version 1.0): A modular high-performance data assimilation framework for an integrated land surface-subsurface model. *Geoscientific Model Development*, *9*(4), 1341–1360. https://doi.org/10.5194/gmd-9-1341-2016, 2016.

Kullback, S., and Leibler, R. A.: On information and sufficiency, *The Annals of Mathematical Statistics, 22*, 79-86, 1951.

Lawrence, D. M., et al.: Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *Journal of Advances in Modeling Earth Systems*, *3*(3), 1–27. https://doi.org/10.1029/2011MS000045, 2011

Li, B., Rodell, M., Zaitchik, B. F., Reichle, R. H., Koster, R. D., and van Dam, T. M.: Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe. *Journal of Hydrology*, *446–447*, 103–115. https://doi.org/10.1016/j.jhydrol.2012.04.035, 2012

Lievens, et al.: Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates. *Geophysical Research Letters*, *44*(12), 6145–6153. https://doi.org/10.1002/2017GL073904, 2017.

Martens, B., Miralles, D. G., Lievens, H., Schalie, R. Van Der, and Jeu, R. A. M. De.: GLEAM v3 : satellite-based land evaporation and root-zone soil moisture, *Geoscientific Model Development, 10,* 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017, 2017

Maxwell, R. M., and N. L. Miller: Development of a Coupled Land Surface and Groundwater Model. *Journal of Hydrometeorology, 6,* 233-247. https://doi.org/10.1175/JHM422.1, 2005

Maxwell, R. M., and Condon, L. E.: Connections between groundwater flow and transpiration partitioning. *Science, 353*, 377-380, https://doi.org/10.1126/science.aaf7891, 2016

Maxwell, R. M., Condon, L. E., and Kollet, S. J.: A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3, 923–937. https://doi.org/10.5194/gmd-8-923-2015, 2015.

Maxwell, R. M., and Kollet, S. J.: Interdependence of groundwater dynamics and land-energy feedbacks under climate change, *Nature Geoscience, 1,* 665–669. https://doi.org/10.1038/ngeo315, 2008.

Maxwell, R. M., Lundquist, J. K., Mirocha, J. D., Smith, S. G., Woodward, C. S., and Tompson, A. F. B.: Development of a Coupled Groundwater–Atmosphere Model. *Monthly Weather Review*, *139*(1), 96–116. https://doi.org/10.1175/2010MWR3392.1, 2011

Moradkhani, H., Hsu, K. L., Gupta, H., and Sorooshian, S.: Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resources Research, 41(5)*, 1–17. https://doi.org/10.1029/2004WR003604, 2005

Nerger, L., and Hiller, W.,: Software for ensemble-based data assimilation systems – implementation strategies and scalability. *Computers and Geosciences. 55,* 110–118. https://doi.org/10.1016/j.cageo.2012.03.026, 2013

Niu, G. Y., Paniconi, C., Troch, P. a., Scott, R. L., Durcik, M., Zeng, X., and Goodrich, D. C.: An integrated modelling framework of catchment-scale ecohydrological processes: 1. Model description and tests over an energy-limited watershed. *Ecohydrology*, *7*(2), 427–439. https://doi.org/10.1002/eco.1362, 2014

Paloscia, S, Pettinato, S., Santi, E., Notarnicola, C., Pasolli, L., and Reppucci, A.: Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation. *Remote Sensing of Environment, 134,* 234-248. https://doi.org/10.1016/j.rse.2013.02.027, 2013.

Pokhrel, P., and Gupta, H. V.: On the use of spatial regularization strategies to improve calibration of distributed watershed models. *Water Resources Research*, *46*(1), 1–17. https://doi.org/10.1029/2009WR008066, 2010

Poterjoy, J., Wicker, L., and Buehner, M.: Progress toward the application of a localized particle filter for numerical weather prediction. *Monthly Weather Review, 147(4)*, 1107–1126. https://doi.org/10.1175/MWR-D-17-0344.1, 2019

Qin, J., Liang, S., Yang, K., Kaihotsu, I., Liu, R., and Koike, T.: Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal. *Journal of Geophysical Research*, *114*(D15), 1–13. https://doi.org/10.1029/2008JD011358, 2009.

Rasmussen, J., Madsen, H., Jensen, K. H., and Refsgaard, J. C.: Data assimilation in integrated hydrological modeling using ensemble Kalman filtering: evaluating the effect of ensemble size and localization on filter performance. *Hydrology and Earth System Sciences*, *19*(7), 2999–3013. https://doi.org/10.5194/hess-19-2999-2015, 2015.

Ridler, M.-E., H. Madsen, S. Stisen,, S. Bircher, and R. Fensholt, Assimilation of SMOS-derived soil moisture in a fully integrated hydrological and soil-vegetation-atmosphere transfer model in Western Denmark, *Water Resour. Res., 50,* 8962 – 8981, https://doi.org/10.1002/2014WR015392, 2014.

Sakamoto, T., et al.: Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. *Remote Sensing of Environement, 109,* 295-313. https://doi.org/10.1016/j.rse.2007.01.011, 2007.

Sawada, Y. and Koike, T.: Simultaneous estimation of both hydrological and ecological parameters in an eco-hydrological model by assimilating microwave signal, *Journal of Geophysical Research – Atmospheres, 119,* 8839-8857, https://doi.org/10.1002/2014JD021536, 2014.

