Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data

E. H. Sutanudjaja,[1] L. P. H. van Beek,[1] S. M. de Jong,[1] F. C. van Geer,[1,3] and M. F. P. Bierkens[1,2]

[1] We explore the possibility of using remotely sensed soil moisture data and in situ discharge observations to calibrate a large-extent hydrological model. The model used is PCR-GLOBWB-MOD, which is a physically based and fully coupled groundwater-land surface model operating at a daily basis and having a resolution of 30 arc sec (about 1 km at the equator). As a test bed, we use the combined Rhine-Meuse basin (total area: about 200,000 km^2), where there are 4250 point-scale observed groundwater head time series that are used to verify the model results. Calibration is performed by simulating 3045 model runs with varying parameter values affecting groundwater head dynamics. The simulation results of all runs are evaluated against the remotely sensed soil moisture time series of SWI (Soil Water Index) and field discharge data. The former is derived from European Remote Sensing scatterometers and provides estimates of the first meter profile soil moisture content at 30 arc min resolution (50 km at the equator). From the evaluation of these runs, we then introduce a stepwise calibration approach that considers stream discharge first, then soil moisture, and finally verify the resulting simulation to groundwater head observations. Our results indicate that the remotely sensed soil moisture data can be used for the calibration of upper soil hydraulic conductivities determining simulated groundwater recharge of the model. However, discharge data should be included to obtain full calibration of the coupled model, specifically to constrain aquifer transmissivities and runoff-infiltration partitioning processes. The stepwise approach introduced in this study, using both discharge and soil moisture data, can calibrate both discharge and soil moisture, as well as predicting groundwater head dynamics with acceptable accuracy. As our approach to parameterize and calibrate the model uses globally available data sets only, it opens up the possibility to set up large-extent coupled groundwater-land surface models in other basins or even globally.

1. Introduction

[2] Groundwater serves as a primary source of drinking water and supplies water for agricultural and industrial activities. During times of drought, groundwater sustains water flows in streams, rivers, lakes, and wetlands, and thus supports ecosystem habitat and biodiversity, while its large natural storage provides a buffer against water shortages. However, groundwater is known as a vulnerable resource. In many areas, it is being consumed faster than naturally replenished [*Rodell et al.*, 2009; *Wada et al.*, 2010]. Given increasing population and heightened variability and uncertainty in precipitation due to climate change, the pressure upon groundwater resources is expected to intensify. Therefore, modeling and predicting groundwater variabilities and changes are imperative.

[3] Yet, the current generation of global-scale hydrological models [e.g., *van Beek et al.*, 2011; *Widén-Nilsson et al.*, 2007; *Liang et al.*, 1994] still does not include a groundwater lateral flow component which is an important link in the hydrological cycle. In addition to river networks, groundwater bodies transport precipitation falling to the earth from higher elevations via lateral fluxes to lower elevations. Water that infiltrates to groundwater bodies subsequently flow to low-pressure areas and discharges in the form of stream base flow, soil evaporation, and plant transpiration. Hence, under close interaction with land surface, groundwater influences root-zone soil moisture, and potentially affect water and energy fluxes between the land and atmosphere [see e.g., *Shukla and Mintz*, 1982; *Koster et al.*, 2004; *Fan and Miguez-Macho*, 2010, 2011; *Alkhaier*

[1]Department of Physical Geography, Faculty of Geosciences, Utrecht University, Utrecht, Netherlands.

[2]Unit Soil and Groundwater Systems, Deltares, Utrecht, Netherlands.

[3]Netherlands Organization for Applied Scientific Research TNO, Utrecht, Netherlands.

Corresponding author: E. H. Sutanudjaja, Department of Physical Geography, Faculty of Geosciences, Utrecht University, POB 80115, NL-3508 TC Utrecht, Netherlands. (E.H.Sutanudjaja@uu.nl)

et al., 2012a]. *Bierkens and van den Hurk* [2007] show that rainfall persistence may be partly explained by groundwater confluence to discharge zones throughout the year to sustain evaporation for longer periods of time. These examples show the importance of including a groundwater flow component into global-scale hydrological models [see also *Fan et al.*, 2007; *Yeh and Eltahir*, 2005a, b]. Such an inclusion is also crucial for the study of hydrological drought as drought propagation in areas with large aquifers cannot be correctly simulated by the current generation of global hydrological models [*van Loon et al.*, 2012].

[4] Most large-scale hydrological models are evaluated using discharge data at catchment outlets [e.g., *Döll et al.*, 2003; *Zaitchik et al.*, 2010; *van Beek et al.*, 2011]. Yet, calibrating hydrological models to discharge data alone does not guarantee the correct simulation of other model states and output variables, such as soil moisture and groundwater. An obvious advantage of having a model that includes groundwater lateral flow is that it simulates groundwater head dynamics. As groundwater head can be measured in the field, we can evaluate and even calibrate the simulated groundwater head dynamics. Such a calibration will greatly complement the traditional discharge calibration. Nevertheless, it is difficult to calibrate large-extent groundwater models due to the sparseness of in situ groundwater head measurement data in many parts of the world. In the limited areas where they are available, head measurements are sparsely located and available over limited spatial extent. Moreover, ground-based groundwater head observations provide only information with limited spatial support (i.e., point scale). Nevertheless, groundwater head spatial variation tends to be smoothly varying over space such that a single point measurement may still be informative for larger areas.

[5] The main strength of spaceborne remote sensing observations is the ability to provide spatially and temporally exhaustive maps of earth surface properties with a near-global coverage. In the last decades, many studies have investigated the possibility of using earth observation for hydrological purposes and spaceborne remote sensing is increasingly being used for mapping hydrological states and fluxes, such as precipitation [e.g., *Kummerow et al.*, 2000], soil moisture [*Kerr et al.*, 2001; *Njoku et al.*, 2003; *Wagner et al.*, 1999b; *Wanders et al.*, 2012], snow cover [*Dankers and de Jong*, 2004; *Immerzeel et al.*, 2009], land surface temperature [*Wan and Li*, 1997], and evaporation [*Bastiaanssen et al.*, 1998a, 1998b; *Su*, 2002; *Mu et al.*, 2007; *Jung et al.*, 2010]. However, their benefits for groundwater hydrology applications are still limited. Up to now, only the space gravity satellite Gravity Recovery and Climate Experiment (GRACE) [*Tapley et al.*, 2004] from the National Aeronautics and Space Administration (NASA) is acknowledged as a groundwater assessment tool, specifically for detecting groundwater storage changes [e.g., *Rodell et al.*, 2009]. A major drawback of GRACE is its coarse resolution of approximately 400 km, severely hampering its regional-scale application. Nevertheless, *Becker* [2006] argues that groundwater behavior may be inferred from remotely sensed surface expressions, such as surface temperature and soil moisture. Indeed, in a recent application *Alkhaier et al.* [2012a, b] show that remotely sensed evaporation and land surface temperature correlate well

with groundwater depth and can be physically connected by means of a soil water and heat flux model.

[6] In this study, we aim to explore the potential of using remotely sensed soil moisture data [see also *Jackson*, 2002]—currently having a better spatial resolution (25–50 km) than the GRACE product—to support groundwater studies. More specifically, we investigate whether the global coverage soil moisture product of Soil Water Index (SWI)—derived by *Wagner et al.* [1999b] based on European Remote Sensing (ERS) scatterometer signals—can be used to calibrate a large-scale coupled groundwater-land surface model of PCR-GLOBWB-MOD [*Sutanudjaja et al.*, 2011; *Sutanudjaja*, 2012]. The model is built by using only global data sets such that it is portable to other areas of the world, including data-poor areas. The underlying idea of this study is that the combination of setting up the model using only global data sets and calibrating it with remote sensing data can make large-scale groundwater modeling feasible throughout the world. As discharge data are commonly used in hydrological model calibration and in principle globally available from the Global Runoff Data Centre (GRDC, http://www.bafg.de/GRDC), we also aim to calibrate the model based on in situ discharge measurements. The combined calibration to soil moisture and discharge can be considered as a complementary to the commonly used multiobjective approaches [e.g., *Gupta et al.*, 1998; *Fenicia et al.*, 2007] that aim to reproduce different parts of discharge (i.e., low and high flow parts of hydrographs) and other studies focusing on various model components, such as evapotranspiration [*Crow et al.*, 2003] and total water storage variations [*Werth et al.*, 2009; *Lo et al.*, 2010].

