Responses to Anonymous Referee 1

We thank the referee reviewer very much for his comprehensive and insightful comments. Those comments are really helpful for us to revise the manuscript.

Before we reply to any specific questions of the comments, we would like to clarify two points.

The first point is that we would like to explain the linear groundwater reservoir in mHM, which was not directly explained in the manuscript since it has been included in the references (Kumar et al., 2013; Samaniego et al., 2010). mHM contains a linear reservoir to generate daily baseflow (please see Figure 1). The generated baseflow of each grid are further routed into streams using Muskingum-Cunge method. In the coupled model mHM#OGS, we take spatially distributed recharge and routed baseflow generated by the linear reservoir, then feed these two boundary sources to GIS2FEM (the coupling interface to convert unit and adjust time step), and then to OGS as upper boundary conditions. The baseflow is still calculated by the linear reservoir in mHM and routed into runoff (please see Figure 1). We have now noticed that the detailed explanation of the linear groundwater reservoir is essential and will include it into the revised manuscript.

The second point is that we are not aiming to develop a fully physically-based model. We are not aiming to study the mechanistic interaction of soil-zone processes and the groundwater heads. Instead, we are aiming to develop an open-source regional-scale model which can predict catchment runoff and groundwater head dynamics simultaneously, while preserves all existing and well-tested mHM features, e.g., the parameterization scheme (Kumar et al., 2013; Rakovec et al., 2016; Samaniego et al., 2010, 2017).

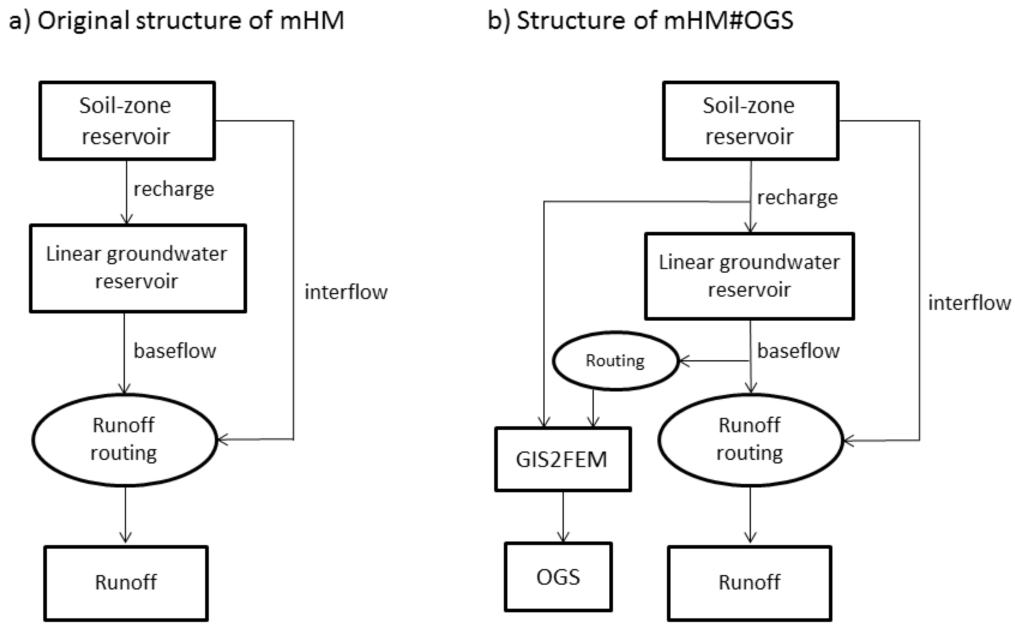

Figure 1 mHM#OGS as an approach into realization of spatially distributed groundwater head.

Major comments

1. The authors present a coupling approach for a land surface hydrologic and ground water flow model, mHM and OGS respectively. The manuscript contains sections on the coupling, model setup over a real catchment and verification of the results. The model coupling is not explained appropriately and it's not clear, whether the coupling approach satisfies the current state-of-the-art published in GMD. Based on the provided explanation, the results cannot be assessed unfortunately.

   Thank you for your comment. Enabling the reader to independently reproduce the results is an important aspect of the publishing process of GMD. To improve that part, we will significantly revise the model section in order to make our approach more clear to the reader and avoid misconception on our work. We will also provide a fully accessible code, a test example together with all needed data in the Github repository.

2. Introduction The introduction is incomplete and misses some of the most important and heavily cited references of integrated models and modeling studies of the terrestrial water cycle. Apparently the authors are not aware of the state-of-the-art. Proper citation of the mentioned models is missing. Is the sole goal of the introduction to promote the work of the co-authors (e.g. statement p 3, l 12-15 and citations throughout)?

Thank you for your insights. We will revise the whole introduction section accordingly and cite all the up-to-date papers properly.

To better convey these points and avoid possible future misunderstanding, we will revise the introduction section in manuscript accordingly. In addition, we further expand our literature review by properly referencing integrated surface/subsurface hydrologic models (ISSHMs) such as InHM (Smerdon et al., 2007; VanderKwaak and Loague, 2001), Parflow (Maxwell et al., 2015), tRIBS ((Ivanov 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), 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), in the revised manuscript. The coupled land surface / groundwater models (CLSGMs) include ParFlow-CLM (Ferguson and Maxwell, 2010; Maxwell and Kollet, 2008; Maxwell and Miller, 2005; Rihani et al., 2010), tRIBS + VEGGIE (Ivanov et al., 2008, 2010), SWAT and MODFLOW (Guzman et al., 2015; Kim et al., 2008), PCR-GLOBWB-MOD (Sutanudjaja et al., 2014), SWMM-OGS (Delfs et al., 2012). Nevertheless, we will revise this section in order to convey comprehensive information of the state-of-the-art science.

Next, we would like to clarify that the within the context of our manuscript a “coupled model” is not the same as a “physically-based” or “mechanistic” integrated model. We will include this point into the storyline in the revised manuscript. What we want to develop is a hybrid model that is using two different modeling paradigms which can be easily applied in regional and continental scale, rather than a mechanistic integrated model. Our reasons for this decision is that more conceptual process-based models like mHM or Noah-MP are good at predicting quantities like discharge but are highly conceptualized and there suffering from interpretability of certain processes (e.g., base flow and interflow components). More mechanistic models like Parflow and HydroGeoSphere are highly interpretable but show consistently worse performance when predicting runoff (Paniconi and Putti, 2015). To the best of our knowledge, the skill of simulating groundwater head dynamics at regional scale of mechanistic models are always neglected and seldom assessed by the data (e.g. GW head, tracer). At the larger scale, the assessment of modeled groundwater heads dynamics can only be found in very few publications (De Graaf et al., 2015; Sutanudjaja et al., 2011).

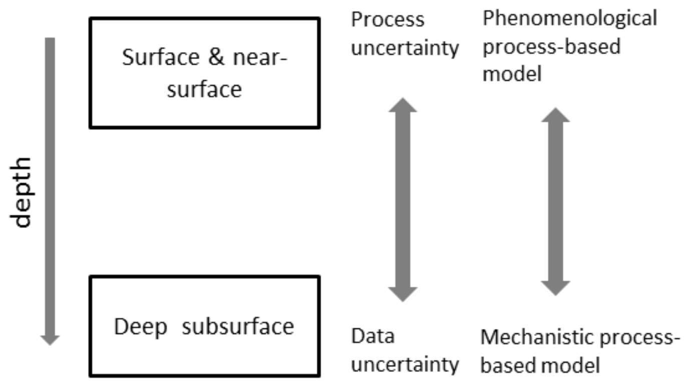

Figure 2 Different questions and challenges in surface and subsurface hydrology.

The above mentioned different abilities of more phenomenological models (e.g., mHM, Noah-MP, etc) vs. the more mechanistic models (e.g., Parflow, Hydrogeosphere, etc.) is caused by the different challenges that are posed by the different compartments of the terrestrial water cycle. One of the main challenges in the surface & near-surface storage is process uncertainty, with the fact that processes like ET, land use, land cover, snow pack, etc. are extremely complex. The process uncertainty decreases as it goes deeper and deeper into the subsurface storage. In subsurface storage, hydrological processes are under Darcy's law and therefore conceptually simpler. Meanwhile, the data uncertainty becomes more significant in deep subsurface storage than in shallow storage (see schematic in Figure 2). Therefore, proper conceptualization is needed in the shallow storage in order to deal with this process uncertainty (please see schematic in Figure 2). Owing to this point, mHM was developed as a bucket-type model to better deal with this process uncertainty by optimally leveraging the information content in the discharge data. On the other hand, OGS is a mechanistic model, i.e., it has a very low process uncertainty but large amount of data uncertainty. It is therefore optimally suited to model processes in the deeper subsurface. To use the strengths and weaknesses of both these modeling concepts, we decided to separate our modeling domain into these two compartments, a strategy that is very common in hydrology (Benettin et al., 2015; Bertuzzo et al., 2013; Botter et al., 2010; van der Velde et al., 2015), and use this different modeling paradigms for each compartment.

We will also add the following two paragraphs into the revised manuscript:

At the larger, i.e., regional scale, most of the mechanistic integrated models are based on a continuity of pressure and flux on the SW/GW interface, while the momentum balance condition is always missing (Paniconi and Putti, 2015). The runoff is generally normalized as "storage-dependent runoff" by solving Richards equation, and the grid-wise generated runoffs are routed by a routing algorithm. These models can principally simulate the dynamic interaction of different processes with SW/GW components, e.g., the interaction of soil moisture and GW head (Cuthbert et al., 2013; Maxwell and Miller, 2005; Rihani et al., 2010; Sutanudjaja et al., 2014), the storage-runoff correlation (Fang and Shen, 2017; Huntington and Niswonger, 2012; Koirala et al., 2014; Liang et al., 2003; VanderKwaak and Loague, 2001), and the dynamical interaction between ET and GW head (Chen and Hu, 2004; Koirala et al., 2014; Yeh and Eltahir, 2005).

In constract to that, in this study, we present a one-way coupling model mHM#OGS and focuses on the representation of Infiltration-Excess Recharge (IER) and Linear Baseflow (LB) through a case study of a mesoscale catchment. The basic scientific question we want to answer is: Can spatially distributed groundwater heads and their dynamics be reasonably captured by expanding on the abilities of a phenomenological model like mHM at the regional scale? Based on the case study, we would conclude that this expansion was successful since in addition to predicting discharge, our coupled model is also able to predict head measurements as well. Since our focus is the predictive accuracy of mHM (compared to interpretability and inference), we consider the physical plausibility of the coupling of recharge and baseflow to be a means to that end and not an end in itself. Improving the plausibility of these processes will, if done right, also lead to higher predictive power. We will elaborate on these points on more details in our answers to Comment 7.

Please check the updated Introduction section in the revised manuscript.

3. Model description Section 2.1 and 2.2 must be expanded. At least, the reader must get some idea about the basic principles that are used to model the different processes mentioned in passing, in order to assess the validity of the coupling. In section 2.3, figure 1b, suggests one-way coupling only i.e. mHm provides “groundwater recharge and base flow as boundary conditions to mHm” (p 3, l 16-17). Since mHm does not include groundwater, how can the calculation of these fluxes be mechanistic (p 3, l 15), because groundwater recharge strongly depends on the dynamics of the water table? Thus, the scarce information provided in this section in combination with the statements in the introduction are misleading to the reader.

This is an important observation by the reviewer, since these sections need to contain the relevant information to enable the reader to replicate our results. To address this current shortcoming, we will expand section 2.1 and 2.2 and make the description more clear. We would like to state our basic coupling principle as the following paragraph and add the two paragraphs into the revised manuscript:

The current mHM#OGS model is a one-way coupling model and focuses on the assessment of infiltration-excess recharge (IER) and linear baseflow (LB).

Considering the different equation systems of two models (ODEs in mHM and PDEs in OGS), the mechanical coupling that fully satisfy conservation of mass, energy and momentum is theoretically impossible. The one-way coupling method can guarantee conservation of mass and was used in this study.

We will also add Figure 1 and its corresponding explanation into the revised manuscript. We believe the readers will get a clear picture of our modeling approach in the revised manuscript.

Please check the updated Section 2.1 and 2.2 in the revised manuscript.

4. Section 2.3.2 with the title "Boundary condition-based coupling" provides the basic equations, yet leaves the reader wondering how the coupling is really done. Something is said about the exchange of fluxes via qe and qe' (p 7, l 3), but these are sources not boundary fluxes. What is equation 2? The upper boundary condition for the groundwater flow model? Shouldn't the coupling be performed via equation 2 as promised in the section title? In addition, the authors state that "the coupling interface converts time series of variables and fluxes to Neumann boundary conditions...". How does that fit in? This reader is left confused.

   Again, this is an important observation by the reviewer. We admit that the qe and qe' in equation 1 is redundant and will confuse the readers. It is the equation 2 that works to connect mHM and OGS. We will delete the equation 1 and revise section 2.3.2 carefully to make sure the "boundary condition-based coupling" is properly presented. With regard to the sentence "the coupling interface converts time series of variables and fluxes to Neumann boundary conditions...", it means that the boundary condition-based coupling is performs by interpolating recharge and baseflow in the interface GIS2FEM, e.g., from coarser grid size in mHM to the finer grid size in OGS.

   Please check the new Section 2.3 in the revised manuscript.

5. Figure 2 is not instructive. What is GIS2FEM doing? Interpolating? How does the coupling work in the vertical direction for each column? As I understand, mHm has a fixed column depth. Can the water table rise into the column along e.g. river corridors? And where does the baseflow go in OGS? How is groundwater storage in mHm (p 7, l 9-10) related to OGS? There is apparently no backward exchange with mHm due to baseflow and exchange with river networks, and no capillary rise. This reader is left confused.

   We appreciate this constructive criticism. Explaining this tool appropriately is indeed necessary for the understanding of the coupling procedure and must therefore not be omitted. GIS2FEM is the model interface which is used to interpolate recharge and baseflow between different grid sizes of two models (p 7, l 19-24).

The baseflow is not determined by OGS. Instead, it is determined by the linear reservoir in mHM and then routed into the streams (see Figure 1). The water table cannot rise into the column along river corridors because we use the linear groundwater storage in mHM to calculate baseflow. The linear reservoir is a simplified reservoir with an overall aim of predicting runoff, whereby the dynamic interaction with groundwater head is conceptualized and simplified in order to keep the robustness of parameterization scheme, which is a unique feature of mHM.

Please check the updated Figure 2 in the revised manuscript.

6. On p 7, l 17-18, what do the authors mean by conversion between volumetric flux, specific flux and water head? Where in the coupling is this conversion required and why does the cell sizes need to be adjusted (there is actual re-gridding going on)?

   Thank you for your questions. The conversion is in terms of unit conversion, e.g., from distributed recharge in mHM (m/s) to volumetric recharge in OGS (m^3/s). There is no re-gridding going on. The boundary fluxes are directly interpolated from mHM to OGS using the interface GIS2FEM.

7. From table 2 it appears that in the author's eyes, coupling and integrated modeling of the terrestrial water cycle simply means to pass groundwater recharge values from a 1D hydrologic land surface scheme to a steady state groundwater flow model and return a head value back as some lower (boundary) condition for the hydrologic scheme (not indicated in figure 1). I feel, in the geosciences, we moved beyond this type of approach quite some time ago.

   Thank you for your comments. We are, however, afraid, that some of the reviewer's comments here are at least in part based on a misunderstanding. We will modify table 2 accordingly so that the right information can be clearly conveyed. The reviewer said "pass groundwater recharge values from a 1D hydrologic land surface scheme to a steady state groundwater flow model and return a head value back as some lower (boundary) condition for the hydrologic scheme", which is unfortunately a misunderstanding. The modeling system is basically one-way pass, which means infiltration-excess recharge (IER) and linear baseflow (LB) are calculated by mHM alone, and then passed to OGS as an upper boundary condition to force the transient groundwater model (please see Figure 1).

   To better motivate this strategy, we would like to elaborate on this decision by continuing the discussion form Comment 2. As mentioned there, we are not aiming to develop a single, seamless, mechanistic, integrated model. Instead, we are trying to establish a "hybrid model" that bridges the gap between two distinct models and makes use of the best of their abilities (see also our answers to Comment 2). These two models have different paradigms and address different challenges; First mHM, which aims for a good prediction ability of discharge across multiple time scales as well as multiple spatial-scale catchments. All of it in a computationally efficient way by using ODE's for each compartment. Second, OGS which solves computationally-expensive PDEs that directly implement flow and transport processes by using modern tools like Finite Element Method (see schematic in Figure 2). In order to achieve a two-way coupling model, strong revisions to the implementation of these tools are necessary that will affect in particular the parametrization process of mHM.

The currently described one-way coupling can be seen as the intermediate move towards such a fully-coupled hybrid model. However, next to leading to such a more thorough coupling, the one-way coupling, described here, has a number of advantages that make it a viable modeling strategy in and of itself. First, the one-way coupling can be regarded as a safe or conservative approach, such that the parametrization process, which is one of its most salient features of mHM, remains fully intact. That way, we do not compromise any of its well-established features, such as calibration of model parameters at different scales and good runoff prediction ability, while getting in addition very good estimates of groundwater storage, flow paths and travel times. The lack of mHM to provide good estimates for these quantities has been noted in the past (see, e.g., Heße et al. 2016; Rakovec et al. 2016) and extends therefore 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 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. Third, even in the future, a one-way coupling would allow to easily expand the predictive power of a mHM catchment model if the practitioners later decide to do so, therefore leaving the option open. 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 (Kumar et al., 2013; Rakovec et al., 2016; Samaniego et al., 2010, 2017).