Sawada, Y., Koike, T., and Walker, J. P.: A land data assimilation system for simultaneous simulation of soil moisture and vegetation dynamics, *Journal of Geophysical Research – Atmospheres*, *120,* 5910-5930, https://doi.org/10.1002/2014JD022895, 2015

Sawada, Y., Nakaegawa, T, and Miyoshi, T.: Hydrometeorology as an inversion problem: Can river discharge observations improve the atmosphere by ensemble data assimilation?, *Journal of Geophysical Research - Atmospheres, 123*, 848-860, https://doi.org/10.1002/2017JD027531, 2018

Sellers, P. J., et al.: A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation. *Journal of Climate.* https://doi.org/10.1175/1520-0442(1996)009<0676:ARLSPF>2.0.CO;2, 1996

Shrestha, P., Sulis, M., Masbou, M., Kollet, S., and Simmer, C.: A Scale-Consistent Terrestrial Systems Modeling Platform Based on COSMO, CLM, and ParFlow. *Monthly Weather Review*, *142*(9), 3466–3483. https://doi.org/10.1175/MWR-D-14-00029.1, 2014

Snyder, C., Bengtsson, T., Bickel, P., and Anderson, J.: Obstacles to High-Dimensional Particle Filtering. *Monthly Weather Review*, *136*(12), 4629–4640. https://doi.org/10.1175/2008MWR2529.1, 2008

Tian, W., Li, X., Cheng, G. D., Wang, X. S., and Hu, B. X.: Coupling a groundwater model with a land surface model to improve water and energy cycle simulation.

Hydrology and Earth System Sciences, *16*(12), 4707–4723. https://doi.org/10.5194/hess-16-4707-2012, 2012.

Van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Science Society of America Journal*, *44*, 892–898, 1980.

Van Leeuwen, P., J.: Particle Filtering in Geophysical Systems, *Monthly Weather Review, 137,* 4089-4114, https://doi.org/10.1175/2009MWR2835.1, 2009

Verbeeck, H., Peylin, P., Bacour, C., Bonal, D., Steppe, K., and Ciais, P.: Fluxes in Amazon forests: Fusion of eddy covariance data and the ORCHIDEE model. Journal of Geophysical Research, 116(G2), 1–19. https://doi.org/10.1029/2010JG001544, 2011.

Vrugt, J. A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M., and Robinson, B. A.: Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Research*, *44*, 1–15. https://doi.org/10.1029/2007WR006720, 2008.

Vrugt, J. A., ter Braak, C. J. F., Diks, C. G. H., and Schoups, G.: Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications. *Advances in Water Resources*, *51*, 457–478. https://doi.org/10.1016/j.advwatres.2012.04.002, 2013.

Williams, J. L., and Maxwell, R. M.: Propagating Subsurface Uncertainty to the Atmosphere Using Fully Coupled Stochastic Simulations. *Journal of Hydrometeorology*, *12*(4), 690–701. https://doi.org/10.1175/2011JHM1363.1, 2011.

Wood, E. F., et al.: Hyperresolution global land surface modeling : Meeting a grand challenge for monitoring Earth's terrestrial water, *Water Resources Research, 47*, https://doi.org/10.1029/2010WR010090, 2011.

Yan, H., and Moradkhani, H.: Combined assimilation of streamflow and satellite soil moisture with the particle filter and geostatistical modeling. *Advances in Water Resources*, *94*, 364–378. https://doi.org/10.1016/j.advwatres.2016.06.002, 2016.

Yang, K., Koike, T., Kaihotsu, I., and Qin, J.: Validation of a Dual-Pass Microwave Land Data Assimilation System for Estimating Surface Soil Moisture in Semiarid Regions. *Journal of Hydrometeorology*, *10*(3), 780–793. https://doi.org/10.1175/2008JHM1065.1, 2009.

Yang, K., Watanabe, T., Koike, T., Li, X., Fujii, H., Tamagawa, K., and Ishikawa, H.: Auto-calibration System Developed to Assimilate AMSR-E Data into a Land Surface Model for Estimating Soil Moisture and the Surface Energy Budget. *Journal of the Meteorological Society of Japan*, *85A*, 229–242. https://doi.org/10.2151/jmsj.85A.229, 2007.

Zhang, D., Madsen, H., Ridler, M. E., Refsgaard, J. C., and Jensen, K. H.: Impact of uncertainty description on assimilating hydraulic head in the MIKE SHE distributed hydrological model, *Advances in Water Resources, 86,* 400-413, https://doi.org/10.1016/j.advwatres.2015.07.018, 2015.

Zhang, F., Snyder, C., and Sun, J.: Impacts of Initial Estimate and Observation Availability on Convective-Scale Data Assimilation with an Ensemble Kalman Filter.

Monthly. Weather. Review., **132**, 1238–1253, https://doi.org/10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2, 2004

Zhang, H., Kurtz, W., Kollet, S., Vereecken, H., and Franssen, H. J. H.: Comparison of different assimilation methodologies of groundwater levels to improve predictions of root zone soil moisture with an integrated terrestrial system model. *Advances in Water Resources*, *111*, 224–238. https://doi.org/10.1016/j.advwatres.2017.11.003, 2018.

Table 1. Configuration of the data assimilation experiments in section 3.1.

	hydraulic conductivity [m/h]	observation's location [m]
LOW_K-UP_O	0.005	2500
LOW_K-DOWN_O	0.005	1500
HIGH_K-UP_O	0.02	2500
HIGH_K-DOWN_O	0.02	1500

Table 2. Configuration of the data assimilation experiments in section 3.2

	overland flows	observing network
noOF_NoDA	none	no data assimilation
noOF_DA_obs1	none	Figure 7a
noOF_DA_obs9	none	Figure 7b
noOF_DA_obs361	none	Figure 7c
OF_NoDA	simulated	no data assimilation
OF_DA_obs1	simulated	Figure 7a
OF_DA_obs9	simulated	Figure 7b
OF_DA_obs361	simulated	Figure 7c

