Response to the referee reveiew by Prof.Dr. Cirpka

February 2018

We thank Prof.Dr. Cirpka very much for his efforts devoted in reviewing our manuscript. His deep thoughts and constructive comments really benefit us a lot. We have revised the manuscript carefully according to the two reviews. We replied (in black) every single question (in blue) raised by the reviewers and itemized them in below.

1 Major Remarks

1. The English of the paper needs severe polishing. I have attached a corrected version, suggesting chances in almost every sentence. I highly recommend that the authors address a native speaker who surely must be available in an institution the size of UFZ before submitting a revision.

   Response: We thank you for your efforts in correcting our language mistakes. We also checked the English carefully with native English speakers. Please feel free to check the revised manuscript.

2. Unfortunately, I must say that the coupling scheme is not described very clearly. In other schemes, like the USGS-based modeling framework GSFLOW, a conceptual hydrological model is used to represent surface and soil processes, while the entire groundwater component is handled by the Darcy-based groundwater model. This is not what the authors do. The conceptual hydrological model mHm has its own groundwater storage (denote x_6) that remains untouched by the coupling. While the authors describe to some extent how mHm-based groundwater recharge is transferred to nodal loads in the groundwater model, it is not clear how groundwater leaves the OpenGeoSys domain. The authors somehow mention that OpenGeoSys and mHm have different representations of streams so that I somehow speculate that the OpenGeoSys model assumes fixed hydraulic heads along rivers, but it is not explained how river stages are calculated within or for OpenGeoSys. And it is also not clear whether any comparisons between the baseflow computed by mHm and the groundwater flux leaving the OpenGeoSys domain via river nodes are performed.

Response: These comments are very important for us to improve the description of coupling scheme. As pointed out by the reviewer, the key difference between our coupling methods and the others, is that we use OpenGeoSys as a post-processor to realize groundwater head, meanwhile we use mHM to determine component-wise water budget. The reviewer raised two important questions: (1) How does OGS calculate the river stages? (2) How to make sure the river discharge from groundwater calculated by OGS is the same with that from mHM?

For the first question, we simply do not calculate the river stage explicitly. Most coupled surface-subsurface hydrological models characterize river-groundwater interaction based on either first-order flow exchange or boundary condition switching [2]. These methods inevitably rely on a parameter set describing geometric, topographic, and hydraulic properties of the stream channel (e.g., river bed conductance, river bed and drain elevations, channel width). Unfortunately, these parameters are essentially unknown at large scale due to the lack of data and the subgrid-scale topography. Due to these limitations, we use an alternative approach which is based on the step-wise routed baseflow estimated by mHM.

Streams in mHM are implicitly defined based on pre-processing of DEM data and a routing scheme, while OGS uses an explicit predefined river geometry. In OGS, each reach of the stream network is defined by a polyline in a geometry file. To coordinate the two different conceptualizations, we developed a model interface RIV2FEM to convert the routed baseflow estimated by mHM to Neumann boundary conditions assigned at stream nodes at the OGS mesh. The illustration of RIV2FEM scheme can be found in Figure 2 of the revised manuscript. Specifically, the step-wise routed baseflow estimated by mHM is transferred to the uniformly disaggregated discharges by distributing it uniformly along the predefined stream network in OGS:

$$\overline{q}_4(t) = \frac{Q_4(t)}{N} \tag{1}$$

where $\overline{q}_4(t)$ denotes the disaggregated discharge assigned at every stream node in OGS at time t [L^3T^{-1}], $Q_4(t)$ denotes the routed baseflow at the outlet of catchment at time t [L^3T^{-1}], N denotes the total number of stream nodes in OGS. The uniformly disaggregated discharges are then assigned to every stream node in OGS to serve as Neumann boundary condition. This approach significantly reduces the number of parameters, avoids the uncertainty caused by the unknown river properties, and is suitable for many real-world applications with scarce data. Moreover, as recharge and baseflow are directly taken from mHM, the mass conservation criteria is naturally satisfied in this approach.

The above statements are also included in the revised manuscript (please refer to P8, L28 - P9, L8).

3. I see the advantages that the authors want to gain by their scheme, but

I do not understand why they don't explicitly write what they are really doing and why. My understanding is that OpenGeoSys is mainly used as a postprocessor of mHm to obtain groundwater tables, whereas the water budget is calculated by mHm alone. The advantage is pretty clear as the entire stochastic calibration procedure of mHm remains intact and the quality of discharge calculations is not hampered by using a computationally expensive groundwater model for baseflow that requires spatially distributed parameters, which are difficult to obtain. Of course, the Darcy-based groundwater model needs to be calibrated to match heads, but that does not affect the water balance. I assess that using the groundwater model as postprocessor has its intellectual elegance, but then I suggest selling the scheme this way. Otherwise the authors run into the problem why they double-account for groundwater in their two model components.

Response: Thank you for your insightful comments. We totally agree with the reviewer on the description of "OpenGeoSys is mainly used as a postprocessor of mHm to obtain groundwater tables, whereas the water budget is calculated by mHm alone". This is a good summary of our coupling approach. In order to better convey what we are doing and why, we modified the paragraph describing the modeling motivation as follows:

"The coupling initiative aims to add additional predictive capability of groundwater heads, which is achieved by OGS, to the existing predictive capability of discharge achieved by mHM. On one hand, mHM is used to estimate step-wise and component-wise water budget through model calibration against discharge. On the other hand, OGS serves as a post-processor of mHM to obtain groundwater heads by using driving forces obtained from mHM. Two model interfaces, namely GIS2FEM and RIV2FEM, have been developed to link the two models by transferring recharge and baseflow from mHM to Neumann boundary conditions in OGS."

The above statements are also included in the revised manuscript (please refer to P8, L3-10).

4. It is also clear that the presented coupled framework cannot overcome difficulties of mHm, or any other similar model, in representing feedbacks of shallow groundwater tables on evapotranspiration, if the shallowness of the groundwater table is caused by lateral groundwater flows. This would require two-way coupling, replacing the groundwater storage of mHm by the Darcy simulator, and joint calibration. Whether the authors will ever implement the two-way coupling, as announced in the outlook of the paper, is doubtful as the modeling and calibration philosophies of the two model components are not particularly compatible.

   Response: We agree with the reviewer that our current coupled model can not overcome the difficulties of typical bucket-type hydrological models. This limitation has been stated in the manuscript (Page 23, Line 25-28): "The main limitation of one-way coupling is that the effects of a shallow

depth to groundwater on actual ET, maintained by lateral groundwater flow, cannot be explicitly addressed."

In terms of the possibility of a two-way coupling, we have observed a successful example, in which a bucket-type hydrological model PCR-GLOBWB is fully coupled to the groundwater simulator MODFLOW [1]. I agree with the reviewer that the two-way coupling is super challenging because of the distinct model concepts. mHM is a "top-down" model which has a strong predictive capability but poor interpretability while OGS is a "bottom-up" model. We still believe that it is feasible to realize the two-way coupling in the future, although it requires restructuring mHM source code. We stated this point in a conservative way in the manuscript:" However, in a future work, we will devote to incorporate a full, two-way coupling using the next version of mHM#OGS model." (P22, L28-29)

5. I am not quite happing with the scheme used to map mHm-based groundwater recharge to OpenGeoSys. In the example application, it seems to be no problem, as apparently the groundwater model simply uses a rectangular grid with finer resolution. (At least the authors don't state otherwise.) In the more general case of mapping gridded fluxes to an unstructured FEM grid, checking for element centroids could be dangerous, and would definitely not ne consistent with the FEM formulation. The consistent formulation would use the weighting functions of the FEM scheme:

$$Q_i^{in} = -\int_{\partial\Omega} W_i(\mathbf{x}) q_{mHm}^{ex}(\mathbf{x}) d(\mathbf{x})$$

in which Q_i^{in} is the nodal load of node i, $\partial\Omega$ is the surface boundary of the FEM domain, W_i is the weighting function of node i, $\mathbf{x}$ is the spatial coordinate on the surface, and q_{mHm}^{ex} is the exchange rate provided by mHm.

Response: Thank you for your insightful comments. However, the mapping method in GIS2FEM is exactly the same one that the reviewer raised. In the manuscript, we stated that GIS2FEM checks the centroid of element, which is not true and will mislead the readers. We apologize for this unclear description. Actually, we checked the nodes in each elements. GIS2FEM searches for the mHM grid cell that the node is located in, and assigns the recharge value of this grid cell to the corresponding node (marked as C^m). After all top surface elements have been processed, GIS2FEM takes subsequently the face integration calculation, by which the specific recharge C^m [LT^{-1}] calculated by mHM is converted into volumetric recharge C^{in} [L^3T^{-1}] and assigned to the corresponding OGS mesh nodes. Specifically, the specific recharge C in a certain element is calculated as:

$$C(\mathbf{x}) = \sum_{j=1}^{N} W_j(\mathbf{x}) C_j^m, \tag{2}$$

where $\mathbf{x}$ is the spatial coordinate on the surface, N is the total number of nodes in a surface element, W_j is the weighting function of node j, C_j^m is the specific recharge at node j calculated by mHM [LT^{-1}]. Then the volumetric recharge C_i^{in} at node i (i is the global node index) is calculated by the face integration calculation:

$$C_i^{in} = -\int_{\partial\Omega} W_i(\mathbf{x}) C(\mathbf{x}) d(\mathbf{x}), \tag{3}$$

where C_i^{in} is the volumetric recharge of node i [L^3T^{-1}], $\partial\Omega$ is the surface boundary of the FEM domain, W_i is the weighting function of node i. This method is clearly shown in line 10434-10510 in file rf_pcs.cpp at `https://github.com/UFZ-MJ/OGS_mHM/blob/master/FEM/rf_pcs.cpp`

The above statements are also included in the revised manuscript (please refer to P9, L13-27).

6. How does OpenGeoSys calculate river stages? This is completely missing, and knowing discharge along the streams is not sufficient for the groundwater model.

   Response: Thank you for your question. We did not calculate river stages. Please find the detailed description of boundary conditions in rivers in the response to question 2.

7. The example application takes place in a karstified aquifer. Muschelkalk is a limestone formation with all features of a karst system. This must not remain unmentioned, as the calibrated hydraulic-conductivity values only hold under the assumption of an equivalent porous medium. Of course, the karst features also hamper the calculation of travel times.

   Response: Thank you again for this important information. A research by Kohlhepp et al.(2017) found that the karstification only occurs at the base of the upper Muschelkalk formation [3]: "Although it is of a karst-fracture type with partially solution enlarged fractures, the karstification and the development of conduits are limited and concentrated at the formations' very base"[3]. He concluded that karstification is not a dominant factor to the permeability as it is limited at the base of upper Muschelkalk [3].

   Nevetheless, we agree that the treatment of karst formation should be mentioned in the manuscript. Accordingly, we have modified the manuscript as follows:"The mo formation has been widely considered as a karstified formation. In this study area, a recent research by [3] has revealed that in the Hainich critical zone, the intense karstification and the conduit are limited at the base of the mo formation. Accordingly, we use the equivalent porous medium approach to characterize the insignificant karst formation. "(P11, L31 - P12, L3).

8. I am very astonished that the Lower Keuper (" Lettenkeuper ") has such high conductivity values given the fact that this formation contains thick clay layers and acts as an aquitard.

Response: Thank you for the question. We use the automatic-calibration code PEST to calibrate the steady state groundwater model by setting a reasonable range of adjustable parameters (in this case, hydraulic conductivity). We set the upper and lower bound of parameters on the basis of the literature [4]. We found that there is a certain degree of parameter uncertainty, which means a range of possible values are compatible with the prediction. Since the Lower Keuper contains confining clay layers that often acts as aquitard, we re-calibrated the model by setting a narrow range of ku. The updated hydraulic conductivity of ku is 2.848e-5. We use the updated parameter values in the revised manuscript. Please check the updated Table 1 (P17) in the revised manuscript.

9. Assuming no-flow boundary-conditions all around the model domain is not reasonable. There must be a lateral groundwater flux leaving the domain in the lowland-part of the aquifer in the Muschelkalk aquifer.

Response: Thank you for your insights. The outer boundary is generally treated as no-flow bouindary except for the northwestern and northeastern edges. We set Dirichlet boundary conditions at the northwestern and northeastern edges on the basis of Sommer et al. [4]. Accordingly, we modified the description of boundary conditions in the manuscript as follows:

"In general, no-flow boundaries are set at the outer perimeters surrounding the basin as well as at the lower aquitard. On the basis of the measurements, a Dirichlet boundary condition is assumed at the northwestern and northeastern edges." (P12, L6-8)

I agree with you that there must be a lateral groundwater flux leaving the domain in the lowland area in the Muschelkalk aquifer. Several studies have revealed that this portion of water is small compared to the portion of baseflow in the study area.

Based on Toth [6], groundwater flow system in a catchment can be divided into local system, intermediate system and regional system. Our study revealed that the dominating flow systems in the study area are local and intermediate systems. The regional system in which the recharge area occupies the water divide and discharge area lies at the bottom of the basin, is negligible. The reason is that the ratio of the depth to impermeable lower boundary to the half-width of the catchment is small (about 0.008) and the general relief of the catchment is small (about 0.01), which means regional flow is almost impossible [6]. In the local and intermediate flow systems in lowland area, the outflow from Muschelkalk formations is negligible compared to the amount of baseflow. The reason is that first, Wechsung [4] and Seidel [5] have demonstrated that the Muschelkalk formations in lowland have significantly lower permeabilities than those in upland. This observation has also been included in our manuscript in Page 11, Line 28-32. The Muschelkalk in lowland acts as aquitard which partly blocks the movement of groundwater from upland to the outlet of

the catchment. Second, the head gradients are directed from southwestern (or southeastern) uplands to the central lowlands. The flow direction is parallel to the direction of bottom of the catchment, which indicates only a small portion of water can flow out at the bottom. Based on above proofs, we can reasonably assume that the outflow in Muschelkalk formations in the vicinity of catchment outlet is less important, and therefore negligible.

2 Minor Remarks

1. See my very extended list of remarks (621 comments, not all of them editorial in nature) in the attachment.

   Response: Thank you again for your detailed review of our manuscript and the insightful comments you raised. We revised the manuscript thoroughly according to each items in the extended list of remarks.

2. I would not use an incomprehensible acronym in the title. My suggestion is: Improved groundwater representation at regional scale by coupling of the mesosclale Hydrological model (mHm) to the groundwater model OpenGeoSys (OGS)

   Response: Thank you very much for this kind suggestion. We have changed the title to "Improved regional scale groundwater representation by the coupling of the mesoscale Hydrologic Model (mHM v5.7) to the groundwater model OpenGeoSys (OGS)".

3. I would not report on version numbers in the abstract.

   Response: We have been informed that the journal GMD requires the specific version number:

   "If the model development relates to a single model then the model name and the version number must be included in the title of the paper. If the main intention of an article is to make a general (i.e. model independent) statement about the usefulness of a new development, but the usefulness is shown with the help of one specific model, the model name and version number must be stated in the title. The title could have a form such as, "Title outlining amazing generic advance: a case study with Model XXX (version Y)"."

4. I know that OpenGeoSys can be used to compute saturate-unsaturated flow, solute transport, heat transport, and geomechanics. However, in the given context, the THMC capabilities are reduced to groundwater hydraulics alone.

   Response: We agree that the THMC capability of OGS does not show up in the manuscript. In the field of hydrogeology, there are nevertheless some important multi-physical processes such as density-dependent flow and thermo-haline flow. OGS is suitable in addressing these problems, and can be easily plugged in current coupling framework in the manuscript.

Nevertheless, we followed the reviewer's suggestion and modified the statement of THMC coupling as follows: "Additionally, OGS demonstrates its capability in addressing thermo-hydro-mechanical-chemical (THMC) coupling processes in large-scale hydrologic cycles (not reflected in this study)" (P22, L4-5).

5. Well identification numbers made of ten digits are totally incomprehensible.

   Response: Thank you. We have removed all those long-digit numbers. Instead, we use numbers with only two digits. Please check Figure 2, 10, and 11 in the revised manuscript.

References

[1] Sutanudjaja, E. H., Van Beek, L. P. H., De Jong, S. M., Van Geer, F. C., & Bierkens, M. F. P. (2014). Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data. Water Resources Research, 50(1), 687–705. http://doi.org/10.1002/2013WR013807

[2] Paniconi, Claudio, and Mario Putti."Physically based modeling in catchment hydrology at 50: Survey and outlook." Water Resources Research 51.9 (2015): 7090-7129.

[3] Kohlhepp, B., Lehmann, R., Seeber, P., Küsel, K., Trumbore, S. E., and Totsche, K. U.: Aquifer configuration and geostructural links control the groundwater quality in thin-bedded carbonate–siliciclastic alternations of the Hainich CZE, central Germany, Hydrol. Earth Syst. Sci., 21, 6091-6116, https://doi.org/10.5194/hess-21-6091-2017, 2017.

[4] Sommer T, Feige H, Klöcking B, Knoblauch S, Maier U, Müller M, Pfützner B, Wechsung F, Clausing T. Die Wirkung des Globalen Wandels im Unstrut-Einzugsgebiet.

[5] Seidel, G., 2004. Geologie von Thüringen. Erdkunde, 58(1).

[6] Toth J. A theoretical analysis of groundwater flow in small drainage basins. Journal of geophysical research. 1963 Aug 15;68(16):4795-812.

Response to comments raised by Edwin Sutanudjaja

We thank Dr. Sutanudjaja very much for his excellent reviews. Below you could find the original comments (marked in blue color) and point-by-point response (marked in black color).

I appreciate that the authors have carefully addressed all the concerns I raised in my review. The writing has been improved. Yet, I still have some minor comments and questions that should be considered:

Page 2 Line 17 (P2L17): … Finally …

P2L24-30: This part, starting from the sentence "For example, models ...", does not flow with the previous sentences. Please consider to rephrase it, Here you may want to consider to discuss several limitations in groundwater model packages (e.g. river, evapotranspiration packages in MODFLOW) for simulating surface water and unsaturated zones.

Response: We have modified the manuscript accordingly. Please check P2, L23-29 in the revised manuscript.