[7] As a test bed, we use the combined Rhine-Meuse basin (see Figure 1). The large size of the study area, covering $\pm 200,000$ km^2, makes it well suited for large-extent modeling studies. The Rhine basin includes the upstream areas in the Alps of Austria and Switzerland and cover large areas in Germany with the outlet located in the town of Lobith on the Dutch eastern border. The Meuse basin used stretches from its headwaters, mainly in France and Belgium, to the town of Borgharen on the Dutch southern border. To simplify this study, the delta area of Rhine-Meuse basin, mainly in Netherlands, is excluded due to strong anthropogenic water management practices (see Figure 1a). Nevertheless, the chosen study area, which is situated in the humid temperate zone of Western Europe, has a good coverage of SWI, discharge and ample groundwater head observations. For this study, there are 4250 point-scale groundwater head observations used for verification. By verifying the calibrated model results to these head observations, we analyze the benefits of using soil moisture and discharge time series, as well as the combination of both, for calibrating the model.

[8] The paper is organized as follows: section 2 briefly explains the coupled groundwater-land surface model used; section 3 describes the calibration procedure; section 4 presents the results and analyses; section 5 provides the discussion in the light of this study; and section 6 summarizes and concludes this paper.

2. Model Structure: PCR-GLOBWB-MOD

[9] Briefly stated, PCR-GLOBWB-MOD, having the spatial resolution of 30 arc sec (30″ × 30″, about 1 km × 1
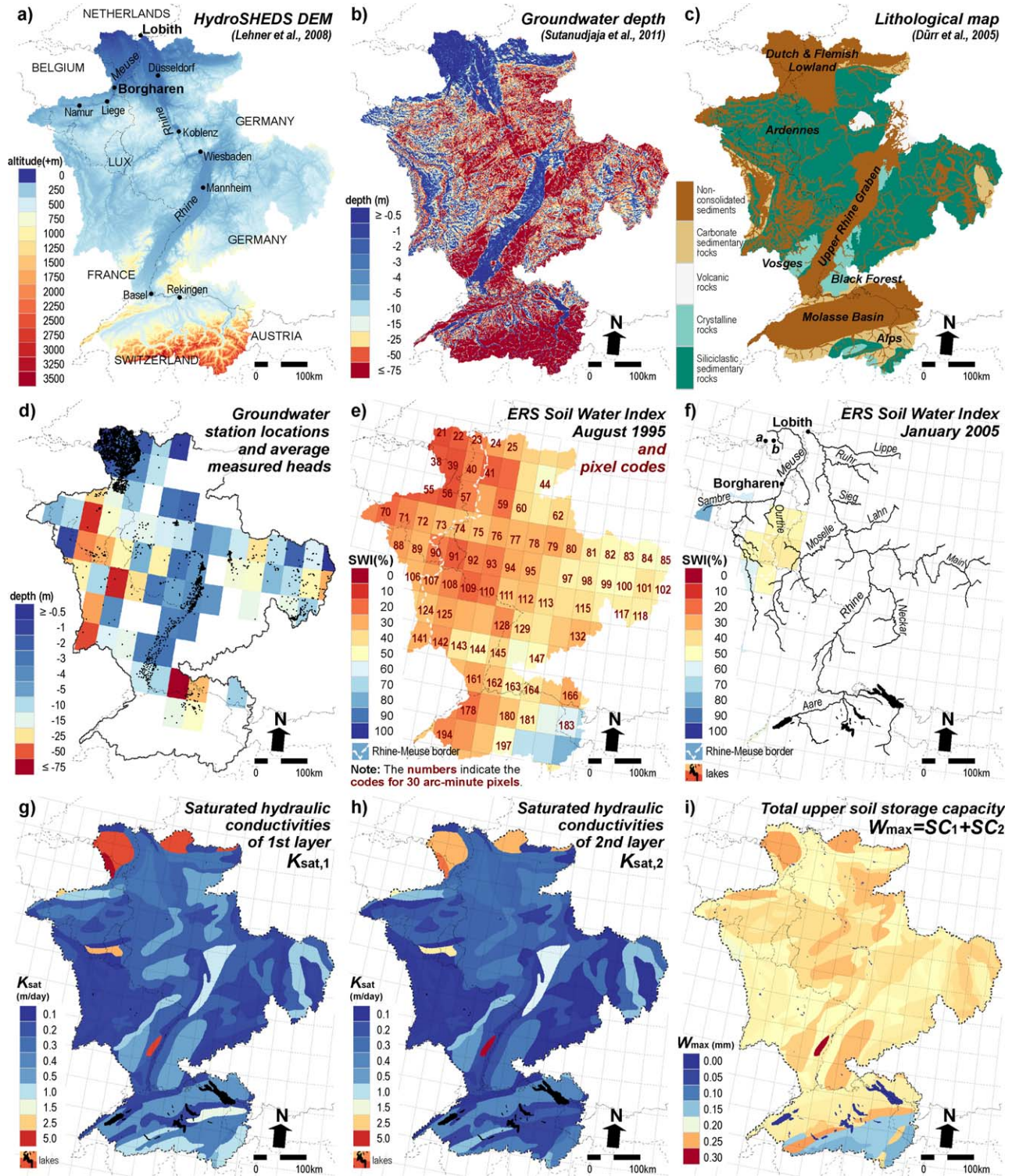
Figure 1. The Rhine-Meuse basin: (a and b) 30 arc sec digital elevation and groundwater depth maps; (c) lithological map; (d) locations of head stations and average heads per pixel; (e and f) 30 arc min SWI fields in August 1995 and January 2005; (g and h) saturated conductivities (K_{sat}) of the first and second soil layers used in the REFERENCE; and (i) total soil water storage capacities ($W_{max} = SC_1 + SC_2$). Figure 1e includes the pixel codes used and Figure 1f also shows the discharge and head stations corresponding to the time series in Figures 4 and 11. White area in Figure 1f indicates missing SWI values due to snow cover and frozen soil.
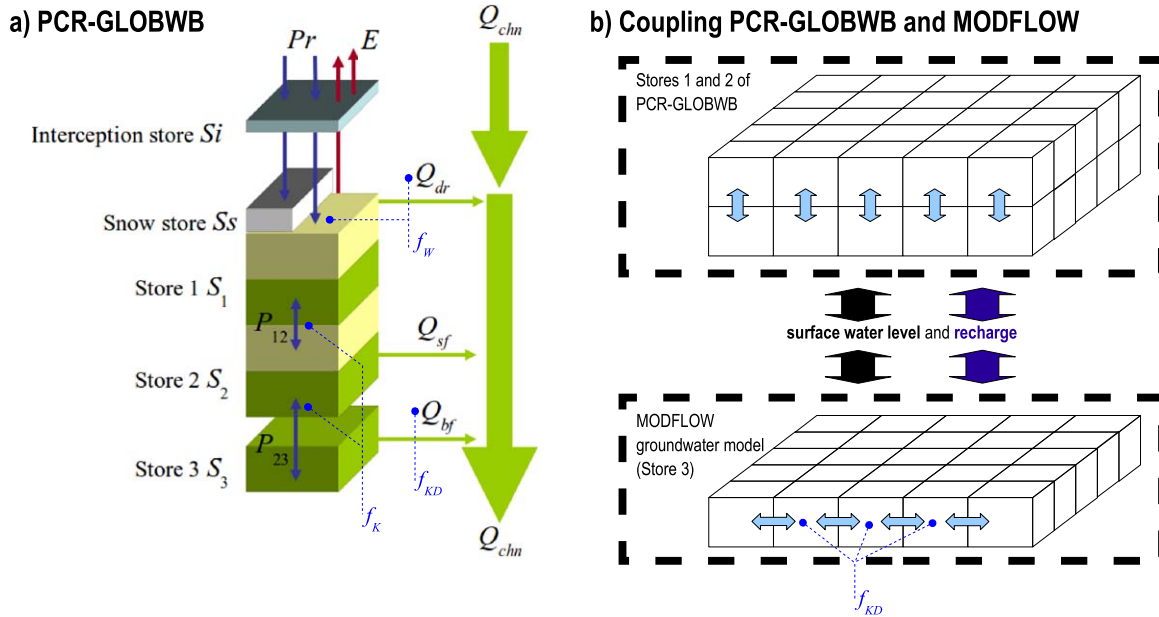
a) PCR-GLOBWB

b) Coupling PCR-GLOBWB and MODFLOW

Figure 2. The model structure and coupling strategy: (a) the land surface model of PCR-GLOBWB [*van Beek and Bierkens*, 2009; *van Beek et al.*, 2011]: (Figure 2a, left) the soil compartment, divided in the two upper soil stores, S_1 and S_2, and the linear groundwater store, S_3, that is replaced by the MOD-FLOW [*McDonald and Harbaugh*, 1988] groundwater model; (Figure 2a, right) the total local gains from all cells are routed along the drainage direction (not shown) to yield the channel discharge, Q_{chn}. (b) The coupling between PCR-GLOBWB and MODFLOW. The latter simulates lateral groundwater flow. In both figures, we indicate the parts that are controlled by the prefactors f_W, f_K, and f_{KD} calibrated in this study (see sections 2 and 3.1).