In addition to improving the predictive power of mHM, OGS is gaining 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 empirical relations derived recharge rates. In the future, we additionally plan to advance in the description of water fluxes between surface and groundwater compartments through the coupled feedback between both simulation tools.To further explain the motivation for the presented one-way coupling, we like to detail some relevant research questions that can now already be answered with our model; Kumar et al (2016) have demonstrated that the Standardized Precipitation Index (SPI) has a limited applicability and low reliability in characterizing groundwater drought. Our model can be a useful tool in predicting groundwater drought & flood under different climate conditions (please check Figure 11 and 13 in the referenced manuscript). Moreover, the coupled model can be used to quantify the catchment scale legacy nitrogen stores in groundwater reservoirs. Recent research shows that a large portion of legacy nitrogen can be older than 10 years (Van Meter et al., 2017). The current version of mHM#OGS fits well with the long-term simulation of nitrogen transport in terrestrial water cycle. The combination of process uncertainty at surface hydrology and data uncertainty at subsurface hydrology is challenging to understand travel time distributions (TTDs) at catchment scale (Benettin et al., 2015; Bertuzzo et al., 2013; Botter et al., 2010; van der Velde et al., 2015). The coupled model mHM#OGS is valuable at TTDs simulations based on the high-reputation of two modeling codes in each other's fields. In addition, field and modeling experiments at large scales suggest that the way bottom boundaries, bedrock interfaces, and other layers are treated will have a large impact on hydrological response (e.g., groundwater heads) (Broda et al., 2011; Buttle and McDonald, 2002; Ebel et al., 2008; Uchida et al., 2002, 2003).

Finally, we would like to conclude by saying that establishing a fully tow-way-coupled hybrid model, which also accounts for dynamic interaction of SW and GW, is a high priority. However, based on the challenges outlined above as well as the problem that such a model would sacrifice some of the predictive power of mHM (e.g., discharge), we consider the present coupling strategy a valuable and viable alternative in its own right, both for the meantime and the future.

We have revised the introduction section of the manuscript accordingly. Please check the illustration of coupling mechanism in page 7, line 12-15 and page8, line 1-23 of the revised manuscript.

8. The description of the study area and model setup, calibration etc. belong into a separate section.
   The results can not be assessed unfortunately, because of the poor explanation of the applied modeling and coupling techniques.

   We appreciate this observation. We have followed the reviewer's suggestion and separate this section into two sections in the revised version of the manuscript. We also provided the source code of the coupled system, the test case along with all needed data in the Github repository in order to facilitate all interested people.

9. Language and grammar require considerable improvement.

Thank you. We have thoroughly revised the manuscript and checked it with a native English speaker.

References

Ajami, H., McCabe, M. F., Evans, J. P. and Stisen, S.: Assessing the impact ofmodel spin-up on surface water-groundwater interactions using an integrated hydrologic model, Water Resour. Res., 50, 1 – 21, doi:10.1002/2013WR014258.Received, 2014.

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

Bertuzzo, E., Thomet, M., Botter, G. and Rinaldo, A.: Catchment-scale herbicides transport: Theory and application, Adv. Water Resour., 52, 232 – 242, doi:10.1016/j.advwatres.2012.11.007, 2013.

Botter, G., Bertuzzo, E. and Rinaldo, A.: Transport in the hydrologic response: Travel time distributions, soil moisture dynamics, and the old water paradox, Water Resour. Res., 46(3), 1–18, doi:10.1029/2009WR008371, 2010.

Broda, S., Paniconi, C. and Larocque, M.: Numerical investigation of leakage in sloping aquifers, J. Hydrol., 409(1), 49 – 61, doi:https://doi.org/10.1016/j.jhydrol.2011.07.035, 2011.

Buttle, J. M. and McDonald, D. J.: Coupled vertical and lateral preferential flow on a forested slope, Water Resour. Res., 38(5), 16 – 18, doi:10.1029/2001WR000773, 2002.

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 Resour. Res., 46(2), 2010.

Chen, X. and Hu, Q.: Groundwater influences on soil moisture and surface evaporation, J. Hydrol., 297(1), 285–300, 2004.

Cuthbert, M. O., MacKay, R. and Nimmo, J. R.: Linking soil moisture balance and source-responsive models to estimate diffuse and preferential components of groundwater recharge, Hydrol. Earth Syst. Sci., 17(3), 1003–1019, doi:10.5194/hess-17-1003-2013, 2013.

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, Environ. Earth Sci., 65(5), 1439–1457, doi:10.1007/s12665-011-1285-4, 2012.

Ebel, B. A., Loague, K., Montgomery, D. R. and Dietrich, W. E.: Physics-based continuous simulation of long-term near-surface hydrologic response for the Coos Bay experimental catchment, Water Resour. Res., 44(7), n/a--n/a, doi:10.1029/2007WR006442, 2008.

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

Ferguson, I. M. and Maxwell, R. M.: Role of groundwater in watershed response and land surface feedbacks under climate change, Water Resour. Res., 46(10), n/a--n/a, doi:10.1029/2009WR008616, 2010.

De Graaf, I. E. M., Sutanudjaja, E. H., Van Beek, L. P. H. and Bierkens, M. F. P.: A high-resolution global-scale groundwater model, Hydrol. Earth Syst. Sci., 19(2), 823–837, doi:10.5194/hess-19-823-2015, 2015.

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

Guzman, J. A., Moriasi, D. N., Gowda, P. H., Steiner, J. L., Starks, P. J., Arnold, J. G. and Srinivasan, R.: A model integration framework for linking SWAT and MODFLOW, Environ. Model. Softw., 73, 103 – 116, doi:10.1016/j.envsoft.2015.08.011, 2015.

Falk Heße, Matthias Zink, Rohini Kumar, Luis Samaniego, and Sabine Attinger: Spatially distributed characterization of soil-moisture dynamics using travel-time distributions. Hydrol. Earth Syst. Sci., 20, 1–22, 2016 doi:10.5194/hess-20-1-2016

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 Resour. Res., 48(11), 1–20, doi:10.1029/2012WR012319, 2012.

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 Sci. Investig. Report., 2013–5159, 2013.

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, Environ. Model. Softw., 61, 39 – 58, doi:10.1016/j.envsoft.2014.06.024, 2014.

Ivanov, V. Y., Bras, R. L. and Vivoni, E. R.: Vegetation-hydrology dynamics in complex terrain of semiarid areas: 1. A mechanistic approach to modeling dynamic feedbacks, Water Resour. Res., 44(3), 2008.

Ivanov, V. Y., Fatichi, S., Jenerette, G. D., Espeleta, J. F., Troch, P. A. and Huxman, T. E.: Hysteresis of soil moisture spatial heterogeneity and the “homogenizing” effect of vegetation, Water Resour. Res., 46(9), n/a--n/a, doi:10.1029/2009WR008611, 2010.

Kim, N. W., Chung, I. M., Won, Y. S. and Arnold, J. G.: Development and application of the integrated SWAT--MODFLOW model, J. Hydrol., 356(1), 1–16, 2008.

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, J. Geophys. Res., 119(1), 75–89, doi:10.1002/2013JD020398, 2014.

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 Zo. J., 8(4), 873–890, 2009.

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, Hydrol. Earth Syst. Sci., 20(3), 1117–1131, doi:10.5194/hess-20-1117-2016, 2016.

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, J. Geophys. Res., 108(Vic), 8613–8629, doi:10.1029/2002JD003090, 2003.

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. Geol. Surv., (Techniques and Methods 6-D1), 240 [online] Available from: http://pubs.er.usgs.gov/publication/tm6D1, 2008.

Maxwell, R. M. and Miller, N. L.: Development of a coupled land surface and groundwater model, J. Hydrometeorol., 6(3), 233–247, doi:10.1175/JHM422.1, 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, Geosci. Model Dev., 8(3), 923–937 [online] Available from: http://www.geosci-model-dev.net/8/923/2015/ (Accessed 3 February 2016), 2015.

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 Biogeochem. Cycles, 31(1), 2–23, doi:10.1002/2016GB005498, 2017.

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, J. Hydrometeorol., 17, 287–307, doi:10.1175/JHM-D-15-0054.1, 2016.

Panday, S. and Huyakorn, P. S.: A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow, Adv. Water Resour., 27(4), 361–382, doi:https://doi.org/10.1016/j.advwatres.2004.02.016, 2004.

Paniconi, C. and Putti, M.: Physically based modeling in catchment hydrology at 50: Survey and outlook, Water Resour. Res., 51(9), 7090 – 7129, doi: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, J. Hydrol., 493(Supplement C), 1–15, doi:https://doi.org/10.1016/j.jhydrol.2013.04.009, 2013.

Qu, Y. and Duffy, C. J.: A semidiscrete finite volume formulation for multiprocess watershed simulation, Water Resour. Res., 43(8), n/a--n/a, doi:10.1029/2006WR005752, 2007.

Rigon, R., Bertoldi, G. and Over, T. M.: GEOtop: A Distributed Hydrological Model with Coupled Water and Energy Budgets, J. Hydrometeorol., 7(3), 371–388, doi: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 Resour. Res., 46(12), 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 Resour. Res., 46(5), n/a-n/a, doi:10.1029/2008WR007327, 2010.

Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., Eisner, S., Schmied, H. M., Sutanudjaja, E. H., Warrach-Sagi, K. and others:

Toward seamless hydrologic predictions across spatial scales, Hydrol. Earth Syst. Sci., 21(9), 4323, 2017.

Shen, C. and Phanikumar, M. S.: A process-based, distributed hydrologic model based on a large-scale method for surface-subsurface coupling, Adv. Water Resour., 33(12), 1524–1541, doi: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 Resour. Res., 43(1), n/a-n/a [online] Available from: http://onlinelibrary.wiley.com/doi/10.1029/2006WR005137/full (Accessed 19 August 2016), 2007.

Spanoudaki, K., Stamou, A. I. and Nanou-Giannarou, A.: Development and verification of a 3-D integrated surface water–groundwater model, J. Hydrol., 375(3), 410–427, doi:https://doi.org/10.1016/j.jhydrol.2009.06.041, 2009.

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, Hydrol. Earth Syst. Sci., 15(9), 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 Resour. Res., 50(1), 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, Groundw. Simulations Group, Univ. Waterloo, Waterloo, 2010.

Uchida, T., Kosugi, K. and Mizuyama, T.: Effects of pipe flow and bedrock groundwater on runoff generation in a steep headwater catchment in Ashiu, central Japan, Water Resour. Res., 38(7), 14–24, doi:10.1029/2001WR000261, 2002.

Uchida, T., Asano, Y., Ohte, N. and Mizuyama, T.: Seepage area and rate of bedrock groundwater discharge at a granitic unchanneled hillslope, Water Resour. Res., 39(1), n/a--n/a, doi:10.1029/2002WR001298, 2003.

van der Velde, Y., Heidbüchel, I., Lyon, S. W., Nyberg, L., Rodhe, A., Bishop, K. and Troch, P. A.: Consequences of mixing assumptions for time-variable travel time distributions, Hydrol. Process., 29(16), 3460–3474, doi:10.1002/hyp.10372, 2015.

VanderKwaak, J. E. and Loague, K.: Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model, Water Resour. Res., 37(4), 999–1013, doi:10.1029/2000WR900272, 2001.

Weill, S., Mouche, E. and Patin, J.: A generalized Richards equation for surface/subsurface flow modelling, J. Hydrol., 366(1 – 4), 9 – 20, doi:10.1016/j.jhydrol.2008.12.007, 2009.

Yeh, P. J. F. and Eltahir, E. A. B.: Representation of water table dynamics in a land surface scheme. Part I: Model development, J. Clim., 18(12), 1861–1880, 2005.

Responses to Anonymous Referee #2

We thank the referee reviewer very much for reading our work and insightful comments. Those comments really let us know the unclear part of our manuscript and help us a lot to improve our manuscript.

Major comments

1. This is a poor paper. The two models the authors have used in their catchment simulation are not described in sufficient detail to enable a reader to understand how all the processes have been implemented. Particularly lacking is how the exchange fluxes are handled. This is surprising given that the focus of the paper is on model coupling (as stated in the title). In addition, the groundwater model is incomplete, as the authors do not describe how a water table is handled in the model, specifically the role of specific yield. This is a major omission, given the influence it has on water table dynamics, a key measure used to assess the model performance. Finally, there is no mention of river geometry (e.g. river width) and how water levels are converted into flows. Given this incompleteness, it is not possible to give any comment on the quality of the model simulations presented in the paper and the author need to address these details in any subsequent resubmission.

   Thank you for your comment. We agree with the reviewer that there are some omissions and over simplistic descriptions in the manuscript, such as the omission of specific yield, the river geometry and coupling mechanism, which are critical and should be revised accordingly. To improve the manuscript, we will significantly revise the manuscript in order to make our approach more clear to the reader and avoid misconception on our work. We will also provide a fully accessible code, a test example together with all needed data in the Github repository.

2. Eq1 refers solely to changes in pressure head being governed by the specific storage coefficient. However, this refers to changes in storage due to water and rock compressibility (see Freeze and Cherry, 1979) and, therefore, is primarily associated with storage change in confined aquifers. In an unconfined aquifer, which is the focus of the paper here, storage changes are largely governed by changes in the water table and the wetting and dewatering of pores, which is typically characterised by the specific yield. It is not clear from the description of model how this is handled. Furthermore, there is no reference to specific yield in the text and,

as this is an important parameter which has a major influence on groundwater dynamics, it's omission makes commenting on the model's performance rather difficult.

Thank you for your important insights. We fully agree that the difference between the specific yield in unconfined aquifer and specific storage in a confined aquifer is theoretically important, and should be addressed in the manuscript. In the current manuscript, we assign a uniform storativity value of 1.0e-5 to all geological layers, which could be specific storage for water table layer and specific storage for confined layer. For the sake of simplicity we did not distinguish those two different parameters. Nevertheless, we will follow the reviewer's suggestion and assign proper specific storage values to corresponding water table layers by referring to literature. We are re-calibrating the model using the specific yield in unconfined aquifers and specific storage in confined aquifers, and will add the updated model settings and performance assessment into the revised manuscript.

Thanks again for pointing out our omission, we really appreciate it. We have updated the equation system and restructured Section 2 in the revised manuscript. Please check the revised Section 2.

3. There are two fluxes qs and qe included in the groundwater continuity equation. Qs is defined as a specified rate source/sink. Presumably, this refers both to abstractions of water from wells as well as recharge from rainfall infiltration in contrast to the flux qe, which is defined as the exchange with surface water. Furthermore, in Eq3. the surface water continuity equation, a flux qe 0 is referred to as the exchange rate with surface water. It is not clear to me what are the differences between these two terms, mainly because, in both cases, no details are given on how these fluxes are calculated. This is particularly problematic, as a key feature of the paper (and referred to in the title) is the coupling between the surface and subsurface models. I would, therefore, have expected to see an equation that includes both p and s showing how the models are explicitly coupled.

Eq 1 and Eq 3 are governing equations of surface water flow and subsurface flow, respectively. The coupling procedures are illustrated as below:

First, the mHM grid cells are artificially classified as soil grid cells (please see brown part in Figure 4) and river grid cells (please see blue part in Figure 4). The classification method was also illustrated in the Section 2.4.3 of manuscript. $q_{s,}$ which represents recharge in the manuscript, is calculated by the water balance equation by removing fast interflow, slow interflow and evapotranspiration from precipitation. Meanwhile, q_e is calculated only at river grid cells. At river grid cells, q_e is calculated as follows:

First, baseflow is generated at every grid cell by a water balance equation combined with a linear groundwater reservoir. The released baseflow by linear groundwater reservoir (please see the detailed description of linear groundwater reservoir in next section) is then routed into the total runoff by means of a Muskingum-Cunge algorithm. The total amount of routed baseflow is then uniformly distributed to every river grid cells. The flux q_e, is equal to the uniformly distributed routed baseflow in each river grid cells.

Using the above scheme, the total water balance is closed because the total amount of groundwater recharge is equal to the total amount of routed baseflow.

The Eq 3 is approximated by using a Muskingum-Cunge method:

$$Q_i^1(t) = Q_i^1(t-1) + c_1\left(Q_i^0(t-1) - Q_i^1(t-1)\right) + c_2\left(Q_i^0(t) - Q_i^0(t-1)\right)$$

with

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

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

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

where

Q_i^0 and Q_i^1 denote the discharge entering and leaving the river reach located on cell i respectively.
$Q_{i'}$ is the contribution from the upstream cell *i'*.
κ is Muskingum travel time parameter.
ξ is Muskingum attenuation parameter.
Δt is time interval.
t Time index for each Δt interval.

The Muskingum parameters, and , are calibrated by matching the historical runoff.
To address this specific comment, we will add Muskingum-Cunge equation after Eq 3 as supplementary information.

Please check the revised Section 2.1 and 2.2.