P3L7: What's wrong with "kinematic wave approximation" for surface flow process? Could you please elaborate? And what did you mean by "1D riil flow"?

Response: The kinematic-wave approximation neglects the dynamic components of flow that are represented by the derivative terms in the more complete form of the momentum equation referred to as the Saint-Venant equations. 1D rill flow means the surface flow problem is simplified from a full 3-D problem to 1-D problem via some assumptions.

P3L27-28: "The process uncertainty decreases as one goes depper into the subsurface storage. In the subsurface storage, hydrological process are relatively well understood …" I am not sure that I can agree with this statement. Please include some references to support this statement.

Response: The process uncertainty in subsurface hydrology has been well demonstrated by the theory of porous media flow [1].

P4L2: "OGS has been demonstrated its ability of dealing with data uncertainty in groundwater aquifers." Please give references.

Reference: We have added a reference accordingly.

Figure 1: What is K that leaves donward from the groundwater part? How do you include it in your coupling strategy?

Reference: the K in Figure 1 represents the original scheme characterizing Karst process in mHM. This feature of mHM is irrelevant to our study. In the case study, we do not consider the Karst process in mHM, thus K is always zero.

P6L22, Equation 22: For clarity, could you please expand the source/sink term qs? I believe that this should consist at least, the recharge C and the baseflow q4.

Section 2.3: While explaining the coupling mechanism, I suggest to include the variable symbols as given in Sections 2.1 and 2.2. An example is given as follows:

Response: thank you for your comments. We have modified the manuscript accordingly.

P7L4: The basic idea is to feed fluxes generated by mHM, e.g. distributed groundwater recharge C and baseflow q4 ...

P8L14-21, step 3: Please clarify that your OGS mesh actually has a regular shape: square 250 m x 250 m (not an irregular ones as suggested in Fig. 2c).

Response: thank you for your comments. We have modified the manuscript accordingly.

P9L1: Did you consider groundwater abstraction in your modelling experiment? Please clarify.

Response: in this case study, we do not consider the groundwater abstraction. Please refer to Section 3.4.

P9L7: So, your OGS mesh is rectangular grid: 250 m x 250 m? (not an irregular ones as suggested in Fig. 2c).

Response: yes, we use a structural 3-D mesh in this study. The resolution of mHM grid cells is 500 m $\times$ 500 m. OGS uses a structured, hexahedral 3-D mesh, with a spatial resolution of 250 m $\times$ 250 m in horizontal direction and 10 m in vertical direction over the whole domain. Please refer to P11, L3-6.

P10, Section 3.2: Please include the number of aquifer layers in the OGS model. 11?

Response: thank you for your comments. We have modified the manuscript accordingly.

P11, Section 3.3: Please clarify to which OGS aquifer layer you force the recharge and baseflow. To the uppermost one only?

Response: yes, we force the recharge and baseflow to the uppermost layer.

P11L21: 1500 mm/month is a very high threshold. Please check.

Response: thank you for pointing out this error. We have modified the statement to: " Only streams with a runoff rate higher than the threshold (in this case study, 0.145 m3/s) are delineated as valid streams." Please refer to P13, L1-3 in the revised manuscript.

P12L23: What are K values? Hydraulic conductivities? Or K values in Fig. 1? Please clarify.

P13L26: What is K? The calibrated K values?

Response: thank you for this suggestion. For sake of consistency, we have removed all the expression of "K values". Instead, we use the alternative expression of "hydraulic conductivities" throughout the manuscript.

P14L8, Figure 8: Please use the same maximum and minimum values on the color scales for all sub-figures in Fig. 8 (page 15) so that they can easily be compared among each other.

Response: we totally agree with this comment. We have re-generated Figure 9 and Figure 10 using the same color scheme with Figure 11. Please refer to the revised manuscript.

Reference:

[1] Darcy, Henry. *Les fontaines publiques de la ville de Dijon: exposition et application...* Victor Dalmont, 1856.

Improved ~~groundwater representation at~~ regional scale ~~(mHM#OGS v1.0) –~~ groundwater representation by the coupling of the mesoscale Hydrologic Model (mHM v5.7) ~~with~~ to the groundwater model OpenGeoSys (OGS)

Miao Jing[1], Falk Heße[1], Rohini Kumar[1], Wenqing Wang[2], Thomas Fischer[2], Marc Walther[2,3], Matthias Zink[1], Alraune Zech[1], Luis Samaniego[1], Olaf Kolditz[2,4], and Sabine Attinger[1,5]

[1]Department of Computational Hydrosystems, UFZ – Helmholtz Centre for Environmental Research, Permoserstr. 15, 04318 Leipzig, Germany
[2]Department of Environmental Informatics, UFZ – Helmholtz Centre for Environmental Research, Permoserstr. 15, 04318 Leipzig, Germany
[3]Institute of Groundwater Management, Technische Universität Dresden, Bergstr. 66, 01069 Dresden, Germany
[4]Applied Environmental Systems Analysis, Technische Universität Dresden, Dresden, Germany
[5]Institute of Earth and Environmental Sciences, University of Potsdam, Karl-Liebknecht-Str. 24–25, 14476 Potsdam, Germany

Abstract.

Most ~~of the current large scale~~ large-scale hydrologic models fall short in reproducing groundwater head dynamics due to their over-simplified representation of groundwater flow. In this study, we aim to extend the applicability of the ~~Mesoscale~~ mesoscale Hydrologic Model (mHM v5.7) to subsurface hydrology ~~through the coupling with a thermo-hydro-mechanical-chemical (THMC)~~ by coupling it with the porous media simulator OpenGeoSys (OGS). The two models are one-way coupled through ~~a model interface~~ model interfaces GIS2FEM and RIV2FEM, by which grid-based vertical fluxes generated by ~~vertical layered reservoirs within mHMrepresenting near-surface hydrological processes~~mHM, are converted into ~~upper surface~~ upper-surface boundary conditions of the groundwater model ~~in~~ OGS. Specifically, the grid-based vertical reservoirs in mHM are completely preserved for ~~land surface fluxes estimation~~ the estimation of land-surface fluxes, while OGS acts as a plug-in to the original mHM modeling framework for groundwater flow and transport modeling. The applicability of the coupled model (mHM#OGS v1.0) is evaluated by a case study in ~~a~~ the central European meso-scale river basin ~~,-~~ Nägelstedt. Different time steps, i.e. ~~,~~daily in mHM and monthly in OGS, are used to account for fast surface flow and slow groundwater flow. Model calibration is conducted following a two-step procedure using discharge and long-term mean of groundwater head measurements, respectively. Based on the model summary statistics, ~~such as~~ namely the Nash–Sutcliffe model efficiency (NSE), ~~Pearson correlation~~ the Pearson correlation coefficient R_{cor}, and ~~inter-quantile~~ the inter-quartile range error QRE, the coupled model is able to satisfactorily represent the dynamics of discharge and groundwater heads at several locations across the study basin. Our exemplary calculations show that the coupled model ~~mHM#OGS v1.0~~ can take advantage of the spatially explicit modeling capabilities of surface and groundwater hydrologic models, and provide us with adequate representation of the spatio-temporal behaviors of

groundwater ~~storages~~ storage and heads, ~~and~~ thus making it ~~the~~ a valuable tool for addressing water resources and management problems.

1 Introduction

Historically, ~~large scale~~ large-scale hydrologic models are ~~commonly~~ developed to predict river discharge. Most of these models use simple bucket-type expressions combined with several vertical water storage layers to describe near-surface water flow (Refsgaard and Storm, 1995; Wood et al., 1997; Koren et al., 2004; Samaniego et al., 2010; Niu et al., 2011). ~~Due~~ Moreover, due to the limitation in computational capability, all traditional hydrologic models simplify water flow processes by ignoring lateral groundwater flow. Thus, such models inevitably fall short of characterizing subsurface groundwater dynamics (Beven et al., 1984; Liang et al., 1994; Clark et al., 2015).

The implicit groundwater representations in traditional hydrologic models ~~show inadequacy~~ are inadequate in many aspects. ~~Water table depth~~ Depth to groundwater has a strong influence on near-surface water processes such as evapotranspiration (Chen and Hu, 2004; Yeh and Eltahir, 2005; Koirala et al., 2014). Moreover, ~~water table fluctuations~~ fluctuations in the water table are known as a factor ~~affecting~~ that affects runoff generation and thus ~~impacting the prediction skill~~ their adequate representation in hydrologic models influences the prediction ability of catchment runoff (Liang et al., 2003; Chen and Hu, 2004; Koirala et al., 2014). Typical hydrologic models ~~also show inadequacy in~~ are also demonstrably inadequate at simulating solute transport and retention at the catchment scale. For example, Van Meter et al. (2017) found that current nitrogen fluxes in rivers can be dominated by groundwater legacies. An over-simplified groundwater representation is inadequate for understanding ~~travel time~~ travel-time distributions (TTDs) at a catchment scale and is therefore incapable of describing such legacy behavior ~~(Benettin et al., 2015; Botter et al., 2010; Benettin et al., 2017)~~(Botter et al., 2010; Benettin et al., 2015, 2017). Moreover, stream-subsurface water interactions may be significant in modulating the human and environmental effects of nitrogen pollution (Azizian et al., 2017). ~~Fianlly~~Finally, to assess the response of groundwater to climate change, a more accurate groundwater representation including lateral subsurface flow is urgently needed (Scibek and Allen, 2006; Green et al., 2011; Ferguson et al., 2016).

~~Parallel to that~~In parallel, numerous groundwater models have been developed, which allow for both steady-state and transient groundwater flow in three dimensions, with complex boundaries and a complex representation of sources and sinks. A variety of numerical codes are available such as MODFLOW (Harbaugh et al., 2000), FEFLOW (Diersch, 2013), and OpenGeoSys (Kolditz et al., 2012). Groundwater models usually ~~contain a mechanistic or physically-based representation of subsurface physics~~represent subsurface flow by Darcy's law substituted into the continuity equation, but fall short in providing good representation of surface and shallow soil processes. For example, models for predicting groundwater storage ~~change~~ changes under either climate change (e.g., global warming) or human-induced scenarios (e.g., agricultural pumping) ~~always~~ often use a constant or linear expression to represent spatially distributed recharge (Danskin, 1999; Selle et al., 2013). ~~The groundwater numerical models may contain some packages or interfaces to simulate surface water or unsaturated zone processes (Harbaugh et al., 2000; Kalbacher et al., 2012; Delfs et al., 2012).~~ Moreover, parameterization of topographical and

geological parameters is a big challenge due to the strong spatial and temporal heterogeneity and lack of data (Moore and Doherty, 2006; Arnold et al., 2009).

In recent years, many integrated surface/subsurface hydrologic models (ISSHMs) have been developed. ISSHMs commonly focus on the comprehensive treatment of both ~~the~~ surface flow processes (e.g., ~~1D or 2D~~ 1-D or 2-D overland flow) and ~~the~~ subsurface flow processes (e.g., ~~1D or 3D~~ 1-D or 3-D Richards flow) using a two-way coupling procedure (Paniconi and Putti, 2015). Some ~~of the highly~~ well- recognized ISSHMs are InHM ~~(Smerdon et al., 2007; VanderKwaak and Loague, 2001)~~(VanderKwaak and Parflow (Maxwell and Miller, 2005; Maxwell et al., 2015), OpenGeoSys ~~(Kolditz et al., 2012; Delfs et al., 2012)~~(Delfs et al., 2012; Kolditz tRIBS (Ivano et al., 2004), CATHY (Camporese et al., 2010), GSFLOW ~~(Hunt et al., 2013; Markstrom et al., 2008), HydroGeoSphere (Hwang et al., 2014; Therrien et al., 2010)~~(Markstrom et al., 2008; Hunt et al., 2013), HydroGeoSphere (Therrien et al., 2010; Hwang et a MIKE SHE (Graham and Butts, 2005), MODHMS (Panday and Huyakorn, 2004; Phi et al., 2013), GEOtop (Rigon et al., 2006), IRENE (Spanoudaki et al., 2009), CAST3M (Weill et al., 2009), PIHM (Kumar et al., 2009; Qu and Duffy, 2007), and PAWS (Shen and Phanikumar, 2010). Although the methods for subsurface flow in ISSHMs are commonly based on saturated/unsaturated groundwater flow equations, the approaches for surface flow are inevitably based on some approximations and conceptualizations (e.g., ~~kinematic wave approximation, 1D rill flow, etc). Besides, the modeling skill of reproducing distributed groundwater head dynamics at regional scale is always neglected and seldom assessed by the data (e.g., groundwater head, tracer, etc).~~ diffusive-wave approximation, 1-D rill flow). The applications of these ISSHMs in the literature ~~are mainly focusing on the field and small watershed scale, while the assessment~~ mainly focus on the field- and small watershed-scale, while assessments of modeled groundwater ~~heads~~ head dynamics at larger scales can only be found in very few publications (Goderniaux et al., 2009; Sutanudjaja et al., 2011). At this larger~~scale~~, i.e., regional~~scale,~~ , scale most of the ISSHMs are based on a continuity of pressure and flux on the surface water/groundwater (SW/GW) interface, while the momentum balance condition is always ~~missing (Paniconi and Putti, 2015). Some of ISSHMs apply a storage-excess runoff generation concept, whereby the runoffs are normalized as storage-excess runoff through solving Richards equation combined with a boundary condition switching method. Then, the generated runoffs are routed into streams by a routing algorithm. These models can simulate~~ neglected (Paniconi and Putti, 2015). ISSHMs relying on shallow-water and the Richards equation assumptions often encounter problems in simulating quick flow dependent on essentially unknown sub-grid-scale topographic variability and subsurface structure (Paniconi and Putti, 2015). Nevertheless, these models are capable of simulating the dynamic interaction of different processes within SW/GW components, e.g., the interaction of soil moisture and groundwater head ~~(Cuthbert et al., 2013; Maxwell et al., 2015; Rihani et al., 2010; Sutanudjaja et al., 2014)~~ (Rihani et al., 2010; Cuthbert et al., 2013; Sutanu well as the storage-runoff correlation (VanderKwaak and Loague, 2001; Liang et al., 2003; Huntington and Niswonger, 2012; Koirala et al., 2014; Fang and Shen, 2017)~~, etc~~.

Typical hydrologic models, ~~like~~ such as mHM (Samaniego et al., 2010; Kumar et al., 2013b), VIC (Liang et al., 1994), and HBV (Lindström et al., 1997), are good at predicting quantities~~like~~ , such as discharge but, as mentioned above, are highly ~~conceptualized and there suffering from interpretability of~~ conceptual and their model results are difficult to interpret with respect to certain processes (e.g., groundwater storage and heads). ~~More~~ The output of more mechanistic ISSHMs, ~~like~~ such as Parflow, CATHY, and HydroGeoSphere, are highly interpretable but show consistently worse performance than typ-

ical hydrologic models when predicting runoff (Gulden et al., 2007; Paniconi and Putti, 2015). ~~These different abilities~~ The differing capabilities of typical hydrologic models ~~vs.~~ in contrast to the more mechanistic ISSHMs are ~~caused by~~ a result of the different challenges that are posed by the ~~different~~ various compartments of the terrestrial water cycle. One of the main challenges in ~~the~~ modeling surface and near-surface storage is process uncertainty, ~~with processes like~~ since processes such as evapotranspiration, land use, land cover, snow pack ~~,~~ are extremely complex and dynamic. The process uncertainty decreases as one goes deeper into ~~the~~ subsurface storage. In ~~the~~ subsurface storage, hydrological processes are relatively ~~well understood~~ well-understood and therefore conceptually simpler (Dagan, 2012). Meanwhile, the data uncertainty becomes more significant in ~~the~~ deep subsurface storage ~~compared to the~~ in comparison to shallow storage. Moreover, a recent study reveals the strong spatial and temporal heterogeneity of processes and properties ~~on~~ at the SW/GW interface, and underlines the importance of quantifying variability across several scales ~~on~~ at the SW/GW interface and its significance to water resources management (McLachlan et al., 2017).

In this study, we therefore coupled the ~~Mesoscale~~ mesoscale Hydrologic Model (mHM v5.7) (Samaniego et al., 2010; Kumar et al., 2013b) with ~~a thermo-hydro-mechanical-chemical (THMC)~~ the porous media simulator OpenGeoSys (OGS) (Kolditz et al., 2012, 2016) with ~~an~~ the overall aim of modelling ~~regional scale~~ regional-scale groundwater flow dynamics. mHM has ~~been demonstrated its preeminence~~ demonstrated its pre-eminence in coping with process uncertainty in the near-surface zone while providing excellent discharge prediction (Huang et al., 2017). On the other hand, OGS has ~~been demonstrated its ability~~ demonstrated its capability of dealing with data uncertainty in ~~groundwater aquifers . With~~ aquifers (Sun et al., 2011; Walther et al., 2012; Selle et al., 2013). Using these two well-tested ~~modeling codesavailable~~codes, we want to answer the following scientific questions: ~~First, can~~ (1) Can spatially distributed groundwater heads and their dynamics be reasonably captured by expanding the ~~abilities~~ capabilities of a surface hydrologic model~~like~~ , such as mHM at the ~~regional scale, all~~ regional-scale, while conserving its excellence in predicting discharge? ~~Second, can~~ (2) Can spatially resolved groundwater recharge estimates, provided by mHM, improve the prediction of head measurements of groundwater models ~~like~~ such as OGS? To answer these questions, we applied the coupled model mHM#OGS v1.0 in a central German meso-scale catchment (850 km^2), and evaluated the model skills using measurements of streamflow and groundwater heads from several wells located in the study area. The ~~, herein, illustrated~~ coupled (surface) hydrologic and groundwater model (mHM#OGS v1.0) presented in this paper is our first attempt toward the development of a large-scale coupled modeling system with the aim to analyze the spatio-temporal variability of groundwater flow dynamics at a regional scale.

The paper is structured as follows. In the next section, we describe the model concept, model structure~~as well as coupling schematic~~, and the coupling scheme. In Section 3.1, the study area and model setup used for illustration in this study are comprehensively ~~illustrated~~described. In Section 4, we present the simulation results of mHM#OGS v1.0 in a catchment in ~~central Germany~~the application. In the Section 5, we discuss the model results as well as advantages and limitations of current modeling approach.