km at the equator) and operating based on daily basis water balance, is the land surface model PCR-GLOBWB [*van Beek and Bierkens*, 2009; *van Beek et al.*, 2011] coupled to a MODFLOW [*McDonald and Harbaugh*, 1988] groundwater model. The coupled model structure is illustrated in Figure 2. The land surface model part conceptualizes the hydrological processes above the surface and in two unsaturated zone soil layers (in which their storages are symbolized as S_1 and S_2 [L] with indexes indicating the sequence of soil layers), while the groundwater model part contains an underlying groundwater layer store (S_3 [L]) conceptualizing deeper saturated lateral flows. The runoff from every cell consists of surface or direct runoff Q_{dr} [$L\ T^{-1}$], interflow Q_{sf} [$L\ T^{-1}$] from the second soil layer and base flow Q_{bf} [$L\ T^{-1}$] from the groundwater layer. The channel discharge Q_{chn} [$L\ T^{-3}$] is calculated by implementing a routing scheme to accumulate the local runoff from all cells along the drainage network. This routing procedure is implemented by considering traveling time through surface water bodies and based on the unit hydrograph method [*Soil Conservation Service* (*SCS*), 1972; *Sólyom and Tucker*, 2004].

[10] In this paper, we briefly explain only the model components that are relevant for the purpose of this calibration study. The following section 2.1 mainly summarizes how direct runoff and vertical fluxes between soil layers are conceptualized, while lateral groundwater flow and base flow related processes are briefly explained in section 2.2. Detailed description of other model components can be found in *Sutanudjaja et al.* [2011] and *Sutanudjaja* [2012].

2.1. The Land Surface Model Part

2.1.1. Direct Runoff and Net Infiltration

[11] As illustrated in Figure 2a, the main storages of a cell in the land surface part of PCR-GLOBWB-MOD are two upper soil stores S_1 and S_2 (Figure 2a) representing the top 30 cm and the following 70 cm of soil (thicknesses $Z_1 \leq 30$ cm and $Z_2 \leq 70$ cm). The model also includes an interception storage and a snow module based on the Hydrologiska Byråns Vattenbalansavdelning (HBV) model [*Bergström*, 1995]. After passing the interception and snow melting-related processes, the basic input to the soil stores is a certain amount of liquid rainfall P_n [$L\ T^{-1}$], which consists of the net rainfall above the interception capacity and melt water from the snow pack.

[12] P_n is later partitioned into direct runoff Q_{dr} and net infiltration to the first soil layer, P_{01} [$L\ T^{-1}$]. The partitioning is done by adopting the improved Arno scheme [*Todini*, 1996; *Hagemann and Gates*, 2003], in which the total soil water storage capacity of a cell is the aggregate of many different soil water storage capacities. Through this scheme, the fraction of saturated soil of a cell, x [–], is given by [*van Beek and Bierkens*, 2009]:

$$x = 1 - \left(\frac{W_{\max} - W_{act}}{W_{\max} - W_{\min}}\right)^{\frac{b}{b+1}} = 1 - \left(\frac{\Delta W_{act}}{\Delta W}\right)^{\frac{b}{b+1}} \tag{1}$$

with

$$\Delta W_{act} = W_{\max} - W_{act} \tag{2}$$

$$\Delta W = W_{\max} - W_{\min} \tag{3}$$

where $W_{\min}$ [L] is the cell-(local)-minimum capacity, W_{act} [L] is the cell-average actual storage, and $W_{\max}$ [L] is the storage capacity of the entire soil profile:

$$W_{\mathrm{act}} = S_1 + S_2 \tag{4}$$

$$W_{\max} = SC_1 + SC_2 \tag{5}$$

with SC [L] indicating the soil water capacity for each

$$Q'_{\mathrm{dr}}\Delta t = \begin{cases} 0 \\[2em] P_{\mathrm{n}}\Delta t - \Delta W_{\mathrm{act}} + \\ \Delta W \left[\left(\frac{\Delta W_{\mathrm{act}}}{\Delta W} \right)^{\frac{1}{b+1}} - \frac{P_{\mathrm{n}}\Delta t}{(b+1)\Delta W} \right]^{b+1} \\[3em] P_{\mathrm{n}}\Delta t - \Delta W_{\mathrm{act}} \end{cases} \quad \begin{array}{l} \text{if } P_{\mathrm{n}}\Delta t + W_{\mathrm{act}} \leq W_{\min} \\[3em] \text{if } W_{\min} < P_{\mathrm{n}}\Delta t + W_{\mathrm{act}} \leq W_{\max} \\[3em] \text{if } P_{\mathrm{n}}\Delta t + W_{\mathrm{act}} > W_{\max} \end{array} \tag{6}$$

with the superscript ′ in Q'_{dr} indicating the estimated amount direct runoff component, without considering the soil permeability.

[14] Equation (6) is only used in a cell with $W_{\min} < W_{\max}$, i.e., a cell with subgrid variability of local soil water capacities. If $W_{\min} = W_{\max}$, it implies that there is no variation of soil water capacities within a cell. In this case, P_n infiltrates if soil is not fully saturated yet, and causes direct runoff if soil is saturated. The estimate of direct runoff for a uniform soil (i.e., $W_{\min} = W_{\max}$) is given as:

$$Q'_{\mathrm{dr}}\Delta t = \begin{cases} 0 & \text{if } P_{\mathrm{n}}\Delta t + W_{\mathrm{act}} \leq W_{\max} \\[1em] P_{\mathrm{n}}\Delta t - \Delta W_{\mathrm{act}} & \text{if } P_{\mathrm{n}}\Delta t + W_{\mathrm{act}} > W_{\max} \end{cases} \tag{7}$$

[15] The amount of infiltration into the first soil layer, P_{01}, is basically the difference between liquid rainfall and direct runoff. However, the infiltration rate P_{01} cannot exceed the saturated conductivity of the first layer, $K_{\mathrm{sat},1}$ [L T^{-1}]. If there is any excess of P_{01} above $K_{\mathrm{sat},1}$, direct runoff is increased. Hence, the partitioning liquid rainfall, P_n, into P_{01} and actual direct runoff component Q_{dr} [L T^{-1}] is given as follows:

$$P_{01}\Delta t = \min \left[P_{\mathrm{n}}\Delta t - Q'_{\mathrm{dr}}\Delta t, K_{\mathrm{sat},1}\Delta t \right] \tag{8}$$

$$Q_{\mathrm{dr}}\Delta t = P_n \Delta t - P_{01}\Delta t \tag{9}$$

2.1.2. Vertical Water Exchange Between Soil and Groundwater Stores

[16] Figure 2a illustrates the conceptualization of vertical fluxes between upper soil stores. In addition to infiltration fluxes P_{01}, changes in soil storages arise due to evaporation E [L T^{-1}] (consisting of bare soil evaporation and plant transpiration) and depletion or specific runoff components

layer. The parameter b [–] in equation (1) defines the distribution of soil water storage capacities within a cell. In this study, the values of b are estimated based on the elevation variability within a cell, such that for a given soil wetness, more runoff is produced in mountainous than flat regions. The value of b ranges from 0.01 for flat areas and up to 0.5 in very rough terrain [see *Hagemann and Gates*, 2003].

[13] Based on equation (1), the following equation (6) estimates the direct runoff from a cell with subgrid variation of soil water capacities (i.e., for the case $W_{\min} < W_{\max}$):

(interflow from S_2 and base flow from S_3), as well as vertical water exchanges between the first and second stores, P_{12} [L T^{-1}], and between the second and groundwater stores, P_{23} [L T^{-1}] (see Figure 2a). P_{12} and P_{23} consist of downward percolation, $P_{1\rightarrow2}$ and $P_{2\rightarrow3}$ [L T^{-1}], and upward capillary rise fluxes, $P_{2\rightarrow1}$ and $P_{3\rightarrow2}$ [L T^{-1}]. These vertical fluxes are driven by the degrees of saturation of both stores: $s_1 = S_1/SC_1$ and $s_2 = S_2/SC_2$ determining the unsaturated conductivity of each soil layer: K_1 (s_1) and K_2 (s_2) [L T^{-1}]. For each layer, this unsaturated conductivity $K(s)$ is calculated as [*Campbell*, 1974]:

$$K(s) = K_{\mathrm{sat}} \times s^{2\beta+3} \tag{10}$$

where the subscript sat indicates saturation, β [–] is an empirical exponent in the matric suction ψ [L] function of *Clapp and Hornberger* [1978]:

$$\psi(s) = \psi_{\mathrm{sat}} \times s^{-\beta} \tag{11}$$

with β varying on average between 4 and 11 over the range from sand to clay.