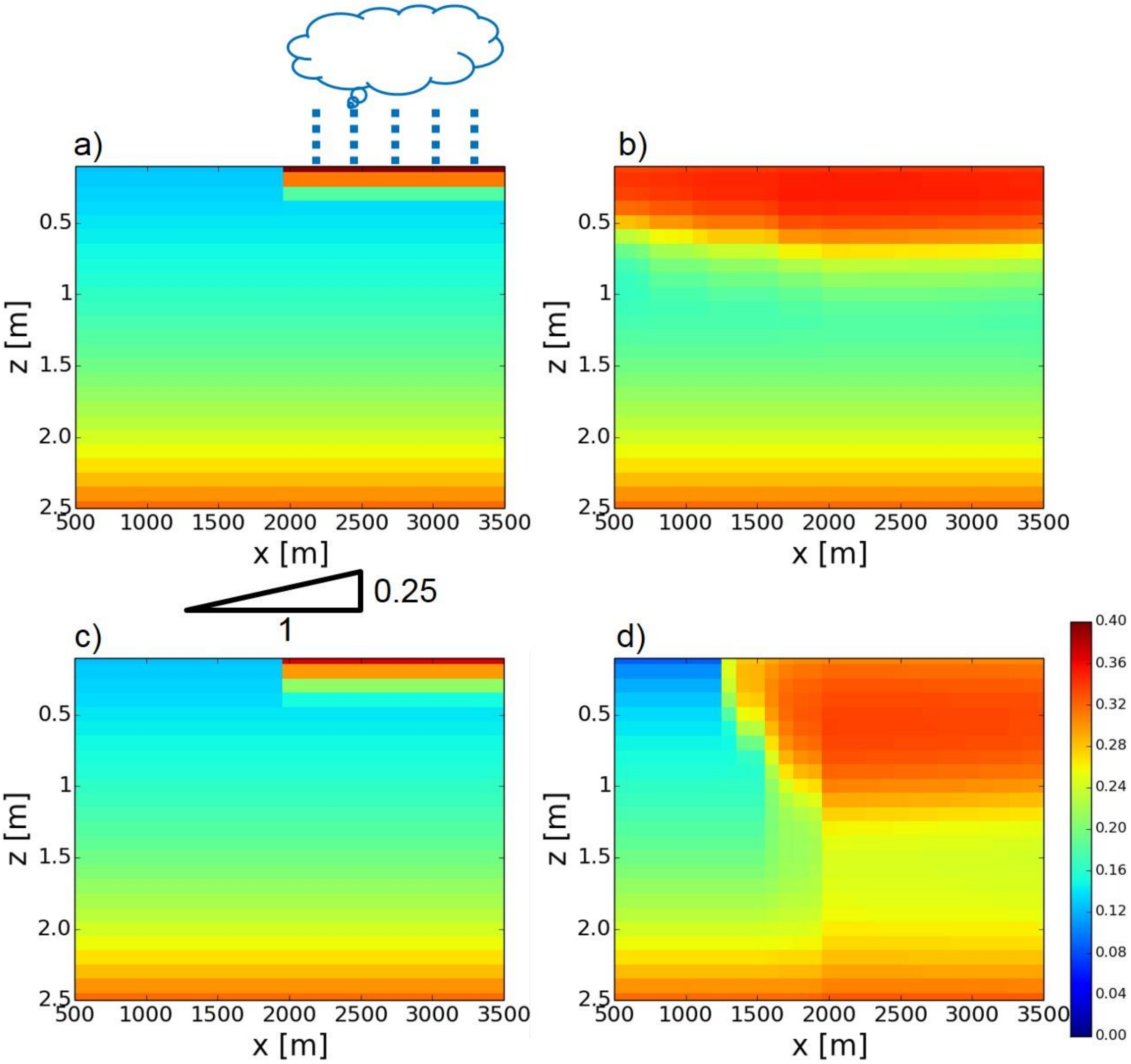

Figure 1. Distributions of volumetric soil moisture simulated by the synthetic reference runs. (a) The distribution of volumetric soil moisture [m^3/m^3] simulated by the LOW_K synthetic reference run at t = 0h. The schematic of the configuration of the synthetic reference runs is also shown (see also section 3). (b) same as (a) but at t = 130h. (c,d) same as (a,b) but for the HIGH_K synthetic reference run.