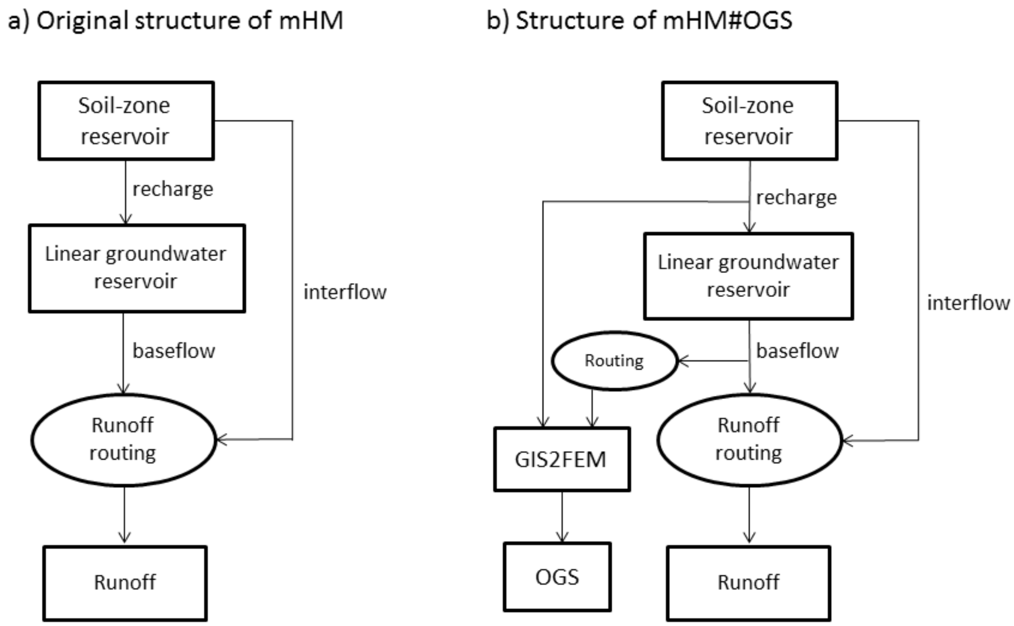

Figure 3 mHM#OGS as an approach into realization of groundwater head dynamics

4. The authors cite Camporese et al. (2010) in their discussion on the two coupling terms, however, there are some important differences. Camporese at al. (2010) do not appear to have an equivalent flux for qe. They have a term in their surface water balance equation that looks to be the equivalent of qe 0 (which they refer to as qs), however, even here the exact definition is not given. Furthermore, Camporese et al. (2010) solve Richards' equation, rather than the saturated groundwater flow equation, in their subsurface model and, therefore, where the water table is below the base of the river, the coupling would be completely different.

As mentioned in the former paragraph, we use boundary condition-based coupling by feeding recharge to the soil grid cells (please see brown part in Figure 4), or feeding distributed routed baseflow plus recharge for the river grid cells (please see blue part in Figure 4). The calculation of distributed routed baseflow has been described in the above section. We would like to add some additional information on linear groundwater reservoir in mHM.

mHM contains a linear reservoir to generate daily baseflow (please see Figure 3**Error! Reference source not found.**). The generated baseflow of each grid are further routed into streams using the Muskingum-Cunge method. In the coupled model mHM#OGS, we take spatially distributed recharge and routed baseflow generated by the linear reservoir, then feed these two boundary sources to GIS2FEM (the coupling interface to convert unit and adjust time step), and then to OGS as upper boundary conditions (see also our answer to Reviewer 1). The baseflow is still calculated by the linear reservoir in mHM and routed into runoff (please see Figure 1).

We have now noticed that the detailed explanation of the linear groundwater reservoir is essential, and have depicted it in the revised manuscript. Please check the Section 2.3 of the revised manuscript.

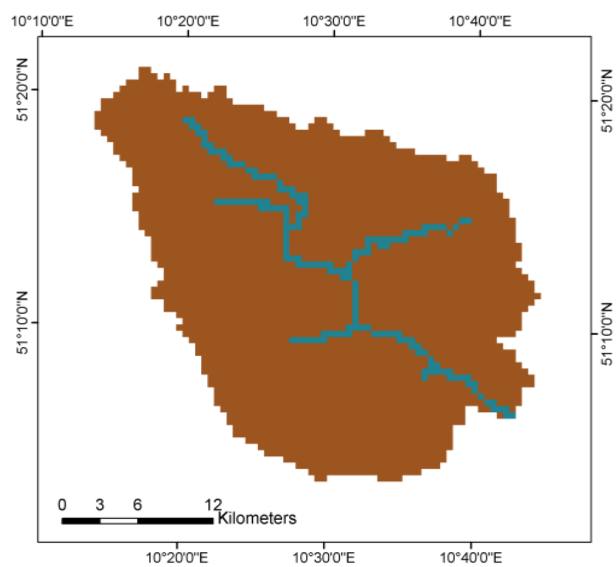

Figure 4 the river grids and soil grids where boundary fluxes are different

5. Finally, in connection with the surface-subsurface coupling, there isn't any reference to river geometry and its role in calculating exchange fluxes and river flow (e.g. as shown in Fig 10).

We appreciate this constructive criticism. To address it, we will add a detailed description of river geometry and its role in exchanging fluxes. The baseflow in Fig 10 is determined by mHM by routing the baseflow generated from linear reservoir of each grid (please see schematic in Figure 1).

The river geometry is displayed in Figure 5 (c) of the manuscript. Within the coupling scheme used in our study, the depth of the river, is irrelevant since we calculate the baseflow directly using mHM's inherent runoff generation and routing scheme (please see the schematic in Figure 3). From the mechanism of mHM as a grid-based hydrologic model, the river network is extracted directly from mHM grid cells and interpolated into OGS upper surface using GIS2FEM, which is the coupling interface. Within that scheme, the width of a river is conceptually equivalent to the width of OGS grid, which is a structured grid with a width of 250 m in the manuscript. The baseflow rate is directly interpolated into surface of OGS mesh and serves as upper boundary condition of the groundwater flow model. The baseflow rate is relevant to the mHM grid size rather than OGS grid size, thus the river geometry, which is mapped in OGS mesh upper surface, has only a minor influence on the catchment scale groundwater dynamics. Since our study focuses on the catchment scale groundwater dynamics rather than the near-field groundwater flow of rivers, the coarse resolution of river network is a simple, however, efficient and reasonable setting. Nevertheless, we can also use an alternative method to portray river network, which is by defining a set of polylines in OGS geometry file. The baseflow rate is then assigned to every node within the polyline by means of linear interpolation. Using this method, the geometry of river can be better represented.

In the revised manuscript, we have incorporated the illustration of river geometry (page 11, line 19-23 and page 12, line 1-2) and modified this section accordingly.

6. Typographic errors:

Eq1. Note z as specified here denotes depth. The description of vertical coordinate is not clear.

Eq. 2, the pressure term, should have a p subscript.

We will take care of these typographic errors and correct them accordingly. Thanks so much.

Responses to Referee Review posted by E.H. Sutanudjaja

We thank the referee reviewer very much for his comprehensive and insightful comments. Those comments are really helpful for us to revise the manuscript.

Below you could find the point-by-point response.

Major comments

10. This paper deals with an effort to couple the regional scale mHM model to a groundwater flow model (i.e. OGS) that can simulate groundwater lateral flow and groundwater head dynamics.
    This topic fits very well to the scope of this journal and I consider this study is an important contribution for regional or large scale hydrological modelling efforts. Currently, there are a still quite limited number of regional (large) scale hydrological models that include lateral groundwater flow component and can simulate groundwater head dynamics. An extension to groundwater head simulation will greatly strengthen the mHM model capabilities, e.g. for enhancing their groundwater drought studies and groundwater transport modelling.
    As a test case, the authors used the Naegelstedt catchment where head observation data are available. They managed to show some convincing validation results of their groundwater modelling result to observation data (e.g. Figure 11). The authors deserved credit for their extensive and successful modelling experiment.

    We appreciate the reviewer's recognition of our manuscript. The reviewer pointed out that "there are a still quite limited number of regional (large) scale hydrological models that include lateral groundwater flow component and can simulate groundwater head dynamics.", which is exactly what we would like to address in the manuscript.

11. However, this paper is still poorly written and therefore it is difficult to comprehend. English must be improved. I strongly recommend that the revised version is checked by an English native speaker.

    Thank you for the suggestion. We have followed the reviewer's suggestions and revised the manuscript carefully to make sure all English expressions are correct. We checked the revised manuscript carefully to make sure the English syntaxes are correct.

12. Details / specific comments:

Page 1, lines 1-2: I suggest to rephrase this sentence. Most hydrological models do include groundwater component, e.g including groundwater (vertical) recharge component and using a linear reservoir concept for groundwater baseflow/discharge. Yet, they hardly include lateral groundwater flow component and simulate groundwater head dynamics.

Thank you for this comment. We have rephrased this sentence accordingly.

13. Page 1, lines 8-9: The sentence (Nested time stepping : : :) does not really flow with the previous ones. Please rephrase. - It will be very informative if the time step lengths used (for both models) are mentioned in the abstract. If I understand correctly, the time step length used for mHM was daily, while OGS used monthly time step. Am I correct?

    This is a good suggestion. Yes, the time step length used for mHM was daily, while OGS used monthly time step. We have rewritten this sentence according to the reviewer's advice. Please see page 1, line 10-11 in the revised manuscript.

14. Page 1, lines 15-16: Please clarify with what you meant by the ' offline coupling method' in your study.

    To avoid any misunderstanding of the phrase "offline coupling method", we replaced the term "offline coupling method" by "one way coupling". For the details of the one way coupling scheme used in the manuscript, please check Section 2.3 in the revised manuscript.

15. Page 1, lines 15-16: How much is the ' little surplus' in your computational cost?

    For each monthly groundwater time step, the simulation time is about 200s. This means the total surplus of the simulation is about 10 hour.

16. Page 2, line 8: : : : ignoring lateral groundwater flow : : :
    Page 2, lines 32-35 and page 3, lines 1-10: Please rewrite this paragraph. I found its sentences (e.g. the first until fourth sentences) do not really flow and connect with each other.

    We have followed the reviewer's suggestions and rewrote the paragraph accordingly. For the details, please check page 2, line 30-35 and page 2, line 1-18.

17. Page 3, line 4: LSM? Common Land Model? I guess that you meant CLM (Community Land Model).

    Thank you for pointing out our misspelling. We have modified the LSM to CLM accordingly.

18. Page 3, line 6: For this study, were you using a similar offline coupling strategy as

used by Sutanudjaja et al. (2011). Did you first run the mHM model for the entire model simulation period (1970-2005?), then use the mHM output to force the OGS model? Please clarify.

We do use a similar coupling strategy as used by Sutanudjaja et al. (2011). We run the coupled model following a four-step procedure:

1) mHM is run independent of OGS to calculate land surface fluxes.
2) After mHM run was finished, the step-wise routed baseflow estimated by mHM are transformed to distributed baseflow along OGS stream network.
3) The distributed groundwater recharge generated from mHM are fed to the coupling interface GIS2FEM, and further transferred to the upper surface boundary conditions of the OGS model.
4) After mHM generated recharge and baseflow were successfully transferred to OGS upper surface boundary conditions, the groundwater model will run subsequently to simulate groundwater flow and transport processes.

Please check page7, line 11-15, and page 8, line 1-24 in the revised manuscript.

19. Page 3, line 8: GSFLOW? What does GSFLOW stand for?

GSFLOW stands for “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)” (Markstrom et al., 2008).

20. Page 3, line 14: What is THMC? I cannot find its long form of this acronym before this line.

THMC is the short form of “Thermo-Hydro-Mechanical-Chemical (coupling)”. THMC modeling is critical in many topics in the field of hydrogeology such as pollutant transport, geothermal heat exchange and seawater intrusion (Kolditz et al., 2012). OpenGeoSys (OGS) project is a scientific open-source initiative for numerical simulation of thermo-hydro-mechanical- chemical (THMC) processes in porous media.

Please see page 6, line 10 of the revised manuscript.

21. Page 3, lines 15-17: Please rephrase this sentence. I am not sure what you meant by ’offline’ coupled here.
Page 3, line 17: : : : an offline coupled model ...

We have followed the reviewer’s comments and rephrased this sentence.

22. Page 4, line 8: So, did you apply MPR for the current study? This is not really clear

for me.

Yes, we applied MPR in the mHM simulation. We also mentioned it in the revised manuscript at page

23. Figure 1: I cannot find the explanations for GOCAD, GO20GS and PEST in the text/paragraph.

    The original Figure 1 includes many external softwares that are dynamically linked to mHM#OGS. These external softwares are not the core of manuscript, so we modified Figure 1 to avoid misunderstanding. For the new figure about the coupling schematic, please check Figure 2 in the revised manuscript.

24. Page 5, line GIS2FEM: What does GIS2FEM stand for?

    The coupling interface GIS2FEM is used to interpolate and transfer mHM grid-based fluxes to OGS nodal flux values. After reading a raster file of mHM generated fluxes, the interface 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 mHM grid cell, the flux of this grid cell is assigned to the corresponding surface element in OGS mesh. After all top surface elements being processed, GIS2FEM will take the face integration calculation, by which the recharge data and baseflow are converted into nodal source terms and assigned to the corresponding OGS mesh nodes.

    Please check page 8, line 16-21 of the revised manuscript.

25. Page 6, lines 7-15: Could you please check this part. I guess that there are some missing lines or sentence here. For example, I cannot find the introduction and explanation for Eq. 2.

    Thank you. We have thoroughly restructured the equations used in the manuscript. Please see the Section 2.1 and 2.2 of the revised manuscript.

26. Page 7, lines 9-10: Due to this liner reservoir conceptualization, I guess that the current coupled model mHM#OGS cannot simulate infiltration from surface water bodies (rivers) to groundwater?

    Yes, you are right. The linear reservoir conceptualization means that the baseflow is from groundwater to the water bodies. We have elaborated this point in the revised manuscript. Please see page 8, line 3-10.

27. Page 9, line 24: What is VTU?

The VTU (equivalent to VTK) data format is the data format for an open-source, freely available software system **Visualization Toolkit** (**VTK**), which is used for 3D computer graphics, image processing and visualization.

28. Section 2.5: Please rewrite this section, particularly to clarify/confirm the following:
- So, you have two scenarios of groundwater modelling: SC1: spatially distributed recharge and SC2: homogeneous recharge - Did you calibrate both scenarios groundwater modelling independently? Or, did you just calibrate SC1 and then using the calibrated SC1 parameters for SC2?

Thank you for this important question. We do groundwater modeling following a two-step procedure. First, we calibrate the steady state groundwater model against the long term mean of groundwater heads. For this step, we calibrate the model separately in SC1 (renamed as mR in the revised manuscript) and SC2 (renamed as RR in the revised manuscript) so that the K values were adjusted to fit the observations. The second step is to run the transient groundwater model using specific yield and specific storage values according to the literature. The K values in SC1 and SC2 are therefore different. We did not calibrate the transient groundwater model. Instead, we performed the sensitivity analysis of recharge scenarios (i.e., mHM generated recharge vs. homogeneous recharge) using the same storage parameters.

Please check the page 12, line 30-31and page 13, line 1-5 of the revised manuscript.

29. Page 16, line 2: Please provide the unit (m2?) for 8625 and 464.74.

We have followed the reviewer's advice and modified this sentence accordingly.

Please check page 13, line 19 of the revised manuscript.

30. Page 16, lines 2-3: What do you mean by the calibration is robust with totally 114 model runs?

It means the objective function in the calibration was successfully converged after a limited number of model runs, which demonstrated the inverse process is well posed.

31. Page 16, lines 3-4: What do you mean by ' convergence criteria relevant to observation'? Please rephrase the sentence.

Thank you. We have rephrased this sentence in the revised manuscript.

32. Page 16, lines 14-16: The sentence does not flow with the previous ones.

We have followed the reviewer's advice and modified this sentence accordingly.

33. Figure 9: I guess this map is for a steady-state condition. Please clarify.

Yes, it is steady-state calibration. We have followed the reviewer's advice and modified this sentence accordingly.

34. Page 16, line 19-20: What do you mean by the last sentence, i.e. the coincidence with Wechsung (2005)? Is it possible to include/visualize some figures from Wechsung (2005)?

Thank you. Wechsung (2005) depicted the regionalized observations of groundwater head in Naegelstedt catchment. The simulated groundwater head depth (Figure 9) shows a good match with the regionalized observations in Wechsung (2005). Unfortunately due to the copyright issue, we cannot include the figure in the manuscript.

35. Figure 10: Could you please also provide other performance metrics, e.g. NSE, KGE? I missed some crucial information, such as the resolution of the forcing data used and the resolution of mHM model used.

Thank you for the insightful comment. We have followed the reviewer's suggestion and include NSE (0.88) in the updated figure (see Figure 6 in the revised manuscript).

The point data at weather stations were subsequently krigged into a 4 km precipitation fields, and then downscaled to mHM grid cells. The resolution of mHM grid cell is 500m.

36. Page 19, line 5: : : : each monitoring well ... (singular)
Page 19, lines 15-17, Page 20: Please check the English. An example: Another reason is that we assigned a homogenous storage coefficient (?) in all aquifers, which an oversimplified setting.

Thank you for this comment. We have followed the reviewer's advices and made the corresponding change. We also carefully checked the English syntax according to the reviewer's insights.

37. Page 21, line 14: Did Kumar et al. (2016) also simulate groundwater heads?

In the study of Kumar et al., 2016, Kumar et al. used the distributed ground head observations in southern Germany and the central Netherlands to reveal the strong spatial variability of groundwater head fluctuations. His study is based on spatially distributed head observations at large scale. Although he did not simulate the groundwater head, his study can still be used as a proof of the strong spatial variability of groundwater heads.

38. Page 22, lines 9-17: For prediction/application in ungauged basins, I believe that hydrogeological characterization (in ungauged basins) still remains as one of the

main challenges.

This is a very insightful comment. We agree with the reviewer that the comprehensive understanding and characterization of hydrogeological processes in ungauged basins is very challenging and remains an open topic. Therefore, we deleted the sentences related to ungauged basin. Please check page 18, line 10-14 in the revised manuscript.

References

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, Environ. Earth Sci., 67(2), 589–599, doi:10.1007/s12665-012-1546-x, 2012.

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, Hydrol. Earth Syst. Sci., 20(3), 1117–1131, doi:10.5194/hess-20-1117-2016, 2016.

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. Geol. Surv., (Techniques and Methods 6-D1), 240 [online] Available from: http://pubs.er.usgs.gov/publication/tm6D1, 2008.