[17] If the degree of saturation of the top layer is higher than that of the underlying store $(s_1 > s_2)$, the percolation rate $P_{1\rightarrow2}$ occurs with the rate of unsaturated conductivity of the first soil layer: K_1 (s_1). If $s_1 < s_2$, capillary rise sustains the soil moisture deficit in the top layer with upward fluxes equal to: $K_2(s_2)$ $(1 - s_1)$.

[18] To calculate the capillary rise $P_{3\rightarrow2}$ for the second soil layer, the method of *Gardner* [1958] and *Eagleson* [1978] is adopted [see also *Soylu et al.*, 2011]. Assuming a steady state condition with a suction head at the surface that is (negatively) large (i.e., dry soil surface), the maximum capillary rise rate $CR_{\max}$ [L T^{-1}], is given as a function of groundwater table position:

$$CR_{max} = K_{sat,2} \left[1 + \frac{3}{2+6/\beta_2} \right] \left(\frac{|\psi_{sat,2}|}{Z_{gw}} \right)^{2+3/\beta_2} \tag{12}$$

where Z_{gw} [L] is the difference between surface level (from the digital elevation model) and groundwater head h [L] (from the MODFLOW groundwater model output, see section 2.2). Equation (12) is used to estimate the maximum $P_{3\rightarrow2}$ and limited by $K_{sat,2}$ if the groundwater is at or above the surface ($Z_{gw} \leq 0$). The model also limits the capillary rise fluxes $P_{3\rightarrow2}$ and $P_{2\rightarrow1}$ such that they do not make the upper soil storages exceed the equilibrium water content, W_{equ} [$L\ T^{-1}$]. The form of this equilibrium soil moisture profile, determined from the balance between the gravity (downward direction) and pressure head gradient (which tends to draw moisture up), is given as [*Clapp and Hornberger*, 1978; *Koster et al.*, 2000]:

$$s_{equ}(z) = \left(\frac{\psi_{sat} - z}{\psi_{sat}} \right)^{(-1/\beta)} \tag{13}$$

where s_{equ} [–] is the degree of saturation at a height z above the water table. The equilibrium storage W_{equ} is determined by integrating s_{equ} from groundwater to surface levels.

2.2. The Groundwater Model Part

[19] The link between the land surface model part of PCR-GLOBWB-MOD—written in the PCRaster scripting language [*Wesseling et al.*, 1996]—and the MODFLOW groundwater model, is provided by a PCRaster-MODFLOW extension developed by *Schmitz et al.* [2009]. Using this extension, the dynamics of groundwater head h and lateral groundwater flow through aquifer bodies are simulated on a daily basis. A single layer MODFLOW groundwater model is constructed and forced with the output from the land surface model part, specifically the net groundwater recharge $(P_{23} = P_{2\rightarrow3} - P_{3\rightarrow2})$ and surface water levels HRIV [L]. The latter is derived from the routed discharge Q_{chn} by assuming channel dimensions and properties based on geomorphological relations to bankfull discharge [*Lacey*, 1930; *Savenije*, 2003]. Within the MODFLOW model, the "recharge" (RCH) package is used to introduce P_{23}, while the "river" (RIV) and "drain" (DRN) packages are used to introduce HRIV as the boundary conditions of the MODFLOW model [see *Sutanudjaja et al.*, 2011].

[20] The implementation of RIV and DRN packages gives the possibility to quantify flow between stream and aquifer, symbolized as $-(qRIV + qDRN)$ [$L\ T^{-1}$] (with the negative sign is introduced as MODFLOW uses a positive sign for flow entering aquifer). The amount of $-(qRIV + qDRN)$, which depends on the difference between groundwater head h and surface water level HRIV, is the main component of the base flow Q_{bf}, especially for channels in flat sedimentary pockets where groundwater flow is relatively slow. However, the magnitude of $-(qRIV + qDRN)$ is too small to satisfy the fast base flow component originated from mountainous areas, where many springs tapping groundwater are located higher up in the valleys. To include this fast base flow component, feeding tributaries to main rivers, it is assumed that the groundwater above the flood plain is drained based on a linear reservoir concept as follows:

$$Q_{bf} = -(qRIV + qDRN) + (JS_{3,fpl}) \tag{14}$$

where $S_{3,fpl}$ [L] is the groundwater storage above the flood plain and J [T^{-1}] is a recession coefficient parameterized based on [*Kraaijenhoff van de Leur*, 1958]:

$$J = \frac{\pi^2(KD)}{4\,Sy\,L^2} \tag{15}$$

with KD [$L^2\ T^{-1}$] and Sy [–] indicating aquifer transmissivities and specific yields, and L [L] indicating average flow lengths.

[21] The consequence of incorporating the fast-response base flow component—represented by the second term of equation (14): $(JS_{3,fpl})$—is that the water balance of the model must be closed by subtracting it from the input of the MODFLOW recharge package, RCH_{inp} [$L\ T^{-3}$]:

$$RCH_{inp} = (P_{23} - JS_{3,fpl})\,A_{cell} \tag{16}$$

where A_{cell} [L^2] is the surface area of each 30 arc sec cell. In equation (16), RCH_{inp} can have negative values, indicating water leaving the groundwater store (S_3) via capillary rise to its overlying soil stores or base flow to surface water bodies.

3. Calibration Procedure

[22] To calibrate the model, we simulated a large number of runs with different parameter values (explained in section 3.1). All runs used the same forcing (section 3.2). Subsequently, using observed discharge and remotely sensed soil moisture data (section 3.3), we evaluated and compared the model performances of all runs. The "best fit" or "optimal" parameter sets were then identified as the "calibrated" parameter sets. As in situ discharge observations are traditionally used to calibrate hydrological models, we identified the calibrated parameter sets based on discharge time series (alone). Also we evaluated the soil moisture fits of all runs to SWI time series and identified the calibrated parameter sets based on (only) these remotely sensed soil moisture time series.

[23] While evaluating the model fits of all runs, we examined how well unique parameter sets could be identified by inspecting the behaviors of various objective functions and their parameter spaces. These analyses were done based on observed discharge and soil moisture SWI time series.

[24] Based on these analyses, we then performed a stepwise calibration procedure to identify the parameter sets that were able to simultaneously reproduce observed discharge and SWI time series. We started this stepwise approach by calibrating all model parameters based on discharge data. Subsequently, the parameters expected to influence the soil moisture performance were recalibrated. By doing this, we could obtain a parameter set that fits well with both observed discharge and SWI time series.

Table 1. Aquifer Parameters for the REFERENCE Run[a]

Lithology/Aquifer Class	Transmissivity KD (m^2/d)
Nonconsolidated sediments	533.52
Carbonate sedimentary rocks	67.17
Volcanic rocks	13.40
Crystalline rocks	0.50
Siliciclastic sedimentary rocks	13.40

[a]The values are adopted from *Gleeson et al.* [2011] who attributed the global lithological map of *Dürr et al.* [2005] with the geometric mean permeability of each class in the map (Figure 1c).

[25] Next, as the last step, we compared the performance of runs with the calibrated parameter sets based on observed discharge, SWI and the combination of both. More specifically, we compared their simulated groundwater head time series to measurements. In this way, we assessed to what extent each observation type (i.e., discharge and SWI) and the combination (of both) is suitable for the calibration.

3.1. Model Parameters and Runs

[26] A 3045 model runs were simulated with different values of minimum (local) soil water capacities W_{min}, upper soil saturated conductivities $K_{sat,1}$ and $K_{sat,2}$, and aquifer transmissivities KD. These various parameters, which are spatially varying over the study area, control different parts of the model (see Figure 2). We used spatially uniform prefactors to vary these parameters over 3045 model runs (see Table 2).

[27] The parameter W_{min}, defined in the improved Arno scheme for conceptualizing the subgrid variation of the soil saturation stores [*Todini*, 1996; *Hagemann and Gates*, 2003; *van Beek and Bierkens*, 2009], controls the partitioning of rainfall into direct runoff and infiltration to the soil (section 2.1.1). For $W_{min} = 0$, direct runoff always occurs for a rainfall event. If $W_{min} > 0$, no direct runoff occurs as long as the total soil storage does not exceed W_{min}. The parameters $K_{sat,1}$ and $K_{sat,2}$, defined in the land surface model part, mainly affect groundwater recharge as they control fluxes between the upper soil stores P_{12} and to the groundwater store P_{23} (section 2.1.2). The transmissivity KD is a MODFLOW groundwater model parameter controlling lateral groundwater flow and also governing the base flow recession coefficient J (section 2.2).

