



## RESEARCH ARTICLE

10.1029/2019WR026004

Hyper-Resolution Continental-Scale 3-D Aquifer  
Parameterization for Groundwater Modeling

## Key Points:

- We integrate local hydrogeological information into a continental-scale aquifer parameterization for the continental United States and Canada
- Aquifer thickness and vertical structure are estimated at a level of detail not available before
- Hydrological modeling confirmed accurate large-scale aquifer parameterization will improve our understanding of the hydrological system

## Supporting Information:

- Supporting Information S1

## Correspondence to:

I. de Graaf,  
inge.de.graaf@hydro.uni-freiburg.de

## Citation:

de Graaf, I., Condon, L., & Maxwell, R. (2020). Hyper-resolution continental-scale 3-D aquifer parameterization for groundwater modeling. *Water Resources Research*, 56, e2019WR026004. <https://doi.org/10.1029/2019WR026004>

Received 19 JUL 2019

Accepted 9 MAR 2020

Accepted article online 4 APR 2020

Inge de Graaf<sup>1,2</sup> , Laura Condon<sup>3</sup> , and Reed Maxwell<sup>2</sup>

<sup>1</sup>Chair of Environmental Hydrological Systems, Faculty of Environmental and Natural Resources, University of Freiburg, Freiburg, Germany, <sup>2</sup>Department of Geology and Geological Engineering, Colorado School of Mines, Golden, CO, USA, <sup>3</sup>Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

**Abstract** Groundwater is the world's most important freshwater resource. Despite this importance, groundwater flow and interactions between groundwater and other parts of the hydrological cycle are often neglected or simplified in large-scale hydrological models. One of the challenges in simulating groundwater flow and continental to global scales is the lack of consistent globally available hydrogeological data. These input data are needed for a more realistic physical representation of the groundwater system, enabling the simulation of groundwater head dynamics and lateral flows. A realistic representation of the subsurface is especially important as large-scale hydrological models move to finer resolutions and aim to provide accurate and locally relevant hydrologic information everywhere. In this study, we aim at improving and extending on current available large-scale data sets providing information of the subsurface. We present a detailed aquifer representation for the continental United States and Canada at hyper resolution (250 × 250 m). We integrate local hydrogeological information, including observations of aquifer layer thickness, conductivity, and vertical structure, to obtain representative aquifer parameter values applicable to the continental scale. The methods used are simple and can be expanded to other parts of the world. Hydrological simulations were performed using the integrated hydrological model ParFlow and demonstrated improved model performance when using the new aquifer parameterization. Our results support that more detailed and accurate aquifer parameterization will advance our understanding of the groundwater system at larger scales.

**Plain Language Summary** Groundwater is the largest available freshwater resource humans can use for drinking and growing food. It is stored beneath our feet in thick layers of sand or rock; we call these layers aquifers. For a realistic estimation of how much groundwater is stored and flows through these aquifers, detailed information on the properties of these layers is essential. However, this information is often missing, or not detailed enough, at the larger scales. In this study, we introduce a new method to obtain this detailed data by including information from observations and local-scale studies what we upscaled to the continent of North America. We estimate aquifer parameters at very high spatial resolution over a large domain. These data were not available previously. We evaluate our newly obtained aquifer parameterization against other existing data sets and tested model sensitivity to different horizontal and vertical resolutions and conductivity values. We used the integrated hydrological model ParFlow. Our results show that estimates of water table depth improved using our new parameterization. The methods we introduce here can be used by other modelers looking to improve their aquifer parameterization. We hope that our study can be used as a road map towards improved and more detailed aquifer parameterization used in current large-scale hydrological models.

## 1. Introduction

Groundwater is an important part of the water cycle. It is widely used for irrigation, drinking water, and industries (Döll et al., 2012; UNESCO, 2010; Wada et al., 2011). During times of droughts, groundwater acts as a natural buffer as it sustains groundwater flow to the river, supporting river low flows, groundwater-fed wetlands, and related ecosystems (de Graaf et al., 2014, 2019) and shallow water tables maintaining evapotranspiration (Chang et al., 2018; Maxwell & Condon, 2016; Shrestha et al., 2018). Groundwater often flows across topographic and administrative boundaries at considerable rates, supporting water budgets of receiving catchments (Schaller & Fan, 2009). With climate change and increasing human water demands,

©2020. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

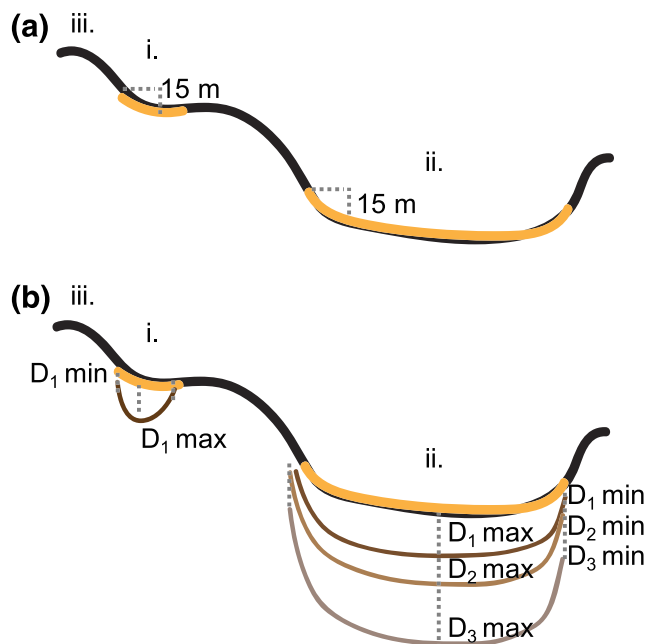
groundwater resources are under continuous pressure. Therefore, a better understanding of the groundwater system and its sensitivity to storage changes, be it from climate or human water use, is critical.

Despite the importance of groundwater and the explicit focus on groundwater recharge and groundwater pumping in global-scale studies (e.g., de Graaf et al., 2014; Döll et al., 2012), most global-scale hydrological models do not simulate groundwater flow. The main reason for this omission is the lack of consistent globally available hydrogeological data that are required for a realistic physical representation of the groundwater system (Gleeson et al., 2014), enabling the simulation of groundwater head dynamics, lateral flows, and the effect of storage changes, be it from climate or groundwater pumping. Also, consistent data are needed for model evaluation and calibration. The data required includes estimates of aquifer thickness, conductivities, and presence of confining layers (de Graaf et al., 2017; Schaller & Fan, 2009). Especially at a finer resolution, the relative importance of horizontal groundwater flux (or lateral flow) will increase in large-scale hydrological assessments (Krakauer et al., 2014).

Currently, few gridded data sets are available that can be used to define hydrogeological parameters. The best currently available global data set on aquifer permeability is GLHYMPS (Gleeson et al., 2014), which was updated to the GLHYMPS2.0 (Huscroft et al., 2018). GLHYMPS is widely used in large-scale groundwater models, although it shows some major inconsistencies like state and country borders, and tends to be biased towards low conductivities, especially for the unconsolidated sediments (Gleeson et al., 2014). The latter was subject of the improvement made in the 2.0 version. When used for groundwater modeling, the inconsistencies and biases impact the reliability of simulated water table depths and fluctuations (de Graaf et al., 2017, 2019; Reinecke et al., 2019). Also, there are a few data sets available providing a depth-to-bedrock estimate (Pelletier et al., 2016; Shangguan et al., 2017), which could be used as a proxy for aquifer thickness. However, both mentioned data sets are limited in depth. For example, Pelletier et al. (2016) estimates average soil and sediment thickness at approximately  $1 \times 1$  km resolution based on topography and limited to a maximum 50 m in depth. Values equal to 50 m could therefore also represent higher thicknesses. Using this limit in a groundwater model is a bit arbitrary, as many aquifer systems are much deeper than 50 m and store groundwater at greater depths as well. Also, groundwater abstractions are often from greater depths, for example, up to 300 m in depth for the United States (Perrone & Jasechko, 2017). The data set of Shangguan et al. (2017) provides an estimate of depth-to-bedrock, equivalent to the thickness of productive aquifers, at  $250 \times 250$  m resolution, based on a purely statistical approach using random forest and gradient boosting tree algorithms. Thickness estimates are limited to 150 m for regions where no borehole data were compiled (assumed for the entire Sahara for example). The maximum predicted thickness is 540 m in this data set. Furthermore, in both depth-to-bedrock estimates, no information on the vertical structure, for example, the presence of confining layers, is provided. In groundwater modeling, especially when impacts of groundwater abstractions are to be simulated and studied, the vertical structure and occurrence of confined and unconfined aquifers is vital information.