Improved representation of groundwater representation at a regional scale (mHM#OGS v1.0) – coupling of mesocale mesoscale Hydrologic Model (mHM v5.7) with 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

Correspondence to: Miao Jing (miao.jing@ufz.de)

Abstract.

Most of the current large scale hydrological models do not contain a physically-based groundwater flow component. The main difficulties in large-scale groundwater modeling include the efficient representation of unsaturated zone flow, the characterization of dynamic groundwater-surface water interaction and the numerical stability while preserving complex physical processes and high resolution. To address these problems, we propose a highly-scalable coupled hydrologic and groundwater model (mHM#OGS) based on the integration of two open-source modeling codes: the mesoscale hydrologic 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 Hydrologic Model (mHM) and the finite element v5.7) to subsurface hydrology through the coupling with a thermo-hydro-mechanical-chemical (THMC) simulator OpenGeoSys (OGS). mHM#OGS is coupled using a boundary condition-based coupling scheme that dynamically links the surface and subsurface parts. Nested time stepping allows smaller time steps for typically faster surface runoff routing in mHM and larger time steps for slower subsurface flow in OGS . mHM#OGS features the coupling interface which can transfer the groundwater recharge and river baseflow rate between mHM and OpenGeoSys. Verification The two models are one-way coupled through a model interface GIS2FEM, by which grid-based fluxes generated by vertical layered reservoirs within mHM representing near-surface hydrological processes, are converted into 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, 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 was conducted using the time-series of observed streamflow and groundwater levels. Moreover, we force the transient model using groundwater recharge in two scenarios: (1) spatially variable recharge based on the mHM simulations, and (2) spatially homogeneous groundwater recharge. The modeling result in first scenario has a slightly higher correlation with

~~groundwater head time-series, which further validates the plausibility of spatial groundwater recharge distribution calculated by mHM in the mesocale. The statistical analysis of model predictions shows a promising prediction ability of the model. The offline coupling method implemented here can reproduce reasonable groundwater head time series while keep a desired level of detail in the subsurface model structure with little surplus in computational cost.~~ (mHM#OGS v1.0) is evaluated by a case study in a 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 Nash–Sutcliffe model efficiency (NSE), Pearson correlation R_{cor}, and inter-quantile 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 ~~can be a valuable tool to assess the effects of variability in land surface heterogeneity, meteorological, topographical forces and geological zonation on the groundwater flow dynamics.~~ 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 and heads, and thus making it the valuable tool for addressing water resources and management problems.

1 Introduction

~~The significance of depiction of terrestrial water cycle as an integrated system has been continuously recognized. Historically, hydrologic models and groundwater models are isolated and developed in parallel, with either near-surface water flow or subsurface water flow being considered. Furthermore,~~ Historically, large scale hydrologic models are commonly developed to predict discharge. Most of these models use simple bucket-type expressions ~~together~~ 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)~~(Due to the limitation in computational capability, all traditional hydrologic models simplify water flow processes by ignoring lateral ~~water flow, thus~~ groundwater flow. Thus, such models inevitably fall short of ~~explicit characterization of the subsurface groundwater head~~ 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 in many aspects. Water table depth 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 ~~fluctuation has been discovered~~ fluctuations are known as a factor affecting runoff generation ~~, thus affects~~ and thus impacting the prediction skill of catchment runoff (Liang et al., 2003; Chen and Hu, 2004; Koirala et al., 2014). Typical hydrologic models also show inadequacy in 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 distributions (TTDs) at catchment scale and is therefore incapable of describing such legacy behavior (Benettin et al., 2015; Botter et al., 2010; Benettin et al., 20 Moreover, stream-subsurface water interactions may be significant in modulating the human and environmental effects of nitrogen pollution (Azizian et al., 2017). ~~To~~ Fianlly, to assess the response of groundwater to climate change, a ~~physically based~~

more accurate groundwater representation including lateral subsurface flow is urgently needed (Scibek and Allen, 2006; Green et al., 2011; Ferguson et al., 2016).

On the other handParallel to that, numerous groundwater models have been developedin parallel. Groundwater models , 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). A constant challenging problem in groundwater modeling is the reasonable characterization of heterogeneous and noncontinuous geological properties (Dagan, 1986; Renard and de Marsily, 1997; Att Groundwater models usually contain a mechanistic or physically-based representation of subsurface physics, but fall short in providing good representation of surface and shallow soil processes. For example, models for predicting groundwater storage change under either climate change (e.g., global warming) or human-induced scenarios (e.g., agricultural pumping) always 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). Those packages always need extra data and right characterization of many Moreover, parameterization of topographical and geological properties. 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).

With the development of computational capability and increasingly importance in responding to climate change, the coupled models coupling two or more hydrological components together have been attracting more and more attentions.The coupled hydrological model highlights the interactions across the shallow soil column and the deep groundwater aquifer. There exist many reviews of the approaches used in coupling surface water–groundwater processes (Brian A. Ebel, 2010; Fleckenstein et al., 2010; Bar In recent years, there have been some efforts towards coupling surface hydrological model with detailed groundwater model. Maxwell and Miller (2005) coupled LSM (Common Land Model) with a variably saturated groundwater model ParFlow as an integrated model, 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 overland flow) and the subsurface flow processes (e.g., 1D or 3D Richards flow) using a two-way coupling procedure (Paniconi and Putti, 2015). Some of the highly recognized ISSHMs are InHM (Smerdon et al., 2007; VanderKwaak and Loague, 2001), Parflow (Maxwell and Miller, OpenGeoSys (Kolditz et al., 2012; Delfs et al., 2012), 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), MIKE SHE (Graham and Butts, 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), 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). The applications of these ISSHMs in the literature are mainly focusing on the field and small watershed scale, while the assessment of modeled groundwater heads 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, 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 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., well as the storage-runoff correlation (VanderKwaak and Loague, 2001; Liang et al., 2003; Huntington and Niswonger, 2012; Koirala et al etc.

Typical hydrologic models, like 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 discharge but, as mentioned above, are highly conceptualized and there suffering from interpretability of certain processes (e.g., groundwater storage and heads). More mechanistic ISSHMs, like Parflow, CATHY, and HydroGeoSphere, are highly interpretable but show consistently worse performance than typical hydrologic models when predicting runoff (Gulden et al., 2007; Paniconi and Putti, 2015). These different abilities of typical hydrologic models vs. the more mechanistic ISSHMs are caused by the different challenges that are posed by the different compartments of the terrestrial water cycle. One of the main challenges in the surface and demonstrated the need for improved groundwater representation in near-surface water schemes. Sutanudjaja et al. (2011) coupled a land surface model PCR-GLOBWB with a groundwater modeling code MODFLOW using offline coupling scheme, and tested the coupled model using a case study in Rhine-Meuse basin. De Graaf et al. (2015) extended this approach to the global scale. Another highly highly developed coupled model is GSFLOW, which is based on integration of the USGS Precipitation-Runoff Modeling System (PRMS) and the USGS Groundwater Flow Model (MODFLOW and MODFLOW-NWT). GSFLOW has been successfully applied to many case studies (Markstrom et al., 2008; Huntington and Niswonger, 2012; Hunt et al., 2013)near-surface storage is process uncertainty, with processes like 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 and therefore conceptually simpler. Meanwhile, the data uncertainty becomes more significant in the deep subsurface storage compared to the shallow storage. Moreover, a recent study reveals the strong spatial and temporal heterogeneity of processes and properties on SW/GW interface, and underlines the importance of quantifying variability across several scales on SW/GW interface and its significance to water resources management (McLachlan et al., 2017).

In this study, we document the development of a coupled surface hydrological and groundwater model (mHM #OGS) with an an therefore coupled the Mesoscale Hydrologic Model (mHM v5.7) (Samaniego et al., 2010; Kumar et al., 2013b) with a thermo-hydro-mechanical-chemical (THMC) simulator OpenGeoSys (OGS) (Kolditz et al., 2012, 2016) with an overall aim of bridging the gap between catchment hydrology and groundwater hydrology at a regional scale . We choose two highly-scalable, open-source codes with high reputations and wide popularities in their corresponding fields: the mesoscale Hydrologic Model mHM (Samaniego et al., 2010; Kumar et al., 2013b; Samaniego et al., 2013) and the THMC simulator OpenGeoSys (Kolditz et al., 2012; V The coupling is achieved by mechanistically accounting for the spatio-temporal dynamics of mHM generated groundwater recharge and baseflow as boundary conditions to the OGS model as a off-line coupled mode. A nested time-stepping approach is used to account for differences in time-scales of modelling regional scale groundwater flow dynamics. mHM has been demonstrated its preeminence in coping with process uncertainty in near-surface hydrological and groundwater processes, i.e., smaller time steps for typically faster surface runoff routing in mHM and larger time steps for slower subsurface flows in OGS. We zone while providing excellent discharge prediction (?). On the other hand, OGS has been demonstrated its ability of dealing with data uncertainty in groundwater aquifers. With these two well-tested modeling codes available, we want to answer the following scientific questions: First, can spatially distributed groundwater heads and their dynamics be reasonably captured by expanding the abilities of a surface hydrologic model like mHM at the regional scale, all while conserving its excellence in predicting discharge? Second, can spatially resolved groundwater recharge estimates provided by mHM, improve the prediction of head measurements of groundwater models like OGS? To answer these questions, we applied the coupled model mHM#OGS in a Central German mesoscale v1.0 in a central German meso-scale catchment (850 km^2), and verify the model functioning evaluated the model skills using measurements of streamflow and groundwater heads from several wells located across in the study area. The, herein, illustrated coupled (surface) hydrological hydrologic and groundwater model (mHM#OGS v1.0) is our first attempt towards 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 , coupling schematic, as well as coupling schematic. In Section 3, the study area and model setup used in this study . In section are comprehensively illustrated. In Section 4, we present the simulation result results of mHM#OGS v1.0 in a catchment in central Germany, where the subsurface properties are well characterized and long-term monitoring of river stage and groundwater level exist. We also assess the effects of different spatial patterns of groundwater recharge to groundwater dynamics. In the last section, conclusion and future work are discussed. In the Section 5, we discuss model results as well as advantages and limitations of current modeling approach.

2 Model description

2.1 mesoscale Hydrologic Model (mHM v5.7)

The mesoscale Hydrologic Model (mHM, www.ufz.de/mhmwww.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 percolationand baseflowand discharge attenuation and , baseflow, discharge attenuation as well as flood routing (Figure 2 b1). The runoff generation applies a robust scheme which routes runoff in upstream cells along river network using the Muskingum-Cunge algorithm. The model is driven by daily meteorological forcings (e.g., precipitation, temperature), and it utilizes observable basin physical properties or signals (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.

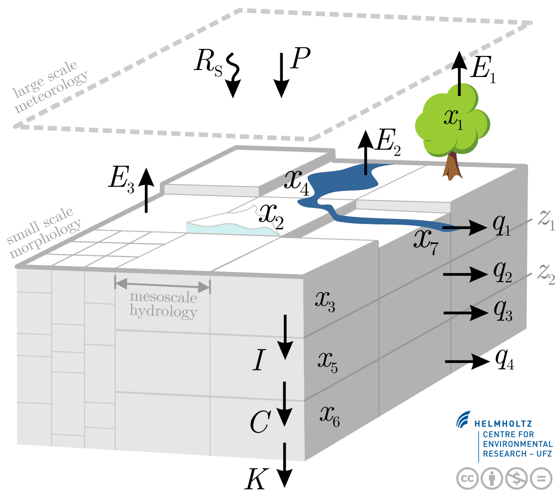

Figure 1. The concept of the mesoscale hydrologic model mHM.

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

Below, we listed the formulas that describe near-surface processes in the deep soil layer and groundwater layer. The comprehensive equation system of mHM can be found in Samaniego et al. (2010). Here, we only listed the equations relevant to this study. In the subsurface reservoir, which is the second layer of the vertical layers (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 depth of water storage in the deep soil reservoir [mm], β_1 is maximum holding capacity of the deep soil reservoir, z_i is depth of subsurface layer i, β_2 is fast recession constant, $q_3(t)$ is slow interflow at time t [mm d^{-1}], β_3 is slow concession constant, β_4 is 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, 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 [mm d^{-1}], β_5 is 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 baseflow [mm d^{-1}], β_6 is baseflow recession rate, x_6 is depth of groundwater reservoir [mm].

The runoff from upstream grid and internal runoff at 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 cell i respectively [mm d^{-1}], $Q_{i'}$ is the contribution from the upstream cell i' [mm d^{-1}], κ is Muskingum travel time parameter, ξ is the Muskingum attenuation parameter, Δt is time interval in hours [hr], t is 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 ~~by~~ 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 (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.

OGS has been successfully applied in different fields ~~like~~ such as water resources management, hydrology, geothermal energy, energy storage, CO_2 storage, and waste deposition (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.

Here we list the governing equations of saturated groundwater flow, which are relevant in this study. They can be expressed as:

$$\partial\psi_p\partial t = -\nabla\cdot\boldsymbol{q} + q_s \quad (9)$$

$\boldsymbol{q} =$

S

$$\partial\psi_p\partial t = -\nabla\cdot\boldsymbol{q} + q_s$$

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

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

2.3 ~~Description of coupled model (mHM#OGS)~~Coupling mechanism

2.3.1 ~~Structure of the coupled model~~

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

- ~~Solve the governing equations for surface water and groundwater flow using a sequential boundary condition switching technique.~~
- ~~Calculate model-wide and detailed water balances in both time and space.~~
- ~~Use nested time stepping method to allow different time steps in surface and subsurface regimes~~
- ~~Allow different cell sizes in the spatial discretization of the grid cell resolution used for mHM and the finite-element solution used for OGS .~~
- ~~Allow different model boundaries definition using standard specified-head, specified-flow, or head-dependent boundary conditions.~~

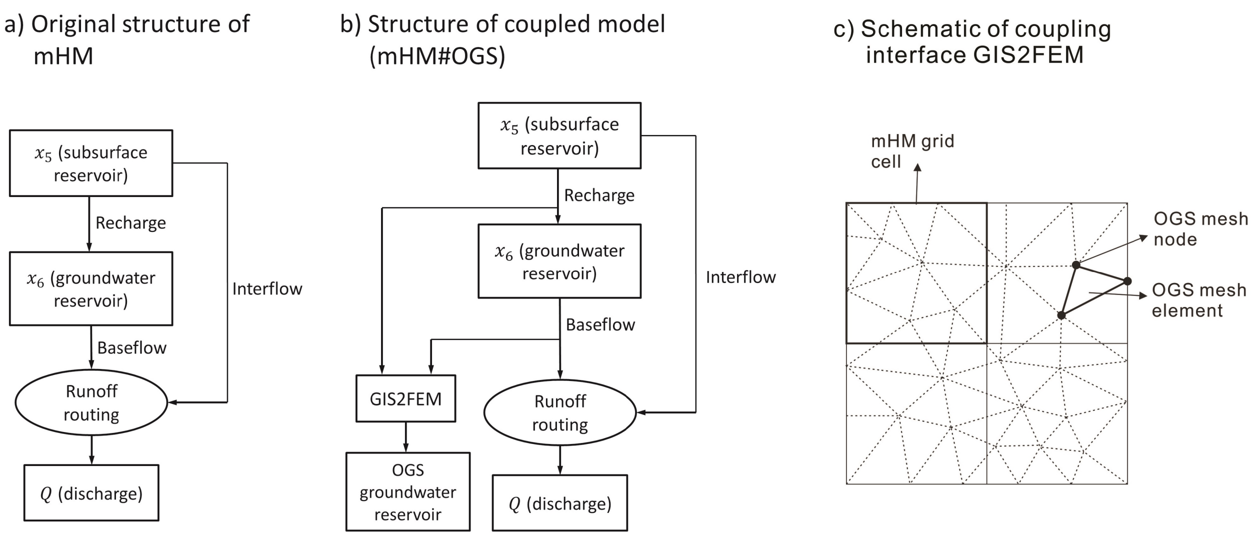

Figure 2. The ~~concept~~ schematic of the coupled model mHM#OGS ~~model~~ v1.0. a) the ~~conceptual representation~~ original structure of ~~hydrological processes in a catchment~~vertical layered reservoir of mHM; b) the ~~schematic used to couple the~~ structure of coupled model (mHM~~and~~ #OGS v1.0). ~~The upper box depicts~~ c) Illustration of data interpolation and transformation through the ~~canopy interception~~coupling interface GIS2FEM. For sake of simplicity, ~~atmospheric forcing~~the figure only displays mHM layers relevant to this study, and neglect the ~~land surface processes represented by other~~ mHM layers (e.~~The lower box depicts the saturated zone represented by the OGS groundwater model; eg.~~, x_1 - x_4)~~the complete workflow including several interfaces with external softwares for data import~~. Grid-based mHM fluxes (e.g., ~~format conversion~~GW recharge and baseflow) are linear interpolated to top surface of OGS mesh, ~~model calibration~~ and ~~water balance check~~further transferred into volumetric values by face integration calculation, which are directly assigned to OGS surface mesh nodes.

~~An integrated workflow for coupled modeling is illustrated in Figure 2c. The entire modeling workflow is separated into three independent parts. The first part is the data preparation and pre-processing part marked by blue background in Figure 2 c.This part includes several codes for data preparation and model pre-processing developed by the mHM and OGS communities (Fischer et al., 2015; Kolditz et al., 2016).The second part is model coupling, which is composed of the respective computations of mHM and~~

The basic idea is to feed fluxes generated by mHM (e.g., distributed groundwater recharge and baseflow) to the OGS ~~and the data communication between two codes (see components marked by red background in Figure 2 c).Technically, we use a self-developed data communication code~~ mesh surface as outer forcings through a coupling interface GIS2FEM ~~to convert data format and exchange information. The detailed illustration of this part is in the following subsection.The third part is the water budget, which is designed to calculate an overall water balance, as well as component-based water budgets for all storages simulated by mHM #OGS (marked by light green background in Figure 2 c).~~(Figure 2). The two models are performed separately and sequentially with usually different temporal (e.g., daily in mHM and weekly or monthly in OGS) and spatial resolutions (e.g., larger grid cell size in mHM and smaller element size in OGS). The original vertical layered reservoirs in mHM, e.g., the soil-zone reservoir, subsurface reservoir and groundwater reservoir are preserved, implying that all well tested features of mHM (e.g., MPR, infiltration-runoff partitioning) are fully preserved in the coupled model.