[28] A REFERENCE run was defined before we started this procedure. Figures 1g and 1h, respectively, show the spatially distributed parameters $K_{sat,1}$ and $K_{sat,2}$ used in the REFERENCE run, for which $W_{min} = 0$ is uniformly introduced within the study area and Table 1 is used to assign

KD based on the lithological map in Figure 1c [*Dürr et al.*, 2005; *Lehner et al.*, 2008; *Sutanudjaja et al.*, 2011; *Gleeson et al.*, 2011].

[29] To vary the parameter values of a run over the others, three nonspatially distributed prefactors are introduced (uniformly within the study area): f_W, f_K, and f_{KD}. Table 2 shows how these prefactors are introduced for 3045 model runs. The prefactors used are referred to the parameters of the REFERENCE run, which has $f_W = 0$, $f_K = 0$, and $f_{KD} = 0$. The prefactor f_W is used to vary W_{min}, the prefactor f_K is introduced to vary simultaneously $K_{sat,1}$ and $K_{sat,2}$, while the prefactor f_{KD} is used to vary KD. All other model parameters remain fixed.

3.2. Forcing Data and Simulation Period

[30] The meteorological forcing data, supplied on a daily basis to force the land surface model, consist of temperature, precipitation, and reference potential evaporation. For all model runs, we used the daily precipitation and temperature input time series obtained by downscaling the monthly CRU-TS2.1 [*Mitchell and Jones*, 2005] data set with the daily ERA-40 reanalysis [*Uppala et al.*, 2005] data (see also *Sutanudjaja et al.* [2011], for detailed methodology]. For the reference potential evaporation, we used the data set of *van Beek* [2008], derived based on the FAO Penman-Monteith method [*Allen et al.*, 1998] using the climatological fields from CRU-TS2.1 and CRU-CL1.0 [*New et al.*, 1999].

[31] In this study, all simulations of PCR-GLOBWB-MOD were performed for the period 1992–2008. As the initial condition for each simulation, the model was spun-up—using the average forcing from the years 1974–2008—until reaching a dynamically steady state. Moreover, to be safe from spin-up problem, we used only the results from the period 1995–2000 for analyzing the model performance. The following section describes the approach for evaluating results of each model runs (based on observed discharge, soil moisture, and groundwater head data).

3.3. Model Evaluation Criteria
3.3.1. Using Discharge Measurement Time Series

[32] The discharge performance of each run was evaluated at daily resolution. The modeled discharge time series, Q_{mod} (i.e., the routed discharge Q_{chn}), were compared with the observations, Q_{obs}, at two downstream locations (see Figure 1f): Lobith (Rhine) and Borgharen (Meuse). In both locations, we determined the *Nash and Sutcliffe* [1970] efficiency coefficients, NS_{eff}:

Table 2. Parameter Values Used in the Calibration Processes[a]

Prefactors	Parameter Values	Number of Discrete Values
$f_W \in \{0, 0.25, 0.5, 0.75, 1\}$	$W_{min} = f_W \times W_{max} = f_W \times (SC_1 + SC_2)$	5
$f_K \in \{-3.5, -3.25, -3, \ldots, 0, \ldots, 3, 3.25, 3.5\}$	$\log(K_{sat,1}) = f_K + \log(K_{sat,1,ref})$	29
	$\log(K_{sat,2}) = f_K + \log(K_{sat,2,ref})$	
$f_{KD} \in \{-2.5, -2.25, -2, \ldots, 0, \ldots, 2, 2.25, 2.5\}$	$\log(KD) = f_{KD} + \log(KD_{ref})$	21
Total number of model runs: 3045		

[a]The subscript ref indicates the REFERENCE run with the prefactors $f_W = 0$, $f_K = 0$, and $f_{KD} = 0$. $K_{sat,1,ref}$ and $K_{sat,2,ref}$ are shown in Figures 1g and 1h, while KD_{ref} is based on Table 1 and Figure 1c. W_{max} is given in Figure 1i. All parameters have spatially distributed values, except f_W, f_K, and f_{KD}.

$$NS_{eff} = 1 - \frac{\sum \left(Q_{obs,t} - Q_{mod,t}\right)^2}{\sum \left(Q_{obs,t} - \overline{Q_{obs}}\right)^2} \qquad (17)$$

with $\overline{Q_{obs}}$ indicating the average of observed data and t indicating the time index. We also calculated the discharge efficiency coefficients NS_{ano} using anomaly time series of the model results and observation data:

$$NS_{ano} = 1 - \frac{\sum \left((Q_{obs,t} - \overline{Q_{obs}}) - (Q_{mod,t} - \overline{Q_{mod}})\right)^2}{\sum \left(Q_{obs,t} - \overline{Q_{obs}}\right)^2} \qquad (18)$$

with $\overline{Q_{mod}}$ indicating the average of modeled time series.

3.3.2. Using Soil Water Index Time Series

[33] The Soil Water Index (SWI) product is derived from European Remote Sensing (ERS) scatterometers and provides spatiotemporal estimates of the first meter profile soil moisture content at about 25–50 km resolution [*Wagner et al.*, 1999b]. SWI [–] time series are derived based on SSM [–] (Surface Soil Moisture) time series measured by the active scatterometers of ERS. The scatterometers retrieve SSM 3–4 times per week, which represent soil moisture content in the top soil layer (as discussed in *Wagner et al.* [1999b]). The retrieval of SSM is based on a change detection approach of backscatter measured by the scatterometers. The approach assumes that backscatter measurements have a linear relationship to soil moisture and are corrected from the vegetation effects by knowing the typical vegetation phenology and how it affects backscatter measurements [*Wagner*, 1998; *Wagner et al.*, 1999a, b]. SSM time series are then derived from the vegetation-corrected backscatter time series as relative saturation with the historically lowest and highest values are assigned to the 0% (dry) and 100% (wet) references, respectively. To derive SWI time series, the following low-pass filter is then applied to SSM time series [*Wagner et al.*, 1999b]:

$$SWI_{d_{SWI}} = \frac{\sum_{d_{SSM}}^{d_{SWI}} \left[SSM_{d_{SSM}} \times e^{-(d_{SWI} - d_{SSM})/T}\right]}{\sum_{d_{SSM}}^{d_{SWI}} \left[e^{-(d_{SWI} - d_{SSM})/T}\right]} \qquad (19)$$
$$\text{for} \quad d_{SSM} \leq d_{SWI}$$

where d_{SWI} and d_{SSM} are the (daily) time indexes of SWI and SSM, and T [T] is the characteristic time length, taken as 20 days as [*Wagner et al.*, 1999b] showed that it provided the best correlation to the field data [see also *Wagner*, 1998].

[34] In their original format, SWI fields have about 25–50 km and 10 day resolution. For this study, the SWI data were resampled to 30 arc min (about 50 km at the equator) and monthly resolution. This is to reduce the number of missing values that often occur due to snow cover or frozen soil condition (as example, see Figure 1f). In this study, the monthly SWI time series are used as the benchmark for evaluating the soil moisture simulation of PCR-GLOBWB-MOD.

[35] For each model run, we compared the time series of the modeled degree of saturation s_{12} [–] of the (entire) upper soil stores of PCR-GLOBWB-MOD to

SWI. The modeled degree of saturation s_{12} is calculated as follows:

$$s_{12} = \frac{S_1 + S_2}{SC_1 + SC_2} \qquad (20)$$

[36] Evaluations s_{12} to SWI were done at monthly resolution. As a measure of performance, we calculated the cross-correlation coefficient ρ_{SM} between SWI and s_{12} time series. This performance indicator—calculated without considering lags—evaluates the timing of modeled to measured time series.

[37] As a measure of accuracy, we evaluated the mean absolute error of modeled soil moisture time series. While evaluating this error, we have to consider that while the simulated s_{12} is given in its absolute saturation value (equation (20)), the remotely sensed data SWI are given in its relative saturation value (see the previous explanation, also *Brocca et al.* [2011]). Moreover, there can be systematic differences or biases between satellite-derived soil moisture time series and soil moisture time series simulated with hydrological models [*Reichle and Koster*, 2004; *Drusch et al.*, 2005]. These systematic differences, attributable to difference in layer thickness or characteristics of data sets (e.g., minimum and/or maximum reference values, as discussed in this study), prevent a statistically optimal analysis and therefore have to be corrected [*Dee and Da Silva*, 1998]. Hence, due to the expected discrepancy between the reference values of both s_{12} and SWI time series, the cumulative distribution function (cdf) matching technique was implemented to rescale s_{12} against SWI time series [see *Reichle and Koster*, 2004; *Liu et al.*, 2011, 2012]. Using this cdf matching procedure, we adjusted the modeled soil moisture time series to the same range and distribution of SWI time series. More specifically, we rescaled the actual values s_{12} to the scaled ("corrected") values s'_{12} [–] such that the cdf curves of s'_{12} and SWI match. After rescaling, we then evaluated the mean absolute error between s'_{12} and SWI time series—symbolized as MAE_{SM-CDF} [–]. Due to the cdf matching, MAE_{SM-CDF} is only a measure of precision as bias has been removed. The correlation coefficient between s'_{12} and SWI time series—symbolized as ρ_{SM-CDF} [–]—was also evaluated.