Another global-scale estimate of aquifer thickness was presented by de Graaf et al. (2017) and includes estimates of aquifer thickness, vertical structure, and aquifer transmissivity. This hydrogeological information was used in a groundwater model to estimate groundwater depths and head declines due to changes in storage (de Graaf et al., 2017), as well as to estimate the impact of groundwater pumping on river flows (de Graaf et al., 2019). Evaluation of the estimated aquifer thickness and groundwater depths showed however that the spatial resolution, which was approximately  $10 \times 10$  km, is too coarse to resolve local aquifer systems in the mountainous systems. For these regions, groundwater table depths and interactions to the surface water are not always simulated accurately (de Graaf et al., 2017, 2019). Furthermore, complex vertical structures of multiple confining layers and confined aquifers are lumped together into one confining layer on top of a confined aquifer. Rivers in confining layers are thus disconnected from the aquifer unless they penetrate through the confining layer. For some intensively abstracted aquifers, groundwater depletion was significantly overestimated compared to observations due to the uncertainty in the used permeability data set (which was GLHYMPS) (de Graaf et al., 2017).

In this study, we address the noted gaps in existing data sets by presenting an estimate of aquifer thickness, vertical structure, and aquifer transmissivity for the continental United States and Canada at high- and hyper-resolution ( $1 \times 1$  km and  $250 \times 250$  m, respectively). We extrapolated data on aquifer thickness and conductivity from aquifer studies in the United States and used a relatively simple method, predominantly based on elevation differences, to delineate local and extensive alluvial aquifers and estimate thicknesses



**Figure 1.** (a) Schematization of delineation of (i) local alluvial aquifers, (ii) extensive alluvial aquifer, and (iii) mountain ranges, using predominantly elevation differences. (b) Schematization of aquifer thickness ( $D$ ) estimation for local alluvial aquifers and vertically stacked layers of extensive alluvial aquifers (indicated by the numbers 1–3).

of these aquifer systems. For the extensive aquifer systems, this estimate entails an approximation of thickness of different vertical stacked layers with varying transmissivity, therewith providing information on the vertical structure of the aquifer system. Aquifer conductivity (needed to calculate transmissivities) values were extrapolated from the same aquifer studies in the United States and assigned to the estimated aquifer layers using regional information as wherever it is available. The approach used to estimate aquifer thickness and transmissivities can be extended to other parts of the world. Although, including more specific information for, for example, a region or aquifer will improve the accuracy of the estimate, for regions where such information is missing, the presented approach provides a useful first-order estimate of the subsurface. Results are evaluated against available data sets of depth-to-bedrock and aquifer thickness at a global scale.

We tested the new aquifer parameterization using the hydrological model ParFlow. The model ParFlow has been used at a larger-scale before (Maxwell et al., 2015) and simulates water table depths and groundwater-surface water interactions with Richards equation. In this study, the model was set up for a test case region in the northern High Plains aquifer. We tested model sensitivity for different parameter settings of conductivity and different horizontal and vertical resolutions. Results are evaluated and compared with a focus on water table depths and groundwater flows. This analysis is limited to a small test case due to the high computational demands and therefore should be seen as a proof of concept that can be used to guide future modeling efforts at larger spatial scales and higher resolutions.

## 2. Methods

The first goal in this study was to develop realistic and detailed hydrogeological parameterization at high- and hyper resolution for continental United States and Canada. For the aquifer thicknesses estimation and the delineation of confining layers, we build on previously published methods applied at the global scale published by de Graaf et al. (2017). Here, major improvements from this data set were obtained by increasing the spatial resolution and by the extrapolation of regional-scale information on thickness, vertical structure, and aquifer conductivities at the continental scale. Details on the estimation of aquifer thickness, vertical structure, and aquifer conductivity are given in sections 2.1 and 2.2

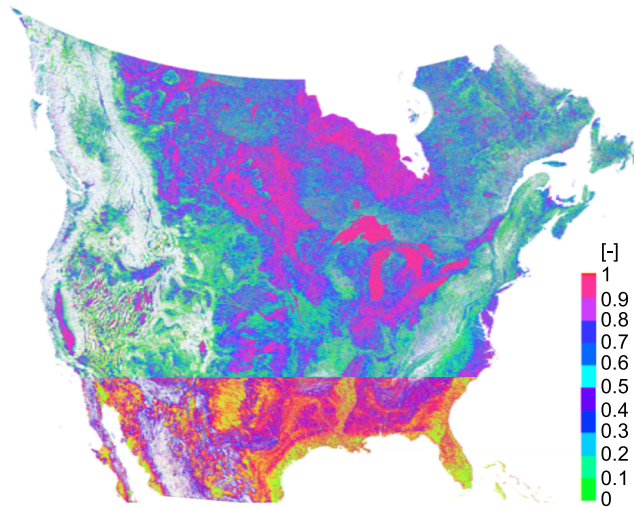
The second goal in this study was to test the new aquifer parameterization hydrologically using the integrated hydrological model ParFlow (Kollet & Maxwell, 2006; Maxwell, 2013) and test the model sensitivity for different conductivity parameterization and horizontal and vertical resolutions. We selected a smaller test case region within the High Plains aquifer to keep computational times reasonable. Also, local data on subsurface permeability as well as observations of water table depths are available, enabling a detailed model evaluation. Details on the model and sensitivity analysis are given in sections 2.3 and 2.4.

### 2.1. Aquifer Thickness

We estimated aquifer thickness and vertical structure at high- and hyper-resolution ( $1 \times 1$  km and  $250 \times 250$  m, respectively). First, we delineated alluvial aquifers and mountain ranges, preliminary based on elevation. We distinguished between (i) local alluvial aquifer systems, present in valleys of all streams, (ii) extensive alluvial aquifer systems, located in the major floodplains of deltas and piedmont belts, and (iii) mountain ranges, where the sediment thickness is negligible (Margat & van der Gun, 2013) (Figure 1a). Second, we estimated the thickness of local systems (i.e., type i) and upper and deeper layers of extensive systems (i.e., type ii) of the alluvial aquifers by applying a simple algorithm predominantly using terrain attributes (Figure 1b). The algorithm is adopted and improved from the studies of de Graaf et al. (2015, 2017).

The following steps were taken:

1. Local alluvial aquifers, extensive alluvial aquifers, and mountain ranges were delineated based on elevation difference between a model cell and the downstream neighboring cell. For this step, terrain elevation



**Figure 2.** Map of calculated spatial distribution of  $F'(x)$ , representing the likelihood of thinner or thicker aquifer layers.

and the drainage direction network were needed. Both these data sets are available at the high- and hyper-resolution from HydroSHEDS (Lehner & Döll, 2004). All cells with an elevation difference of 15 m or less compared to their downstream neighboring cell were classified as alluvial aquifer cells. Cells with an elevation difference of more than 15 m were classified as mountain ranges (Figure 1a) (For evaluation, see section 3.1). Moreover, we made sure that aquifer cells have at least one neighboring aquifer cell.