2.3.1 Boundary condition-based coupling

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

The subsurface flow equation is solved in OGS.OGS applies a standard (centered) Galerkin finite element method to discretize PDEs. Here we list the governing equations of groundwater flow in saturated zones used in this study. It can be expressed as:

$$S_s \frac{\partial \psi_p}{\partial t} = -\nabla \cdot \boldsymbol{q} + q_s + q_e$$

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

where S_s is the specific storage coefficient 1/L, ψ_p is the pressure head in porous media L, t is timeT, $\boldsymbol{q}$ is the Darcy flux (LT^{-1}), q_s is a specified rate source/sink (T^{-1}), q_e is the exchange rate with the surface water (T^{-1}), K_s is the saturated hydraulic conductivity tensor LT^{-1}, z is the vertical coordinate L.

$$-K_s \nabla(\psi - z) \cdot \boldsymbol{n} = q_{bc}$$

on Γ, where $\boldsymbol{n}$ is the outer norm of the boundary surface

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

   The surface flow over the catchment in either the hillslopes or open channels can be expressed using a kinematic-wave equation (an approximation of Using a gridded observations of meteorological forcings (precipitation, temperature, and potential evapotranspiration), the Saint-Venant equations). The kinematic-wave equation for flow in streams can be expressed as:

$$\frac{\partial \psi_s}{\partial t} = \nabla \cdot (\psi_s \boldsymbol{v}) + q'_e$$

   where t is time T, $\boldsymbol{v}$ is the averaged velocity vector LT^{-1}, ψ_s is the surface water depth L, q'_e is the exchange rate with subsurface water LT^{-1}.In mHM, a Muskingum-Cunge method is used to solve Eq.(**??**) (Te Chow, 1988).

   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 kinematic-wave approximation assumes the gravitational forces are balanced by frictional forces such that:

$$S_o = S_f$$

where S_o is the slope of the channel -, and S_f is the friction slope of the channel -. This assumption is used because the potential areas of application of this model would hardly exhibit abruptly hydrographs with supercritical flows. 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.

Manning's equation (Te Chow, 1988) is used to calculate the depth averaged velocity using discharge:

$$\boldsymbol{v} = \frac{\sqrt{S_f}}{m}\psi_s^{2/3}$$

where m is the Manning roughness coefficient $L^{-1/3}T$. As an empirical formula, the Manning's formula has been widely used in surface water flow models.

2. After mHM run was finished, the step-wise routed baseflow estimated by mHM are transformed to distributed baseflow along OGS stream network.

The source terms q_e and q_e' in Eq. (2.2) and Eq. (**??**) explicitly represent the communication of water between the surface and subsurface flow compartments (Camporese et al., 2010). The surface-to-subsurface flux q_e is determined after solving the surface routing equation Eq. (**??**) for the following feed to the subsurface flow equation Eq. (2.2), meanwhile the subsurface-to-surface feed q_e' is determined by solving the subsurface flow equation for the following feed to the surface water equation. Note that the time step of subsurface flow is always larger than that of surface flow. Note that poor water balance might occur if more than 30–50 surface time steps exist in one subsurface time step (Camporese et al., 2010). Here we use a nested time stepping method to avoid the water balance problem (see the following paragraph). Besides, we use the original linear groundwater storage in mHMto calculate daily q_e'. The stream conceptions within mHM and OGS are slightly different, in terms that the stream in mHM is based on pre-processing of DEM data and a routing scheme, while in OGS by an explicit predefined river geometry. In OGS, each reach of the stream network is defined by a polyline in an OGS geometry file. To coordinate the two different conceptions, we transfer the total routed baseflow (estimated by mHM) to distributed baseflow along OGS streams by distributing it uniformly along the predefined stream network in OGS. The detailed description of stream network in OGS can be found in Section 3.3. 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.

Transfer of groundwater recharge from mHM grid cells to OGS nodes using the model interface GIS2FEM.

Nested time stepping is adopted in this study in order to calculate fast surface and slow subsurface flow simultaneously. As reported by Cunge (1969), the Muskingum-Cunge used to solve surface flow equations is unconditionally stable in case some perquisites being meet, such as proper grid size and time step size. The subsurface solver in OGS is however, implicit in time and limited by less restrictive precision constraints. We use a nested time stepping method to calculate daily surface processes and monthly subsurface processes in a sequential manner. This strategy automatically fits to any stepping size difference, and avoids water balance error in flux exchange between two modules.

Conversions between volumetric flux (L^3/T), specific flux (L/T), and water head (L)are performed by adjusting different time steps or cell sizes. Specifically, the time series of groundwater recharge obtained from mHM were

3. The distributed groundwater recharge generated from mHM are fed to the model interface of mHM#OGS coupling interface GIS2FEM, and further 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 to OGS nodal flux values. After reading a raster file of groundwater rechargemHM generated fluxes, the interface assign the proper recharge GIS2FEM interpolates the flux value to the top surface elements of OGS mesh by checking the coordinates of the centroid of each top surface elementthe OGS mesh. For each surface element, if its centroid is within a grid cellof the raster file, the value the range of mHM grid cell, the flux of this grid cell is assigned to the surface elementcorresponding surface element in OGS mesh. After all top surface elements have been processed, the elements that have been assigned with the recharge values are involved in the being processed, GIS2FEM will take the face integration calculation, whereby by which the recharge data is and baseflow are converted into nodal source terms (see details in Figure **??**). The parameter set used in this study is shown in Table **??**. and assigned to the corresponding OGS mesh nodes (Figure 2c).

Table of hydrological parameters used in this study. Symbol Description Source or estimation method Values Unit K_s saturated hydraulic conductivities calibrated values (Table 1) distributed m/s S_s aquifer specific storage coefficient Table of Batu (1998) 1.0×10^{-5} - S_f channel longitudinal slope DEM distributed - m Manning roughness coefficient best guess estimate 0.045 m$^{-1/3}$s

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

The two models are coupled in a sequential manner by fed fluxes and variables from one model to another at every subsurface time step. Technically, the coupling interface converts time-series of variables and fluxes to Neumann boundary conditions, which can be directly read by OGS . The modified OGS source code can produce raster files containing the time-series of flow-dependent variables and volumetric flow rates with the same resolution of mHM grid cells which can be directly read by mHM. The detailed workflow of the coupling technique is shown in Table **??**.

Description of computational sequence for mHM#OGS using a sequential coupling scheme **Sequence No. Computation** 1 **Initialize, assign, and read —** Run mHM and OGS initialize procedures, OGS assign, read and prepare parameters and subroutines for later simulation.2 **Compute near-surface hydrologic processes in mHM —** Calculate near-surface processes such as snow melting, evapotranspiration, fast interflow, slow interflow, groundwater recharge , and surface runoff for each grid cell.3 **Compute the long-term mean of land-surface and soil-zone hydrologic processes —** Compute the long-term mean of the entire simulation period and write them as a set of raster files.4 **Transfer surface-to-subsurface exchange rates to OGS —** Transfer surface-to-subsurface volumetric flow rates needed for computing saturated flow as Neumann boundary conditions in OGS.5 **Steady-state calibration —** Run OGS-only steady-state simulation using boundary conditions given by mHM in step 4. The calibrated K field is fed to transient model. The steady-state groundwater head serves as an initial condition of the transient mHM#OGS modeling.6 **Start transient simulation of mHM#OGS** — Sequence through coupled mHM and OGS components.7 **Compute stepwise near-surface flow and storage in mHM** — Compute spatially-distributed daily near-surface processes.8 **Transfer primary variables and volumetric flow rates to OGS** — The same as step 4, except this step generate time-dependent raster files of flow rates. 9 **Solve the groundwater flow equation** — Calculate groundwater heads and groundwater flow velocity field in study region.10 **Compute budgets** — Run water budget package to check overall water balance as well as time-dependent water budgets in each storages.11 **Write results** — Output the simulations results. 12 **End of simulation** — Close all input files and processes.

2.4 Study area and model setup

3 Study area and model setup

We use a mesoscale catchment upstream of Naegelstedt catchment meso-scale catchment (about 850 km^2) upstream of the Nägelstedt gauge located in central Germany , with a drainage area of about 845 km^2 to verify to establish and assess our model (see Figure 3). The Naegelstedt Nägelstedt catchment comprises the headwaters of the Unstrut river basin. The Unstrut river basin is a sedimentary basin of Unstrut river, a left tributary of the Saale. It was is selected in this study because there are many groundwater monitoring wells operated by Thuringian State office for the Environment and Geology (TLUG) and the Collaborative Research Center AquaDiva (Küsel et al., 2016). Morphologically, the terrain elevation within the catchment is in a range of 164 m and 516 m, whereby the higher regions are in the west and south as part of the forested hill chain of the Hainich (see Figure 3). The Naegelstedt Nägelstedt catchment is one of the most intensively used agricultural regions in Germany. In terms of 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 and transport area. The mean annual precipitation in this area is about 660 mm.

3.0.1 Meteorological forcings and morphological properties

In this study, mHM runs are 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 spin-up period. OGS is run for the period from January 1, 1975 to December 30, 2005. mHM is run with a daily time step, while OGS is run with a monthly time step. The resolution of mHM grid cells is 500 m × 500 m. The spatial resolution of OGS mesh is set to 250 m × 250 m in horizontal direction and 10 m in 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 to run both models are detailed in the following sections.

3.1 Meteorological forcings and morphological properties

We started the modeling by performing the daily simulation of mHM to calculate near-surface and soil zone hydrological processes. Several resolutions ranging from 200 m to 2 km are applied in mHM to account different scale of spatial heterogeneity. Heße et al. (2017) have already established the mHM simulation over the study area. The meteorological and

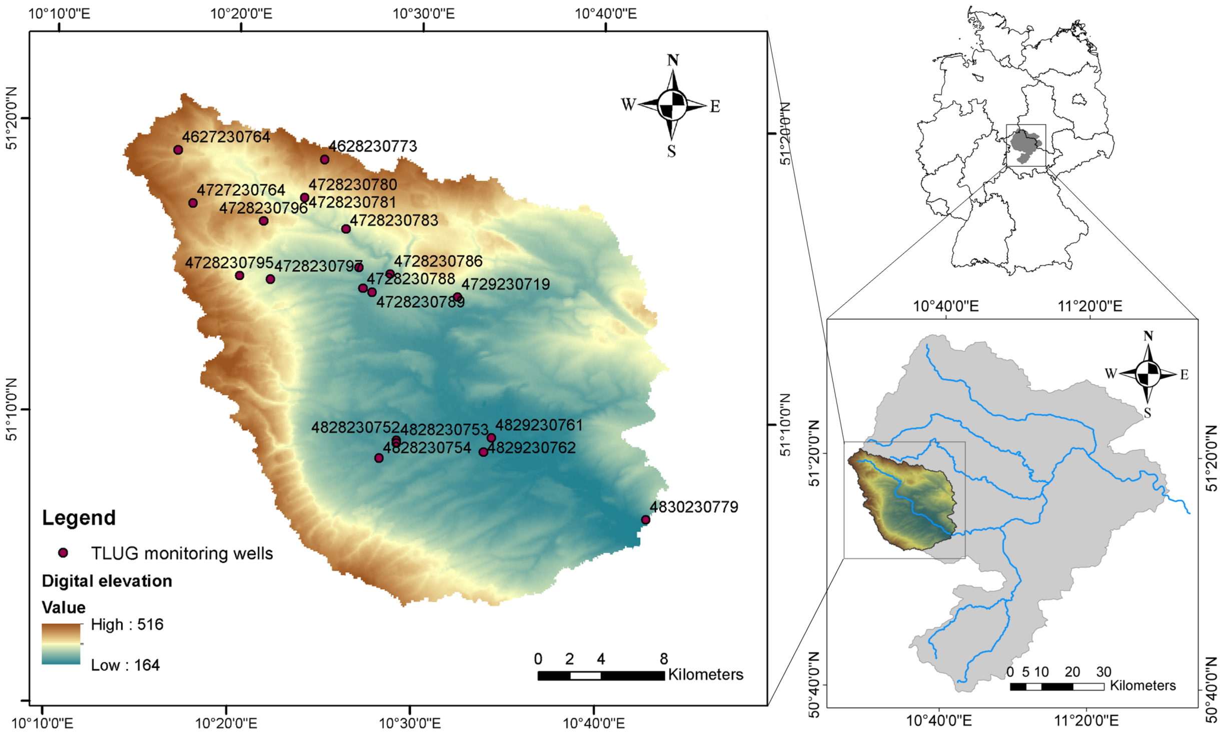

Figure 3. The Naegelstedt Nägelstedt catchment used as the test catchment for this model. The left map shows elevation and locations of monitoring wells used in this study. The lower right map shows the relative location of Naegelstedt Nägelstedt catchment in Unstrut basin. The upper right map shows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

morphological settings in this paper are the same with his work. For the detailed settings of meteorological forcings and morphological properties The mHM model is forced by daily meteorological forcings, including distributed precipitation and atmosphere temperature. The spatial patterns of precipitation and atmosphere temperature were based on point measurements of precipitation and atmosphere temperature at weather stations from the German Meteorological Service (DWD) . The point data at weather stations were subsequently krigged into a 4 km precipitation fields, 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) data, which is the basis for deriving properties like slope, river beds, 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, please refer to Heße et al. (2017). 2000 and 2005); and discharge data at the outlet of catchment.

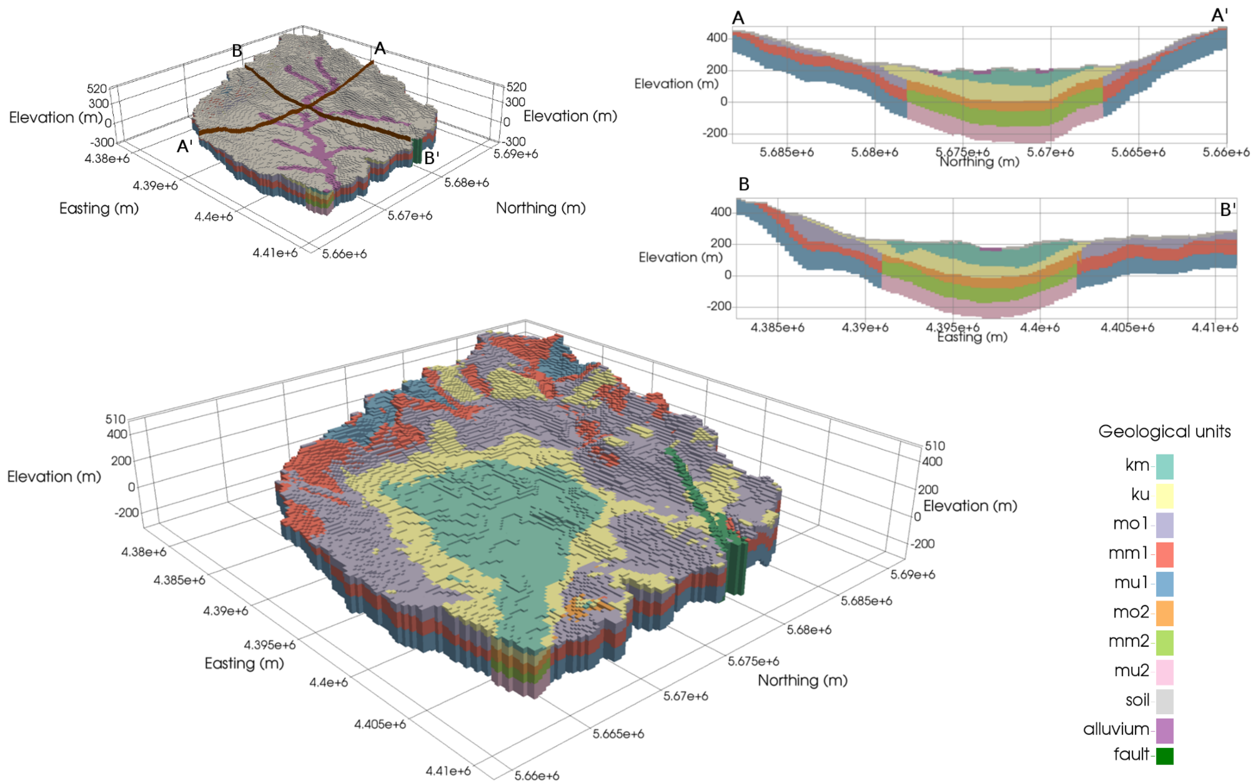

Figure 4. Three-dimensional and cross ~~Section~~ section view of hydrogeologic zonation in the ~~Naegelstedt~~ Nägelstedt catchment. The upper left figure shows the complete geological characterization and zonation including alluvium and soil zone. The upper right figure shows the geological characterization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

3.1.1 ~~Aquifer properties and meshing~~

3.2 Aquifer properties

~~Typical distributed hydrological models use shallow soil profiles or extended soil profile to present groundwater storage. Here, we use a spatially distributed~~ We used a stratified aquifer model to explicitly present ~~groundwater storage. We set up this spatially distributed aquifer model through geological modeling~~ heterogeneous distribution of hydraulic properties (e.g., K value, specific yield, specific storage). The stratified aquifer model is based on well log data and geophysical data from Thuringian State office for the Environment and Geology (TLUG). To convert the data format, we ~~use~~ used the workflow developed by ~~Fischer et al. (2015) to convert~~ Fischer et al. (2015), by which the complex 3D geological model were converted into open-source ~~VTU~~ VTK format file that can be ~~used~~ read by OGS. ~~Model elements of OGS were set to a 250 m × 250 m horizontal resolution and a 10 m vertical resolution over the whole model domain.~~

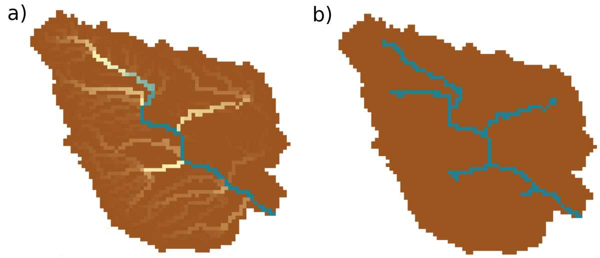

Figure 5. Illustration of stream network used in this study. a) ~~Stream~~ Original stream network based on ~~long-term average of accumulated routed~~ the streamflow routing algorithm of mHM; b) ~~Stream~~ Processed stream network ~~whereby long-term averaged accumulated monthly streamflow rate is above 1000 mm; c) Stream network whereby long-term averaged accumulated monthly streamflow rate is above 1500 mm, which is also the default setting~~ that are used in this study~~; d) Stream network whereby long-term averaged accumulated monthly streamflow rate is above 2000~~ . The small tributaries where the runoff rates are below 1500 mm/month have been cut out from the original stream network.