3.3.3. Using Groundwater Head Measurement Time Series

[38] The modeled groundwater head time series were compared to head observations from the uppermost aquifer. There are 4250 stations used in this study. Evaluations were done at monthly resolution. In each station, we calculated the correlation coefficient between the monthly observed and simulated time series—symbolized as ρ_{HEAD}. Also, we calculated the mean absolute error of predicted time series. Note that biases between observed groundwater head time series are expected, especially due to the discrepancies in the resolution and elevation references of 30 arc sec resolution model results and point-scale measurement data [see *Sutanudjaja et al.*, 2011]. Hence, to evaluate the error of simulated head dynamics, we preferred to calculate the mean absolute error MAE_{ano} using the anomaly time series of the model results and observation data, i.e., deviations from the time series mean values.
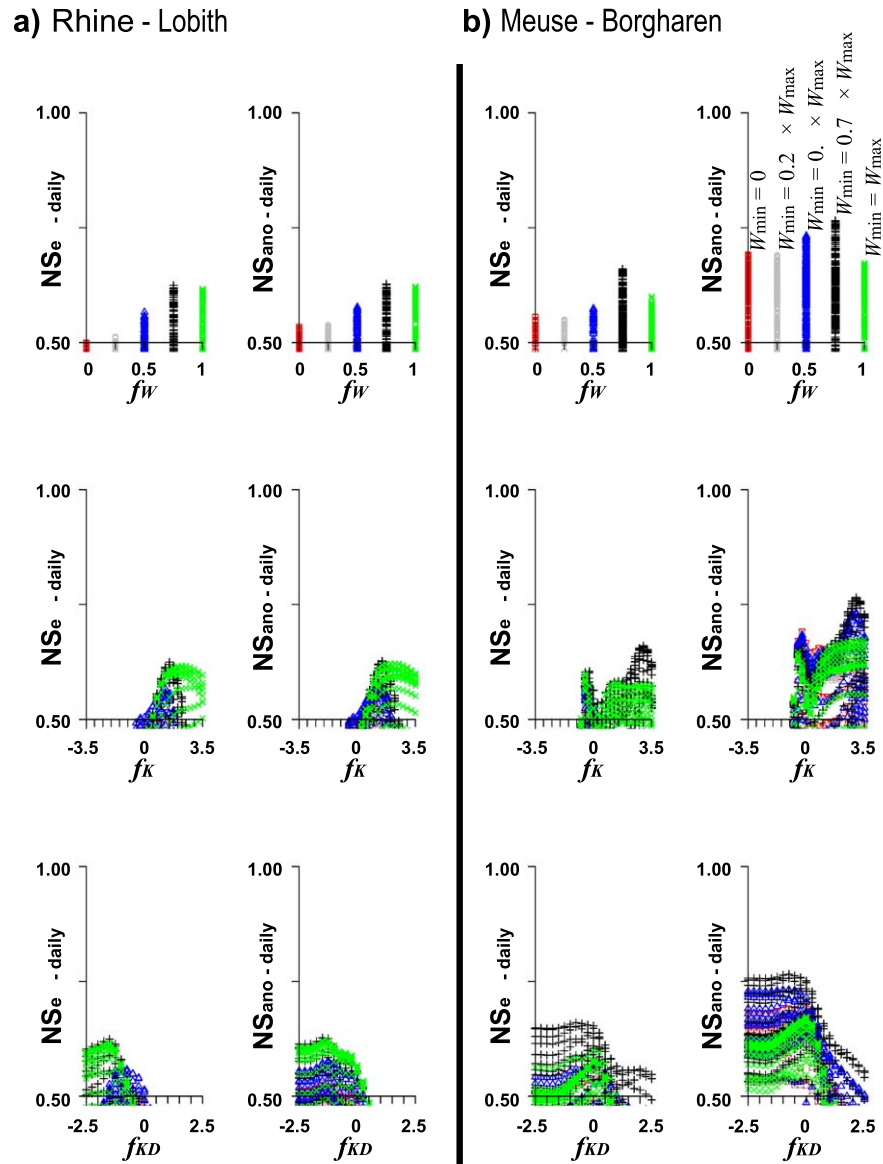
Figure 3. Scatterplots of discharge performance indicators NS_{eff} and NS_{ano} (y axes) based on the daily observations in (a) Lobith, Rhine basin and (b) Borgharen, Meuse basin versus prefactors f_W, f_K, and f_{KD} (x axes). In each column, the three graphs plot the same performance on the y axes using different parameters on the x axes. The prefactor f_W controls the partitioning of rainfall into direct runoff and infiltration to the soil (section 2.1.1), while the prefactor f_K affects groundwater recharge (section 2.1.2) and the prefactor f_{KD} controls lateral groundwater flow and base flow (section 2.2). Different colors and dot shapes indicate different f_W.

4. Results

4.1. Calibration Using Discharge Data

[39] To summarize the results of all 3045 runs, Figures 3a and 3b present the scatterplots between the prefactors (given in x axes: f_W, f_K, and f_{KD}) and the discharge performance indicators (y axes: NS_{eff} and NS_{ano}), for the Rhine basin and Meuse basin, respectively. For the Rhine (Figure 3a), we find that the scatterplots of NS_{eff} and NS_{ano} are similar. Yet, for the Meuse (Figure 3b), the values of NS_{ano} are higher than NS_{eff}. This difference results from a bias between simulated and observed time series that can be explained by the fact that the model did not include the

water diversion from the Meuse River to sustain the navigation function of the Scheldt River, as reported by *de Wit* [2001].

[40] The scatterplots in Figures 3a and 3b show that all prefactors are identifiable. We can readily distinguish parameter ranges leading to better and worse performance, as well as their global optimal values. A combination of high f_W, high f_K, and low f_{KD} provides the best performance in terms of discharge. For the Rhine, the optimal discharge performance is obtained with $f_W = 0.75$, $f_K = 1.5$, and $f_{KD} = -1.5$. The run with this parameter set reproduces the observed discharge at Lobith well, as plotted in Figure 4a (at monthly resolution: $NS_{eff} = 0.76$ and $NS_{ano} = 0.77$) and
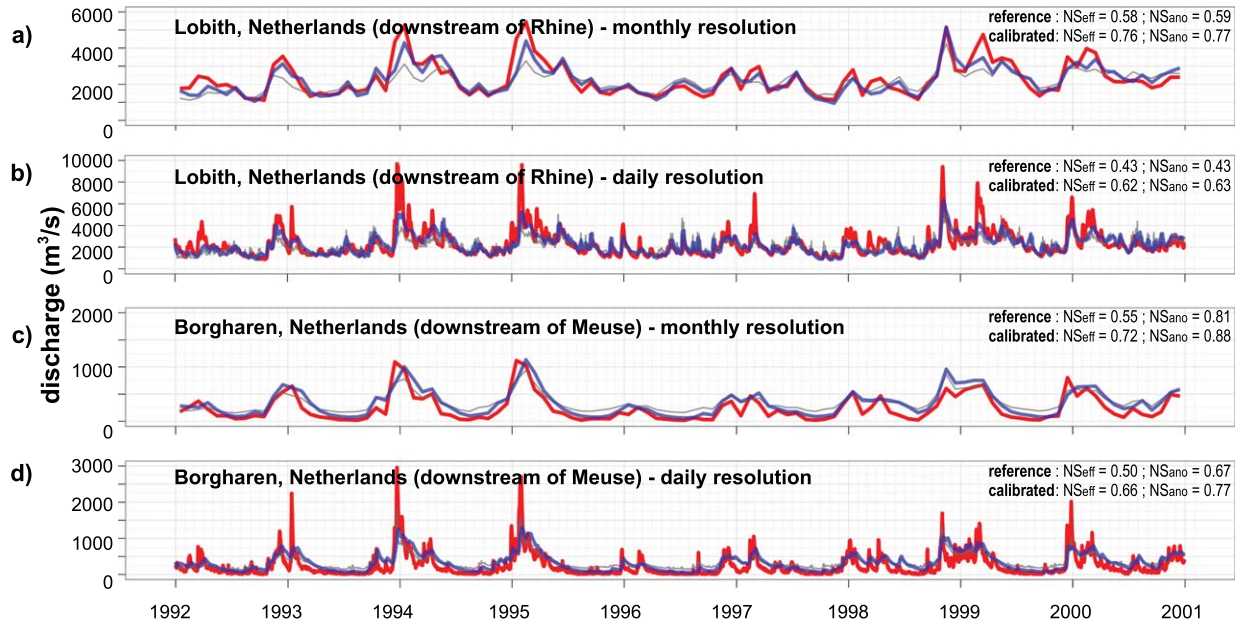
Figure 4. Comparisons between measured discharge (red) and simulated discharge (gray and blue) time series. The gray lines are from the REFERENCE (uncalibrated run) and the blue lines are from the calibrated time series, which are taken from the runs that yield the best simulations based on discharge observations near the basin outlets (see section 4.1).