2. By definition, alluvial aquifers are linked to fluvial systems and deltas. Sediments are deposited perpendicular to the main gradient (constituting the transversal axis of the basin) with sediment grain sizes and volumes decreasing at a greater distance away from the transversal axis. Therefore, we used the relative elevation from the main gradient as a measure of proximity to the river measured along the transversal axis and as an indicator of the associated depth (following de Graaf et al., 2015). We standardized the relative elevation and used this to define the distribution of aquifer thickness using a log-normal distribution, assuming thickness is non-negative and positively skewed (following de Graaf et al., 2015). In more detail, this was done using the following procedure:

First, for each cell-location  $x$ , classified as an alluvial aquifer in step 1, a measure expressing the relative elevation difference was calculated:

$$F'(x) = 1 - \frac{F(x) - F_{min}}{F_{max} - F_{min}} \quad (1)$$

where  $F(x)$  is the elevation difference between a cell at location  $x$  and the neighboring downstream cell.  $F_{min}$  and  $F_{max}$  are the minimum and maximum difference following from the distinction between alluvial aquifers and mountain ranges: 0 m and 15 m, respectively. This  $F'(x)$  (Figure 2) leads to a thick layer close to the river which thins further away from the stream towards the edge of the alluvial aquifer (Figure 1b).

Second, the associated  $z$ -score was calculated as

$$Z(x) = G^{-1}(F'(x)) \quad (2)$$

where  $G^{-1}$  is the inverse standard normal distribution.

3. Statistics of local alluvial aquifer thickness and shallow layers of extensive alluvial aquifers were obtained from local- to regional-scale groundwater studies in the United States (covering 23 different aquifer systems, available from Miller (2000)). As a measure of the difference between aquifer systems, we determined for each study the average thickness. This resulted in a range of averaged thicknesses between 25 and 135 m. Assuming that the thickness of alluvial aquifer layers is log-normally distributed (i.e., aquifer layers cannot be thinner than 0 m), aquifer thicknesses can be described using the average thickness of the  $\ln$ -transformed thickness  $\ln D$ . This  $\ln D$  was chosen uniformly over the domain and sampled from the range of thicknesses:

$$\overline{\ln D} = U(25 : 135) \quad (3)$$

where  $U$  indicates the random value. Moreover, as a measure of the variation of thickness within an aquifer system, the average coefficient of determination ( $Cv_{\overline{\ln D}}$ ) was determined from the same aquifer studies. This coefficient of variation was fixed when calculating the aquifer thickness distribution over the domain.

4. Spatial distribution of aquifer thickness was generated, assuming a log-normal distribution with random average  $\ln D$ , sampled from  $U(25 : 135)$  with a fixed  $Cv_{\overline{\ln D}}$  and using the standard normal ordinate  $Z(x)$  which is based on the topographical controls within each delineated aquifer:

$$Y(x) = \overline{\ln D}(1 + Cv_{\overline{\ln D}}Z(x)) \quad (4)$$

$$D(x) = e^{Y(x)}. \quad (5)$$

The parameters  $Y(x)$  and  $D(x)$  are random as  $\ln D$  is random, while spatial variation is driven by  $Z(x)$ , reflecting the likelihood of a thick aquifer layer. Here, 20 equally likely maps of aquifer thickness were estimated by choosing average aquifer thickness randomly from  $U(25 : 135)$ . The map that matches the reported values best is presented as the final result in this paper.

5. Lastly, statistics of the thickness of deeper, vertically stacked aquifer layers, of extensive alluvial aquifers (type ii), were obtained from the same local- to regional-scale groundwater studies in the United States. Extensive alluvial aquifers are found for major deltas, piedmont belts, coastal areas, and subsidence zones and were delineated assuming that these aquifers are represented by more extended flat areas. The same procedure as described in steps 2–4 was used to estimate the thickness of the different layers (Figure 1b). In this study, we estimated three vertical stacked layers; a shallow aquifer layer as described above, a deeper layer which can be a confining unit or a deeper aquifer layer, and a third layer that represents all layers below the second estimate layer part of the productive aquifer until the first impermeable layer. In the estimation of the layer thicknesses, the  $F^I(x)$  (equation (1)) was recalculated for extensive alluvial aquifer systems, and  $\ln D$  was sampled from the range of thickness corresponding the layer and based on data obtained from the aquifer studies.

To evaluate the reliability of this study's aquifer-thickness estimate, the spatial distribution of aquifers and thickness estimates are evaluated against available global and continental data sets of depth-to-bedrock.

## 2.2. Aquifer Conductivity

Groundwater flow volumes are dependent on aquifer transmissivities ( $m^2/d$ ), calculated as the conductivity ( $m/d$ )  $\times$  thickness ( $m$ ) of the aquifer (layer). Data on aquifer conductivities are very limited at a larger scale. Currently, only two data sets on conductivity are available: GLHYMPS (Gleeson et al., 2014), providing permeability of the upper 100 m of the subsurface, and GLHYMPS2.0 (Huscroft et al., 2018), an updated version of the first data set providing permeability values for unconsolidated sediments and upper bedrock layers. These data sets are a combination of global maps of surface lithology and a compilation of 250 model studies with calibrated values for different lithologies.

Based on previous model experiences (Condon & Maxwell, 2014; de Graaf et al., 2017, 2019) and test runs done in this study, we have strong evidence that current global-scale subsurface permeability values, presented in GLHYMPS, tend to be biased towards lower permeability (also supported by the discussion of the GLHYMPS data sets Gleeson et al., 2014; Huscroft et al., 2018). Also, the global permeability maps suffer from inconsistencies, for example, over country and state borders were permeabilities shifts up to an order of magnitude. The GLHYMPS2.0 provides permeabilities that are higher than the original values but still show the same inconsistencies as it uses the same surface lithology map (GLiM Hartmann & Moosdorf, 2012). These inconsistencies have a major impact on model results of water table depths, water table fluctuations, and head declines when storage is changed (de Graaf et al., 2017; Maxwell et al., 2015; Reinecke et al., 2019).

In this study, we extrapolated aquifer permeability values from reported data available from aquifer studies in the United States (Miller, 2000). The data covered unconsolidated sediments, consolidated sediments, and carbonate rocks. For the unconsolidated and consolidated values, we distinguish between fine-, mixed-, and course-grained (see Table 1 and Figure A1). In total, data from 15 studies were used. The estimates are evaluated against both GLHYMPS data sets.

For the calculation of aquifer transmissivity, we used the new estimated thicknesses and conductivity values. We did not use the subsurface lithology map (i.e., GLiM) to define where, for example, unconsolidated

**Table 1**

*Estimated Conductivity Values Using Information From Aquifer Studies, Compared to Available Global-Scale Estimates*

Hydrogeology	This study's conductivity (m/d)	GLHYMPS1.0 conductivity (m/d)	GLHYMPS2.0 conductivity (m/d)
Course-grained unconsolidated sediments	30	11	26
Mixed-grained unconsolidated sediments	3	0.09	2
Fine-grained unconsolidated sediments	0.03	0.01	0.007
Coarse-grained siliciclastic sediments	1.5	0.27	-
Mixed-grained siliciclastic sediments	0.3	0.0005	-
Carbonates	43	1.4	-

*Note.* Only the values of hydrogeological classes that changed compared to GLHYMPS are given (for the other hydrogeological classes the original GLHYMPS values were assumed). As GLHYMPS2.0 only provides new estimates for unconsolidated sediments, only these values are given here.

sediments are found (like done in de Graaf et al., 2015, 2017 and Hutcroft et al., 2018), but assumed all top layers of the aquifer consist of unconsolidated sediments. Therewith, we overcome the inconsistencies present in the lithology map and both GLHYMPS maps. However, we used information in the lithology map on grain size (fine-mixed-coarse), as this information is often missing, and assigned conductivity values accordingly. For the deeper layers, if information on the hydrogeological class was available, the corresponding conductivity value was assigned. If no information was available, the value for fine-grained unconsolidated sediments was assigned in case of a confining layer, and the value of mixed-grained unconsolidated was assigned in case of a deeper aquifer layer. If information was available of rock type below local alluvial aquifers, the corresponding conductivity value was assigned. If information was missing and for mountain ranges, the original value of GLHYMPS was assigned. The total transmissivity of all aquifer layers was calculated. The total transmissivity of extensive aquifers was calculated as the sum of transmissivities. The estimated aquifer transmissivities were evaluated against the estimate using the conductivity values of GLHYMPS (a similar approach currently used in, e.g., de Graaf et al., 2015, and Reinecke et al., 2019).