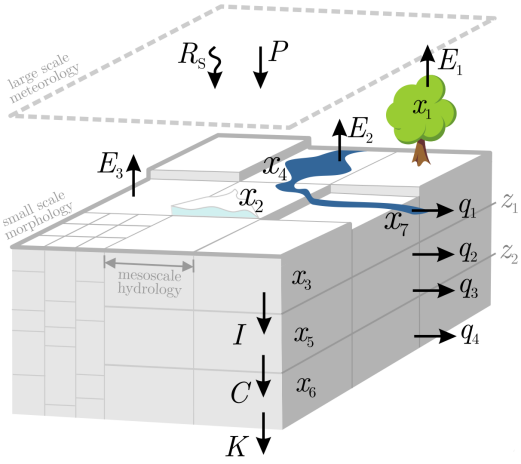

Figure 1. The concept of the mesoscale hydrologic model, mHM.

2 Model description

2.1 mesoscale Hydrologic Model (mHM~~v5.7~~)

The mesoscale Hydrologic Model (mHM, www.ufz.de/mhm) is a spatially explicit distributed hydrologic model that uses grid cells as a primary modeling unit, and accounts for the following processes: canopy interception, snow accumulation and melting, soil moisture dynamics, infiltration and surface runoff, evapotranspiration, subsurface storage and discharge generation, deep percolation, baseflow, discharge attenuation, as well as flood routing (Figure 1). The runoff generation applies a robust scheme which routes runoff in upstream cells along river ~~network~~ networks using the Muskingum-Cunge algorithm. The model is driven by daily meteorological forcings (e.g., precipitation, temperature), and ~~it utilizes observable basin~~ utilizes observable physical properties or signals of the basin (e.g., soil textural, vegetation, and geological properties) to infer the spatial variability of the required parameters. mHM is an open-source project written in Fortran 2008. Parallel versions of mHM ~~are available based on OpenMP concepts~~ using OpenMP concepts are available.

A unique feature of mHM is the application of Multiscale Parameter Regionalization (MPR). The MPR method accounts for subgrid ~~variabilities of catchment physical characteristics~~ variability in physical characteristics of the catchment such as topography, ~~terrain,~~ soil and vegetation. The ~~model is flexible~~ MPR methodology facilitates the flexibility of the model for hydrological simulations at various spatial scales ~~due to~~ by applying the MPR methodology ~~(Samaniego et al., 2010; Kumar et al., 2013a, b; Rakovec~~ ~~Within mHM , three levels are differentiated~~ (Samaniego et al., 2010; Kumar et al., 2013a, b; Rakovec et al., 2016a, b; Samaniego et al., 20 mHM differentiates three levels to better represent the spatial variability of state and ~~inputs~~ input variables. The effective parameters ~~in~~ at different spatial scales are dynamically linked by a physically-based upscaling scheme. ~~Detailed~~ A detailed description of MPR, as well as the formulations governing hydrological processes~~could be referred to in~~ , are given by Samaniego et al. (2010) and Kumar et al. (2013b).

Below, we ~~listed the formulas~~ list the equations that describe near-surface processes in the deep soil ~~layer and groundwater layer~~and groundwater layers. The comprehensive ~~equation system of~~ system of equations of mHM can be found in Samaniego et al. (2010). Here, we only listed the equations ~~relevant to this study~~for the coupling to OpenGeoSys. In the subsurface reservoir, which is the second ~~layer of the vertical layers~~ vertical layer (x_5 in Figure 1), interflow is partitioned into fast interflow (q_2) and slow interflow (q_3):

$$q_2(t) = max\left\{I^2(t) + x_5(t-1) - \beta_1(z_2 - z_1),\ 0\right\}\beta_2 \tag{1}$$

$$q_3(t) = \beta_3(x_5(t-1))^{\beta_4} \tag{2}$$

where $q_2(t)$ is fast interflow at time t ~~([mm d^{-1}])~~, I is the infiltration capacity [mm~~d^{-1}~~], x_5 is the water depth of water storage in the deep soil reservoir [mm], β_1 is the maximum holding capacity of the deep soil reservoir, z_i is depth of subsurface layer i, β_2 is the fast recession constant, $q_3(t)$ is slow interflow at time t [mm d^{-1}], β_3 is the slow concession constant, and β_4 is the exponent that quantifies the degree of non-linearity of the cell response.

The groundwater recharge is equivalent to the percolation to the groundwater reservoir (the third ~~of the vertical layers~~vertical layer, see x_6 in Figure 1). The groundwater recharge $C(t)$ can be expressed by

$$C(t) = \beta_5 x_5(t-1) \tag{3}$$

where $C(t)$ is ~~groundwater recharge at cell i~~ the groundwater recharge in cell i [mm d^{-1}], and β_5 is the effective percolation rate.

In the groundwater reservoir, baseflow is generated following a linear relationship between storage and runoff:

$$q_4(t) = \beta_6 x_6(t-1) \tag{4}$$

where $q_4(t)$ is the baseflow [mm d^{-1}], β_6 is the baseflow recession rate, and x_6 is depth of groundwater reservoir [mm].

The runoff from upstream grid ~~and internal runoff at~~ cells and the internal runoff in cell i are routed into streams using the Muskingum algorithm:

$$Q_i^1(t) = Q_i^1(t-1) + c_1(Q_i^0(t-1) - Q_i^1(t-1)) + c_2(Q_i^0(t) - Q_i^0(t-1)) \tag{5}$$

with

$$Q_i^0(t) = Q_{i'}(t) + Q_{i'}^1(t) \tag{6}$$

$$c_1 = \frac{\Delta t}{\kappa(1-\xi) + \frac{\Delta t}{2}} \tag{7}$$

$$c_2 = \frac{\frac{\Delta t}{2} - \kappa\xi}{\kappa(1-\xi) + \frac{\Delta t}{2}} \tag{8}$$

where Q_i^0 and Q_i^1 denote the runoff entering and leaving the river reach located ~~on~~ in cell i, respectively [mm d^{-1}], $Q_{i'}$ is the contribution from the upstream cell ~~i'~~ i' [mm d^{-1}], κ is the Muskingum travel time parameter, ξ is the Muskingum attenuation parameter, Δt is ~~time interval in hours~~ the time step-size [~~hr~~h], and t is the time index for each Δt interval.

2.2 OpenGeoSys (OGS)

OpenGeoSys (OGS) is an open-source project with the aim of developing robust numerical methods for the simulation of Thermo-Hydro-Mechanical-Chemical (THMC) processes in porous and fractured media. OGS is written in C++ with a focus on the finite element analysis of coupled multi-field problems. Parallel versions of OGS ~~are available~~ based on both MPI and OpenMP concepts are available (Wang et al., 2009; Kolditz et al., 2012; Wang et al., 2017). To date, two OGS versions are available~~for use. These are~~: OGS5 (https://github.com/ufz/ogs5) and OGS6 (https://github.com/ufz/ogs). In this study, the term "OpenGeoSys (OGS)" represents OGS5 if ~~there is no special instruction~~not stated otherwise.

OGS has been successfully applied in different fields, such as water resources management, hydrology, geothermal energy, energy storage, CO_2 storage, and waste ~~deposition~~ disposal (Kolditz et al., 2012; Shao et al., 2013; Gräbe et al., 2013; Wang et al., 2017). In the field of hydrology / hydrogeology, OGS has been applied to ~~various topics such as~~ regional groundwater flow and transport (Sun et al., 2011; Selle et al., 2013), contaminant hydrology (Beyer et al., 2006; Walther et al., 2014), reactive transport (Shao et al., 2009; He et al., 2015), and sea water intrusion (Walther et al., 2012), ~~etc~~among others.

~~Here we list the governing equations of saturated groundwater flow , which are relevant in this study. They can be expressed as~~Saturated groundwater flow follows the continuity equation and Darcy's law:

$$S\frac{\partial \psi_p}{\partial t} = -\nabla \cdot \boldsymbol{q} + q_s \tag{9}$$

$$\boldsymbol{q} = -K_s \nabla(\psi_p - z) \tag{10}$$

where S is specific storage ~~for confined aquifer and specific yield for unconfined aquifer~~ coefficient in confined aquifers, and the specific yield in unconfined aquifers [1/L], ψ_p is the pressure head in ~~porous media~~ the porous medium [L], t is time[T], $\boldsymbol{q}$ is the ~~Darcy flux (LT^{-1})~~specific discharge or Darcy velocity [LT^{-1}], q_s is ~~a~~ the volumetric source/sink term ~~(T^{-1})~~[T^{-1}], K_s is the saturated hydraulic conductivity tensor [LT^{-1}], ~~z~~ and z is the vertical coordinate [L].

The stream network is normally represented by a set of polylines in the geometry file of OGS. In the case of a 3-D model, a common way to set up the polyline system is to utilize the mapping tool embedded in OGS source code, by which the shape file obtained from GIS software representing streams can be easily mapped onto the upper surface of OGS mesh and converted into a set of polylines. Each reach of the stream network can be represented by one polyline or several continuous polylines, depending on the demand of the user. Each polyline consists of a set of continuous mesh nodes, upon which Dirichlet, Neumann or Robin boundary conditions can be applied.

2.3 Coupling mechanism

The coupled model mHM#OGS v1.0 is developed to simulate SW/GW flow in one or more catchments by simultaneously calculating flow across the land surface and within ~~subsurface materials~~the groundwater. mHM#OGS v1.0 simulates flow within three hydrological regions. The first region is limited by the upper ~~bound of~~ boundary of the plant canopy and the lower ~~bound~~ boundary of the soil zone bottom. The second region includes open-channel water, such as streams. The third region is

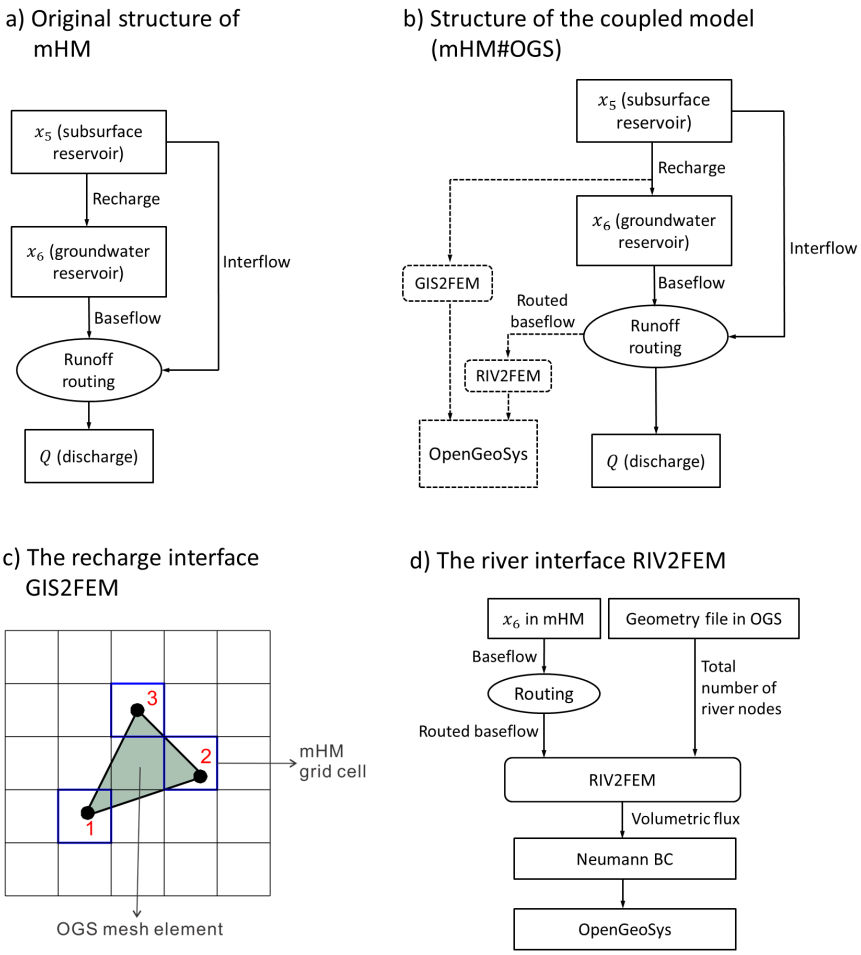

Figure 2. ~~The schematic~~ Schematic of the coupled model mHM#OGS v1.0. a) ~~the original~~ Original structure of ~~vertical~~ the vertically layered reservoir of mHM~~;~~. b) ~~the structure~~ Structure of the coupled model (mHM#OGS v1.0). c) Illustration of data interpolation and transformation through the coupling interface GIS2FEM. d) Scheme of the river interface RIV2FEM. For the sake of simplicity, the figure only displays mHM layers relevant to this study ~~,~~and ~~neglect~~ neglects the other mHM layers (i.e. ~~g.,~~ x_1 - x_4). ~~Grid-based~~ In Figure 2c, the grid-based mHM fluxes (e.g., ~~GW~~ recharge~~and baseflow~~) are ~~linear~~ linearly interpolated to the top surface of the OGS mesh, and further transferred into volumetric values ~~by face integration calculation, which are~~ and directly assigned to ~~OGS~~ the surface mesh nodes of the OGS grid.

the ~~saturated groundwater~~ water-saturated aquifer. mHM is used to simulate the processes in the first and second regions, while OGS is used to simulate the hydrological processes in the third region.

The ~~basic idea is to feed fluxes generated by mHM(e. g. , distributed groundwater~~ coupling initiative aims to add additional predictive capability of groundwater heads, which is achieved by OGS, to the existing capability of predicting discharges that is achieved by mHM. mHM is used to estimate step-wise and component-wise a water budget through model calibration against discharge. In contrast, OGS serves as a post-processor to obtain groundwater heads by using mHM simulated recharge and

baseflow ~~) to the OGS mesh surface as outer forcings through a coupling interface~~ as driving forces. Two model interfaces, namely GIS2FEM ~~(Figure 2).~~ and RIV2FEM, have been developed to link the two models by transferring recharge and baseflow from mHM to Neumann boundary conditions in OGS. The two models are ~~performed~~ executed separately and sequentially~~with usually~~ , typically with different temporal (e.g., daily in mHM and weekly or monthly in OGS) and spatial resolutions (e.g., ~~larger grid cell size~~ rectangular, structured grids with coarse resolution in mHM and smaller~~element size~~ , potentially unstructured grids with fine resolution in OGS). The original ~~vertical~~ vertically layered reservoirs in mHM, ~~e.g.,~~ namely the soil-zone reservoir ~~, subsurface reservoir and groundwater reservoir~~ and the subsurface reservoir are preserved, implying that all ~~well tested~~ well-tested features of mHM (e.g., MPR, infiltration-runoff partitioning) are ~~fully preserved~~ retained in the coupled model.

To illustrate the coupling mechanism in detail, we itemized the coupling workflow ~~in~~ below.

1. mHM is run ~~independent~~ independently of OGS to calculate land surface fluxes.

   Using ~~a gridded observations of~~ gridded meteorological forcings (precipitation, temperature, and potential evapotranspiration), the grid-based ~~distributed~~ infiltration rates (e.g., groundwater recharge) and runoff components (e.g., interflow, baseflow) are ~~thereby~~ estimated and saved as mHM output files. The original linear groundwater reservoir (depth x_6 in Figure 1) is used to estimate baseflow. Moreover, MPR is used in the calibration process such that subgrid variabilities can be validly calculated. The ~~details of physical basis and parameterization scheme in mHM can be found in Section 2.1. The~~ spatially distributed groundwater recharge and total routed baseflow are written into raster files for later use.

2. After ~~mHM run was~~ the mHM run has finished, the step-wise routed baseflow estimated by mHM ~~are~~ is transformed to distributed ~~baseflow along OGS stream network~~river discharges along streams as represented in OGS.

   ~~The stream conceptions within mHM and OGS are slightly different, in terms that the stream in mHM is~~ Most physically-based ISSHMs characterize river-groundwater interaction based on either first-order flow exchange or boundary condition switching (Paniconi and Putti, 2015). However, this approach inevitably relies on a parameter-set describing geometric, topographic, and hydraulic properties of the stream channel (e.g., river bed conductance, river bed and drain elevations, channel width). Unfortunately, these parameters are essentially unknown at a large scale due to the lack of data and the subgrid-scale variability of these parameters. Due to these limitations, we use an alternative approach which is based on the routed baseflow estimated by mHM.

   mHM and OGS conceptualize streams differently: streams in mHM are implicitly defined based on pre-processing of ~~DEM~~ digital elevation model (DEM) data and a routing scheme, while ~~in OGS by~~ OGS uses an explicit predefined river geometry. In OGS, each reach of the stream network is defined by a polyline in ~~an~~ the OGS geometry file. To coordinate the two different ~~conceptions, we transfer the total routed baseflow (~~approaches, we developed a model interface, RIV2FEM, to convert the routed baseflow estimated by mHM ~~) to distributed baseflow along OGS streams by~~ to Neumann boundary conditions assigned at stream nodes of the OGS mesh (Figure 2d). Via RIV2FEM, the routed baseflow estimated by mHM is transferred to the uniformly disaggregated discharges by distributing it uniformly along

the predefined stream network in OGS ~~. The detailed description of stream network in OGScan be found in Section 3.4~~(Figure 2d):

$$\overline{q}_4(t) = \frac{Q_4(t)}{N} \tag{11}$$

where $\overline{q}_4(t)$ denotes the disaggregated discharge assigned at every stream node in OGS at time t [L^3T^{-1}], $Q_4(t)$ denotes the routed baseflow at the outlet of catchment at time t [L^3T^{-1}], N denotes the total number of stream nodes in OGS. The uniformly disaggregated discharges are then assigned to every stream node in OGS to serve as the Neumann boundary condition (Figure 2d). This approach ~~is based on the fact that due to lack of data (e.g., river bed conductance, tracer tracking, etc), the spatial pattern of baseflow along streams is uncertain. Based on the mass conservation criteria, we made the assumption that baseflow is uniformly distributed along streams such that the step-wise water balance is guaranteed~~significantly reduces the number of parameters, avoids the uncertainty caused by the unknown river properties, and is suitable for many real-world applications that suffer from scarce data. Moreover, as recharge and baseflow are directly taken from mHM, the mass conservation criterion is naturally satisfied in this approach.