The dominant sediments in the study site are Muschelkalk (Middle Triassic) and Keuper (Middle and Late Triassic). Younger deposits from Tertiary and Quaternary are less important for the large scale hydrogeology of the basin. The Keuper deposits mainly lie in the center of the Unstrut basin and act as permeable shallow aquifers. In ~~Naegelstedt~~ the Nägelstedt catchment, the Keuper deposits are further classified into two geological sub-units: Middle Keuper (km) and low Keuper (ku) (see Figure 4). The Muschelkalk is marked by a prevailing marine environment and is subdivided into three sub-units Upper Muschelkalk (mo), Middle Muschelkalk (mm) and low Muschelkalk (mu). According to previous geological survey (Seidel, 2004), even the same sub-unit of Muschelkalk have a diverse hydraulic properties depending on their positions and depths. They are further divided into sub-units with higher ~~permeability (see~~ permeabilities, which are mo1, mm1 and mu1 ~~in~~ (Figure 4), and sub-units with lower ~~permeability (see~~ permeabilities, which are mo2, mm2 and mu2 ~~in~~ (Figure 4). The uppermost layer with a depth of 10 m is set as a soil layer ~~, whereby the hydraulic properties are set the same with mHM setting (see soil in~~ (Figure 4). A alluvium layer is set along the mainstream and major tributaries representing granite and stream deposits (Figure 4).

3.2.1 ~~Boundary conditions~~

3.3 Boundary conditions

Based on the steep topography along the watershed divides, groundwater is assumed to be naturally ~~seperated~~ separated and not able to pass across the boundaries of the watershed. No-flow boundaries ~~were~~ are imposed at the outer perimeters surrounding the basin as well as at the lower aquitard.

The stream network ~~were delineated using a digital elevation model (DEM)-based pre-processor, and then clipped based on a threshold correlated with field observations. In general, all streams are regarded as perennial in this study, except for those in the mountainous region where flow is intermittent. Stream network is determined based on~~ is delineated by processing a grid-based runoff raster file generated by mHM. The grid-based runoff is converted to a valid stream network compatible with OGS. The necessity of transferring mHM runoff raster file to OGS stream network has been elaborated in Section 2.3. Particularly in this case study, we cut out the small intermittent tributaries by setting a threshold value of long-term ~~average of accumulated routed streamflow(see Figure 5). We cut off the small tributaries by which long-term average of accumulated routed streamflow is below a threshold~~averaged routed runoff. Only streams with a runoff rate higher than the threshold (in this case study, 1500 mm/month) are delineated as valid streams. In other words, we neglect the intermittent ~~streamflow~~ streams to the upper stream reaches (~~see the lower left graph in~~ Figure 5). The preprocessed stream network consists of a main stream and four tributaries (Figure 5b). Each stream is defined as a polyline in OGS geometry file and comprises many consecutive nodes in OGS mesh. As mentioned in Section 2.3, uniformly distributed baseflow rates were subsequently assigned to every OGS mesh nodes within the stream network.

3.4 Calibration procedure

~~Calibration of the integrated model were conducted using two different conceptual scenarios in order to test the effect of spatial heterogeneity evaluated by Multiscale parameter regionalization (MPR) method in mHM. For the first conceptual scenario (SC1), spatial variability of physiographic characteristics are characterized by MPR method in mHM. The heterogeneous groundwater recharge distribution is determined within the MPR framework. For the second scenario (SC2), we kept the total amount of groundwater recharge of the catchment the same with SC1 in every time step, but use the homogeneous distributed groundwater recharge. SC2 does not consider the complex spatial variability of groundwater recharge caused by variations in climatic conditions, land use, topography and geological heterogeneity.~~

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

For the first step, mHM ~~was~~ is calibrated independent of OGS for the period from 1970 ~~-~~ to 2005 by matching observed runoff at the outlet of the catchment. The first 5 years are used as “spin-up” period to set up initial conditions in 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 ~~calibrated mHM model is also verified by measurements from a single eddy-covariance measurement station in the study area. The calibration goodness was~~ calibration goodness is handled by means of calculating the Nash-Sutcliffe efficient (NSE):

$$NSE = 1 - \frac{\sum_{i=1}^{n} |(h_m - h_s)|_i^2}{\sum_{i=1}^{n} |(h_m - \bar{h}_m)|_i^2} \tag{10}$$

where h_s is the simulated groundwater head [m], $\bar{h}_m$ is the mean of measured groundwater head [m].

~~In~~ For the second step, ~~OGS is run independent of mHM in steady-state using~~ the steady state groundwater model is calibrated to match the long-term ~~averaged outer forcings. Spatially distributed but long-term average recharge~~ mean of groundwater observations. The long term mean of recharge and baseflow estimated by mHM ~~were fed as~~ are fed to the steady-state ~~boundary condition. The long-term average baseflow rate estimated by mHM simulations were also used as boundary~~

~~condition at stream beds. The~~ groundwater model as 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 ~~to seek for a~~ 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 objective function, which is the sum of weighted squared residuals of long-term mean of modeled groundwater heads and observed groundwater heads in this case. Goodness of fit between the simulated and observed long-term average groundwater levels ~~was assessed by~~ is assessed by the root-mean-square error (RMSE).

~~For the third step, we conduct a 30 years transient simulation using the calibrated K values. The modeling results are compared with historical~~

3.5 Model evaluation and sensitivity analysis

Besides the observed discharge at the catchment outlet, we also used observed groundwater head time series ~~. Calibrated K values from~~ 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 model calibration. Meanwhile, the steady-state ~~groundwater model are used as the parameter set for transient groundwater model . The groundwater head distribution in steady-state model~~ groundwater hydrographs are used as ~~initial condition of the transient model. The modeling result verification was approached on the basis of the goodness of cross-correlation~~ the initial condition for the transient groundwater model. For evaluating the predictive ability in terms of groundwater heads, the Pearson correlation coefficient R_{cor} and the ~~(relative)~~ inter-quantile range error QRE ~~$_{7525}$ of the simulated and observed groundwater levels~~are used as two summary statistics. The (relative) inter-quantile range error QRE ~~$_{7525}$~~ is defined by:

$$QRE_{7525} = \frac{IQ^{md}_{7525} - IQ^{dt}_{7525}}{IQ^{dt}_{7525}} \quad (11)$$

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

We also sought to quantify the effect of a spatially-distributed recharge on the simulated groundwater heads. 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. The sensitivity analysis follows a two-step workflow. First, we calibrated the steady-state groundwater models in two recharge scenarios independently. Second, we conducted transient simulations by assigning the same values of storage parameters and compared their corresponding performances in two recharge scenarios. Despite the different spatial pattern of recharge and K values, all model parameters (e.g., specific yields, specific storage) and model inputs in the two scenarios are identical. The R_{cor} and QRE are used to assess model performances in the two respective simulations.

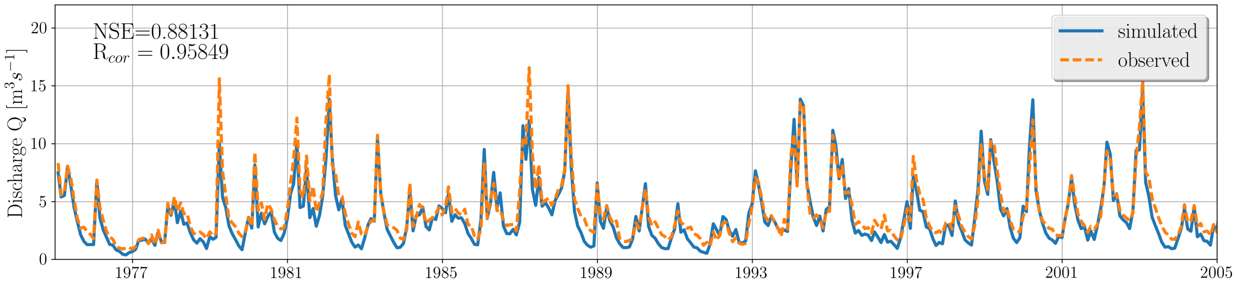

Figure 6. ~~Mean monthly water balance of groundwater over the Naegelstedt catchment. a) Boxplot indicates spread, skewness,~~ Observed and ~~outliers of groundwater recharge and groundwater~~ simulated monthly discharge ~~. b) Histogram indicates~~ at the ~~distribution~~ outlet of ~~groundwater balance~~Nagelstedt catchment.~~c) Monthly time series of groundwater recharge and baseflow.~~

4 Results

4.1 ~~Spatial-temporal dynamics of recharge and baseflow~~Calibration

~~Spatial distributions of groundwater recharge in Naegelstedt catchment (unit: mm/month) (a) during early spring, (b) late spring, and (c) winter of year 2005.~~ As the first step, mHM is calibrated independent of OGS. Monthly discharge data from January 1975 to December 2004 are used for model calibration. The calibration results indicate that mHM is capable to reasonably reproduce 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 measured at eddy-covariance stations inside this area, also shows quite reasonable correspondence to the modeled estimation (Heße et al., 2017).

Subsequently, steady state groundwater model is calibrated against long term mean of groundwater heads using PEST (Doherty et al., 1994). Table 1 shows the calibrated hydraulic conductivities of each 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 function of calibration, which is the sum of squared weighted residuals, converged from the initial value of 8625 m^2 to 464.74 m^2 after 114 total model runs.

Broadly speaking, the calibration results suggest that the groundwater model can plausibly reproduce the finite numbers of observed groundwater heads within the catchment. Figure 7 shows the 1-to-1 plot of simulated and observed groundwater heads (locations of those wells are shown in Figure 3). It can be observed that the model is capable of reproducing spatially-distributed groundwater heads in 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 $\pm$6 m. Nevertheless, there are some monitoring wells where the predictions are biased from observations due to the complex local geological formations around monitoring wells. No further attempt was made to add more model complexity to improve model-to-measurement match.

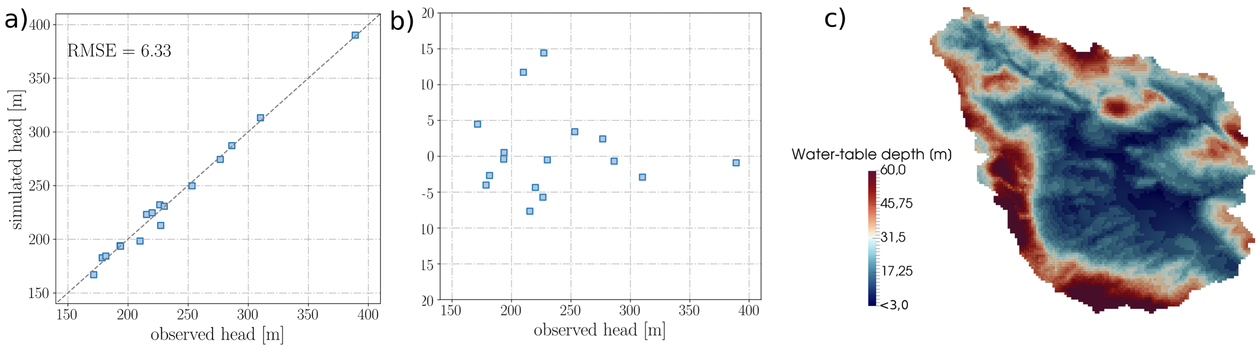

Figure 7. Illustration of steady-state groundwater model calibration and simulatied heads. (a) Observed and simulated groundwater head (including RMSE) ; (b) Difference between simulated and observed head related to the observaed head values; c) Simulated long term mean water table depth over Nägelstedt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

Simulated water table depth over the whole catchment using the calibrated K values is shown in Figure 7c. Broadly speaking, the calibrated model reasonably reproduced spatial groundwater table distribution. Groundwater table depth is as large as above 40 m in the higher southwestern and northern mountainous areas, whereas 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 depending on the sporadic, irregular, and complex features of storm rainfall occurrences, geological structure, and morphological features. The temporal and spatial variability of groundwater recharge and baseflow is estimated by mHM calculation with a period of 30 years from 1975 to 2005.

Figure 8 shows the spatial variability of groundwater recharge in three months: the early spring (March) (Figure 8a), late spring (May) (Figure 8b), and winter (January) (Figure 8c). The results indicate that the largest groundwater recharge may occur at mountainous areas. ~~Greatest~~ Largest recharge occurs in the upstream bedrock areas where dominant sedimentary is Muschelkalk with a relatively low hydraulic conductivity. The ~~greatest~~ largest point-wise monthly groundwater recharge varies from 26 mm in early spring, to 51 mm in late spring ~~, to~~ and 14 mm in winter. ~~We have also~~ Besides, we evaluated the plausibility of groundwater recharge simulated by mHM with other reference datasets. ~~On~~ At the large scale, the simulated groundwater recharge from mHM agrees quite well with ~~the estimation~~ estimates from the Hydrological Atlas of Germany ~~(please refer to Kumar et al. (2016))~~(?).

~~Figure 9 shows the boxplot and histogram of normalized monthly groundwater income and outcome over the whole catchment. We do not include the human effects (e.g., pumping, abstractions and irrigation) . Therefore, baseflow to streams is considered as the only source of groundwater discharge. The boxplot~~ The boxplot (Figure 9a) shows the degree of spread and skewness of the ~~distribution~~ distributions of the monthly groundwater recharge and groundwater discharge. It ~~shows that~~ reveals the

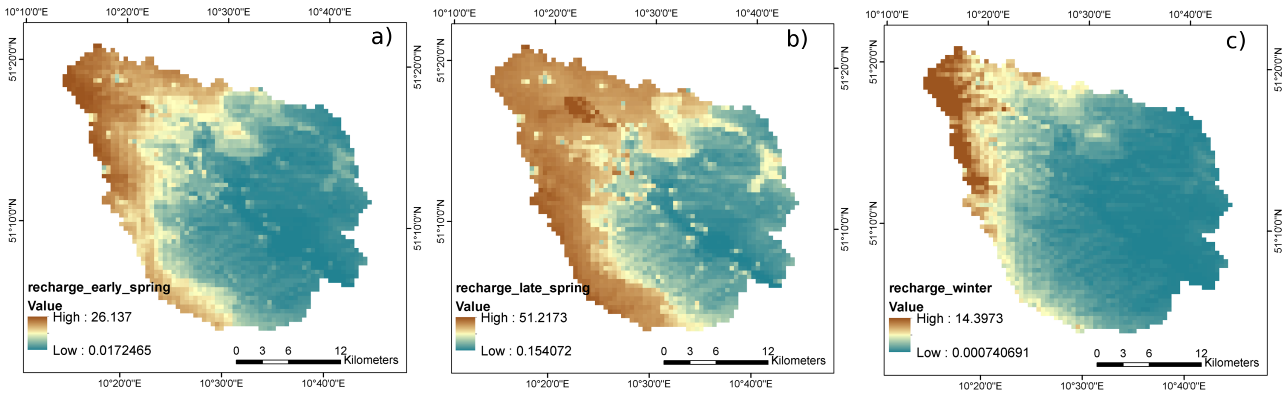

Figure 8. ~~Illustration~~ Spatial distributions of ~~steady-state calibration results.~~ groundwater recharge in Nägelstedt catchment (unit: mm/month) (a) ~~Observed and simulated groundwater head~~during early spring, ~~including RMSE and R_{cor};~~ (b) ~~Difference between simulated~~ late spring, and ~~observed head related to the observaed head values.~~ (c) winter of year 2005.

long-term mean ~~value of monthly groundwater recharge and discharge~~ of groundwater inflow and outflow are balanced with the ~~value of approximately~~ same monthly value of 8 mm/month. Due to the numerical error, a tiny difference of 2% between groundwater recharge and baseflow is observed in the boxplot. ~~This slight bias is~~ Nevertheless, we consider this bias to be within an acceptable interval. ~~The figure also shows that the spread of groundwater recharge is wider than the baseflow, which demonstrate the buffering effect of groundwater storage.~~ Figure 9b shows the distribution of monthly groundwater recharge and monthly baseflow. The figure indicates that the distribution ~~pattern~~ of monthly groundwater recharge is skewed to the right, whereas the distribution ~~pattern~~ of monthly baseflow is unimodal. ~~It also indicates that the~~ Figure 9c depicts the time series of groundwater recharge and baseflow, which further demonstrates that the deviation of monthly groundwater recharge ~~distribution has a higher deviation than baseflow distribution.~~ is larger than the baseflow. This phenomenon further reveals the significant buffering effect of the linear groundwater storage in mHM.