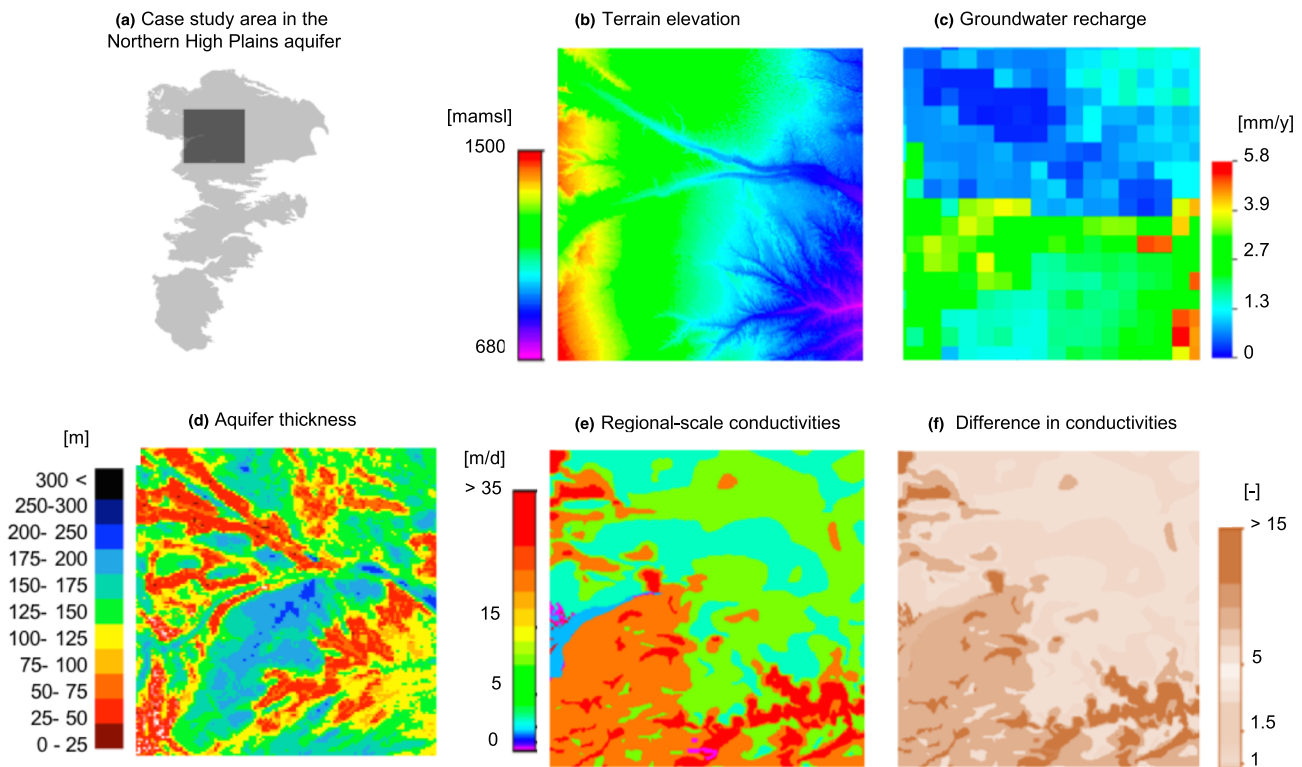
### 2.3. Hydrological Model

Second, we tested the sensitivity of simulated water table heads and fluxes to different parameter settings for conductivity, vertical discretization, and horizontal resolution by running the hydrological model ParFlow under different parameter scenarios. In this study, we use ParFlow to simulate saturated flow in three dimensions, using a terrain-following grid. The grid describes in this study shallow and deeper aquifer systems and bedrock layers. The model simulates subsurface flow solving Richards' equations. ParFlow allows the interconnected groundwater-surface water systems to evolve dynamically based on the governing equations and system properties and does not use simplifying assumptions like more conceptual-based models would do (e.g., predefine the stream network and/or simplified relations between stream head and baseflow). This approach, however, requires a high amount of computational time and memory storage in comparison to more conceptually based models. Details on the case study formulated here are given below. For more details on the model ParFlow (i.e., equations, code, etc.), we refer to the ParFlow manual (Maxwell et al., 2019).

### 2.4. Test Case Region and Model Runs

For the test case, we selected a region of 25 × 250 km in the Northern High Plains (Figure 3a). The High Plain aquifer generally consists of alluvial sediments on top of sedimentary hardrock. The thickness of the top layers varies between a few meters to hundreds of meters where ancient valleys got filled. The depth of the alluvial sediments is generally greater in the Northern High Plains to the middle and south (Gutentag et al., 1984). Our thickness estimate ranges between 10 and 300 < m (Figure 3d). Groundwater in the High Plains aquifer generally flows from west to east following the general terrain (Gutentag et al., 1984) (Figure 3b). In the Northern High Plains, water table depths are reported to range in depth between 0 and 120 m deep (McGuire, 2014).

We formulated two scenarios for the aquifer conductivities: (1) using this study's conductivity values, available at the continental scale at high- and hyper-resolution (hereafter called continental conductivities), and (2) using a data set of hydraulic conductivity contours and polygons for the Northern High Plains (Cederstrand & Becker, 1998) (Figure 3e; hereafter called regional conductivities). This Northern High Plain data set provided spatial heterogeneous data for the upper alluvial aquifer layer only, a value of the underlying hard rock was not given. Most of the regional conductivity values are 1.5 to 3 times higher than the



**Figure 3.** (a) Spatial boundary of the High Plains aquifer, with the case study location indicated by the dark grey box. Model inputs: (b) terrain elevation (meters above sea level, obtained from Lehner & Döll, 2004), (c) groundwater recharge input, obtained from Condon and Maxwell (2014), (d) estimated aquifer thickness (see also Figure 4), (e) regional conductivities of the top aquifer layer (obtained from Cederstrand & Becker, 1998), (f) fractional difference between this study's estimated continental conductivities and the regional conductivities.

continental conductivity (Figure 3f, calculated as regional conductivities/continental conductivities). In the selection of the case study area, we limited ourselves to regions where local data sets are freely available and ready to use (as a shapefile, or raster) and describe aquifer conductivities over at least an area of  $10 \times 10$  km. Even in the United States, this kind of data is limited.

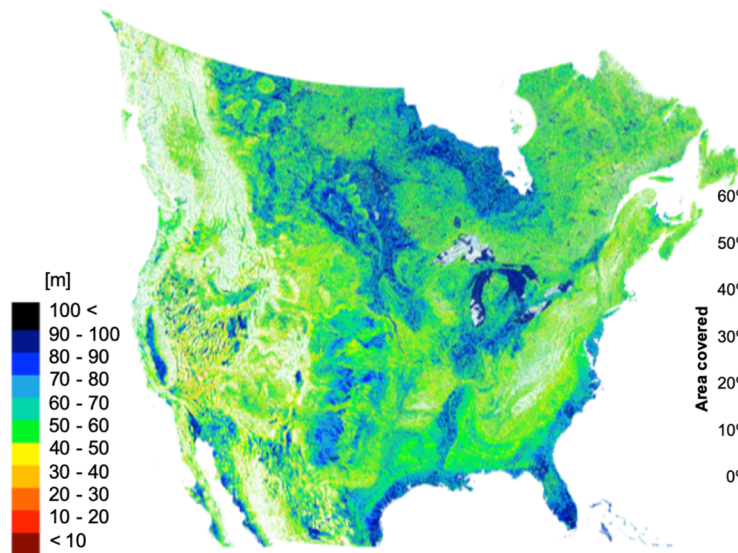
We formulated two scenarios defining the level of detail on which vertical structure is described: using six or 12 model layers. The number of layers was chosen so that the estimated vertical structure of the aquifer was not averaged out because of the terrain-following grid used by ParFlow. This means that the model's vertical discretization should not be too coarse that thinner layers are lumped together with thicker layers so that important information is lost (e.g., confining layers between two permeable layers or thin local aquifer systems). On the other hand, adding more layers will increase computational time. The formulated scenarios leave us with a minimum of 375,000 grid cells and a maximum of 12,000,000 grid cells. For comparison, the latter is almost twice as much as the current ParFlow domain covering the largest part of the continental United States, run at a  $1 \times 1$  km resolution and one aquifer layer (Maxwell et al., 2015). For each model layer, average conductivities were estimated.