Figure 4b (at daily resolution: $NS_{eff} = 0.62$ and $NS_{ano} = 0.63$), except some extreme peaks which cannot be simulated properly. This lack of fit may be due to errors in the forcing data and other parameters, such as the snow module and soil thickness parameters which were not calibrated. For the Meuse, the optimal parameter set is a combination of $f_W = 0.75$, $f_K = 3$, and $f_{KD} = -0.75$. The modeled discharge time series of this run at Borgharen, given in Figure 4c (monthly resolution, $NS_{eff} = 0.72$ and $NS_{ano} = 0.88$) and Figure 4d (daily resolution, $NS_{eff} = 0.66$ and $NS_{ano} = 0.77$), indicate good model performance. As expected, the major drawback of this run is that the low flow events during the summer cannot be properly simulated as the model does not incorporate water diversion from the river.

[41] We also performed a single model run that reproduces well the discharge data of both rivers, simultaneously. For this run, the combination of $f_W = 0.75$, $f_W = 1.5$, and $f_{KD} = -1.5$ was assigned for the Rhine basin (which gives the highest NS_{eff} and NS_{ano} at Lobith), while the combination of $f_W = 0.75$, $f_K = 3$, and $f_{KD} = -0.75$ was assigned for the Meuse basin (providing the highest NS_{eff} and NS_{ano} at Borgharen). This combined set—considered as the calibrated prefactors based on discharge data—gives discharge time series which are identical to the best runs discussed previously and presented in Figure 4 (see also Table 4).

4.2. Calibration Using Soil Water Index

[42] In Figures 5a and 5b, we plotted the simulated soil moisture time series of the REFERENCE run against SWI time series for pixel 21, located in the low-lying areas (see Figure 1e, for the pixel location). In Figure 5a, the original simulated soil moisture time series s_{12} are plotted, while

the rescaled time series s'_{12} (after cdf matching technique implemented, see section 3.3.2) are plotted in Figure 5b. In both figures, the observed SWI time series are also given. Comparing both figures, we observe that the cdf matching technique effectively removes the biases between the simulated and observed soil moisture time series. The soil moisture performance indicators are given in every chart in Figure 5. From all 3045 model runs, we identified the best run that yields the highest ρ_{SM} for pixel 21 and plotted it in Figure 5c.

[43] The results from all 3045 model runs are summarized in Figure 6 showing the scatterplots between the prefactors used to vary parameter values (x axes: f_W, f_K, and f_{KD}) and soil moisture performance indicators (y axes: objective functions ρ_{SM}, $\rho_{SM\text{-}CDF}$, and $MAE_{SM\text{-}CDF}$) for pixel 21. For all objective functions ρ_{SM}, $\rho_{SM\text{-}CDF}$, and $MAE_{SM\text{-}CDF}$, the prefactors f_W and f_{KD} are not identifiable. Hence, it can be concluded that the parameters W_{min} and KD cannot be calibrated if we rely on SWI data alone. The prefactor f_K is somewhat better identifiable. This implies that SWI may be used to indirectly tune groundwater recharge by calibrating the upper soil saturated hydraulic conductivities, K_{sat}. Yet, there are two distinct local optima observed in Figure 6: $f_K = -2$ (global optimum) and $f_K = 2.75$ (local optimum). The occurrence of these two optima indicates that the calibration of K_{sat} is not trivial if we only calibrate the model output to monthly SWI time series. Other observations, such as discharge or groundwater head measurements, are needed to further constrain this parameter.

[44] The results found for pixel 21 are exemplary for the other pixels located in the low-lying areas, specifically the Dutch and Flemish lowlands, which are sedimentary basin
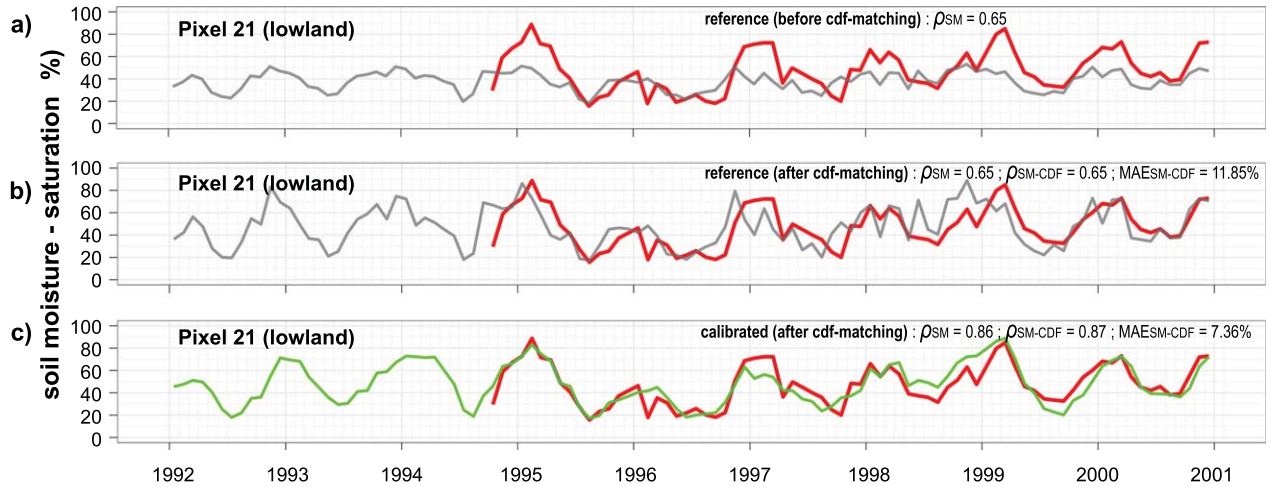
Figure 5. Comparisons between remotely sensed soil moisture data of SWI (red) and simulated soil moisture (degree of saturation) from PCR-GLOBWB-MOD (gray and green). Time series given are for pixel 21 (see Figure 1e, for the pixel location): (a and b) from the REFERENCE run (gray); (c) from the calibrated or best run (green) that yields the highest correlation ρ_{SM} to SWI data at pixel 21 (see section 4.2). The simulated soil moisture time series in Figure 5a is plotted before the cdf matching technique implemented, while the simulated time series in Figures 5b and 5c are rescaled based on the cdf matching technique (section 3.3.2).

areas having shallow groundwater tables (see Figures 1b and 1c). As illustration for other pixels, Figures 7a and 7b show the scatterplots between the soil moisture performance indicators ρ_{SM} and f_K for pixels 197 and 112 (see Figure 1e, for the pixel location). Pixel 197 is an example for a location in the mountainous regions, while pixel 112 is representative for the Upper Rhine Graben area, which is an intermediate site between low-lying and mountainous regions.

[45] Figure 7b shows that in mountainous regions model performances are low ($\rho_{SM} \leq 0.29$). A possible explanation for this reduced performance is the fact that scatterometer signals from mountainous regions often contain artifacts affecting the accuracy of SWI [*Wagner*, 1998; *Wagner et al.*, 1999a]. The scatterplots of f_W and f_{KD} are not presented in Figure 7 as they are hardly identifiable for all pixels in the study area (like the case for pixel 21 in Figure 6). Moreover, as the performance indicator, we plotted only ρ_{SM} as there are strong and significant correlations among the objective functions used to evaluate soil moisture performance, as illustrated in Figure 8. These correlations are found at all pixels. Figure 6 shows this in addition to Figure 8. Hence, most analyses and discussions in this paper are simplified by focusing only on ρ_{SM}.