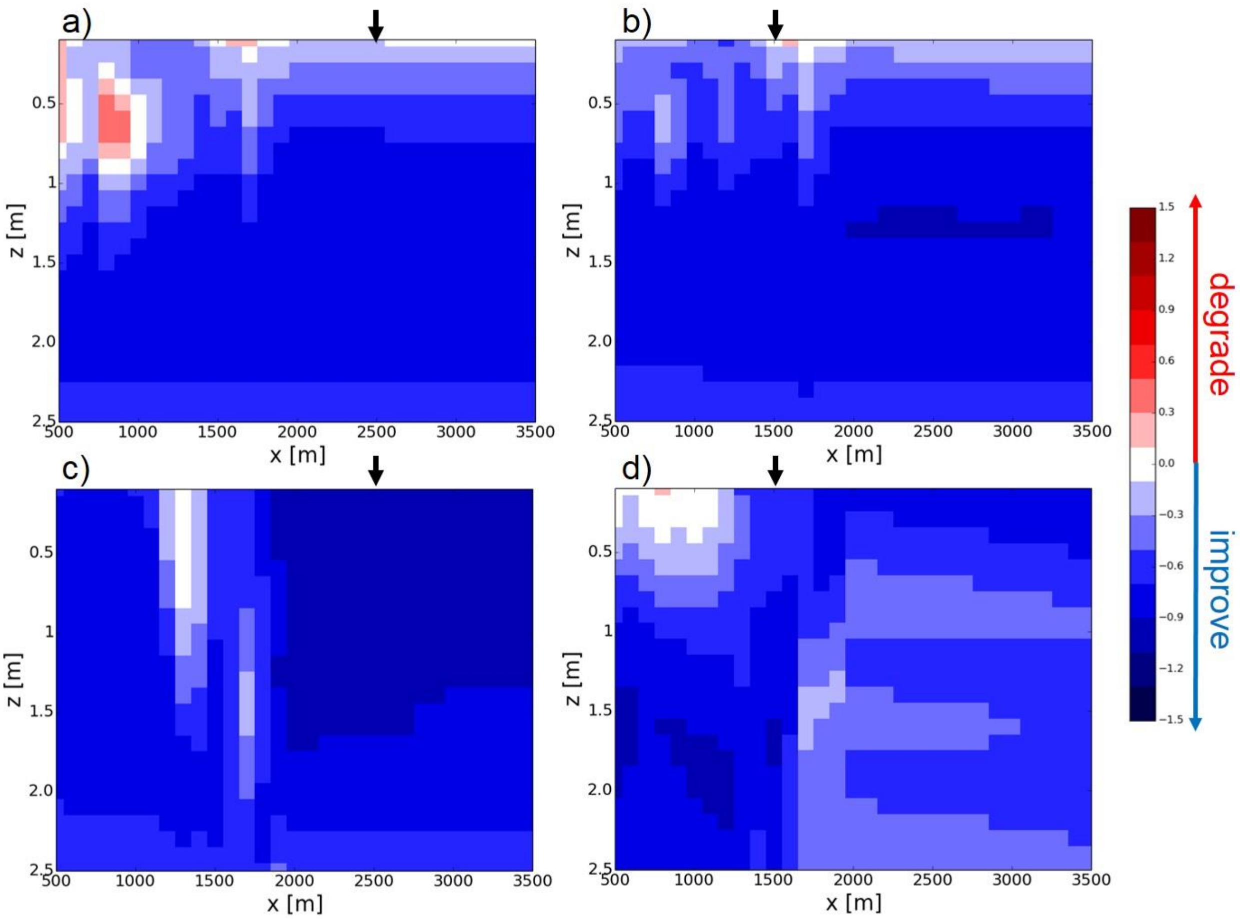

Figure 2. The improvement rates of the (a) LOW_K-UP_O, (b) LOW_K-DOWN_O, (c) HIGH_K_UP_O, (d) HIGH_K-DOWN_O experiments (see Table 1 and section 3). Black arrows show the locations of the soil moisture observations in each experiment.

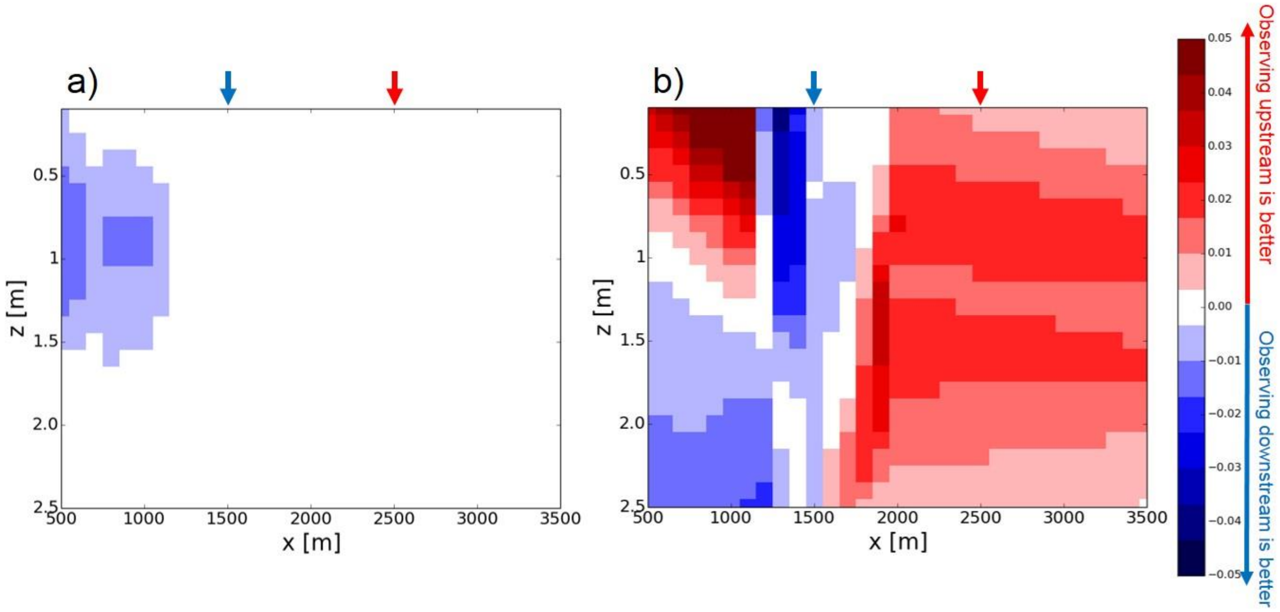

Figure 3. (a) The difference of time-mean RMSEs between the LOW_K-UP_O and LOW_K-DOWN_O experiments (see Table 1 and section 3). Red (blue) color indicates that the observations in the upper (lower) part of the slope reduce time-mean RMSE by data assimilation better than those in the lower (upper) part of the slope (see also arrows which are the locations of the observations). (b) same as (a) but for the difference between the HIGH_K-UP_O and HIGH_K-DOWN_O experiments.