In total, five runs were done: at the  $1 \times 1$  km spatial resolution, two conductivity estimates times two vertical discretizations. The total vertical domain was 300 m. We forced the model with simulated groundwater recharge obtained from the previous study of Maxwell et al. (2015) (Figure 3c). This groundwater recharge estimate is available at the continental scale at a  $1 \times 1$  km resolution. Because of the high computational demand, we run the model at the  $250 \times 250$  m resolution once, using the best performing parameter set (discussed in section 2.5). All runs are done in steady state under natural conditions.

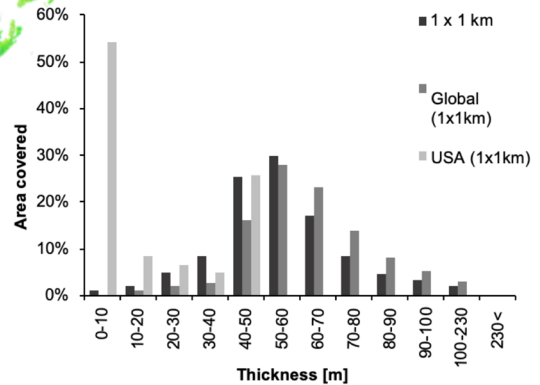
### 2.5. Evaluation Model Results

Simulated water table depths were validated against long-term averaged piezometer data (available from Fan et al., 2013). When more than one observation were available in a grid cell, the average depth was used. The water table head (WTH) was evaluated instead of water table depth (WTD), as WTH measures the

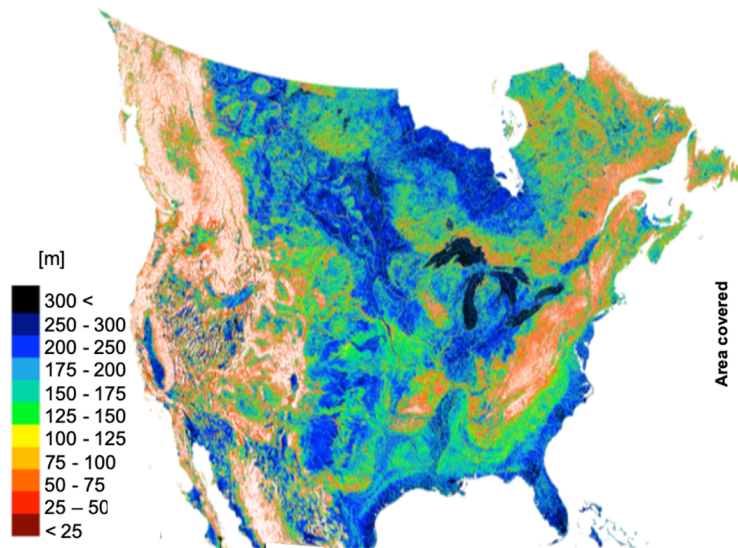
(a) Thickness shallow aquifers and top layers



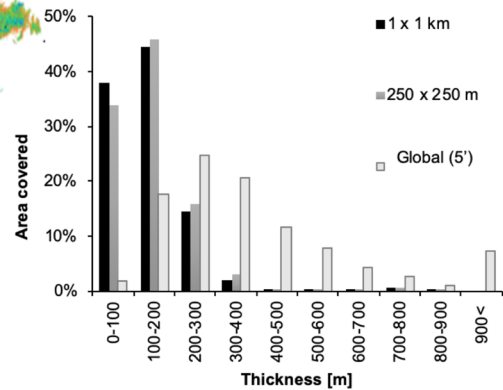
(b) Comparison of estimated thickness of shallow aquifers and top layers



(c) Total aquifer thickness



(d) Comparison of total estimated aquifer thickness



**Figure 4.** Estimated aquifer thickness of (a) shallow aquifers and top layers and (c) total aquifer thickness at the  $250 \times 250$  m resolution. The estimate of shallow aquifers and top layers is compared to the global scale estimate of (Pelletier et al., 2016) (global,  $1 \times 1$  km). (b) The estimate of total aquifer thickness is compared to the global scale estimate of de Graaf et al. (2015) (global,  $5'$ ) (d). In both maps, mountain ranges, where no aquifer thickness is estimated, are shown in white.

potential energy driving flow and therefore is physically more meaningful. The coefficients of determination ( $R^2$ ) and regression ( $\alpha$ ) were calculated for every run and were used as the criteria for selecting the best-performing parameter set.

Also, to evaluate model outcomes relative residuals in WTD were estimated, calculated as  $(WTD_{observed} - WTD_{simulated}) / WTD_{observed}$ . This evaluation provides insight into the spatial model performance and the spatial difference between the runs.

### 3. Results and Discussion

#### 3.1. Aquifer Thickness

Figure 4 shows the estimated aquifer thickness of local aquifers and upper layers of extensive aquifers (Figure 4a) and the total estimated aquifer thickness (Figure 4c) at the  $250 \times 250$  m resolution (results at the  $1 \times 1$  km resolution are given in the supporting information Figure S2). This study's estimates were compared



to previously published estimates of shallow layers and the total thickness of extensive aquifers (Figures 4b and 4d).

The maps show that local alluvial aquifers higher up in the mountains, as well as larger aquifer systems, are captured at a high level of detail. At the  $250 \times 250$  m resolution, about 82% of the North American land area is covered by an alluvial aquifer. The estimated thickness of these layers varies between  $<10$  and 230 m. Of the estimated alluvial aquifers, 67% is delineated as an extensive alluvial aquifer, covering 95% of the total area described as unconsolidated and semi-consolidated principal aquifers in the USGS groundwater atlas (Miller, 2000). The estimated thickness of these extensive systems varies between 10 and 350 m. At the  $1 \times 1$  km resolution, similar results are found, indicating consistency at different spatial scales (Figures 4b, 4d, and A2), although slightly thicker layers are simulated at the  $250 \times 250$  m resolution.

The spatial distribution and thickness estimates of previously published estimates of shallow aquifers (Miller & White, 1998; Pelletier et al., 2016) and deeper aquifers (de Graaf et al., 2015) show, in general, a similar pattern of thick layers for floodplains and piedmont belts and thinner layers higher up in the mountains for local aquifer systems (see Figure A3). However, the estimates differ compared to this study, as also presented in the histograms (Figures 4b and 4d). The data set of Pelletier et al. (2016) limits its maximum estimated thickness to 50 m. Also, this data set shows an accumulation of 1 m thickness values. This study does not need such preliminary assumptions and captures a more realistic range of aquifer thicknesses. The study of de Graaf et al. (2015) used a similar algorithm and also used data from the United States (however less detailed), but estimates aquifer thickness at a coarser resolution. Because of the coarse resolution, local alluvial aquifers higher up in the mountains are not captured. At the fine resolution, these aquifers are included and thus provide a more realistic aquifer parameterization for these regions. Also, due to the coarse resolution, total estimated thicknesses estimated by de Graaf et al. (2015) are much deeper.