4.3 ~~Model evaluation using discharge & groundwater heads~~

~~We use the scenario with distributed groundwater recharge calculated by mHM (SC1) as the default scenario, and the scenario with homogeneous groundwater recharge (SC2) as the reference scenario in this study. All calibrated parameters' values are in SC1 by default. For mHM calibration, the calibration result is good with R_{cor}>0.9 for the monthly discharge simulation at the Naegelstedt station (see Figure 6). Other fluxes like evaportransporation measured at eddy-covariance stations inside this area, also shows quite reasonable correspondence to the modeled estimation (please refer to Heße et al. (2017)).~~

4.3 Model evaluation against dynamic groundwater heads

~~Table 1 shows the calibrated hydraulic conductivities of each geological units. The upper and lower limit of each parameter was determined according to literature (Seidel, 2004; Wechsung, 2005). The calibrated conductivities in each geological zones~~

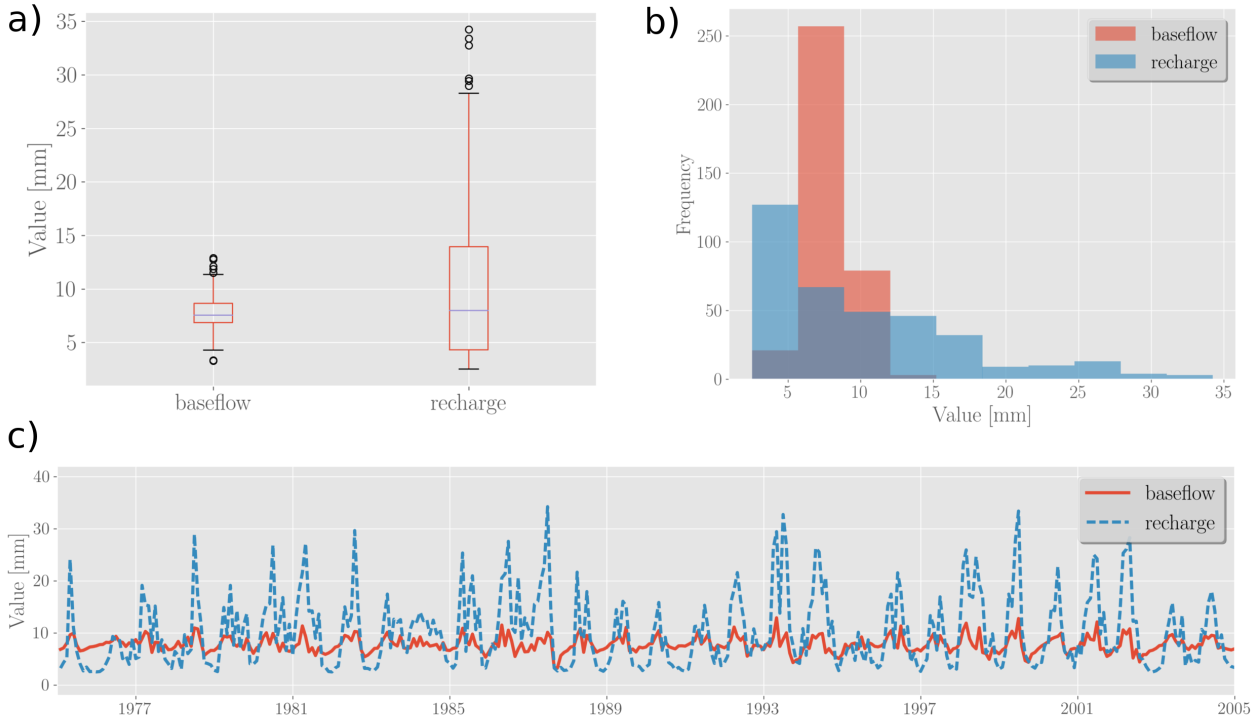

Figure 9. Analysis of groundwater inflow (recharge) and outflow (baseflow) over the Nägelstedt catchment. a) Boxplot indicates spread, skewness, and outliers of groundwater recharge and baseflow. b) Histogram indicates the distribution of groundwater balance components. c) Monthly time series of groundwater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

are displayed in Table 1. The objective function of calibration, which is represented by sum of squared weighted residuals ϕ, converged from the initial value of 8625 to 464.74 after 114 total model runs. The calibration of hydraulic conductivities is robust with totally 114 model runs. However, in spite of convergence criteria relavant to observation, there is still uncertainty around the calibrated "best-fit-value" of hydraulic conductivities.

The steady-state groundwater model calibration result shows that the groundwater model can plausibly reproduce the finite numbers of observed groundwater heads within the catchment. Figure 7 shows the 1-to-1 plot of simulated and observed groundwater heads using different recharge scenario SC1 and SC2, respectively (locations of those wells are shown in Figure 3). It can be observed that the model is capable of reproducing spatially-distributed groundwater heads in a wide range, with low RMSE values of 6.22 m in SC1 and 10.14 m in SC2, respectively. There are certain differences between simulated heads and observed heads. This difference is caused by many possible reasons, such as the limited spatial resolution, uniform meshing, or over-simplified geological zonation. A smaller RMSE in SC1 indicates that mHM is able to capture spatial heterogeneity and produce a more realistic groundwater recharge distribution. The errors in simulated heads (lower two graphs in Figure 7) show that most of simulated head errors are within an interval of ± 6 m in SC1, while most of errors are within a range of ± 10

Table 1. Estimates of Main hydraulic properties used for calibrated steady-state groundwater model in Naegelstedt catchmentmHM#OGS v1.0.

Geological units	Hydraulic conductivity [m/s]			Specific yield [-]	Specific storage [m^{-1}]
	Lower limit	Upper limit	Calibrated value [m/s]		
Middle Keuper (km) km	1.0×10^{-6}	5.5×10^{-3}	1.106×10^{-3} 1.493×10^{-5}	log-transformed -	1×10^{-6}
Lower Keuper (ku) ku	1.0×10^{-7}	3.4×10^{-3}	3.027×10^{-4} 5.164×10^{-4}	log-transformed -	1×10^{-6}
Upper Muschelkalk 1 (mo1)	8.0×10^{-8}	2.0×10^{-3}	4.230×10^{-6} 4.030×10^{-5}	log-transformed 0.10	1×10^{-6}
Middle Muschelkalk 1 (mm1)	1.0×10^{-7}	9.0×10^{-4}	6.390×10^{-7} 9.372×10^{-6}	log-transformed -	1×10^{-6}
Lower Muschelkalk 1 (mu1)	5.0×10^{-9}	2.0×10^{-4}	1.962×10^{-6} 2.297×10^{-6}	log-transformed -	1×10^{-6}
Upper Muschelkalk 2 (mo2)	1.0×10^{-8}	5.0×10^{-4}	4.230×10^{-6} 4.030×10^{-6}	tied with mo1 -	1×10^{-6}
Middle Muschelkalk 1 (mm2)	3.0×10^{-8}	9.0×10^{-5}	6.390×10^{-8} 9.372×10^{-7}	tied with mm1 -	1×10^{-6}
Lower Muschelkalk 1 (mu2)	5.0×10^{-10}	2.0×10^{-5}	1.962×10^{-7} 2.297×10^{-7}	tied with mu1 -	1×10^{-6}
soil	- 5.0×10^{-5}	- 1.0×10^{-2}	- 3.068×10^{-4}	tied with mHM 0.10	-
alluvium	1.0 4.0×10^{-4}	1.0×10^{-2}	1.026×10^{-3}	log-transformed 0.18	-

m in SC2. Nevertheless, there are still some flawed points where the prediction is biased. Adding more model complexity to improve the match between simulation and observation was avoided due to the large spatial scale, the limited spatial resolution of mesh, and the noise of groundwater head data (i.e., with a various time span from 10 years to 30 years).

Simulated water-table depth over Naegelstedt catchment.

Simulated water-table depth over the whole catchment using the calibrated K is shown in Figure ??. The simulation provides a reasonable distribution of the hydraulic head. Groundwater-table depth is as large as above 40 m in the higher southwestern and northern mountainous areas, whereas less than 5 m in the central lowlands from simulation result. This simulation result is coincident with regionalized observations of groundwater head (Wechsung, 2005).

Observed and modeled monthly streamflow at the outlet of Naegelstedt catchment.

4.4 Transient state groundwater simulation under different recharge scenarios

In this subsection, the model skill scores under different recharge scenarios were assessed by comparing the simulated groundwater head time series to the observations at 19 monitoring wells . In order to better display the difference in model results between SC1 and SC2, we plotted head observations at several monitoring wells at the catchment were used to evaluate the model performance. We analyze the discrepancies between the modeled groundwater heads and their corresponding observed values in their anomalies by removing long-term mean values $\overline{h}_{mod}$ and $\overline{h}_{obs}$. Four model skill scores including the mean value, the median value, the Pearson correlation coefficient R_{cor}, and the inter-quantile range error QRE are used to judge the model performance.

Figure 10 presents observed and simulated groundwater heads for the period 1975-2005 in SC1. Five out of 19 monitoring Five wells with different geological and morphological types were properties were randomly chosen as samples to test the effectiveness of our model. Well Specifically, well 4728230786 is located at northern upland and near the mainstream, whereas well 4828230754 is located at the southwestern lowland. Both of those two wells show promising simulation results As can be observed from Figure 10, they provide good fits between simulated heads and observed heads with the R_{cor} of 0.81247 and 0.752790.87 and 0.76, and the QRE of -11.555% and -22.792-23.34% and -1.65%, respectively. Well 4628230773 is located in lower Keuper sediment, while well 4728230781 is located at upper Muschelkalk sediment. For those two monitoring wells, simulation results are highly correlated with the observations with the R_{cor} of 0.72716 and 0.82345 0.71 and 0.81 in spite of their different geological properties (Figure 10). The simulation result at monitoring well 4728230783 located at northern mountainous area also has a high correlation with the observation with a R_{cor} of 0.7894800.85 (Figure 10). In general, the model is capable of capturing the historical trend of groundwater dynamics, even though the mean value values of simulation and observation values may differ slightlymay mismatch to some extent. Due to the limitation of limited spatial resolution and homogeneous K in each geological unit, this difference complex hydrogeological locality, this degree of discrepancy is acceptable.

4.4 Model sensitivity to different recharge scenarios

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 wet season (August 2002); f) Monthly mean groundwater head distribution in dry season (August 2003).

To compare the As described in Section 3.5, we set up a reference recharge scenario (RR) in which the recharge is homogeneously distributed in space to assess the effect of spatial recharge pattern towards groundwater hydrographs. For the purpose of showing discrepancies of model results between two recharge scenarios, we drew the barcharts compared the values of R_{cor} and |QRE| at each monitoring wells using two recharge scenarios, respectively (see well between mR and RR (Figure 11). The mean value and the median value of the basin scale R_{cor} and QRE are also calculated and shown in Figure 11. Figure 11(a) a indicates that the correlation with observations of simulations using SC1 between observations and simulations using mR is higher than that using SC2RR, with the averaged R_{cor} of 0.703 and 0.685vale 0.704 and 0.677, respectively. The standard deviation of R_{cor} in SC1 is 0.109, which is 13% smaller than 0.125 in SC2. Considering deviations are nearly the same in both scenarios. Considering that the only difference between SC1 and SC2 mR and RR is the spatial distribution pattern of recharge, the heterogeneous groundwater recharge estimated using mHM can be verified as a better evaluation than the homogeneous spatially distributed recharge. The relative difference we do conclude that the inclusion of spatially-distributed recharge provides a solid improvement. The individual discrepancies of R_{cor} between SC1 and SC2 is moderate, which indicates spatial characterization recharge distribution might has a less important influence to groundwater

~~dynamics in the study area. This phenomenon is highly related to the coarsest resolution describing meteorological forcings (e.g., precipitation) among three spatial levels used in mHM.~~ at well levels between the two scenarios are moderate.

Figure 11b shows the ~~distribution of the absolute value~~ absolute values of inter-quantile range error (|QRE|~~in SC1 and SC2. It can be found in Figure 11 that the distribution pattern~~) in simulations under two recharge scenarios (mR and RR), respectively. We found that the uncertainty of |QRE| is ~~more complicated~~ higher than R_{cor}~~. We can see that~~ , e.g., the |QRE| in two wells are abnormally higher than the other wells. ~~This indicates~~ The higher values of wells 4728230789 and 4627230764 may be caused by the proximity to model bounds, as the two wells locate near either river or outer bound. This phenomenon reveals the accurate quantification of ~~amplitude in particular~~ fluctuation amplitude in certain locations are difficult due to the complex ~~local hydrogeological properties. Another reason is we assign a homogeneous storage in all groundwater aquifers, which might be a over-simplified setting.~~ hydrogeological locality of individual wells. Nevertheless, 16 out of 19 wells ~~show a~~ have low inter-quantile range errors, with |QRE| ~~within~~ values within an interval of ±40% in ~~both scenarios~~mR scenario. We also ~~observed that the mean~~ observe a smaller mean value of |QRE| in ~~SC1 is 26.93%, which is smaller than 28.24% in SC2~~mR than in RR. The standard deviation of |QRE| in ~~SC1~~ mR is also slightly smaller than ~~SC2. Notice that those~~ RR. Those 19 monitoring wells cover geological zones of alluvium, Keuper, and Muschelkalk, and ~~ranges~~ range from high mountains to lowlands all over the catchment. These ~~facts evidence~~ results point towards the promising modeling capability of the model and highlight the ~~slightly better match in SC1~~moderately better simulation-to-observation match in mR.

~~The coupled model also shows its potential in predicting groundwater flood and drought.~~

Figure 12 displays the seasonality of groundwater heads over the whole catchment by means of calculating the long-term mean groundwater heads in spring, summer, autumn and winter, respectively. It indicates that in general, the possibility of groundwater flood ~~in spring & summer is~~ event in spring and summer are higher than in autumn ~~&~~ and winter. However, ~~the spatial variability within groundwater flood & drought event is significant. The groudwater head~~ there can be observed a strong spatial variability over different morphological classes and within different geological groups. For example, the fluctuation amplitudes of groundwater heads in northern, eastern and southeastern mountainous areas ~~tend to fluctuate more wildly~~ are larger than the central plain areas. ~~This phenomenon is consistent with the fact that the mountainous areas have a larger recharge rate than the plain areas.~~ Considering the need ~~of~~ for predicting groundwater flood ~~&~~ and drought in extreme climate events, we select a meteorologically wet month (August 2002) and a meteorologically dry month (August 2003), and show the groundwater heads variation in these months. Figure 12e ~~) and f)~~ and Figure12f show two scenes of groundwater head ~~variation~~ variations in wet season and dry season, respectively. ~~The~~ In general, the groundwater heads in ~~wet season is~~ the wet season are higher than the long term mean values (~~see~~ Figure 12e). The variation of groundwater heads in the dry season, however, shows a strong spatial variability. ~~The~~ Such a strong spatial variability of groundwater heads variation has also been ~~found in Kumar et al. (2016).~~reported, e.g., by Kumar et al. (2016).

5 Discussion and conclusions

A coupled hydrologic Our simulation results demonstrate that the coupled model mHM#OGS is proposed and applied it in a mesoscale catchment in central Germany. A boundary condition-based off-line coupling method is applied to depict the dynamic flow exchanges between the surface and subsurface water regimes. This coupling method, together with nested time stepping, allow the surface and subsurface parts to be solved sequentially, keep computational surplus to a minimum, while avoid possible water balance problem in the flow exchange processes.The result shows a promising prediction capability in surface and subsurface water modeling via calibration and comparison to groundwater time series. The SC1 using spatially heterogeneous groundwater recharge distribution is more plausible than SC2 which uses spatially homogeneous groundwater recharge. The v1.0 can reproduce groundwater heads 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 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 highlight demonstrate the successful application of Multiscale Parameter Regionalization (MPR) method in characterizing the well-established mHM in estimating spatial heterogeneous groundwater recharge and baseflow at regional scale. In the spatial scale of 10^3 km^2 (the scale in this study), the MPR method shows a moderate improvement in groundwater rechargerepresentation. Note that MPR distributed recharge estimated by mHM shows its priority over the reference homogeneous recharge. The mHM has been successfully applied at a larger scale over Europe (Thober et al., 2015; Kumar et al., 2 The effectiveness of MPR to characterize groundwater dynamics at larger scales (Thober et al., 2015; Kumar et al., 2013b; **?**; Rakovec et al The successful application of the coupled model in this study suggests a huge potential of extending the applicability of mHM#OGS v1.0 to a larger scale (e.g., 10^4 –10- 10^6 km^2) even global scales is still unknown and needs to be explored. Moreover, MPR has been proved its capability to produce better runoff prediction in cross-validated locations (e.g., ungauged basins) (Samaniego et al., 2010). To date, the effectiveness of MPR in unguaged basins has not been verified by groundwater head dynamics. In the next step, we may use the groundwater time series to test the effectiveness of MPR in ungauged basins using the coupled model mHM#OGS . or even global scale.