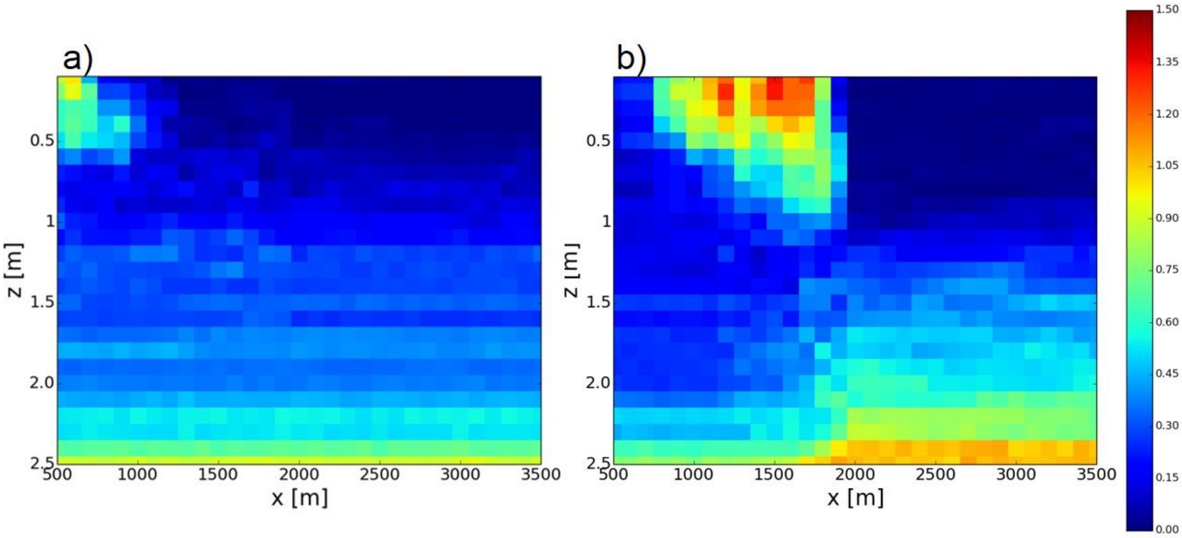

Figure 4. The Kullback-Leibler divergence of the NoDA experiment generated by (a) the LOW_K reference and (b) the HIGH_K reference at t = 130h (see also Figure 1b and 1d).

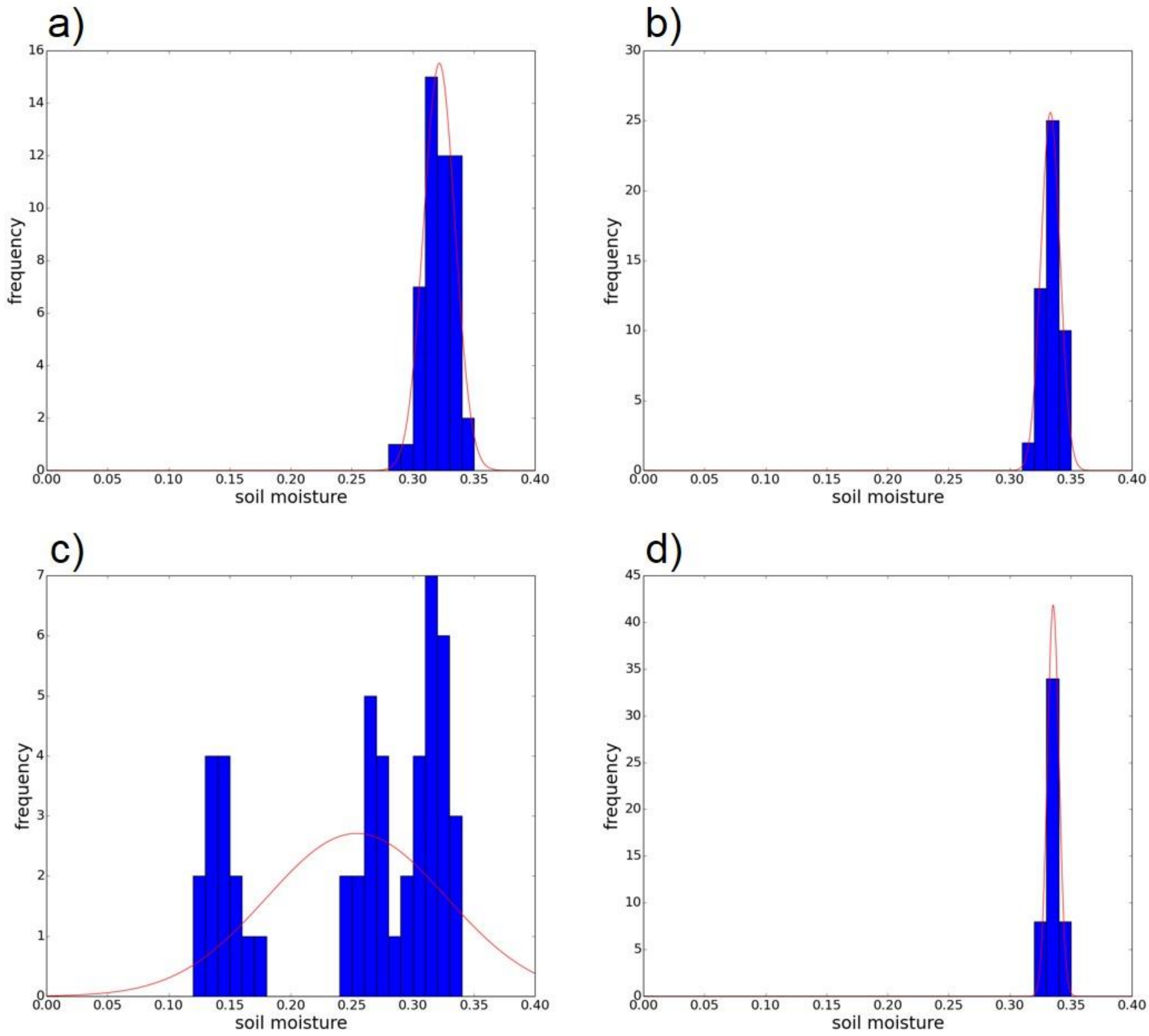

Figure 5. (a) The histogram (blue bars) of the volumetric soil moisture simulated by the NoDA experiment (see section 3) with the LOW_K reference at x=1500m, z=0.5m, and t=130h (see also Figure 4). Red line shows the Gaussian distribution with the mean and variance sampled by the ensemble. (b) same as (a) but at x=2500m, z=0.5m, and t=130h. (c) same as (a) but for the HIGH_K reference. (d) same as (c) but at x=2500m, z=0.5m, and t=130h.