3. The distributed groundwater recharge generated ~~from mHM are~~ by mHM is fed to the coupling interface GIS2FEM, and ~~further~~ then transferred to the upper surface boundary conditions of the OGS model.

   The coupling interface GIS2FEM is used to interpolate and transfer mHM grid-based ~~fluxes~~ recharge to OGS nodal ~~flux values. After reading a raster file of mHM generated fluxes, the interface~~ recharge values. GIS2FEM interpolates the flux value to the top surface elements of the OGS mesh. ~~For each surface element, if its centroid is within the range of~~ The detailed workflow is:

   - GIS2FEM reads the raster file generated by mHM and the mesh file of OGS.
   - In the case of a 3-D mesh, GIS2FEM extracts the upper surface of the OGS mesh. For each of the nodes on this surface, GIS2FEM searches for the mHM grid cell ~~, the flux~~ that the node is located in, and assigns the recharge value of this grid cell ~~is assigned~~ to the corresponding ~~surface element in OGS mesh.~~ node (marked as C^m).
   - After all top surface elements ~~being~~ have been processed, GIS2FEM ~~will take~~ undertakes the face integration calculation, by which the ~~recharge data and baseflow are converted into nodal source terms~~ specific recharge C^m [LT^{-1}] calculated by mHM is converted into volumetric recharge C^{in} [L^3T^{-1}] and assigned to the corresponding OGS mesh nodes (Figure 2c). Specifically, the specific recharge C in a certain element is calculated as:

     $$C(\mathbf{x}) = \sum_{j=1}^{N} W_j(\mathbf{x}) C_j^m, \tag{12}$$

     where $\mathbf{x}$ is the spatial coordinate on the surface, N is the total number of nodes in a surface element, W_j is the weighting function of node j, C_j^m is the specific recharge at node j calculated by mHM [LT^{-1}]. Then the volumetric

recharge C_i^{in} at node i (i is the global node index) is calculated by the face integration calculation:

$$C_i^{in} = -\int_{\partial\Omega} W_i(\mathbf{x}) C(\mathbf{x}) d(\mathbf{x}), \tag{13}$$

where C_i^{in} is the volumetric recharge of node i [L^3T^{-1}], $\partial\Omega$ is the surface boundary of the FEM domain, W_i is the weighting function of node i.

4. After ~~mHM generated~~ the mHM-generated recharge and baseflow ~~were successfully transferred to boundary conditions on upper-surface of~~ have been transferred into boundary conditions at the upper surface of the OGS mesh, the groundwater model ~~will run subsequently to simulate~~ is run to simulate the groundwater flow and transport~~processes~~.

3 ~~Study area and model setup~~Example application

3.1 Study area and model setup

We use a meso-scale catchment (about 850 km^2) upstream of the Nägelstedt gauge located in central Germany to ~~establish and assess our~~ test our coupled model (Figure 3). The Nägelstedt catchment comprises the headwaters of the Unstrut ~~river basin. The Unstrut river basin is a sedimentary basin of Unstrut river, a left~~ River, a tributary of the ~~Saale. It is selected in this study because there are many~~ River Saale. We selected this study area because many of the groundwater monitoring wells ~~operated by Thuringian State office~~ in the area are operated by the Thuringian State Office for the Environment and Geology (TLUG) and the Collaborative Research Center AquaDiva (Küsel et al., 2016). ~~Morphologically, the terrain~~ The elevation within the catchment ~~is in a range of~~ ranges between 164 m and 516 m, whereby the higher regions are in the west and south ~~as part of~~ and belong to the forested hill chain of the Hainich (Figure 3). The Nägelstedt catchment is one of the most intensively used agricultural regions in Germany. In terms of drinking water supply, about 70% of the water requirement is satisfied by groundwater (Wechsung, 2005). About 17% of the land in this region is forested area, 78% is covered by crop and grassland, and 4% is ~~housing~~ urban and transport area. The mean annual precipitation in this area is about 660 mm.

In this study, mHM runs ~~are~~ were executed for a time period of 35 years (from January 1, 1970 to December 30, 2004), with the period 1970 - 1974 being used ~~as a~~ for spin-up~~period. OGS is~~. OGS was run for the period from January 1, 1975 to December 30, 2005. mHM ~~is~~ was run with a daily time step, while OGS ~~is~~ was run with a monthly time step. The resolution of mHM grid cells is 500 m $\times$ 500 m. ~~The~~ OGS uses a structured, hexahedral 3-D mesh, with a spatial resolution of ~~OGS mesh is set to~~ 250 m $\times$ 250 m in the horizontal direction and 10 m in the vertical direction over the whole domain. The ~~fluxes are interpolated from coarser mHM grid to finer OGS surface element through GIS2FEM (Figure 2c). The~~ detailed input data and ~~parameter set~~ parameter-set to run both models are detailed in the following sections.

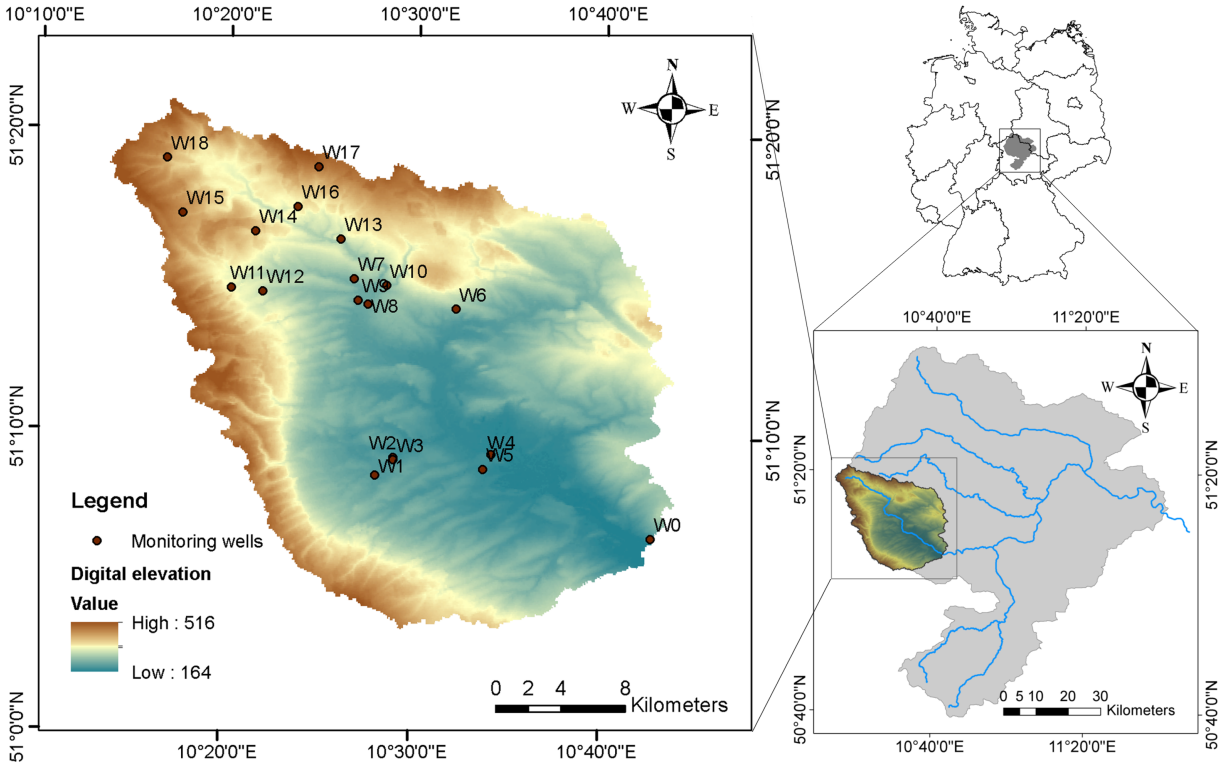

Figure 3. The Nägelstedt catchment used as the test catchment for this ~~model~~study. The ~~left~~ left-hand map shows elevation and locations of monitoring wells used in this study. The lower ~~right~~ right-hand map shows the relative location of Nägelstedt catchment in the Unstrut ~~basin~~Basin. The upper ~~right~~ right-hand map shows the location of the Unstrut ~~basin~~ Basin in Germany.

3.2 Meteorological forcings and morphological properties

We started the modeling by performing the daily simulation of mHM to calculate near-surface hydrological processes. The mHM model is forced by daily meteorological ~~forcings~~conditions, including distributed precipitation and ~~atmosphere~~ atmospheric temperature. The spatial patterns of precipitation and ~~atmosphere~~ atmospheric temperature were based on point measurements of precipitation and ~~atmosphere~~ atmospheric temperature at weather stations from the German Meteorological Service (DWD). The point data at weather stations were subsequently ~~krigged~~ kriged into a 4 km precipitation ~~fields~~field, and then downscaled to mHM grid cells. Moreover, the potential ET was quantified based on the method from Hargreaves and Samani (1985). Other datasets used in mHM are the ~~digital elevation model (DEM)~~ DEM data, which is the basis for deriving properties ~~like~~ such as slope, river beds, and flow direction; soil and geological maps, and meta-data such as sand and clay contents, bulk density, and dominant geological types; CORINE land-cover information (in the years 1990, 2000 and 2005); and discharge data at the outlet of the catchment.

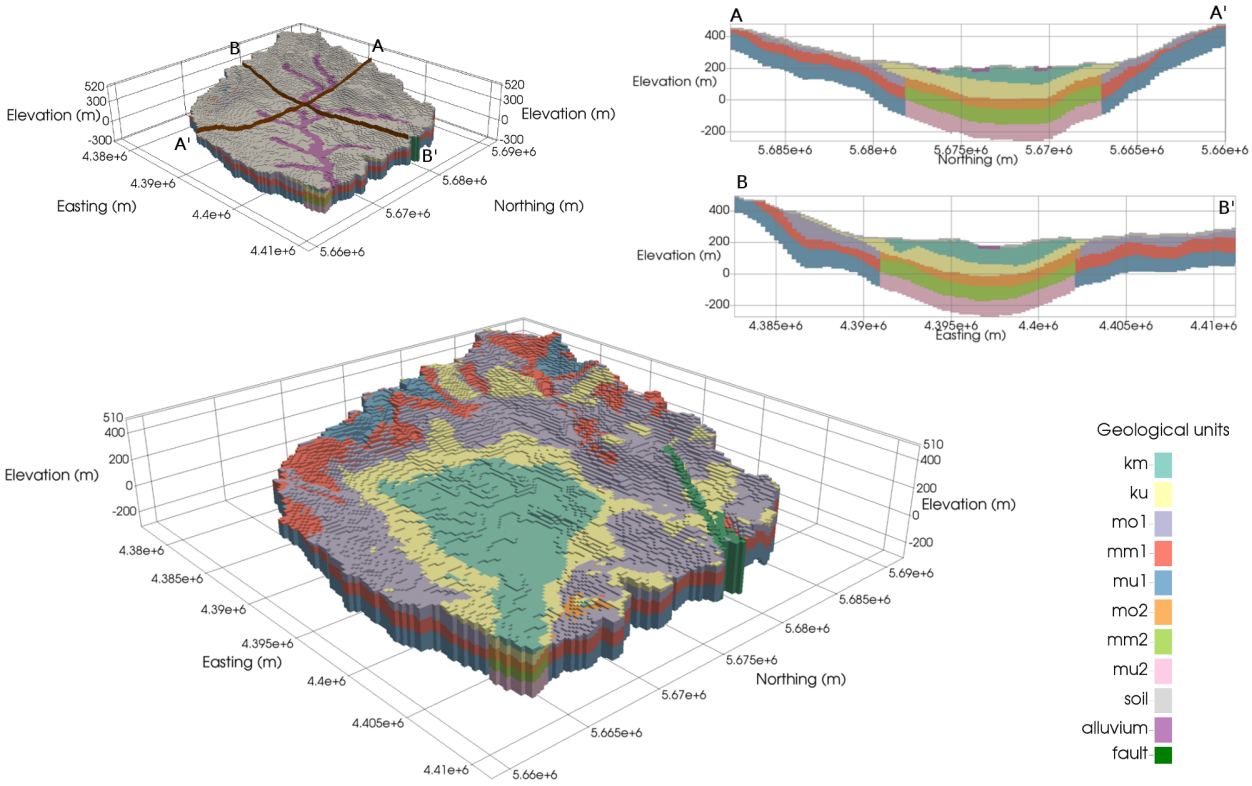

Figure 4. Three-dimensional and ~~cross section view~~ cross-sectional views of the hydrogeologic zonation in the Nägelstedt catchment. The upper ~~left~~ left-hand figure ~~shows~~ highlights the ~~complete geological characterization and zonation including~~ distribution of alluvium and soil ~~zone~~zones. The upper ~~right~~ right-hand figure shows ~~the geological characterization along~~ two ~~cross sections~~vertical geological cross-sections. The lower map shows the detailed zonation of geological sub-units beneath the soil zone and alluvium.

3.3 Aquifer properties

We used a stratified aquifer model to explicitly ~~present~~ represent the heterogeneous distribution of hydraulic properties (~~e.g., K value,~~ hydraulic conductivity, specific yield, and specific storage). The stratified aquifer model is based on well log data and geophysical data ~~from Thuringian State office~~ obtained from the Thuringian State Office for the Environment and Geology (TLUG). ~~To convert the data format, we~~ We used the workflow developed by Fischer et al. (2015) to convert the data format, by which the complex ~~3D geological model were converted into~~ 3-D geological model was converted into the open-source VTK format file that can be directly read by OGS.

The ~~dominant sediments~~ major stratigraphic units in the study site are the Muschelkalk (Middle Triassic) and ~~Keuper (Middle and Late~~ the Keuper (Upper Triassic). Younger ~~deposits from~~ Tertiary and Quaternary deposits are less important for the ~~large scale~~ large-scale hydrogeology of the basin. The Keuper deposits mainly lie in the center of the Unstrut ~~basin~~ Basin and act as permeable shallow aquifers. In the Nägelstedt catchment, the Keuper deposits are further ~~classified~~ subdivided into

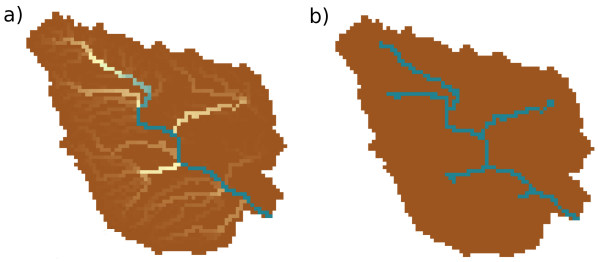

Figure 5. Illustration of the stream network used in this study. a) Original stream network based on the streamflow routing algorithm of mHM; b) Processed stream network that ~~are~~ was used in this study. The small tributaries where the runoff rates are below ~~1500 mm/month~~ the threshold have been ~~cut out~~ removed from the original stream network.

two geological sub-units: Middle Keuper (km) and ~~low~~ Lower Keuper (ku) (see Figure 4). The Muschelkalk is marked by a prevailing marine environment and is subdivided into three sub-units the Upper Muschelkalk (mo), Middle Muschelkalk (mm~~) and low~~ , dolomites and residues of eroded salt layers) and Lower Muschelkalk (mu, limestones). According to previous geological ~~survey (Seidel, 2004), even the same sub-unit of Muschelkalk have a diverse~~ surveys (Seidel, 2004), the sub-units of the Muschelkalk have varying hydraulic properties depending on their positions and depths. They are further divided into sub-units with higher permeabilities ~~, which are~~ (mo1, mm1 and mu1~~(Figure 4) ,~~) and sub-units with lower permeabilities ~~, which are~~ (mo2, mm2 and mu2) (Figure 4). The Upper Muschelkalk (mo) has been widely considered as a karstified formation. Recent research by Kohlhepp et al. (2017) has revealed that in the Hainich Critical Zone, the intense karstification and the conduit are limited at the base of the mo formation. Accordingly, we use the equivalent porous medium approach to characterize the Upper Muschelkalk. The uppermost layer with a depth of 10 m is set as a soil layer (Figure 4). A high-permeability alluvium layer is set along the mainstream and major tributaries ~~representing~~ to represent granite and stream deposits (Figure 4).

3.4 Boundary conditions

Based on the steep topography along the watershed divides, groundwater is assumed to be naturally separated and ~~not able~~ unable to pass across the boundaries of the watershed. ~~No-flow boundaries are imposed~~ In general, no-flow boundaries are set at the outer perimeters surrounding the basin as well as at the lower aquitard. On the basis of the measurements, a Dirichlet boundary condition is assumed at the northwestern and northeastern edges.

The stream network ~~is~~ was delineated by processing a grid-based runoff raster file generated by mHM. The grid-based runoff ~~is~~ was converted to a valid stream network compatible with OGS. The necessity of transferring the mHM runoff raster file to the OGS stream network has been elaborated in Section ~~??~~2.3. Particularly in this case study, we ~~cut out~~ removed the small intermittent tributaries by setting a threshold value of long-term averaged routed runoff. Only streams with a runoff rate higher than the threshold (in this case study, ~~1500 mm~~0.145 m^3/~~month~~s) are delineated as valid streams. In other words, we ~~neglect~~ neglected the intermittent streams to the upper stream reaches (Figure 5). The preprocessed stream network consists of a main stream and four tributaries (Figure 5b). ~~Each~~ The reach of each stream is defined as a polyline in ~~OGS geometry~~

~~fileand comprises many consecutive nodes in OGS mesh. As mentioned in Section ??, uniformly distributed baseflow rates were subsequently~~ a geometry file. As illustrated in Section 2.3, uniformly disaggregated groundwater discharges processed by the interface RIV2FEM were assigned to every OGS mesh ~~nodes~~ node within the stream network.

3.5 Calibration procedure

The calibration of the coupled model ~~mHM#OGS v1.0~~ follows a two-step procedure.