The convincing results of this study provide a new possibility in improving classic demonstrate a viable strategy for improving classic meso- to large-scale disbributed distributed hydrologic models, such as the current unmodified version of mHM (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 distributed hydrologic models do not include the function of calculating spatial-temporal groundwater heads , therefore are possibility of calculating spatio-temporal groundwater heads and are therefore not able to reasonably represent groundwater head and storage dynamics in their groundwater regime. This may be insignificant in global scale hydrologic modeling, which always has a coarser resolution of 25-50 km. compartment. The physical representation of groundwater flow isneeded in future global hydrologic model with finer spatial resolutions down to 1 km. Moreover, the inclusion of groundwater model OGS in the coupled model is particularly significant for areas with large sedimentary basins or deltas (e. g., the sedimentary basins of Mekong, Danube, Yangtze, Amazon, and Ganges-Brahmaputra Rivers). , 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 & drought in extreme climate events. and 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; **?**).

Due to the prediction capability of mHM in ungauged basins, the coupled model is also capable of predicting groundwater flood & drought in ungauged basins, which is quite valuable due to the lack of comprehensive groundwater observations at regional scales. Providing For example, building on previous work by Heße et al. (2017)in , who calculated Travel Time Distributions (TTDs) using mHM, we can now expand the range of their work to the complete hydrologic cycle beneath atmosphere, which is important due to the pollutantlegacy in groundwater storage. for comprehensively understanding particle (e.g., pollutant) transport behavior and historical legacy in soil zone and groundwater storage (**??**). mHM#OGS v1.0 fits well with the long-term simulation of nitrogen transport in 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 under different meteorological forcings, which allows the decent study comprehensive evaluation of hydrologic response to climate change (e.g. global warming). Besides, the versatility of OGS also offers the possibility to address Thermo-Hydro-Mechanical-Chemical (THMC) coupling processes in large-scale hydrologic cycles, 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; **?**).

We realize that there are several limitations in the current model . The first one is the fact that we use an off-line coupling scheme instead of a full coupling scheme between mHM and In addition to improving the predictive abilities of mHM, we could 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 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 this study, we focus our efforts on extending the mHM applicability in surface hydrology to subsurface hydrology in a simple one-way coupling . Consequently, we do not account for the explicit feedback from OGS to mHM, although it might be less important related to the large subsurface time step. This approach has certain advantages of 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 features of mHM, remains fully intact. That way, we do not compromise any of its well-established features, such as calibration of model parameters at different scales and good runoff prediction ability, while getting in addition very good estimates of groundwater storage, flow paths and travel times. The lack 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. (2016)) and extends therefore 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 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. Third, even in the future, a one-way coupling would allow to easily expand the predictive power of a mHM catchment model if the practitioners later decide to do so, therefore leaving the option open. Finally, one-way coupling takes less computational consumption and achieves better numerical stability~~, and fits perfectly on the long-term large-scale groundwater modeling in this study. In the~~ . 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 ~~may~~ will try to incorporate ~~the fullcoupling scheme in~~ a full, two-way coupling using the next version of mHM#OGS model. Via ~~the~~ such a full coupling scheme, the dynamic interactions between overland flow and groundwater flow, ~~and~~ as well as between soil moisture dynamics and groundwater dynamics ~~are explicitly accounted~~can be explicitly modeled and investigated. This approach is open to a broader spectrum of calibration options, such as calibration using remotely sensed soil moisture data. ~~The second one is that we do not use the parallel computing. Although the whole simulation is conducted on the EVE linux cluster at UFZ, which is a high performance computing platform, we do not use the distributed computing to reduce computational efforts. In the future, a parallel version of~~

In conclusion, we can state that the coupled model mHM#OGS v1.0 fully preserves the predictive capability of discharge of mHM. In addition, it proves to be capable of reproducing groundwater head dynamics. The simulation results show a promising prediction ability via calibration and comparison to 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 ~~is needed to reduce computation time for the computationally-expensive full coupling procedure.~~v1.0 is a valuable tool in coping with 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 can be freely acquired via the following link: https://github.com/UFZ-MJ/OGS_mHM/tree/master/UTL/GIS2FEM.

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. ~~Miao Jing was funded by the Chinese Scholarship Council (CSC). We kinkly~~ 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.

Attinger, S.: Generalized coarse graining procedures for flow in porous media, Computational Geosciences, 7, 253–273, 2003.

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.

Barthel, R. and Banzhaf, S.: Groundwater and Surface Water Interaction at the Regional-scale – A Review with Focus on Regional Integrated Models, Water Resources Management, doi:10.1007/s11269-015-1163-z, http://link.springer.com/10.1007/s11269-015-1163-z, 2015.

Batu, V.: Aquifer hydraulics: a comprehensive guide to hydrogeologic data analysis, vol. 1, John Wiley & Sons, 1998.

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.

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.

Brian A. Ebel: First-order exchange coefficient coupling for simulating surface water–groundwater interactions: parameter sensitivity and consistency with a physics-based approach, Hydrological Processes, 2274, 2267–2274, doi:10.1002/hyp, http://jamsb.austms.org.au/courses/CSC2408/semester3/resources/ldp/abs-guide.pdf, 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.

Cunge, J. A.: On the subject of a flood propagation computation method (Musklngum method), Journal of Hydraulic Research, 7, 205–230, 1969.

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.: Statistical theory of groundwater flow and transport: Pore to laboratory, laboratory to formation, and formation to regional scale, Water Resources Research, 22, 1986.

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.

De Graaf, I. E. M., Sutanudjaja, E. H., Van Beek, L. P. H., and Bierkens, M. F. P.: A high-resolution global-scale groundwater model, Hydrology and Earth System Sciences, 19, 823–837, doi:10.5194/hess-19-823-2015, 2015.

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.

Fleckenstein, J. H., Krause, S., Hannah, D. M., and Boano, F.: Groundwater-surface water interactions: New methods and models to improve understanding of processes and dynamics, Advances in Water Resources, 33, 1291–1295, doi:10.1016/j.advwatres.2010.09.011, http://dx.doi.org/10.1016/j.advwatres.2010.09.011, 2010.

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 graoundwater 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., Prykhodko, V., and Attinger, S.: Assessing the validity of a lower-dimensional representation of fractures for numerical and analytical investigations, Advances in Water Resources, 56, 35–48, doi:10.1016/j.advwatres.2013.03.003, 2013.

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.

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.

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.

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., 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, 2016.

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

Renard, P. and de Marsily, G.: Calculating equivalent permeability: a review, Advances in Water Resources, 20, 253–278, doi:10.1016/S0309-1708(96)00050-4, 1997.

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., and Zink, M.: Implications of parameter uncertainty on soil moisture drought analysis in Germany, Journal of Hydrometeorology, 14, 47–68, 2013.

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.

Te Chow, V.: Applied hydrology, Tata McGraw-Hill Education, 1988.

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.

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.

Zech, A., Zehner, B., Kolditz, O., and Attinger, S.: Impact of Heterogeneous Permeability Distribution on the Groundwater Flow Systems of a Small Sedimentary Basin, Journal of Hydrology, 532, 90–101, doi:10.1016/j.jhydrol.2015.11.030, http://linkinghub.elsevier.com/retrieve/pii/S0022169415009117, 2015.

Zink, M., Samaniego, L., Kumar, R., Thober, S., Mai, J., Schäfer, D., and Marx, A.: The German drought monitor, Environmental Research Letters, 11, 074 002, 2016.

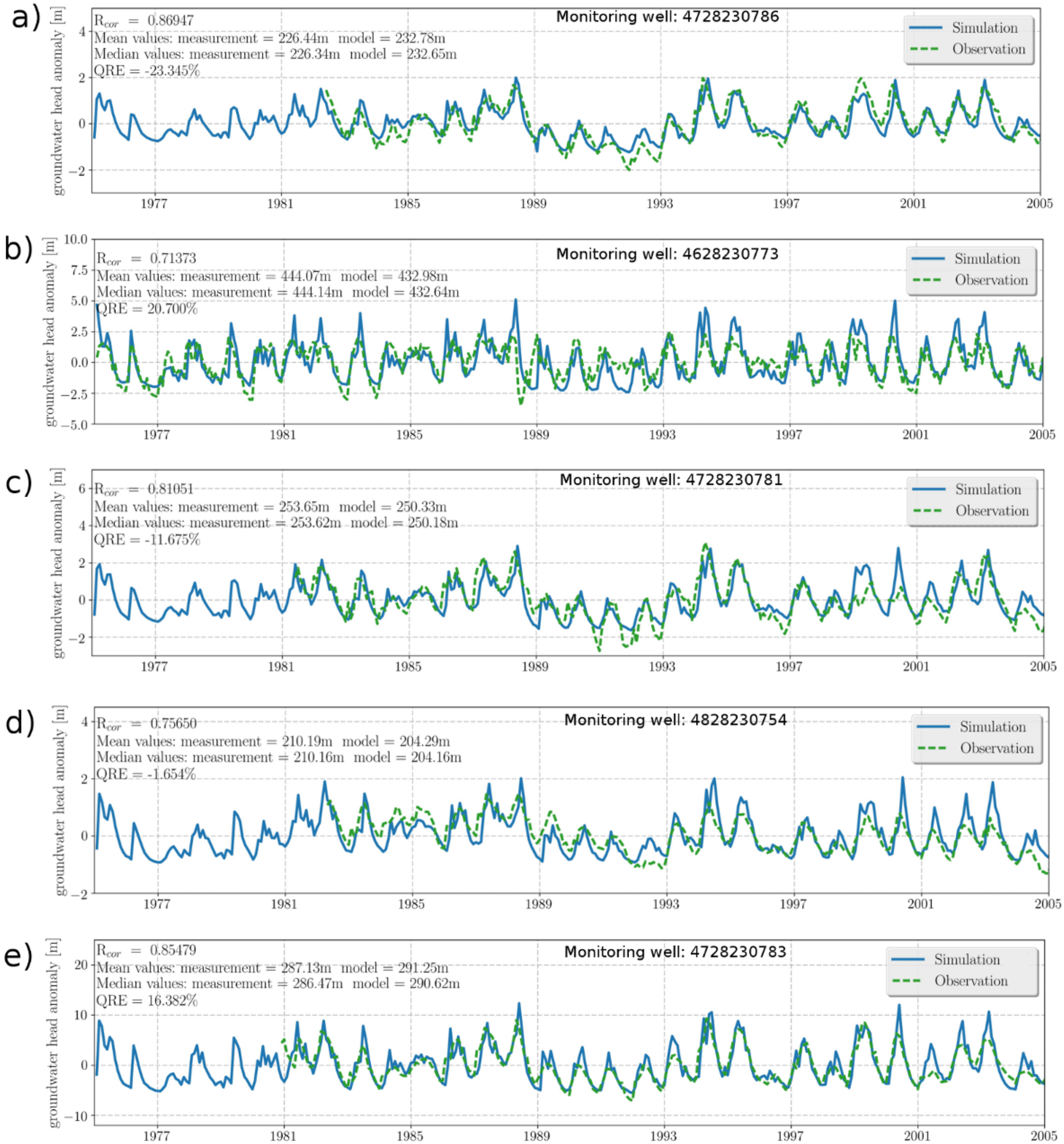

Figure 10. The comparison between measurement data (green dashed line) and model output of groundwater head anomaly ~~in SC1~~ (blue solid line). (a) Monitoring well 4728230786 located at upland near stream. (b) Monitoring well 4628230773 located at mountainous area. (c) Monitoring well 4728230781 located at a hillslope at northern upland. (d) Monitoring well 4828230754 located at lowland. (e) Monitoring well 4728230783 located at northern mountain.

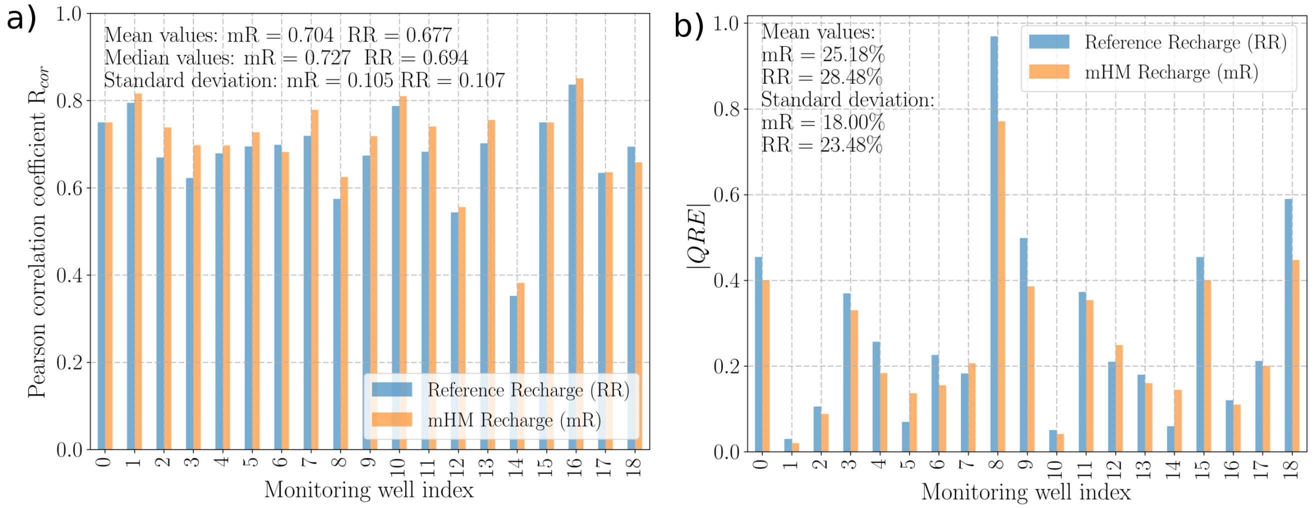

Figure 11. Barplots of a) the Pearson correlation coefficient R_{cor} and b) absolute inter-quantile range error |QRE| in all monitoring wells in two recharge scenarios. Each bar corresponds to an individual monitoring well in the following order: 0 - 4830230779, 1 - 4828230754, 2 - 4828230752, 3 - 4828230753, 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.

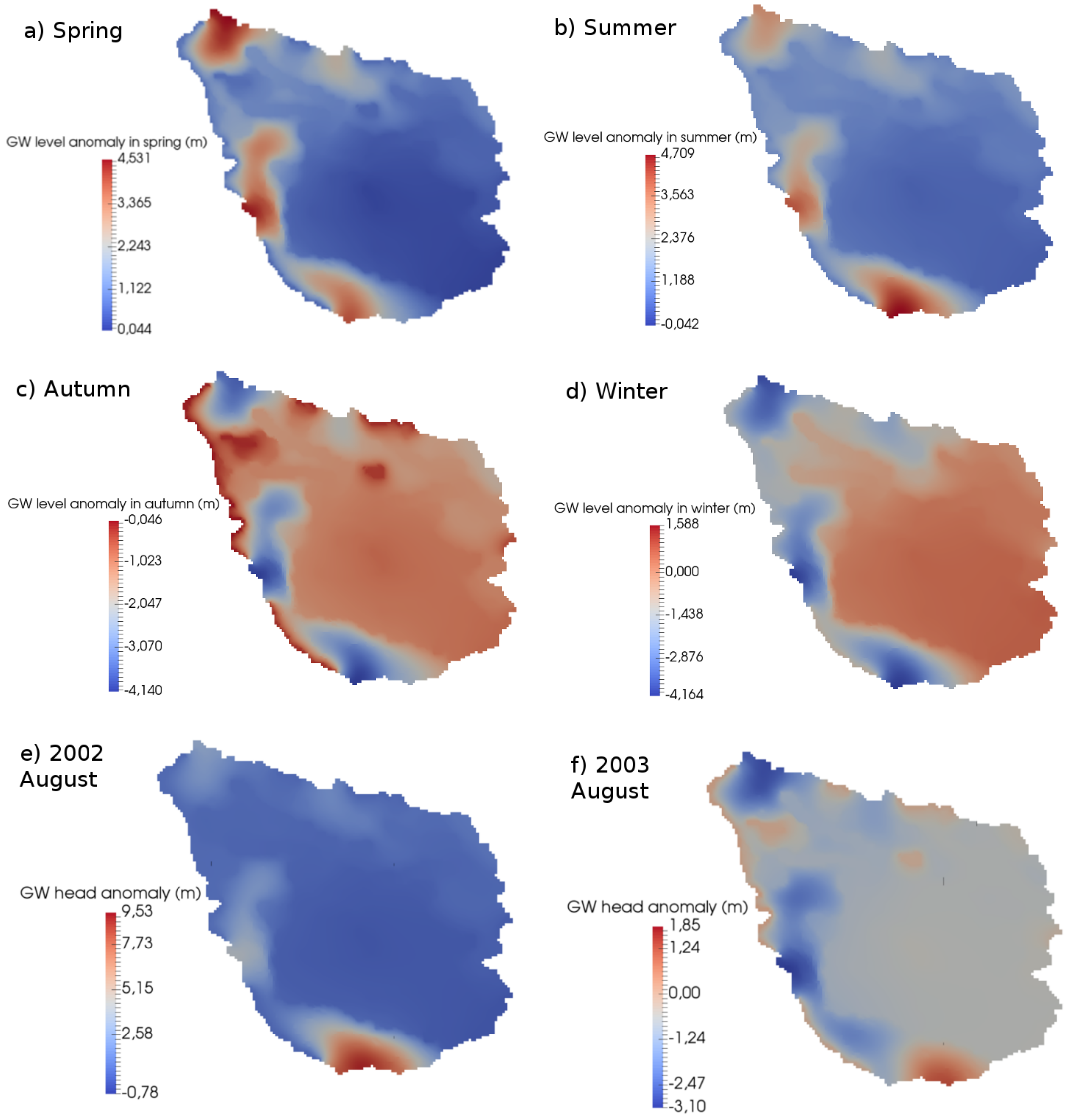

Figure 12. Seasonal variation of spatially-distributed groundwater heads by their anomalies after removing the long-term mean groundwater heads (unit: m). a) Long-term mean groundwater head distribution in spring; b) Long-term mean groundwater head distribution in summer; c) Long-term mean groundwater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