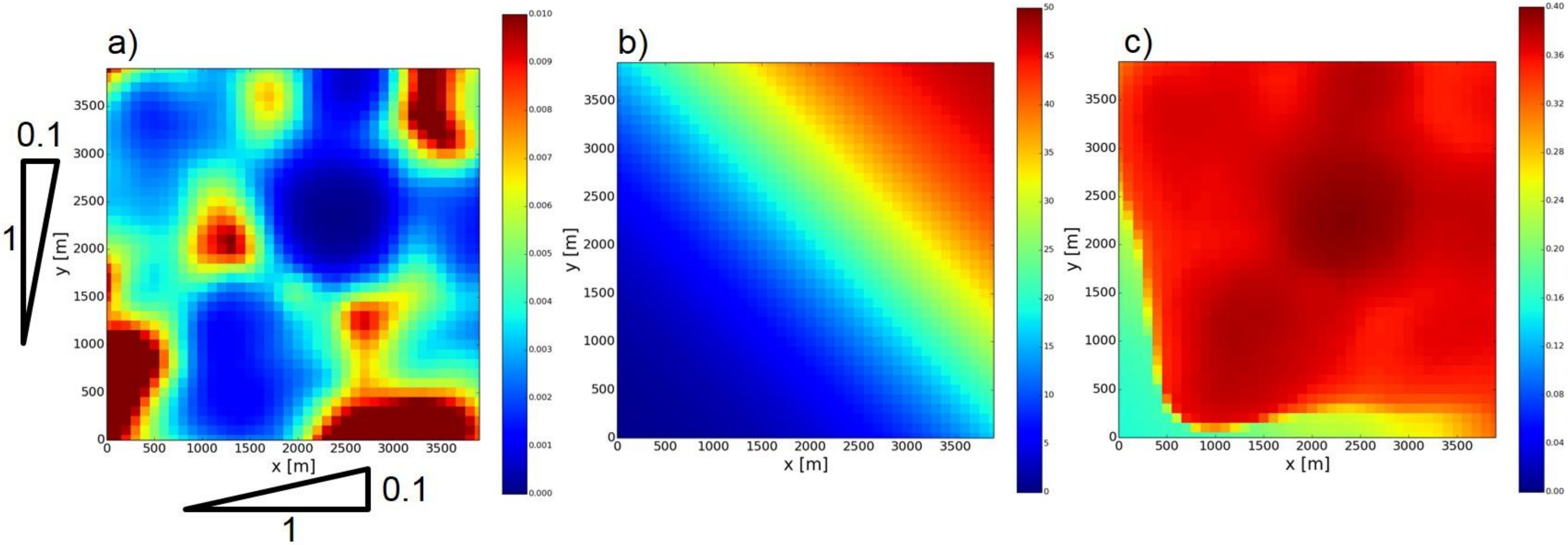

Figure 6. (a) Distribution of surface saturated hydraulic conductivity [m/h] in the synthetic reference. (b) Distribution of rainfall rate [mm/h] in the synthetic reference. (c) Surface volumetric soil moisture [m^3/m^3] at t = 5 [h] in the synthetic reference.

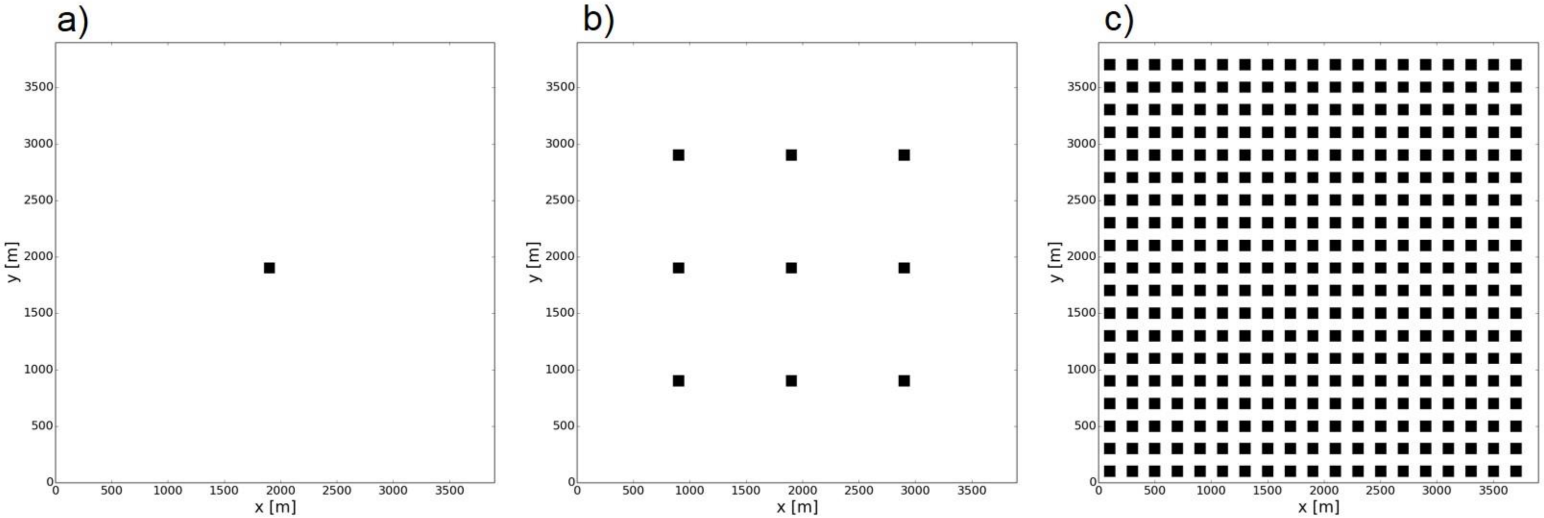

Figure 7. Observing networks. Black boxes are observed grids. (a) obs1, (b) obs9, (c) obs361 See also section 3.2.1.