~~For~~ In the first step, mHM ~~is~~ was calibrated independent of OGS for the period from 1970 to 2005 by matching the observed runoff at the outlet of the catchment. The first 5 years ~~are used as "~~were used as spin-up ~~"~~ period to set up initial conditions in the near-surface soil zone. ~~The calibration workflow is a consecutive workflow where the parameters which affect the potential evapotranspiration, soil moisture, runoff and shallow subsurface flow were first calibrated until convergence criteria was matched. The calibration goodness is handled by means of calculating~~ The calibration quality is quantified by the Nash-Sutcliffe ~~efficient~~ coefficient of efficiency (NSE):

$$NSE = 1 - \frac{\sum_{i=1}^{n} |(h_m - h_s)|_i^2}{\sum_{i=1}^{n} |(h_m - \bar{h}_m)|_i^2} \frac{\sum_{i=1}^{n} |(q_m - q_s)|_i^2}{\sum_{i=1}^{n} |(q_m - \bar{q}_m)|_i^2} \tag{14}$$

where ~~h_s~~ q_s is the simulated ~~groundwater head~~ discharge [~~m~~L^3T^{-1}], ~~$\bar{h}_m$ is the~~ q_m is the measured discharge [L^3T^{-1}], and $\bar{q}_m$ is the mean of measured ~~groundwater head~~ discharge [~~m~~L^3T^{-1}].

~~For~~ In the second step, the ~~steady state groundwater model is~~ steady-state groundwater model in OGS was calibrated to match the long-term mean of ~~groundwater observations. The long term~~ observed groundwater levels. The long-term mean of recharge and baseflow estimated by mHM ~~are~~ were fed to the steady-state groundwater model as Neumann boundary conditions. ~~Meanwhile, the groundwater levels obtained from a couple of monitoring wells are averaged over the whole simulation period. The calibration of the steady-state groundwater model aims for the most plausible distribution of hydraulic conductivities. The intervals (i.e., upper and lower bounds) of adjustable parameters are taken from the literature (Wechsung, 2005; Seidel, 2004). Model-to-measurement matching is implemented by minimizing the~~ The calibration was performed using the software package PEST (Doherty et al., 1994). The model parameters were adjusted within a fixed interval until the value of objective function, which is the sum of weighted squared residuals of ~~long-term mean of modeled groundwater heads~~ modeled and observed groundwater heads~~in this case. Goodness of fit between the simulated and observed long-term average groundwater levels~~ , was minimized. Specifically, the intervals of adjustable parameters were taken from the literature (Wechsung, 2005; Seidel, 2004), and the weights assigned to each observation were set uniformly to 1. The calibration result is assessed by the root-mean-square error (RMSE).

3.6 Model evaluation and sensitivity analysis

~~Besides the observed discharge at the catchment outlet, we also used observed groundwater head time series~~ We used the time series of groundwater levels in 19 monitoring wells ~~distributed over the catchment to evaluate model performance. The K values are set to the optimized values from the steady state modelcalibration~~to evaluate the predictive capability of the transient

model. In the transient model, hydraulic conductivities are obtained from the calibrated steady-state model. Meanwhile, the initial condition of the groundwater head is directly taken from the result of steady-state ~~groundwater hydrographs are used as the initial condition for the transient groundwater model. For evaluating the predictive ability in terms of groundwater heads, the~~ model. The Pearson correlation coefficient R_{cor} and the ~~inter-quantile~~ inter-quartile range error QRE are used as two summary statistics to evaluate the predictive capability. The (relative) ~~inter-quantile~~ inter-quartile range error QRE is defined by:

$$QRE = \frac{IQ^{md}_{7525} - IQ^{dt}_{7525}}{IQ^{dt}_{7525}} \tag{15}$$

where IQ^{md}_{7525} and IQ^{dt}_{7525} are the ~~inter-quantile ranges of the time-series of modeling result~~ inter-quartile ranges of simulations and observations, respectively.

We ~~also~~ sought to quantify the ~~effect of a spatially-distributed recharge on the simulated groundwater heads~~sensitivity of groundwater flows to the different spatial pattern of recharge. For this purpose, ~~we set up a reference scenario by spatially homogenizing mHM generated recharge . To distinguish the two recharge scenarios, we use the abbreviation "RR" to represent reference recharge scenario, and "mR" to represent mHM recharge scenario.~~ a uniform recharge scenario was established as the reference scenario. The sensitivity analysis follows a two-step workflow. First, we calibrated the steady-state groundwater models ~~in~~ for the two recharge scenarios independently. Second, we conducted transient simulations by assigning the same values of storage parameters~~and compared~~ , and then observed their corresponding performances in two recharge scenarios. ~~Despite the different spatial pattern of recharge and K~~ With the exception of recharge scenario and hydraulic-conductivity values, all model parameters (e.g., specific ~~yields,~~ yield and specific storage) and ~~model inputs in the two scenarios are identical~~inputs are set to be identical in both scenarios. The ~~R_{cor} and~~ Pearson correlation coefficient R_{cor}, and the inter-quartile range error QRE are used as two summary skill scores to assess model performances in the two ~~respective simulations~~recharge scenarios.

4 Results

4.1 Calibration

As the first ~~step~~part of calibration, mHM is calibrated ~~independent of OGS. Monthly dischargedata from January 1975 to December 2004 are used for model calibration~~against discharge. The calibration results ~~indicate that mHM is capable to reasonably reproduce~~ demonstrate the predictive capability of mHM in reproducing the time series of catchment discharge (Figure 6). The Nash–Sutcliffe model efficiency coefficient (NSE) is 0.88, while the Pearson correlation coefficient is 0.96~~(Figure 6)~~. Other fluxes~~like evaportransporation~~ , such as evapotranspiration measured at eddy-covariance stations inside this area, also shows quite reasonable correspondence to the modeled ~~estimation~~ estimate (Heße et al., 2017).

~~Subsequently, steady-state~~ In the second step, the steady-state groundwater model is calibrated against ~~long-term~~ the long-term mean of groundwater heads~~using PEST (Doherty et al., 1994)~~. Table 1 shows the calibrated hydraulic conductivities ~~of each~~ in each of the geological units. The ~~upper and lower limit of each parameter are defined according to literature (Seidel, 2004; Wechsung, 2005). The calibrated conductivities in each geological zone are displayed in Table 1. The~~ objective

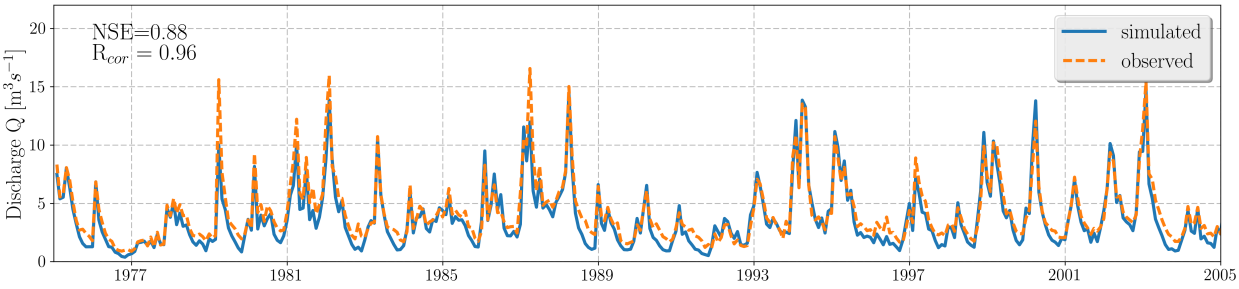

Figure 6. Observed and simulated monthly discharge at the outlet of ~~Nagelstedt~~ the Nägelstedt catchment.

function of calibration, which is the sum of squared weighted residuals, converged from ~~the~~ an initial value of 8625 m^2 to 464.74 m^2 after ~~114 total~~ a total of 114 model runs.

Broadly speaking, the ~~calibration results suggest that the groundwater~~ steady-state model can plausibly reproduce the finite numbers of observed groundwater heads ~~within~~ in the catchment. Figure 7 shows the ~~1-to-1~~ one-to-one plot of simulated and observed groundwater heads (locations of those wells are shown in Figure 3). ~~It can be observed that~~ In general, the model is capable of reproducing spatially-distributed groundwater heads ~~in~~ over a wide range, with an overall RMSE ~~value of 6.33 m. The errors in simulated heads (Figure 7b) show that most of simulated head errors are within an interval of ±~~of 6.45 m. Most of the discrepancies between individual observations and simulations are within a reasonable range (i.e., less than 6 m). Nevertheless, ~~there are~~ some monitoring wells ~~where the predictions are biased from observations~~ show larger discrepancies between observations and simulations (i.e., greater than 6 m), which is due to the ~~complex local geological formations around monitoring wells. No~~ unknown local geological properties and subgrid-scale topographies. For the sake of simplicity, no further attempt was made to add more model complexity to improve ~~model-to-measurement match~~the model fit.

~~Simulated water table depth~~ The simulated depth to groundwater over the whole catchment using the calibrated ~~K~~ hydraulic-conductivity values is shown in Figure 7c. Broadly speaking, the calibrated model reasonably ~~reproduced~~ reproduces the spatial groundwater table distribution. Groundwater ~~table depth is as large as above~~ depth varies between greater than 40 m in the higher southwestern and northern mountainous areas, ~~whereas~~ to less than 5 m in the central lowlands~~from simulation results~~. The plausibility of steady-state simulation results can ~~further~~ be assessed through regionalized observations of groundwater heads (Wechsung, 2005).

4.2 Spatio-temporal patterns of recharge and baseflow

Groundwater recharge has ~~an arbitrary behaviour~~ a spatially variable and dynamic behavior depending on the sporadic, irregular, and complex features of ~~storm rainfall occurrences~~precipitation, geological structure, and morphological features. The temporal and spatial variability of groundwater recharge and baseflow is estimated by mHM ~~calculation with~~ over a period of 30 years from 1975 to 2005.

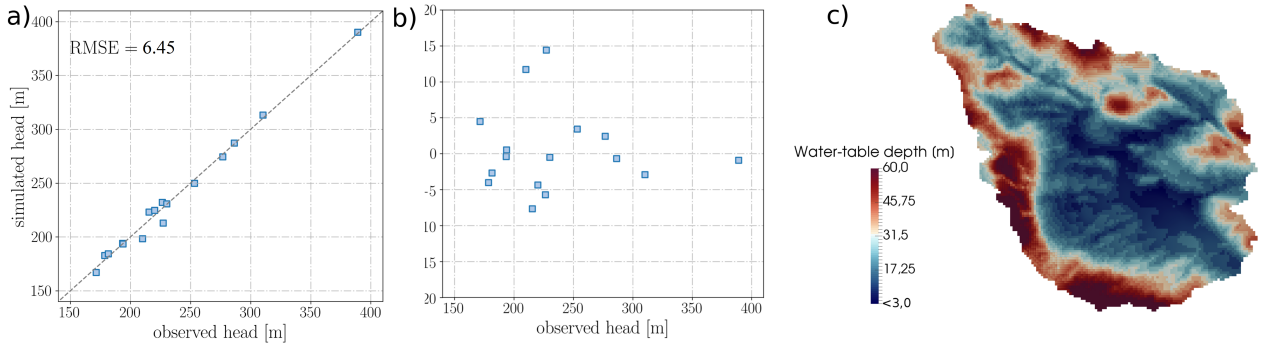

Figure 7. Illustration of steady-state groundwater model calibration and ~~simulatied~~ simulated heads. (a) Observed and simulated groundwater head (including RMSE); (b) Difference between simulated and observed head related to the ~~observaed~~ observed head values; c) Simulated long term mean water table depth ~~over~~ across the Nägelstedt catchment.

Figure 8 shows the spatial variability of groundwater recharge in three months: ~~the early spring (March)~~ March (Figure 8a), ~~late spring (May)~~ May (Figure 8b), and ~~winter (January)~~ January (Figure 8c). The results indicate that the ~~largest groundwater recharge may occur at mountainous areas. Largest recharge occurs~~ location of highest recharge rate is in the upstream ~~bedrock areas where dominant sedimentary is Muschelkalk with a relatively low hydraulic conductivity. The largest point-wise~~ mountainous areas where the Muschelkalk aquifer outcrops, but varies in different seasons. The maximum value of monthly groundwater recharge varies from 26 mm in ~~early spring~~March, to 51 mm in ~~late spring~~ May and 14 mm ~~in winter. Besides, we~~ January. We also evaluated the plausibility of groundwater recharge simulated by mHM ~~with~~ through comparison to other reference datasets. At the ~~large scale, the simulated groundwater recharge from~~ large-scale, the groundwater recharge simulated by mHM agrees quite well with estimates from the Hydrological Atlas of Germany (Zink et al., 2017).

~~The boxplot (~~Figure 9a ~~) shows the degree of spread and skewness of the distributions of the~~ shows the distribution of monthly groundwater recharge and ~~groundwater discharge. It reveals the long-term mean of groundwater inflow and outflow are balanced with the same monthly~~ monthly baseflow. Over the entire year, groundwater inflow (recharge) and outflow (baseflow) are balanced, exhibiting a mean value of 8 mm/month. ~~Due to the numerical error, a tiny difference of~~ The difference between the two values is merely 2~~% between groundwater recharge and baseflow is observed in the boxplot. Nevertheless, we consider this bias to be within an acceptable interval. Figure 9b shows the distribution of monthly groundwater recharge and monthly baseflow. The figure~~%. The figure, however, indicates that the distribution of monthly groundwater recharge is skewed to the right, whereas the distribution of monthly baseflow is ~~unimodal~~more peaked. Figure 9~~c~~ b depicts the time series of groundwater recharge and baseflow, which further demonstrates that the deviation of monthly groundwater recharge is larger than the baseflow. This phenomenon further reveals the significant buffering effect of the linear groundwater storage in mHM.

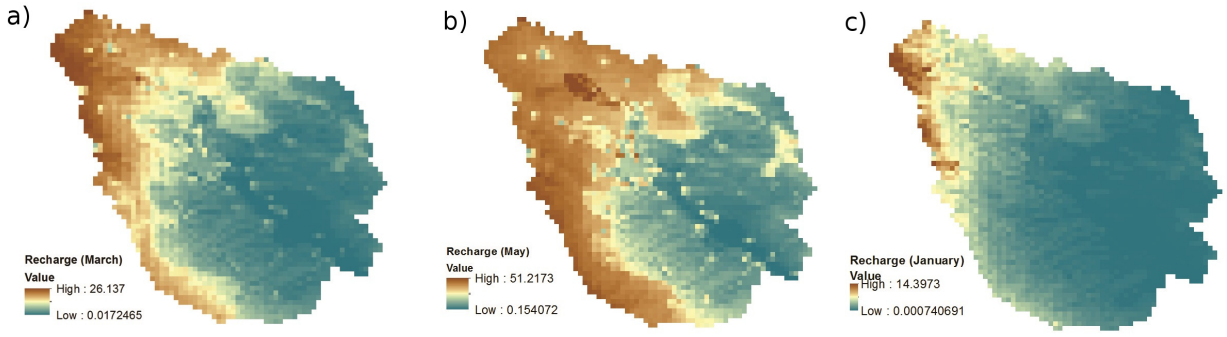

Figure 8. Spatial distributions of groundwater recharge in the Nägelstedt catchment (unit: mm/month) (a) ~~during early spring,~~ in March (b) ~~late spring~~in May, and (c) ~~winter of year~~ in January 2005.

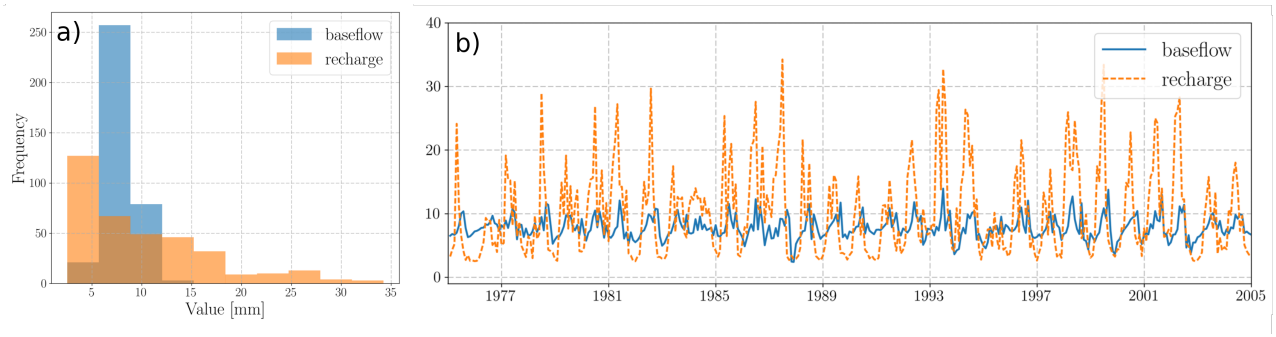

Figure 9. Analysis of groundwater inflow (recharge) and outflow (baseflow) over the Nägelstedt catchment. a) ~~Boxplot indicates spread, skewness, and outliers~~ Distribution of groundwater ~~recharge and baseflow. b) Histogram indicates the distribution of groundwater~~ balance components. ~~c~~b) Monthly time series of groundwater recharge and baseflow.

4.3 Model evaluation against dynamic groundwater heads

In this subsection, the head observations ~~at~~ of several monitoring wells ~~at~~ in the catchment were used to evaluate the model performance. We ~~analyze the~~ analyzed discrepancies between the modeled ~~groundwater heads and their corresponding observed values in their anomalies by removing~~ and observed groundwater heads by subtracting long-term mean values, $\overline{h}_{mod}$ and $\overline{h}_{obs}$. Four ~~model skill~~ model-skill scores including the mean value, the median value, the Pearson correlation coefficient R_{cor}, and the ~~inter-quantile range error~~QRE inter-quartile range error, QRE, are used to ~~judge~~ evaluate the model performance.