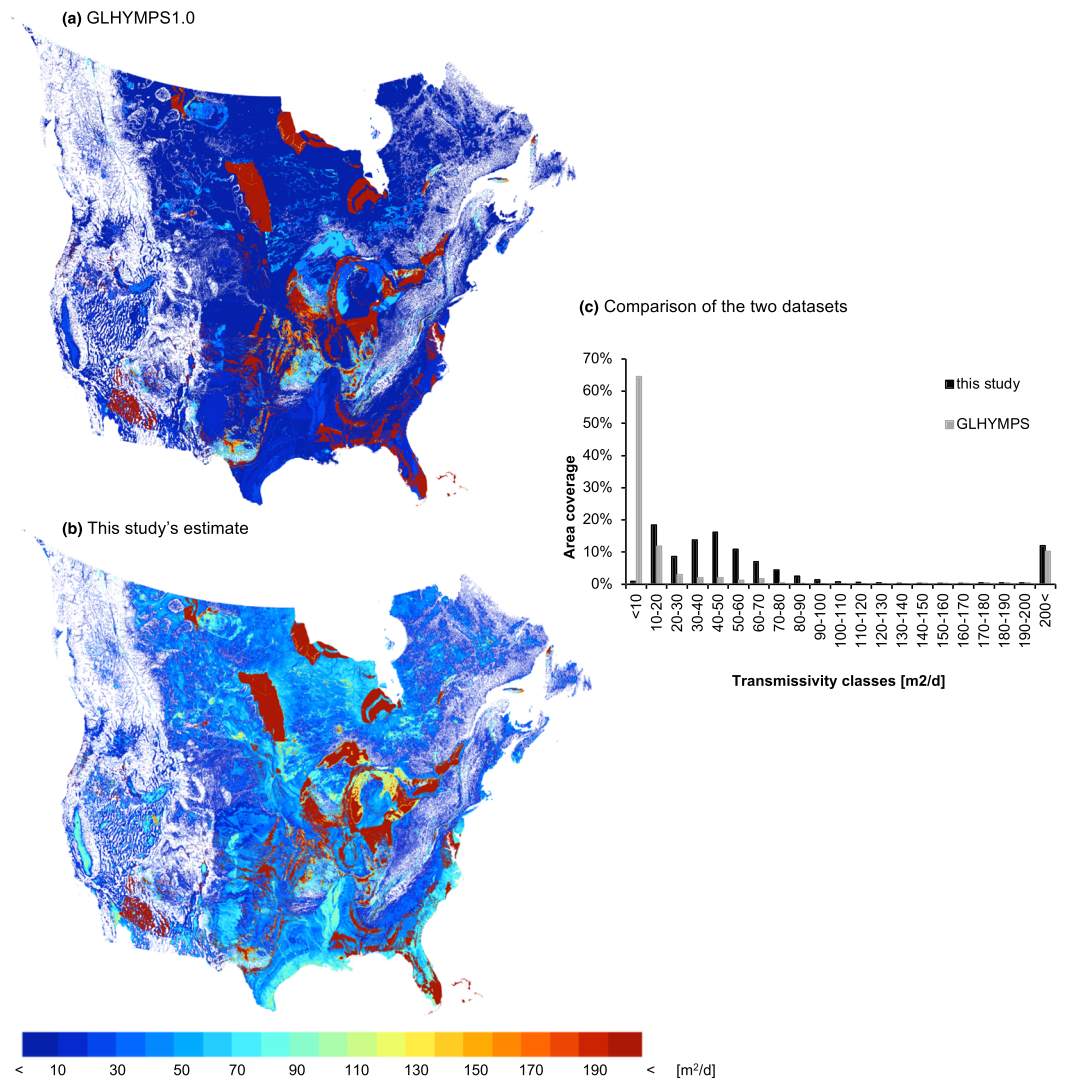
### 3.2. Aquifer Conductivity and Transmissivity

In this study, we extrapolated local- and regional-scale data on aquifer permeabilities from aquifer studies in the United States (available from the USGS). We used the same hydrogeological classification as used in the global permeability data set GLHYMPS (Gleeson et al., 2014) and provide new conductivity estimates (calculated from permeability) based on the data provided by the aquifer studies. These estimates were compared to the conductivities presented in GLHYMPS and the updated version GLHYMPS2.0 (Huscroft et al., 2018), the only two permeability data set available at the global scale (Table 1).

Including local data led to estimated conductivity values higher than the GLHYMPS and more similar to GLHYMPS2.0 and reported aquifer specific values (Figure A1). The conductivity values of unconsolidated sediments, siliciclastic sediments, and carbonates were updated representing more permeable rocks (Table 1). This is in line with the conclusion drawn from previous groundwater modeling experiences, that the values of GLHYMP1.0 are biased towards the lower end (Condon & Maxwell, 2014; de Graaf et al., 2015, 2017). Higher conductivity values will likely lead to a more realistic simulation of groundwater head dynamics and storage changes (de Graaf et al., 2017).

For the volume of groundwater flow aquifer transmissivity ( $\text{m}^3/\text{d}$ ) is an important parameter. The more detailed aquifer thickness estimate and updated conductivity values provided in this study will most likely lead to more realistically simulated groundwater volume fluxes when used in groundwater modeling over larger domains. Figure 5 shows the difference in estimated aquifer transmissivities when using this study's estimated aquifer thicknesses and the conductivity values obtained from GLHYMPS or this study's conductivity values) as assigned to the different aquifer layers (Table 1). In Figure 5, total aquifer transmissivity is presented.

From the maps and histogram (Figure 5), it can be seen that this study's conductivity values and vertical structure lead to higher estimated aquifer transmissivities and that the spatial distribution is more diverse compared to the estimate using GLHYMPS, showing more regional detail. For example, more local detail is obtained, and inconsistencies are overcome over the High Plains aquifer, the coastal aquifer systems, and for local aquifer systems present in valleys higher up in the mountains. However, some inconsistencies present in GLHYMPS are still visible, for example, for the Canadian-U.S. border and regions where information on the underlying bedrock was missing and high conductivities of GLHYMPS were assumed. A clear example of the latter are the regions with carbonated rocks, for example, in the northeast of the United States.



**Figure 5.** Estimated aquifer transmissivities calculated as (a) this study's estimated total aquifer thickness  $\times$  conductivities calculated from permeabilities in GLHYMPS (following the methodology of de Graaf et al., 2015) and as (b) total aquifer transmissivity calculated as the sum of transmissivities calculated per aquifer layer assigning conductivities as described in section 2.2. In this study, transmissivities are higher and more realistically distributed (c). Cells without an aquifer thickness are shown in white.

Potentially, these inconsistencies could be minimized if accurate data, specifically for these regions, would be included in the conductivity estimation.

### 3.3. Test Case: Evaluation of Water Table Depth and Groundwater Outflow

We first run the model ParFlow over the small test case area at the  $1 \times 1$  km resolution, describing the vertical structure of the aquifer in six or 12 model layers, using this study's conductivity estimates (i.e., continental conductivity) or the local conductivity data for the northern High Plains (i.e., local conductivity). Model results of water table heads were compared to piezometer observations, and  $R^2$  and  $\alpha$  were calculated (Table 2). Also, the percentages of simulated open water cells were compared between the different runs. Based on these statistics, the run using 12 layers and continental conductivities performed best. At the  $250 \times 250$  m resolution, we used therefore 12 layers and continental conductivities as well and calculated  $R^2$  and  $\alpha$  (Table 2). The estimated water table depths of the latter run are presented in Figure 6a.

The high  $R^2$ 's and  $\alpha$ 's show that simulated WTH match observed WTH well. Somewhat surprisingly, the difference between the different model runs is minimal. When the vertical discretization increased from six to 12 layers, model performance increased slightly, as well as when continental conductivities were used