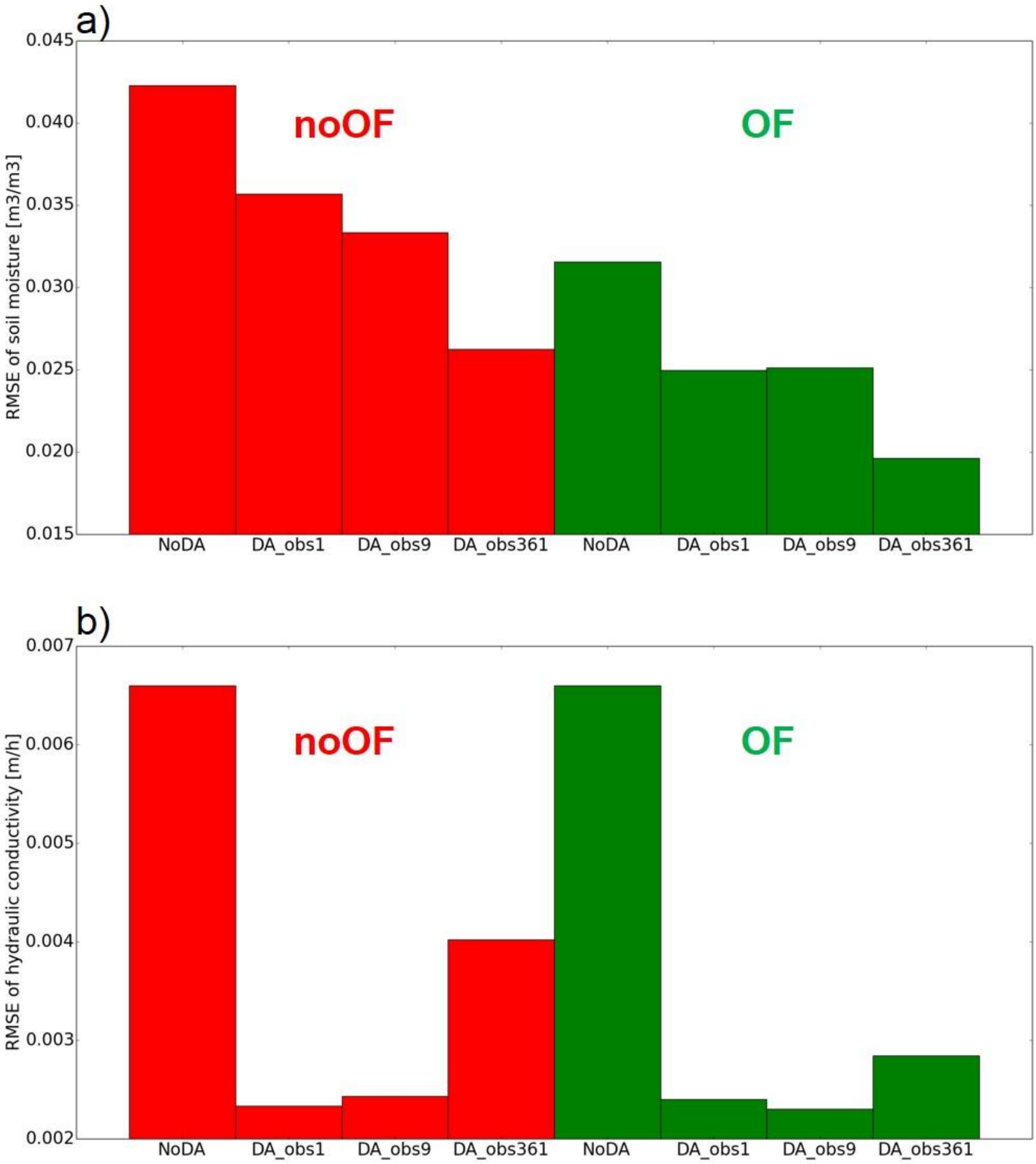

Figure 8. Time-mean RMSEs of the estimation of (a) soil moisture and (b) hydraulic conductivity. Red and green bars are results of the noOF and OF configuration, respectively (see section 3.2.1 and Table 2).