~~Five~~ Six wells with different geological and morphological properties were ~~randomly~~ chosen as samples to ~~test the effectiveness of our model~~ exhibit the model performance (Figure 10). Specifically, well ~~4728230786 is located at northern upland and near the mainstream~~W10 is located in the northern uplands and is near the main stream, whereas well ~~4828230754 is located at the southwestern lowland~~W1 is located in the southwestern lowlands. As can be observed from Figure 10, they provide good fits between simulated ~~heads~~ and observed heads~~with the~~ , with a R_{cor} of 0.87 and 0.76, and ~~the~~ QRE of -23.34% and -1.65%,

Table 1. Main hydraulic properties used ~~for mHM#OGS v1.0~~in the case study under the default mHM-generated recharge scenario.

Geological units	Hydraulic conductivity [m/s]			Specific yield [-]	Specific storage [m^{-1}]
	Lower limit	Upper limit	Calibrated value [m/s]		
km	1.0×10^{-6}	5.5×10^{-3}	~~1.493×10^{-5}~~ 1.844×10^{-5}	-	1×10^{-6}
ku	1.0×10^{-7}	~~3.4×10^{-3}~~ 3.4×10^{-4}	~~5.164×10^{-4}~~ 2.848×10^{-5}	-	1×10^{-6}
mo1	8.0×10^{-8}	2.0×10^{-3}	~~4.030×10^{-5}~~ 3.570×10^{-5}	0.10	1×10^{-6}
mm1	1.0×10^{-7}	9.0×10^{-4}	~~9.372×10^{-6}~~ 3.594×10^{-5}	-	1×10^{-6}
mu1	5.0×10^{-9}	2.0×10^{-4}	~~2.297×10^{-6}~~ 6.202×10^{-6}	-	1×10^{-6}
mo2	1.0×10^{-8}	5.0×10^{-4}	~~4.030×10^{-6}~~ 3.570×10^{-6}	-	1×10^{-6}
mm2	3.0×10^{-8}	9.0×10^{-5}	~~9.372×10^{-7}~~ 3.594×10^{-6}	-	1×10^{-6}
mu2	5.0×10^{-10}	2.0×10^{-5}	~~2.297×10^{-7}~~ 6.202×10^{-7}	-	1×10^{-6}
soil	5.0×10^{-5}	1.0×10^{-2}	~~3.068×10^{-4}~~ 6.617×10^{-5}	0.10	-
alluvium	~~4.0×10^{-4}~~ 4.0×10^{-5}	~~1.0×10^{-2}~~ 1.0×10^{-2}	~~1.026×10^{-3}~~ 3.219×10^{-4}	0.18	-

Table 2. Hydraulic properties used in the uniform-recharge scenario (RR).

Geological units	km	ku	mo1	mm1	mu1	mo2	mm2	mu2	soil	alluvium
Hydraulic conductivity [m/s]	5.023 $\times 10^{-5}$	6.216 $\times 10^{-5}$	8.608 $\times 10^{-5}$	2.990 $\times 10^{-5}$	5.316 $\times 10^{-6}$	8.604 $\times 10^{-6}$	2.997 $\times 10^{-6}$	5.317 $\times 10^{-7}$	5.239 $\times 10^{-5}$	7.302 $\times 10^{-4}$

respectively. Well ~~4628230773~~ W17 is located in ~~lower Keuper sediment~~the Lower Keuper unit, while well ~~4728230781 is located at upper Muschelkalk sediment. For those~~ W16 is located in the upper Muschelkalk formation. In these two monitoring wells, ~~simulation results~~ the simulations are highly correlated with ~~the observations with the R_{cor} of~~ observations with high values of R_{cor} (0.71 and ~~0.81~~ 0.82), in spite of their different geological properties (Figure 10). The simulation result at monitoring well ~~4728230783 located at~~ W13 (located in the northern mountainous area~~also has~~) also exhibits a high correlation with the observation with a R_{cor} of 0.85 (Figure 10). In general, the model is capable of capturing the historical ~~trend~~ trends of groundwater dynamics, even though the mean values of ~~simulation and observation may mismatch~~ simulations and observations may deviate to some extent. Due to the limited spatial resolution and complex hydrogeological ~~locality~~structure, this degree of discrepancy is acceptable.

4.4 Model sensitivity to different recharge scenarios

As described in Section 3.6, ~~we set up~~ a reference recharge scenario (RR)~~in which the recharge is homogeneously distributed in space~~ , i.e., a spatially uniform recharge scenario, is set up to assess the effect of spatial ~~recharge pattern towards groundwater hydrographs.~~ patterns of recharge on groundwater heads. In this uniform recharge scenario, RR, the steady-state groundwater model, was re-calibrated using the long-term mean of spatially uniform recharge (Table 2). For the purpose of showing dis-

crepancies ~~of model results~~ between two recharge scenarios, we compared the values of R_{cor} and |QRE| at each monitoring well between ~~mR and~~ the spatially distributed recharge, mR, and the uniform recharge, RR (Figure 11). The mean value and the median value of R_{cor} and QRE ~~are~~ were also calculated and are shown in Figure 11. Figure 11a indicates that the correlation between observations and simulations using the spatially distributed recharge mR is higher than that using the uniform recharge RR, with ~~the averaged vale~~ mean values of 0.704 and 0.677, respectively. The standard deviations are nearly the same in both scenarios. Considering that the only difference between ~~mR and RR is the spatial pattern of recharge , we do conclude that the inclusion of~~ the two recharge scenarios is their spatial patterns, we conclude that accounting for spatially-distributed recharge provides a ~~solid improvement . The individual discrepancies of R_{cor} at well levels between the two scenarios are moderate.~~ moderate improvement in the model.

Figure 11b shows the absolute values of ~~inter-quantile~~ inter-quartile range error (|QRE|) in simulations ~~under~~ using the two recharge scenarios (mR and RR)~~, respectively~~. We found that the ~~uncertainty~~ deviation of |QRE| is ~~higher~~ significantly larger than R_{cor}, ~~e.g.~~ i.e., the |QRE| in two wells are abnormally higher than the other wells. The higher values of ~~wells 4728230789 and 4627230764~~ |QRE| at W8 and W18 may be caused by ~~the~~ their proximity to model ~~bounds~~boundaries, as the two wells ~~locate near either river or outer bound. This phenomenon reveals the~~ are located either near a river or near the catchment perimeter. This deviation indicates that accurate quantification of ~~fluctuation amplitude in certain locations are difficult~~the amplitude of head fluctuations at certain locations is difficult, which may be due to the ~~complex hydrogeological locality of individual wells~~proximity of boundaries or complex local topography and geology. Nevertheless, 16 out of 19 wells ~~have low inter-quantile~~ exhibit low inter-quartile range errors, with the values of |QRE| ~~values within an interval~~ in a range of $\pm 40\%$ in the spatially distributed mR scenario. We also observe a smaller mean ~~value of |QRE| in mR than in RR. The~~ and standard deviation of |QRE| in ~~mR is also slightly smaller than RR. Those~~ the spatially distributed mR than in the uniform scenario, RR. The 19 chosen monitoring wells cover ~~geological zones of~~ the geological units of the alluvium, Keuper, and Muschelkalk, and range from high mountains to lowlands ~~all over~~ across the catchment. These results ~~point towards~~ demonstrate the promising modeling capability of the model and highlight the moderately better ~~simulation-to-observation match in mR~~historical matching when using a spatially distributed pattern of groundwater recharge.

Figure 12 ~~displays~~ illustrates the seasonality of groundwater heads ~~over the whole catchment by means of calculating the long-term mean groundwater heads in~~ by showing the spatial distribution of groundwater heads averaged over the spring, summer, autumn~~and winter~~, and winter seasons, respectively. ~~It indicates that in general, the possibility of groundwater flood event in spring and summer are higher than in autumn and winter. However, there can be observed a~~ A strong spatial variability ~~over different morphological classes and within different geological groups~~can be observed. For example, the fluctuation amplitudes of groundwater heads in the northern, eastern, and southeastern mountainous areas are larger than ~~the central plain areas. Considering the need for predicting groundwater flood and drought in~~ in the central plains area. In order to illustrate predicted groundwater levels and droughts caused by extreme climate events, we ~~select~~ selected a meteorologically wet month (August 2002) and a meteorologically dry month (August 2003), and show the ~~groundwater heads variation in these months. Figure~~ corresponding variations of groundwater heads in Figures 12e and ~~Figure12fshow two scenes of groundwater head variations in wet season and dry season, respectively~~12f. In general, the groundwater heads in the wet season are higher than

the long term mean values (Figure 12e). The variation of groundwater heads in the dry season, however, shows a strong spatial variability. Such a strong spatial variability of groundwater heads variation has also been reported ~~, e.g.,~~ by Kumar et al. (2016).

5 Discussion and conclusions

Our simulation results demonstrate that the coupled model mHM#OGS v1.0 can ~~reproduce groundwater heads~~ generally reproduce groundwater-head dynamics very well~~in general~~. It is also able to reasonably reproduce fluctuation amplitudes of groundwater heads, although with less accuracy. Compared to the good predictive capability of capturing the general trend behavior, the amplitude of head time series is hard to reproduce. This might be ~~due to the fact that the~~ because local geological formations in the vicinity of monitoring wells ~~(e.g, small apertures and rock cracks)~~ may significantly alter local groundwater flow behavior, and thus further affect groundwater head fluctuations.

The results of this study demonstrate the successful application of the well-established ~~mHM in estimating spatial~~ hydrologic model, mHM, in estimating spatially heterogeneous groundwater recharge and baseflow at a regional scale. ~~In the~~ At a spatial scale of 10^3 km^2 (the scale in this study), the distributed recharge estimated by mHM ~~shows its priority over the reference~~ is clearly superior to using homogeneous recharge. ~~The~~ mHM has been successfully applied at ~~a larger scale over Europe (Thober et al., 2015; Kumar et al., 2013b; Zink et al., 2017; Rakovec et al., 2016b)~~the continental scale covering entire Europe (Thober et al., 2015; Kumar et al., 2013b; Rakovec et al., 2016b; Zink et al., 2017). The successful application of the coupled model in this study suggests a huge potential ~~of~~ for extending the applicability of mHM#OGS v1.0 to a ~~larger scale~~ larger-scale (e.g., 10^4 - 10^6 km^2) or even a global scale.

The results of this study demonstrate a viable strategy for improving classic meso- to large-scale distributed hydrologic models, such as the current version of mHM ~~(Samaniego et al., 2010; Kumar et al., 2013b)~~(Samaniego et al., 2010; Kumar et al., 2013b), VIC (Liang et al., 1994), PCR-GLOBWB (Van Beek and Bierkens, 2009), WASMOD-M (Widén-Nilsson et al., 2007). ~~Those~~ These distributed hydrologic models do not ~~include the possibility of calculating~~ calculate spatio-temporal groundwater heads and are therefore ~~not able~~ unable to represent groundwater head ~~and storage~~ dynamics in their groundwater compartment. The physical representation of groundwater flow is, however, relevant in future regional-scale and possibly global hydrologic models to accurately determine travel times, solute export from catchments, and water quality in rivers (Botter et al., 2010; Benettin et al., 2015; Van Meter et al., 2017). The coupled model mHM#OGS v1.0 also ~~provides a potential in predicting groundwater flood and~~ offers the potential for predicting groundwater drought in analyzing the dynamic behavior of groundwater heads. Thus, it could be a useful tool for understanding groundwater anomalies under extreme climate conditions (Kumar et al., 2016; Marx et al., 2017).

For example, building on previous work ~~by~~ of Heße et al. (2017), who calculated Travel Time Distributions (TTDs) using mHM, we can now expand the range of their work to the complete ~~hydrologic cycle beneath atmosphere~~critical zone, which is important for comprehensively understanding particle (e.g., pollutant) transport behavior and the historical legacy in soil zone and groundwater storage ~~(Beniston et al., 2014; Basu et al., 2010).~~ (Basu et al., 2010; Beniston et al., 2014). mHM#OGS v1.0 fits well with the long-term simulation of nitrogen transport in the terrestrial water cycle~~based on the high-reputation~~

~~of two modeling codes in each other's fields~~. The coupled model is also able to evaluate surface water and groundwater storage ~~change~~ changes under different meteorological ~~forcings~~forcing conditions, which allows the comprehensive evaluation of hydrologic response to climate ~~change~~ changes (e.g., global warming). ~~Besides, the versatility of OGS also offers the possibility to address Thermo-Hydro-Mechanical-Chemical~~ Additionally, OGS demonstrates its capability in addressing thermo-hydro-mechanical-chemical (THMC) coupling processes in large-scale hydrologic cycles (not reflected in this study), which is significant for a wide range of real-world applications, including ~~land subsidence, agricultural irrigation,~~ nutrient circulation, salt water intrusion, drought, and heavy metal transport (Kalbacher et al., 2012; Selle et al., 2013; Walther et al., 2014, 2017).

In addition to improving the predictive abilities of mHM, we ~~could~~ can also demonstrate some improvements for the groundwater model OGS. Our results showed a modest improvement using mHM generated recharge compared to a simpler, uniform recharge rate. We currently gain a strong advantage for the description of the top boundary condition, i.e. the recharge, which is temporal and spatially variable through the input of mHM. Even more, the recharge fluxes provided are based on mHM~~'~~'s phenomenological process description, which significantly better describes the surface level recharge fluxes than common approaches through recharge rates derived by empirical relations derived recharge rates.~~In the future, we will additionally advance in the description of water fluxes between surface and groundwater compartments through the coupled feedback between both simulation tools.~~

~~For~~ In this study, we ~~focus~~ have focused our efforts on extending the ~~mHM applicability in~~ applicability of mHM from surface hydrology to subsurface hydrology ~~in~~ by a simple one-way coupling. Consequently, we do not account for any feedback between river and groundwater head fluctuations. While being a simplification of reality, this approach has certain advantages. First, the one-way coupling can be regarded as a conservative approach, such that the parametrization process, which is one of the most ~~salient~~ significant features of mHM, remains ~~fully intact. That~~ intact. In this way, we do not compromise any of its well-established features, such as the calibration of model parameters at different scales and ~~good runoff predictionability, while getting~~ accurate runoff prediction, while obtaining in addition very good estimates of groundwater ~~storage~~heads, flow paths, and travel times. The ~~lack~~ inability of mHM to provide good estimates for these quantities has been noted in the past (see, e.g., ~~Heße et al. (2017); Rakovec et al. (2016b)) and extends therefore~~ Rakovec et al. (2016b); Heße et al. (2017)) and our work therefore extends the predictive abilities of mHM. Second, using such a one-way coupling will allow users of mHM to simply extend currently established catchment models and ~~extend~~ enhance their abilities in the aforementioned way. Using a more sophisticated two-way coupling, would ~~mean that user would have to re-establish these models almost from scratch~~entail users having to rebuild their models almost entirely. Third, ~~even in the future,~~ a one-way coupling ~~would allow to easily expand~~ allows for ready future expansion of the predictive power of ~~a~~ an mHM catchment model~~if the practitioners later decide to do so, therefore leaving the option open~~, should the need arise. Finally, one-way coupling takes less computational consumption and achieves better numerical stability than two-way coupling. In short, unlike ~~a~~ two-way coupling, the one-way coupling described here allows the user to expand the abilities of mHM without sacrificing any of its well-known and well-established properties. However, in a next step, we will ~~try~~ devote to incorporate a full, two-way coupling using the next version of the mHM#OGS model. ~~Via such a full coupling scheme,~~ The main limitation of one-way coupling is that the

effects of a shallow depth to groundwater on actual ET, maintained by lateral groundwater flow, cannot be explicitly addressed. However, the dynamic interactions between overland flow and groundwater flow, as well as between soil moisture dynamics and groundwater dynamics can be explicitly modeled and investigated using a full coupling scheme. This approach is open to a broader spectrum of calibration options, such as calibration ~~using~~ by remotely sensed soil moisture data.

In conclusion, we can state that the coupled model mHM#OGS v1.0 ~~fully preserves~~ retains the predictive capability of ~~discharge of mHM~~ mHM for discharge volumes. In addition, it ~~proves to be~~ is capable of reproducing groundwater head dynamics. The simulation results ~~show a promising prediction abilityvia~~ indicate a promising predictive ability, confirmed by calibration and comparison to observed discharge and groundwater heads. Based on the historical match of discharge and groundwater heads in the case study, we ~~would~~ conclude that the coupled model mHM#OGS v1.0 is a valuable tool ~~in coping with~~ for addressing many challenging problems in the field of water management, including pollutant transport and legacy, climate change, and groundwater ~~flood and~~ drought.

Code and data availability. The mesoscale Hydrologic Model mHM (current release: 5.7) is an open-source community software and can be accessed from several mirrored repositories: SVN: http://www.ufz.de/index.php?en=40114; GitLab: https://git.ufz.de/mhm; GitHub: https://github.com/mhm-ufz. The modified source code of OGS5 can be freely acquired via the following link: https://github.com/UFZ-MJ/OGS_mHM.git. The model interface GIS2FEM and RIV2FEM can be freely acquired via the following link: https://github.com/UFZ-MJ/OGS_mHM/tree/master/UTL/GIS2FEM and https://github.com/UFZ-MJ/OGS_mHM/tree/master/RIV2FEM.

The input files of the case study in Nägelstedt catchment can be found in the Github repository: https://github.com/UFZ-MJ/OGS_mHM/tree/master/test_case. The dataset used in the case study can be found in the Github repository: https://github.com/UFZ-MJ/OGS_mHM/tree/master/data.

Acknowledgements. This research received funding from the Deutsche Forschungsgemeinschaft via Sonderforschungs- bereich CRC 1076 AquaDiva. We kindly thank Sabine Sattler from Thuringian State office for the Environment and Geology (TLUG) for providing the geological data. We kindly thank our data providers: the German Weather Service (DWD), the Joint Research Center of the European Commission, the Federal Institute for Geosciences and Natural Resources (BGR), the European Environmental Agency (EEA), the European Water Archive (EWA), and the Global Runoff Data Centre (GRDC).

References

Arnold, S., Attinger, S., Frank, K., and Hildebrandt, A.: Uncertainty in parameterisation and model structure affect simulation results in coupled ecohydrological models, Hydrology and Earth System Sciences, 13, 1789, 2009.