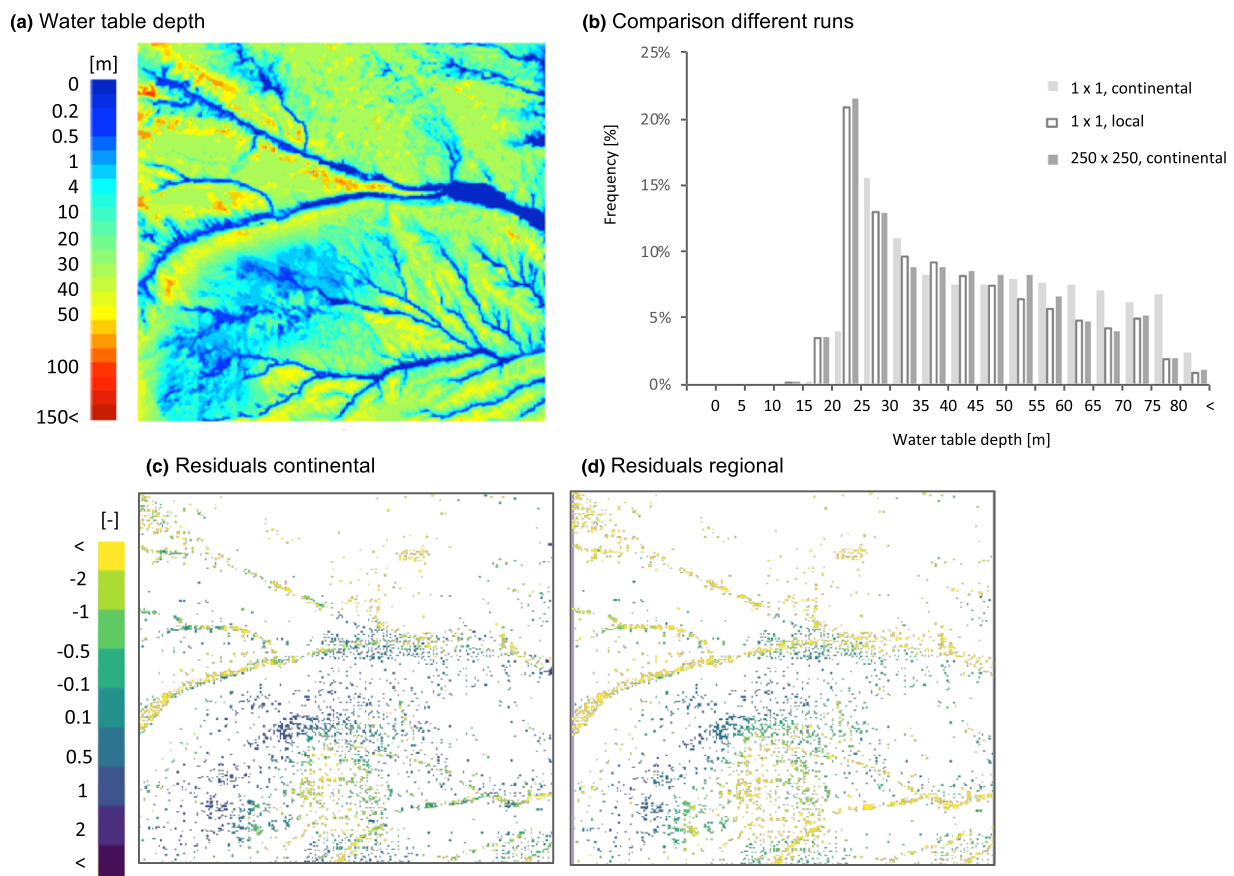
**Table 2**  
Statistics of Model Performance of All Runs Done

	1 × 1 km					
	Continental conductivities			Local conductivities		
	$R^2$	$\alpha$	% surface water cells	$R^2$	$\alpha$	% surface water cells
6 layers	0.98	0.97	6	0.97	0.99	10
12 layers	0.99	0.98	4	0.98	0.99	6
	250 × 250 m					
12 layers	0.99	0.97	2			

Note. Simulated water table heads were compared to observed water table heads. Additionally, the percentage of surface water cells (i.e.,  $WTD \leq 0$ ) is given.

instead of local conductivities. These conclusions are confirmed by the histogram of Figure 6b, showing the distribution of simulated water table depths of the three runs using 12 layers to describe the vertical structure of the aquifer. However, using continental conductivities at the  $1 \times 1$  km resolution clearly leads to deeper water tables (deeper than 50 m) compared to the other two runs. Using continental conductivity values at the  $250 \times 250$  m resolution resulted in a similar distribution of water table depths as simulated in the run using local conductivities at  $1 \times 1$  km resolution, although the latter run clearly estimated fewer water tables in the range 45–60 m below surface. The slightly lower model performance of the local conductivity run shows that these local values do not represent aquifer properties accurately at the larger scale.

Simulated water table depths (WTD) were compared to observed WTD, and residuals were calculated (as  $(WTD_{simulated} - WTD_{observed}) / WTD_{observed}$ ). Figures 6c and 6d show the maps of these residuals at  $1 \times 1$  km



**Figure 6.** (a) Simulated water table depths (wtd) at the  $250 \times 250$  m resolution, using 12 layers to describe the vertical structure and using continental conductivities. Differences between the runs are small (b) and showing similar patterns of residuals (c and d).

resolution. The maps of residuals show a similar pattern for both runs; however, local differences exist. First of all, underestimated WTD (i.e., simulated WTD are too shallow; residual  $< 0$ ) in both runs can be seen for simulated drainage cells. It is very likely that the observations in these cells do represent water table depths close to the river, whereas in the model, a whole cell gets a WTD value above the surface. Also, open water is not represented in the data set, as observed WTD of 0 m does not exist. In the run using local conductivities, this overestimation in WTD seems to be larger and simulated WTD of the run using continental conductivities match observations close to the main drainage better. In the regional case, more open water cells were simulated (6% compared to 4%) leading to more shallow water tables also next to those drainage cells. Second, for both runs, WTDs are overestimated (i.e., simulated WTD are too deep; residual value  $> 0$ ) for the region where recharge is lowest (see Figure 3c). The spatial uncertainty in recharge is most likely driving here. Residuals in the local conductivity estimate are slightly smaller for part of this region most likely caused by the shallower water tables closer to the main drainage cells. Compared to previous estimates of absolute residuals, this study's results are much smaller and fall on average between  $-10$  and  $+10$  m.

Evaluation of the water balance showed that in the regional conductivity run, almost twice as much surface water runoff was simulated than in the continental conductivity runs. In the regional run, more recharged groundwater is drained within the test area indicating the groundwater and surface water systems are more connected. We saw this already by the slightly more overestimated WTD close to the main drainage cells, although in both model runs, the main gradient of groundwater flow is from west to east and deep groundwater flow paths transport large volumes of water.

#### 4. Conclusion

In this paper, we presented a detailed aquifer representation for the continental United States and Canada at hyper-resolution ( $250 \times 250$  m). We integrated local hydrogeological information on aquifer thickness, vertical structure, and conductivity into a continental-scale aquifer parameterization that can be used for groundwater modeling. Compared to previously published estimates, we estimated aquifer thickness and vertical structure at a level of detail not obtained before. Our thickness estimates are predominantly based on elevation differences and do not use preliminary assumptions on minimum or maximum depth. Aquifer transmissivity values were calculated using updated values of conductivity, which were obtained from aquifer studies available in the United States. The conductivity values were assigned to the estimated aquifer layers and thus not limited by a lithology map (as done in previous studies). The methods used are simple, showed to be consistent at different horizontal resolutions, and can be expanded to other parts of the world. Although, including more specific information for, for example, a region or aquifer will improve the accuracy of the estimation for that region or aquifer, for regions where this information is missing, the presented approach provides a useful first-order estimate of the subsurface.

We tested our new aquifer parameterization hydrologically by running the hydrological model ParFlow. Although we were only able to test over a small test case, due to the high computational demands of ParFlow, the hydrological test enhanced our understanding of the sensitivity of the groundwater system to different parameter settings. The model results showed us that more accurate conductivity values applicable at the larger scale are key for more reliable simulation of the groundwater system, as well as vertical representation of the aquifer system. However, our model test is limited in its size, and results might not be the same for other hydrogeological settings. Currently, we performed our tests in steady state and did not study the difference in groundwater head fluctuations due to different parameter settings. In future work, the dynamics of estimated groundwater depth should be studied as well, to fully understand the need for more detailed aquifer parameterization in large-scale hydrological modeling. Future work then also focus on the impact of groundwater abstractions. Our presented aquifer parameterization distinguishes between confined and unconfined aquifer systems and thus can be used to study the impacts of groundwater abstractions.

We hope that our research and the methods we presented here will serve as a base for future studies. Future work should be targeted on improving current large-scale estimates of aquifer thickness and conductivities to support detailed large-scale groundwater modeling and evaluation of large-scale subsurface parameterization and sensitivity of model results. Integrating local-scale information into continental-scale estimates, as presented here, is one way forward.

### Acknowledgments

The new data sets presented in this paper, that is, aquifer thickness and aquifer transmissivities at  $1 \times 1$  km and  $250 \times 250$  m resolution, are available through <https://doi.org/10.25739/e2fw-fe79>. ParFlow specific input files to run the test case are available from the corresponding author. This work was supported by the U.S. Department of Energy Office of Science, Offices of Advanced Scientific Computing Research and Biological and Environmental Sciences IDEAS project.