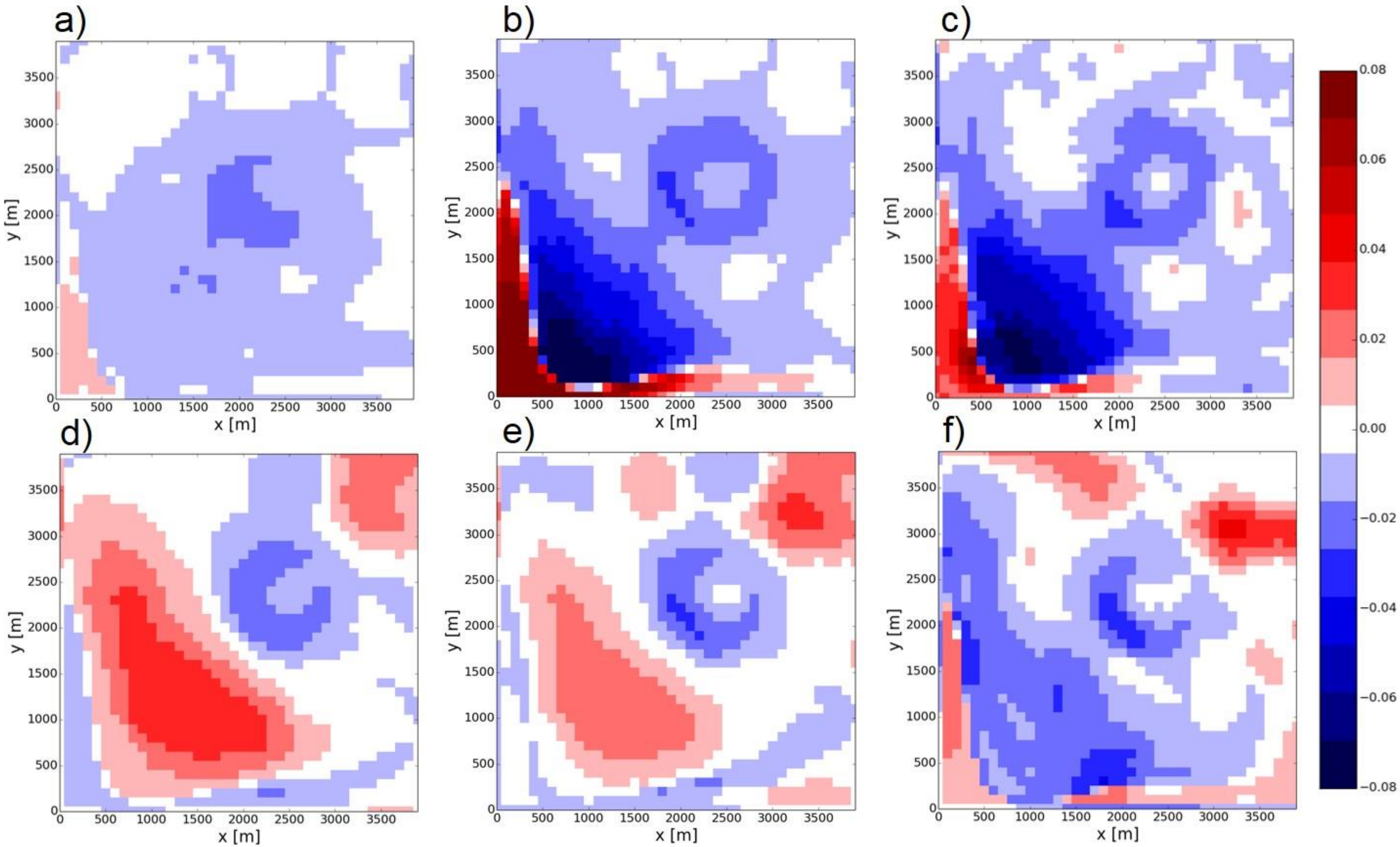

Figure 9. Differences of time-mean soil moisture RMSEs between the DA experiments and the OF_NoDA experiment. (a) OF_DA_obs1, (b) OF_DA_obs9 (c) OF_DA_obs361 (d) noOF_DA_obs1, (e) noOF_DA_obs9, (f) noOF_DA_obs361.

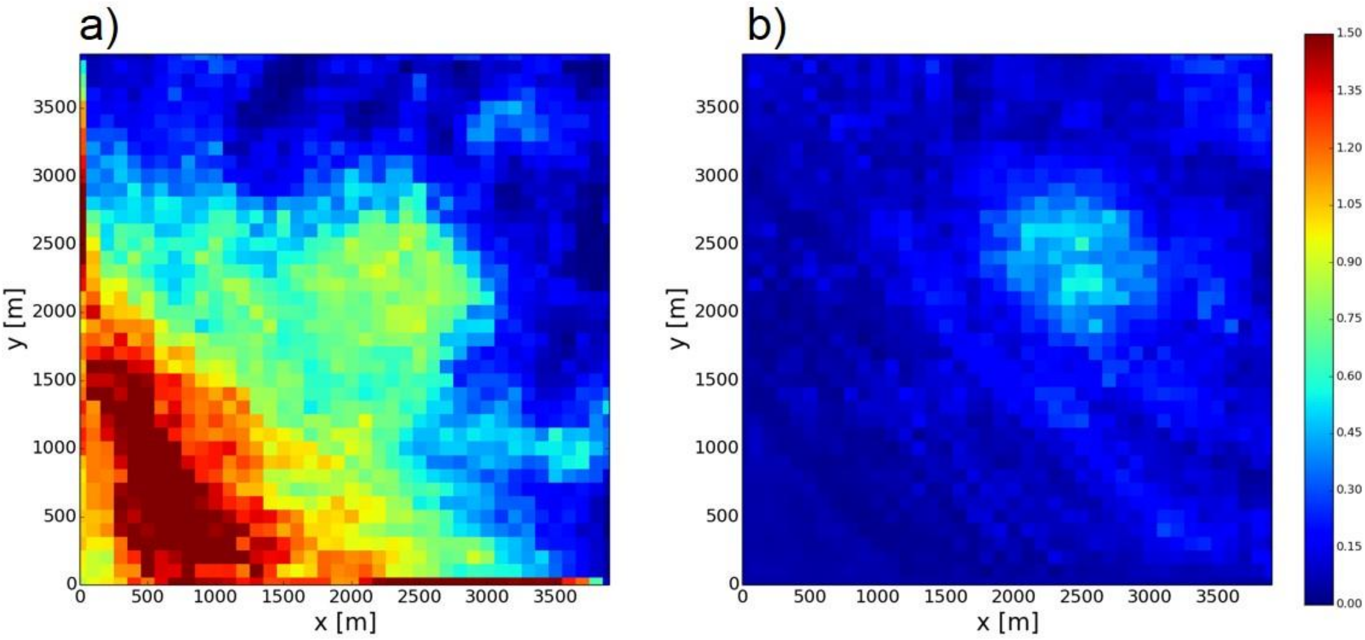

Figure 10. The Kullback-Leibler divergence of ensemble members generated by the (a) OF_NoDA and (b) noOF_NoDA experiments at t = 4 [h].