Azizian, M., Boano, F., Cook, P. L. M., Detwiler, R. L., Rippy, M. A., and Grant, S. B.: Ambient groundwater flow diminishes nitrate processing in the hyporheic zone of streams, Water Resources Research, 53, 3941–3967, doi:10.1002/2016WR020048, http://dx.doi.org/10.1002/2016WR020048, 2017.

Basu, N. B., Destouni, G., Jawitz, J. W., Thompson, S. E., Loukinova, N. V., Darracq, A., Zanardo, S., Yaeger, M., Sivapalan, M., Rinaldo, A., et al.: Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity, Geophysical Research Letters, 37, 2010.

Benettin, P., Kirchner, J. W., Rinaldo, A., and Botter, G.: Modeling chloride transport using travel time distributions at Plynlimon, Wales, Water Resources Research, pp. 3259–3276, doi:10.1002/2014WR016600, 2015.

Benettin, P., Soulsby, C., Birkel, C., Tetzlaff, D., Botter, G., and Rinaldo, A.: Using SAS functions and high-resolution isotope data to unravel travel time distributions in headwater catchments, Water Resources Research, 53, 1864–1878, doi:10.1002/2016WR020117, http://dx.doi.org/10.1002/2016WR020117, 2017.

Beniston, J. W., DuPont, S. T., Glover, J. D., Lal, R., and Dungait, J. A.: Soil organic carbon dynamics 75 years after land-use change in perennial grassland and annual wheat agricultural systems, Biogeochemistry, 120, 37–49, 2014.

Beven, K., Kirkby, M., Schofield, N., and Tagg, A.: Testing a physically-based flood forecasting model (TOPMODEL) for three UK catchments, Journal of Hydrology, 69, 119–143, 1984.

Beyer, C., Bauer, S., and Kolditz, O.: Uncertainty assessment of contaminant plume length estimates in heterogeneous aquifers, Journal of contaminant hydrology, 87, 73–95, 2006.

Botter, G., Bertuzzo, E., and Rinaldo, A.: Transport in the hydrologic response: Travel time distributions, soil moisture dynamics, and the old water paradox, Water Resources Research, 46, 1–18, doi:10.1029/2009WR008371, 2010.

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, 2010.

Chen, X. and Hu, Q.: Groundwater influences on soil moisture and surface evaporation, Journal of Hydrology, 297, 285–300, 2004.

Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., Hooper, R. P., Kumar, M., Leung, L. R., Mackay, D. S., Maxwell, R. M., Shen, C., Swenson, S. C., and Zeng, X.: Improving the representation of hydrologic processes in Earth System Models, Water Resources Research, 51, 5929–5956, doi:10.1002/2015WR017096, http://dx.doi.org/10.1002/2015WR017096, 2015.

Cuthbert, M., Mackay, R., and Nimmo, J.: Linking soil moisture balance and source-responsive models to estimate diffuse and preferential components of groundwater recharge, Hydrology and Earth System Sciences, 17, 1003–1019, 2013.

Dagan, G.: Flow and transport in porous formations, Springer Science & Business Media, 2012.

Danskin, W. R.: Evaluation of the hydrologic system and selected water-management alternatives in the Owens Valley, California, vol. 2370, US Department of the Interior, US Geological Survey, 1999.

Delfs, J. O., Blumensaat, F., Wang, W., Krebs, P., and Kolditz, O.: Coupling hydrogeological with surface runoff model in a Poltva case study in Western Ukraine, Environmental Earth Sciences, 65, 1439–1457, doi:10.1007/s12665-011-1285-4, 2012.

Diersch, H.-J.: FEFLOW: finite element modeling of flow, mass and heat transport in porous and fractured media, Springer Science & Business Media, 2013.

Doherty, J. et al.: PEST: a unique computer program for model-independent parameter optimisation, Water Down Under 94: Groundwater/Surface Hydrology Common Interest Papers; Preprints of Papers, p. 551, 1994.

Fang, K. and Shen, C.: Full-flow-regime storage-streamflow correlation patterns provide insights into hydrologic functioning over the continental US, Water Resources Research, pp. 1–20, doi:10.1002/2016WR020283, http://doi.wiley.com/10.1002/2016WR020283, 2017.

Ferguson, I. M., Jefferson, J. L., Maxwell, R. M., and Kollet, S. J.: Effects of root water uptake formulation on simulated water and energy budgets at local and basin scales, Environmental Earth Sciences, 75, 1–15, 2016.

Fischer, T., Naumov, D., Sattler, S., Kolditz, O., and Walther, M.: GO2OGS 1.0: A versatile workflow to integrate complex geological information with fault data into numerical simulation models, Geoscientific Model Development, 8, 3681–3694, doi:10.5194/gmd-8-3681-2015, 2015.

Goderniaux, P., Brouyère, S., Fowler, H. J., Blenkinsop, S., Therrien, R., Orban, P., and Dassargues, A.: Large scale surface-subsurface hydrological model to assess climate change impacts on groundwater reserves, Journal of Hydrology, 373, 122–138, doi:10.1016/j.jhydrol.2009.04.017, http://dx.doi.org/10.1016/j.jhydrol.2009.04.017, 2009.

Gräbe, A., Rödiger, T., Rink, K., Fischer, T., Sun, F., Wang, W., Siebert, C., and Kolditz, O.: Numerical analysis of the groundwater regime in the western Dead Sea escarpment, Israel+ West Bank, Environmental earth sciences, 69, 571–585, 2013.

Graham, D. N. and Butts, M. B.: Flexible, integrated watershed modelling with MIKE SHE, Watershed models, 849336090, 245–272, 2005.

Green, T. R., Taniguchi, M., Kooi, H., Gurdak, J. J., Allen, D. M., Hiscock, K. M., Treidel, H., and Aureli, A.: Beneath the surface of global change: Impacts of climate change on groundwater, Journal of Hydrology, 405, 532–560, 2011.

Gulden, L. E., Rosero, E., Yang, Z. L., Rodell, M., Jackson, C. S., Niu, G. Y., Yeh, P. J. F., and Famiglietti, J.: Improving land-surface model hydrology: Is an explicit aquifer model better than a deeper soil profile?, Geophysical Research Letters, 34, 1–5, doi:10.1029/2007GL029804, 2007.

Harbaugh, B. A. W., Banta, E. R., Hill, M. C., and Mcdonald, M. G.: MODFLOW-2000 , The U .S . Geological Survey modular graound-water model — User guide to modularization concepts and the ground-water flow process, U.S. Geological Survey, p. 130, http://www.gama-geo.hu/kb/download/ofr00-92.pdf, 2000.

Hargreaves, G. H. and Samani, Z. A.: Reference crop evapotranspiration from temperature, Applied engineering in agriculture, 1, 96–99, 1985.

He, W., Beyer, C., Fleckenstein, J. H., Jang, E., Kolditz, O., Naumov, D., and Kalbacher, T.: A parallelization scheme to simulate reactive transport in the subsurface environment with OGS#IPhreeqc 5.5.7-3.1.2, Geoscientific Model Development, 8, 3333–3348, doi:10.5194/gmd-8-3333-2015, 2015.

Heße, F., Zink, M., Kumar, R., Samaniego, L., and Attinger, S.: Spatially distributed characterization of soil-moisture dynamics using travel-time distributions, Hydrology and Earth System Sciences, 21, 549–570, doi:10.5194/hess-21-549-2017, http://www.hydrol-earth-syst-sci.net/21/549/2017/, 2017.

Huang, S., Kumar, R., Flörke, M., Yang, T., Hundecha, Y., Kraft, P., Gao, C., Gelfan, A., Liersch, S., Lobanova, A., Strauch, M., van Ogtrop, F., Reinhardt, J., Haberlandt, U., and Krysanova, V.: Evaluation of an ensemble of regional hydrological models in 12 large-scale river basins worldwide, Climatic Change, 141, 381–397, doi:10.1007/s10584-016-1841-8, https://doi.org/10.1007/s10584-016-1841-8, 2017.

Hunt, R. J., Walker, J. F., Selbig, W. R., Westenbroek, S. M., and Regan, R. S.: Simulation of Climate - Change effects on streamflow, Lake water budgets, and stream temperature using GSFLOW and SNTEMP, Trout Lake Watershed, Wisconsin, USGS Scientific Investigations Report., pp. 2013–5159, 2013.

Huntington, J. L. and Niswonger, R. G.: Role of surface-water and groundwater interactions on projected summertime streamflow in snow dominated regions: An integrated modeling approach, Water Resources Research, 48, 1–20, doi:10.1029/2012WR012319, 2012.

Hwang, H. T., Park, Y. J., Sudicky, E. A., and Forsyth, P. A.: A parallel computational framework to solve flow and transport in integrated surface-subsurface hydrologic systems, Environmental Modelling and Software, 61, 39–58, doi:10.1016/j.envsoft.2014.06.024, http://dx.doi.org/10.1016/j.envsoft.2014.06.024, 2014.

Ivano, V. Y., Vivoni, E. R., Bras, R. L., and Entekhabi, D.: Catchment hydrologic response with a fully distributed triangulated irregular network model, Water Resources Research, 40, n/a—-n/a, doi:10.1029/2004WR003218, http://dx.doi.org/10.1029/2004WR003218, 2004.

Kalbacher, T., Delfs, J. O., Shao, H., Wang, W., Walther, M., Samaniego, L., Schneider, C., Kumar, R., Musolff, A., Centler, F., Sun, F., Hildebrandt, A., Liedl, R., Borchardt, D., Krebs, P., and Kolditz, O.: The IWAS-ToolBox: Software coupling for an integrated water resources management, Environmental Earth Sciences, 65, 1367–1380, doi:10.1007/s12665-011-1270-y, 2012.

Kohlhepp, B., Lehmann, R., Seeber, P., Küsel, K., Trumbore, S. E., and Totsche, K. U.: Aquifer configuration and geostructural links control the groundwater quality in thin-bedded carbonate–siliciclastic alternations of the Hainich CZE, central Germany, Hydrology and Earth System Sciences, 21, 6091–6116, doi:10.5194/hess-21-6091-2017, https://www.hydrol-earth-syst-sci.net/21/6091/2017/, 2017.

Koirala, S., Yeh, P. J.-F., Hirabayashi, Y., Kanae, S., and Oki, T.: Global-scale land surface hydrologic modeling with the representation of water table dynamics, Journal of Geophysical Research: Atmospheres, 119, 75–89, 2014.

Kolditz, O., Bauer, S., Bilke, L., Böttcher, N., Delfs, J. O., Fischer, T., Görke, U. J., Kalbacher, T., Kosakowski, G., McDermott, C. I., Park, C. H., Radu, F., Rink, K., Shao, H., Shao, H. B., Sun, F., Sun, Y. Y., Singh, A. K., Taron, J., Walther, M., Wang, W., Watanabe, N., Wu, Y., Xie, M., Xu, W., and Zehner, B.: OpenGeoSys: an open-source initiative for numerical simulation of thermo-hydro-mechanical/chemical (THM/C) processes in porous media, Environmental Earth Sciences, 67, 589–599, doi:10.1007/s12665-012-1546-x, http://link.springer.com/10.1007/s12665-012-1546-x, 2012.

Kolditz, O., Shao, H., Wang, W., and Bauer, S.: Thermo-hydro-mechanical-chemical processes in fractured porous media: modelling and benchmarking, Springer, 2016.

Koren, V., Reed, S., Smith, M., Zhang, Z., and Seo, D.-J.: Hydrology laboratory research modeling system (HL-RMS) of the US national weather service, Journal of Hydrology, 291, 297–318, 2004.

Kumar, M., Duffy, C. J., and Salvage, K. M.: A Second-Order Accurate, Finite Volume–Based, Integrated Hydrologic Modeling (FIHM) Framework for Simulation of Surface and Subsurface Flow, Vadose Zone Journal, 8, 873, doi:10.2136/vzj2009.0014, https://www.soils.org/publications/vzj/abstracts/8/4/873, 2009.

Kumar, R., Livneh, B., and Samaniego, L.: Toward computationally efficient large-scale hydrologic predictions with a multiscale regionalization scheme, Water Resources Research, 49, 5700–5714, doi:10.1002/wrcr.20431, 2013a.

Kumar, R., Samaniego, L., and Attinger, S.: Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations, Water Resources Research, 49, 360–379, 2013b.

Kumar, R., Musuuza, J. L., Van Loon, A. F., Teuling, A. J., Barthel, R., Ten Broek, J., Mai, J., Samaniego, L., and Attinger, S.: Multiscale evaluation of the Standardized Precipitation Index as a groundwater drought indicator, Hydrology and Earth System Sciences, 20, 1117–1131, doi:10.5194/hess-20-1117-2016, 2016.

Küsel, K., Totsche, K. U., Trumbore, S. E., Lehmann, R., Steinhäuser, C., and Herrmann, M.: How deep can surface signals be traced in the critical zone? Merging biodiversity with biogeochemistry research in a central German Muschelkalk landscape, Frontiers in Earth Science, 4, 32, 2016.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, Journal of Geophysical Research: Atmospheres, 99, 14 415–14 428, 1994.

Liang, X., Xie, Z., and Huang, M.: A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model, Journal of Geophysical Research, 108, 8613–8629, doi:10.1029/2002JD003090, http://www.agu.org/pubs/crossref/2003/2002JD003090.shtml, 2003.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, Journal of hydrology, 201, 272–288, 1997.

Markstrom, S. L., Niswonger, R. G., Regan, R. S., Prudic, D. E., and Barlow, P. M.: GSFLOW—Coupled Ground-Water and Surface-Water Flow Model Based on the Integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005), U.S. Geological Survey, p. 240, http://pubs.er.usgs.gov/publication/tm6D1, 2008.

Marx, A., Kumar, R., Thober, S., Zink, M., Wanders, N., Wood, E. F., Ming, P., Sheffield, J., and Samaniego, L.: Climate change alters low flows in Europe under a 1.5, 2, and 3 degree global warming, Hydrology and Earth System Sciences Discussions, 2017, 1–24, doi:10.5194/hess-2017-485, https://www.hydrol-earth-syst-sci-discuss.net/hess-2017-485/, 2017.

Maxwell, R. M. and Miller, N. L.: Development of a coupled land surface and groundwater model, Journal of Hydrometeorology, 6, 233–247, doi:10.1175/JHM422.1, isi:000230393600001, 2005.

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, Geoscientific Model Development, 8, 923–937, http://www.geosci-model-dev.net/8/923/2015/, 2015.

McLachlan, P., Chambers, J., Uhlemann, S., and Binley, A.: Geophysical characterisation of the groundwater–surface water interface, Advances in Water Resources, 109, 302 – 319, doi:https://doi.org/10.1016/j.advwatres.2017.09.016, http://www.sciencedirect.com/science/article/pii/S0309170817304463, 2017.

Moore, C. and Doherty, J.: The cost of uniqueness in groundwater model calibration, Advances in Water Resources, 29, 605–623, doi:10.1016/j.advwatres.2005.07.003, 2006.

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., and Others: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, Journal of Geophysical Research: Atmospheres, 116, 2011.

Panday, S. and Huyakorn, P. S.: A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow, Advances in Water Resources, 27, 361–382, doi:https://doi.org/10.1016/j.advwatres.2004.02.016, http://www.sciencedirect.com/science/article/pii/S030917080400017X, 2004.

Paniconi, C. and Putti, M.: Physically based modeling in catchment hydrology at 50: Survey and outlook, Water Resources Research, 51, 7090–7129, doi:10.1002/2015WR017780, http://dx.doi.org/10.1002/2015WR017780, 2015.

Phi, S., Clarke, W., and Li, L.: Laboratory and numerical investigations of hillslope soil saturation development and runoff generation over rainfall events, Journal of Hydrology, 493, 1–15, doi:https://doi.org/10.1016/j.jhydrol.2013.04.009, http://www.sciencedirect.com/science/article/pii/S0022169413002813, 2013.

Qu, Y. and Duffy, C. J.: A semidiscrete finite volume formulation for multiprocess watershed simulation, Water Resources Research, 43, n/a—-n/a, doi:10.1029/2006WR005752, http://dx.doi.org/10.1029/2006WR005752, 2007.

Rakovec, O., Kumar, R., Attinger, S., and Samaniego, L.: Improving the realism of hydrologic model functioning through multivariate parameter estimation, Water Resources Research, 52, 7779–7792, 2016a.

Rakovec, O., Kumar, R., Mai, J., Cuntz, M., Thober, S., Zink, M., Attinger, S., Schäfer, D., Schrön, M., and Samaniego, L.: Multiscale and multivariate evaluation of water fluxes and states over European river basins, Journal of Hydrometeorology, 17, 287–307, 2016b.

Refsgaard, J. C. and Storm, B.: Mike SHE, Computer models of watershed hydrology, 1, 809–846, 1995.

Rigon, R., Bertoldi, G., and Over, T. M.: GEOtop: A Distributed Hydrological Model with Coupled Water and Energy Budgets, Journal of Hydrometeorology, 7, 371–388, doi:10.1175/JHM497.1, https://doi.org/10.1175/JHM497.1, 2006.

Rihani, J. F., Maxwell, R. M., and Chow, F. K.: Coupling groundwater and land surface processes: Idealized simulations to identify effects of terrain and subsurface heterogeneity on land surface energy fluxes, Water Resources Research, 46, 1–14, doi:10.1029/2010WR009111, 2010.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resources Research, 46, 2010.

Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., Eisner, S., Schmied, H. M., Sutanudjaja, E. H., Warrach-Sagi, K., et al.: Toward seamless hydrologic predictions across spatial scales, Hydrology and Earth System Sciences, 21, 4323, 2017.

Scibek, J. and Allen, D.: Modeled impacts of predicted climate change on recharge and groundwater levels, Water Resources Research, 42, 2006.

Seidel, G.: Geologie von Thüringen, Erdkunde, 58, 2004.

Selle, B., Rink, K., and Kolditz, O.: Recharge and discharge controls on groundwater travel times and flow paths to production wells for the Ammer catchment in southwestern Germany, Environmental Earth Sciences, 69, 443–452, doi:10.1007/s12665-013-2333-z, 2013.