### References

- Cederstrand, J., & Becker, M. (1998). Digital map of hydraulic conductivity for the High Plains Aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming (Open-File Report 98-548). <https://doi.org/10.3133/ofr98548>
- Chang, L. L., Dwivedi, R., Knowles, J. F., Fang, Y. H., Niu, G. Y., Pelletier, J. D., & Meixner, T. (2018). Why do large-scale land surface models produce a low ratio of transpiration to evapotranspiration? *Journal of Geophysical Research: Atmospheres*, *123*, 9109–9130. <https://doi.org/10.1029/2018jd029159>
- Condon, L. E., & Maxwell, R. M. (2014). Groundwater-fed irrigation impacts spatially distributed temporal scaling behavior of the natural system: A spatio-temporal framework for understanding water management impacts. *Environmental Research Letters*, *9*(3), 34009. <https://doi.org/10.1088/1748-9326/9/3/034009>
- de Graaf, I. E. M., Gleeson, T., Van Beek, L. P. H., Sutanudjaja, E. H., & Bierkens, M. F. P. (2019). Environmental flow limits to global groundwater pumping. *Nature*, *574*, 90–94.
- de Graaf, I. E. M., Sutanudjaja, E. H., van Beek, L. P. H., & Bierkens, M. F. P. (2015). A high-resolution global-scale groundwater model. *Hydrology and Earth System Sciences*, *19*(2), 823–837. <https://doi.org/10.5194/hess-19-823-2015>
- de Graaf, I. E. M., van Beek, R. L., Gleeson, T., Moosdorf, N., Schmitz, O., Sutanudjaja, E. H., & Bierkens, M. F. (2017). A global-scale two-layer transient groundwater model: Development and application to groundwater depletion. *Advances in Water Resources*, *102*, 53–67. <https://doi.org/10.1016/j.advwatres.2017.01.011>
- de Graaf, I. E. M., van Beek, L., Wada, Y., & Bierkens, M. (2014). Dynamic attribution of global water demand to surface water and groundwater resources: Effects of abstractions and return flows on river discharges. *Advances in Water Resources*, *64*, 21–33. <https://doi.org/10.1016/j.advwatres.2013.12.002>
- Döll, P., Hoffmann-Dobrev, H., Portmann, F., Siebert, S., Eicker, A., Rodell, M., & Scanlon, B. (2012). Impact of water withdrawals from groundwater and surface water on continental water storage variations. *Journal of Geodynamics*, *59-60*, 143–156. <https://doi.org/10.1016/j.jog.2011.05.001>
- Fan, Y., Li, H., & Miguez-Macho, G. (2013). Global patterns of groundwater table depth. *Science*, *339*(6122), 940–943. <https://doi.org/10.1126/science.1229881>
- Gleeson, T., Moosdorf, N., Hartmann, J., & van Beek, L. P. H. (2014). A glimpse beneath Earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity. *Geophysical Research Letters*, *41*, 3891–3898. <https://doi.org/10.1002/2014gl059856>
- Gutentag, E., Heimes, F., Krothe, N., Luckey, R., & Weeks, J. (1984). Geohydrology of the High Plains Aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming (Professional Paper 1400). <https://doi.org/10.3133/pp1400b>
- Hartmann, J., & Moosdorf, N. (2012). The new global lithological map database GLiM: A representation of rock properties at the Earth surface. *Geochemistry, Geophysics, Geosystems*, *13*, Q12004. <https://doi.org/10.1029/2012gc004370>
- Huscroft, J., Gleeson, T., Hartmann, J., & Börker, J. (2018). Compiling and mapping global permeability of the unconsolidated and consolidated Earth: GLobal HYdrogeology MaPS 2.0 (GLHYMPS 2.0). *Geophysical Research Letters*, *45*, 1897–1904. <https://doi.org/10.1002/2017gl075860>
- Kollet, S. J., & Maxwell, R. M. (2006). Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. *Advances in Water Resources*, *29*(7), 945–958. <https://doi.org/10.1016/j.advwatres.2005.08.006>
- Krakauer, N. Y., Li, H., & Fan, Y. (2014). Groundwater flow across spatial scales: Importance for climate modeling. *Environmental Research Letters*, *9*(3), 34,003. <https://doi.org/10.1088/1748-9326/9/3/034003>
- Lehner, B., & Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, *296*(1–4), 1–22. <https://doi.org/10.1016/j.jhydrol.2004.03.028>
- Margat, J., & van der Gun, J. (2013). *Groundwater around the world: A geographic synopsis*. Boca Raton, FL: CRC Press. ISBN 9781138000346.
- Maxwell, R. M. (2013). A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling. *Advances in Water Resources*, *53*, 109–117. <https://doi.org/10.1016/j.advwatres.2012.10.001>
- Maxwell, R. M., & Condon, L. E. (2016). Connections between groundwater flow and transpiration partitioning. *Science*, *353*(6297), 377–380.
- Maxwell, R. M., Condon, L. E., & Kollet, S. J. (2015). 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*(3), 923–937. <https://doi.org/10.5194/gmd-8-923-2015>
- Maxwell, R. M., Kollet, S. J., Smith, S. G., Woodward, C. S., Falgout, R. D., Ferguson, I. M., & Kulkarni, K. B. (2019). ParFlow user's manual (*Integrated GroundWater Modeling Center*).
- McGuire, V. L. (2014). Water-level changes and change in water in storage in the High Plains aquifer, predevelopment to 2013 and 2011–13: Scientific Investigations Report. <https://doi.org/10.3133/sir20145218>
- Miller, D., & White, R. (1998). A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. *Earth Interactions*, *2*(2), 1–26. [https://doi.org/10.1175/1087-3562\(1998\)002<0001:acusms>2.3.co;2](https://doi.org/10.1175/1087-3562(1998)002<0001:acusms>2.3.co;2)
- Miller, J. A. (Ed.) (2000). *Groundwater atlas of the United States* (p. 404). U.S. Geological Survey.
- Pelletier, J. D., Broxton, P. D., Hazenberg, P., Zeng, X., Troch, P. A., Niu, G., & Gochis, D. (2016). A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling. *Journal of Advances in Modeling Earth Systems*, *8*, 41–65. <https://doi.org/10.1002/2015ms000526>
- Perrone, D., & Jasechko, S. (2017). Dry groundwater wells in the western United States. *Environmental Research Letters*, *12*, 104,002. <https://doi.org/10.1088/1748-9326/aa8ac0>
- Reinecke, R., Foglia, L., Mehl, S., Herman, J. D., Wachholz, A., Trautmann, T., & Döll, P. (2019). Spatially distributed sensitivity of simulated global groundwater heads and flows to hydraulic conductivity, groundwater recharge and surface water body parameterization. *Hydrology and Earth System Sciences*, *23*, 4561–4582. <https://doi.org/10.5194/hess-23-4561-2019>
- Schaller, M. F., & Fan, Y. (2009). River basins as groundwater exporters and importers: Implications for water cycle and climate modeling. *Journal of Geophysical Research*, *114*, D04103. <https://doi.org/10.1029/2008JD010636>
- Shangguan, W., Hengl, T., Mendes de Jesus, J., Yuan, H., & Dai, Y. (2017). Mapping the global depth to bedrock for land surface modeling. *Journal of Advances in Modeling Earth Systems*, *9*, 65–88. <https://doi.org/10.1002/2016ms000686>
- Shrestha, P., Kurtz, W., Vogel, G., Schulz, J. P., Sulis, M., Hendricks Franssen, H. J., & Simmer, C. (2018). Connection between root zone soil moisture and surface energy flux partitioning using modeling, observations, and data assimilation for a temperate grassland site in Germany. *Journal of Geophysical Research: Biogeosciences*, *123*, 2839–2862. <https://doi.org/10.1029/2016jg003753>

UNESCO (2010). The United Nations world water development report (*Technical Report*): Unesco.

Wada, Y., Van Beek, L. P. H., & Bierkens, M. F. P. (2011). Modelling global water stress of the recent past: On the relative importance of trends in water demand and climate variability. *Hydrology and Earth System Sciences*, *15*(12), 3785–3805. <https://doi.org/10.5194/hess-15-3785-2011>