Shao, H., Dmytrieva, S. V., Kolditz, O., Kulik, D. A., Pfingsten, W., and Kosakowski, G.: Modeling reactive transport in non-ideal aqueous–solid solution system, Applied Geochemistry, 24, 1287 – 1300, doi:https://doi.org/10.1016/j.apgeochem.2009.04.001, http://www.sciencedirect.com/science/article/pii/S0883292709000985, 2009.

Shao, H., Nagel, T., Roßkopf, C., Linder, M., Wörner, A., and Kolditz, O.: Non-equilibrium thermo-chemical heat storage in porous media: Part 2–A 1D computational model for a calcium hydroxide reaction system, Energy, 60, 271–282, 2013.

Shen, C. and Phanikumar, M. S.: A process-based, distributed hydrologic model based on a large-scale method for surface-subsurface coupling, Advances in Water Resources, 33, 1524–1541, doi:10.1016/j.advwatres.2010.09.002, http://dx.doi.org/10.1016/j.advwatres.2010.09.002, 2010.

Smerdon, B. D., Mendoza, C. A., and Devito, K. J.: Simulations of fully coupled lake-groundwater exchange in a subhumid climate with an integrated hydrologic model, Water Resources Research, 43, n/a–n/a, http://onlinelibrary.wiley.com/doi/10.1029/2006WR005137/full, 2007.

Spanoudaki, K., Stamou, A. I., and Nanou-Giannarou, A.: Development and verification of a 3-D integrated surface water–groundwater model, Journal of Hydrology, 375, 410–427, doi:https://doi.org/10.1016/j.jhydrol.2009.06.041, http://www.sciencedirect.com/science/article/pii/S0022169409003795, 2009.

Sun, F., Shao, H., Kalbacher, T., Wang, W., Yang, Z., Huang, Z., and Kolditz, O.: Groundwater drawdown at Nankou site of Beijing Plain: model development and calibration, Environmental Earth Sciences, 64, 1323–1333, doi:10.1007/s12665-011-0957-4, http://link.springer.com/10.1007/s12665-011-0957-4, 2011.

Sutanudjaja, E. H., Van Beek, L. P. H., De Jong, S. M., Van Geer, F. C., and Bierkens, M. F. P.: Large-scale groundwater modeling using global datasets: A test case for the Rhine-Meuse basin, Hydrology and Earth System Sciences, 15, 2913–2935, doi:10.5194/hess-15-2913-2011, 2011.

Sutanudjaja, E. H., Van Beek, L. P. H., De Jong, S. M., Van Geer, F. C., and Bierkens, M. F. P.: Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data, Water Resources Research, 50, 687–705, doi:10.1002/2013WR013807, 2014.

Therrien, R., McLaren, R. G., Sudicky, E. A., and Panday, S. M.: HydroGeoSphere: A three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport, Groundwater Simulations Group, University of Waterloo, Waterloo, ON, 2010.

Thober, S., Kumar, R., Sheffield, J., Mai, J., Schäfer, D., and Samaniego, L.: Seasonal Soil Moisture Drought Prediction over Europe Using the North American Multi-Model Ensemble (NMME), Journal of Hydrometeorology, 16, 2329–2344, doi:10.1175/JHM-D-15-0053.1, https://doi.org/10.1175/JHM-D-15-0053.1, 2015.

Van Beek, L. and Bierkens, M. F.: The global hydrological model PCR-GLOBWB: conceptualization, parameterization and verification, Utrecht University, Utrecht, The Netherlands, 2009.

Van Meter, K. J., Basu, N. B., and Van Cappellen, P.: Two centuries of nitrogen dynamics: Legacy sources and sinks in the Mississippi and Susquehanna River Basins, Global Biogeochemical Cycles, 31, 2–23, doi:10.1002/2016GB005498, 2017.

VanderKwaak, J. E. and Loague, K.: Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model, Water Resources Research, 37, 999–1013, doi:10.1029/2000WR900272, 2001.

Walther, M., Delfs, J.-O., Grundmann, J., Kolditz, O., and Liedl, R.: Saltwater intrusion modeling: Verification and application to an agricultural coastal arid region in Oman, Journal of Computational and Applied Mathematics, 236, 4798 – 4809, doi:http://dx.doi.org/10.1016/j.cam.2012.02.008, http://www.sciencedirect.com/science/article/pii/S0377042712000659, fEMTEC 2011: 3rd International Conference on Computational Methods in Engineering and Science, May 9–13, 2011, 2012.

Walther, M., Solpuker, U., Böttcher, N., Kolditz, O., Liedl, R., and Schwartz, F. W.: Description and verification of a novel flow and transport model for silicate-gel emplacement, Journal of contaminant hydrology, 157, 1–10, 2014.

Walther, M., Graf, T., Kolditz, O., Liedl, R., and Post, V.: How significant is the slope of the sea-side boundary for modelling seawater intrusion in coastal aquifers?, Journal of Hydrology, 551, 648 – 659, doi:https://doi.org/10.1016/j.jhydrol.2017.02.031, http://www.sciencedirect.com/science/article/pii/S0022169417301105, investigation of Coastal Aquifers, 2017.

Wang, W., Kosakowski, G., and Kolditz, O.: A parallel finite element scheme for thermo-hydro-mechanical (THM) coupled problems in porous media, Computers & Geosciences, 35, 1631–1641, doi:10.1016/j.cageo.2008.07.007, http://linkinghub.elsevier.com/retrieve/pii/S0098300409000065, 2009.

Wang, W., Kolditz, O., and Nagel, T.: Parallel finite element modelling of multi-physical processes in thermochemical energy storage devices, Applied Energy, 185, 1954–1964, 2017.

Wechsung, F.: Auswirkungen des globalen Wandels auf Wasser, Umwelt und Gesellschaft im Elbegebiet, vol. 6, Weißensee Verlag, 2005.

Weill, S., Mouche, E., and Patin, J.: A generalized Richards equation for surface / subsurface flow modelling, Journal of Hydrology, 366, 9–20, doi:10.1016/j.jhydrol.2008.12.007, http://dx.doi.org/10.1016/j.jhydrol.2008.12.007, 2009.

Widén-Nilsson, E., Halldin, S., and Xu, C.-y.: Global water-balance modelling with WASMOD-M: Parameter estimation and regionalisation, Journal of Hydrology, 340, 105–118, 2007.

Wood, E. F., Lettenmaier, D., Liang, X., Nijssen, B., and Wetzel, S. W.: Hydrological modeling of continental-scale basins, Annual Review of Earth and Planetary Sciences, 25, 279–300, 1997.

Yeh, P. J. and Eltahir, E. A.: Representation of water table dynamics in a land surface scheme. Part I: Model development, Journal of climate, 18, 1861–1880, 2005.

Zink, M., Kumar, R., Cuntz, M., and Samaniego, L.: A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty., Hydrology & Earth System Sciences, 21, 2017.

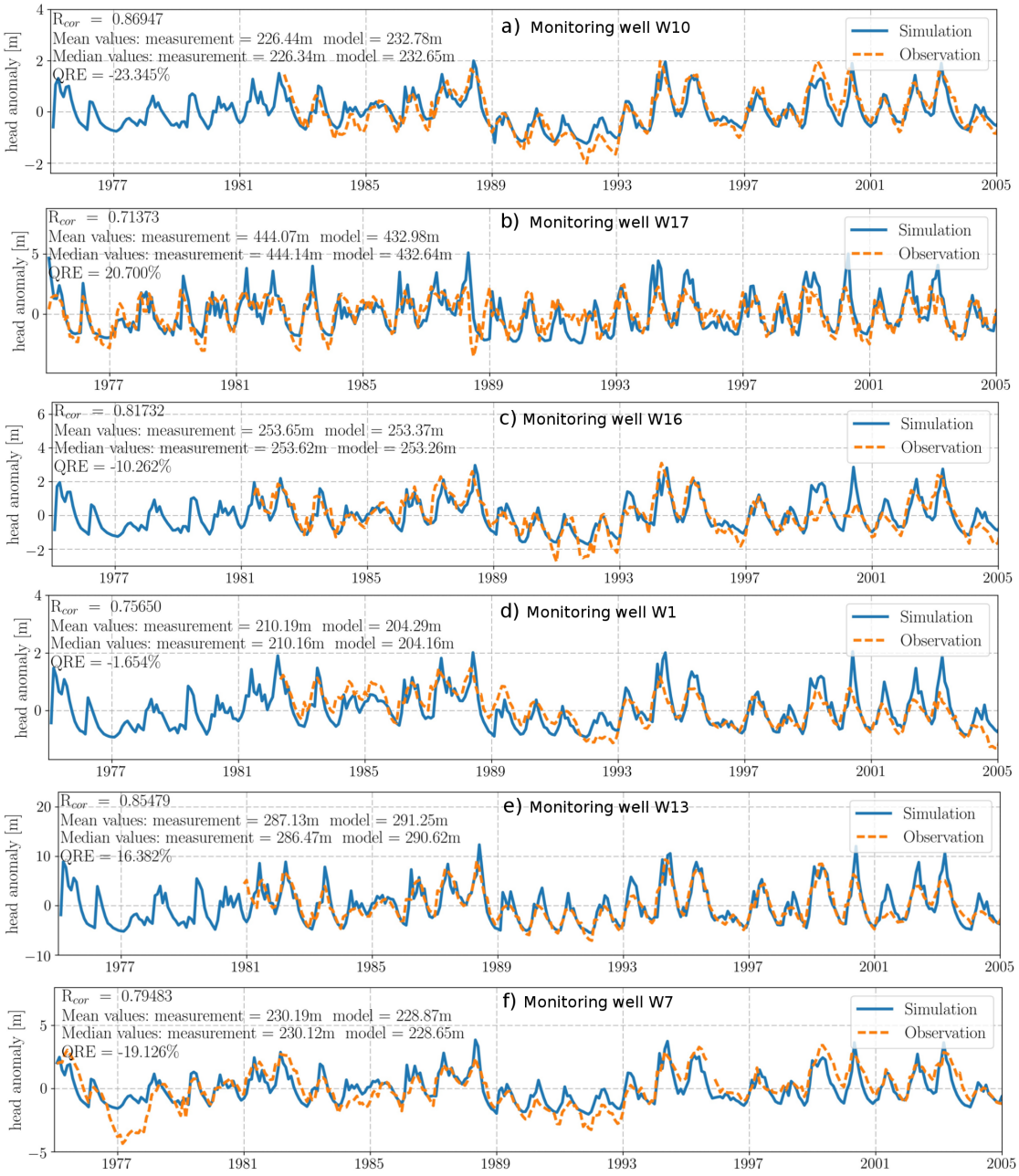

Figure 10. ~~The comparison~~ Comparison between ~~measurement data~~ measured (green dashed line) and ~~model output of~~ simulated groundwater head ~~anomaly~~ anomalies (blue solid line). (a) ~~Monitoring well 4728230786~~ W10 is located ~~at upland~~ in uplands, near a stream. (b) ~~Monitoring well 4628230773~~ W17 is located ~~at~~ in a mountainous area. (c) ~~Monitoring well 4728230781~~ W16 is located at a hillslope ~~at~~ in the northern ~~upland~~uplands. (d) ~~Monitoring well 4828230754~~ W1 is located ~~at lowland~~in the lowlands. (e) ~~Monitoring well 4728230783~~ W13 is located ~~at~~ in the northern ~~mountain~~mountains. (f) W7 is located in the uplands.

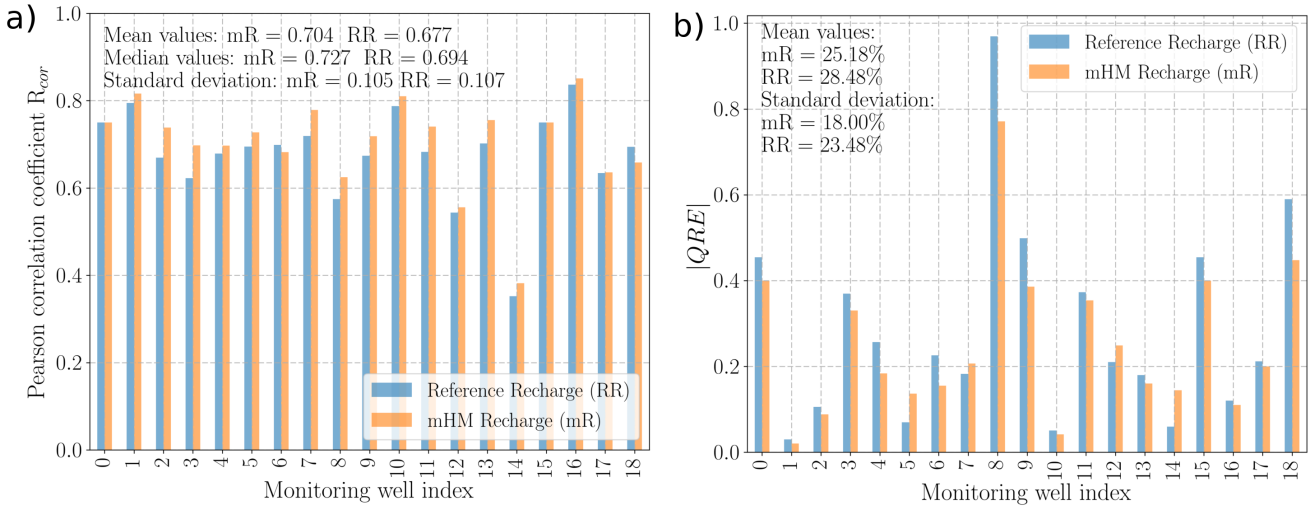

Figure 11. Barplots of a) the Pearson correlation coefficient R_{cor} and b) absolute ~~inter-quantile~~ inter-quartile range error |QRE| in all monitoring wells in two recharge scenarios.~~Each bar corresponds to an individual monitoring well in the following order: 0 - 4830230779, 1 - 4828230754, 2 - 4828230752, 3 - 4828230753, 4 - 4829230761, 5 - 4829230762, 6 - 4729230719, 7 - 4728230785, 8 - 4728230789, 9 - 4728230788, 10 - 4728230786, 11 - 4728230795, 12 - 4728230797, 13 - 4728230783, 14 - 4728230796, 15 - 4727230764, 16 - 4728230781, 17 - 4628230773, 18 - 4627230764.~~

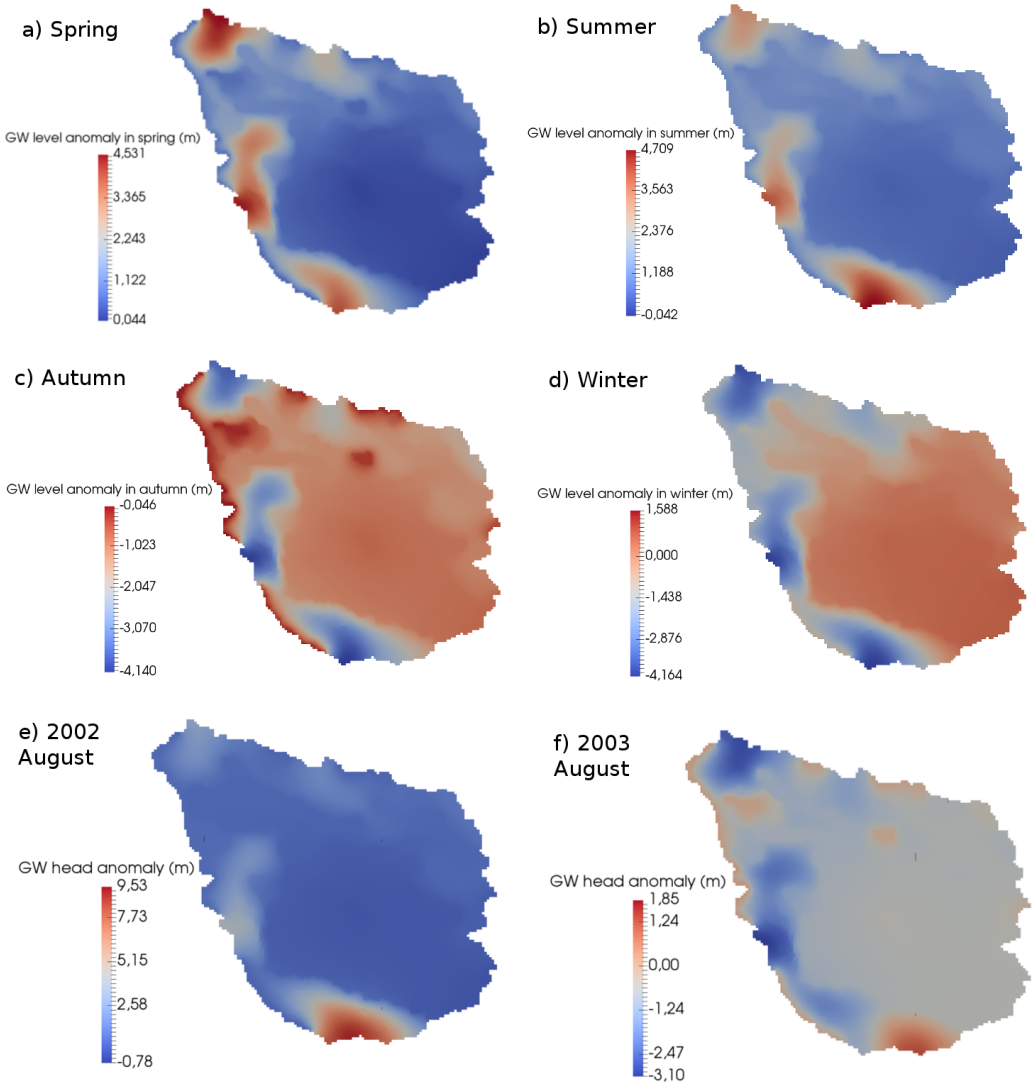

Figure 12. Seasonal variation of spatially-distributed groundwater heads by their anomalies after removing the long-term mean groundwater heads (unit: m). a) Long-term mean groundwater head distribution in spring; b) Long-term mean groundwater head distribution in summer; c) Long-term mean groundwater head distribution in autumn; d) Long-term mean groundwater head distribution in winter; e) Monthly mean groundwater head distribution in the wet season (August 2002); f) Monthly mean groundwater head distribution in the dry season (August 2003).