[46] Figure 7a shows that the performance in an intermediate site (e.g., pixel 112) is somewhat better than that in mountainous regions, but still lower than the performance in low-lying regions (Figure 6). Prefactor f_K is difficult to identify. This is corroborated by Figure 9 that shows the prefactor spaces of pixel 112. In Figure 9, the magnitude of ρ_{SM}—indicated by the sizes of circles—is plotted for all combinations of prefactors. Figures 9a–9e show ρ_{SM} for all combinations f_K and f_{KD}, with f_W set to defined values (i.e., $f_W = 0$ in Figure 9a, $f_W = 0.25$ in Figure 9b, $f_W = 0.5$ in Figure 9c, $f_W = 0.75$ in Figure 9d, $f_W = 1$ in Figure 9e, respectively). Prefactor f_K is identifiable for $f_W \leq 0.5$. For

$f_W > 0.5$, f_K becomes less identifiable. Overall, f_K is hardly identifiable. The latter conclusion is corroborated by Figure 9f that shows the maximum values of ρ_{SM}—also indicated by the sizes of circles—from all combinations of f_K and f_{KD} in Figures 9a–9e. In Figure 9f, the values of f_W—indicated by the subscript *opt*—are varying for every combination of f_K and f_{KD}. Similarly to Figure 9f, the prefactor space of f_K and f_W are given in Figure 9g, while Figure 9h shows the prefactor space of f_W and f_{KD}. Figure 9g indicates that there is mutual dependence between increasing f_K and f_W for high ρ_{SM}. This suggests that there are multiple solutions leading to an equal performance. This equifinality is also apparent from Figure 9h that shows that the prefactors f_W and f_{KD} are hardly identifiable.

[47] To illustrate the study area-wide soil moisture performance, we plotted the study area averages of ρ_{SM} against the prefactor f_K in Figure 7c. Figure 7c shows relatively weak performance ($\rho_{SM} < 0.65$). This low performance is expected as the study area (Figure 1) is dominated by mountainous areas, where SWI data most likely contain artifacts. From the results of 3045 model runs, we identified one single run with the "best" soil moisture performance, i.e., the run with the highest study area-average value of ρ_{SM} (Figure 1). This run may be considered as the calibrated run based on SWI data alone. Having identified this "best" run for soil moisture, we then checked the discharge performance of this run. Unfortunately, this results in relatively low discharge performance, with $NS_{eff} = 0.01$ and $NS_{ano} = 0.04$ for its daily discharge simulation at Lobith and $NS_{eff} = 0.13$ and $NS_{ano} = 0.25$ for its daily discharge simulation at Borgharen. These efficiency coefficients are much lower than the ones for the run calibrated using (only) the discharge data (see Figure 4 and section 4.1).

[48] Table 3 summarizes the identifiability of each parameter based on each observation type: in situ discharge data (section 4.1) and remotely sensed soil moisture data of
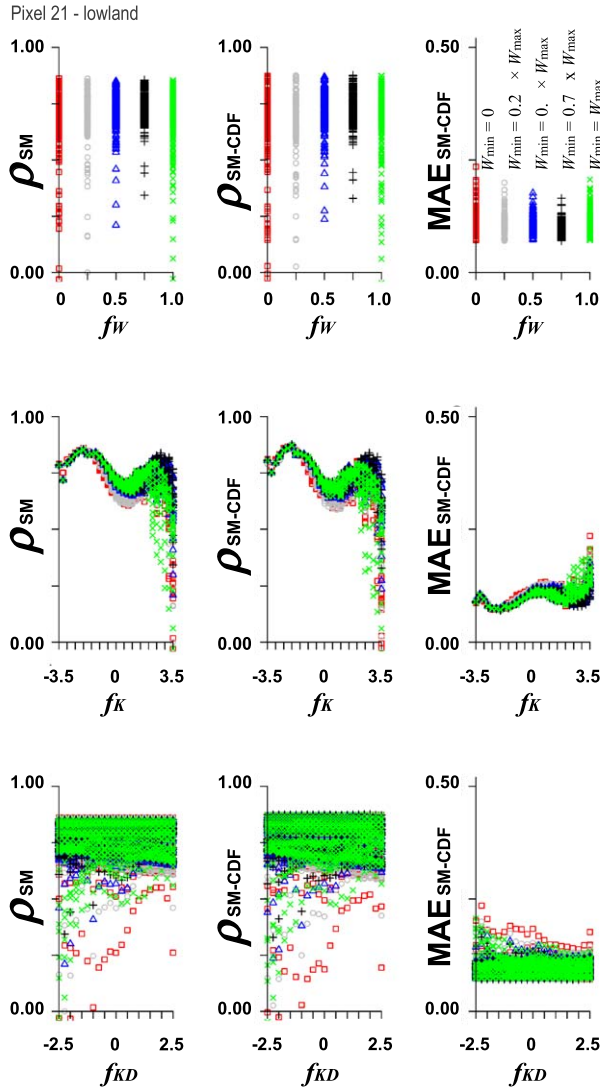
Figure 6. Scatterplots of soil moisture performance indicators ρ_{SM}, $\rho_{SM\text{-}CDF}$, and $MAE_{SM\text{-}CDF}$ versus prefactors f_W, f_K, and f_{KD} for pixel 21 (see Figure 1e for the pixel location). Different colors and dot shapes indicate different f_W.
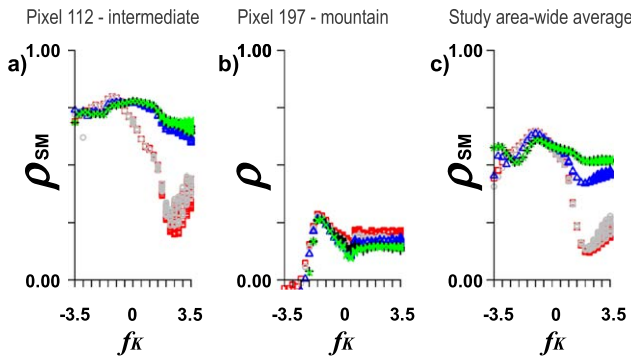

Figure 7. Scatterplots of soil moisture performance indicators ρ_{SM} versus prefactor f_K for pixels (a) 112, (b) 197 (see Figure 1e, for the pixel locations), and (c) using study area-average values.
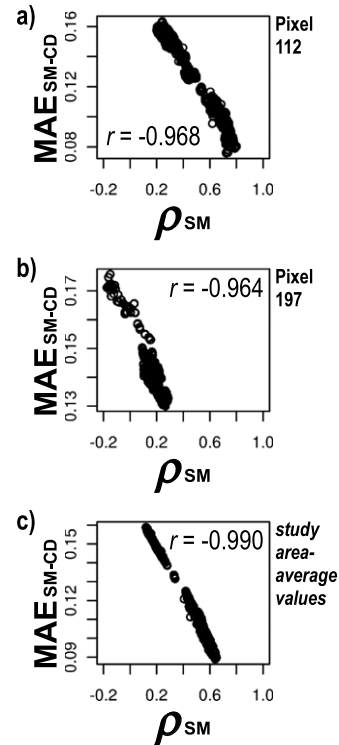
Figure 8. Scatterplots between soil moisture performance indicators ρ_{SM} and $MAE_{SM\text{-}CDF}$, showing their significant correlations (all p values < 0.001), shown for pixels (a) 112, (b) 197 (see Figure 1e, for the pixel locations), and (c) using study area-average values.

SWI (section 4.2). In this table, if identifiable, the optimal prefactor values are listed. The highest performance indicators (from all 3045 model runs) are also given in the table.

4.3. Step-Wise Calibration: Using Discharge and Soil Water Index Time Series

[49] The results and analyses discussed in sections 4.1 and 4.2 are summarized in Table 3 showing the parameter identifiabilities and optimal values according discharge and SWI data. Learning from these results, a stepwise procedure combining discharge and soil moisture calibration was introduced and explained as follows.

[50] We started this multiobjective calibration by adopting the parameter sets identified during the discharge calibration: $f_W = 0.75$ for the entire study area; $f_K = 1.5$ and $f_K = -1.5$ for the Rhine basin; and $f_K = 3$ and $f_{KD} = -0.75$ for the Meuse basin (see Table 3). As f_K is identifiable from soil moisture, we recalibrated the prefactor f_K based on the objective function ρ_{SM} such that the pixel-scale values of ρ_{SM} were optimized. In this procedure, we assigned different f_K for different half arc-degree pixels. However, the recalibration of f_K was limited to pixels that have $\rho_{SM} \geq 0.75$ from any of 3045 model runs performed previously (e.g., pixels 21, 25, 38, 39, and 112, as indicated in Table 3). For pixels with all runs having $\rho_{SM} < 0.75$ (such as pixels 117 and 197, Table 3), we suspected that SWI data contain artifacts (e.g., due to the influence of mountainous landscape) and decided to fix f_K based on the
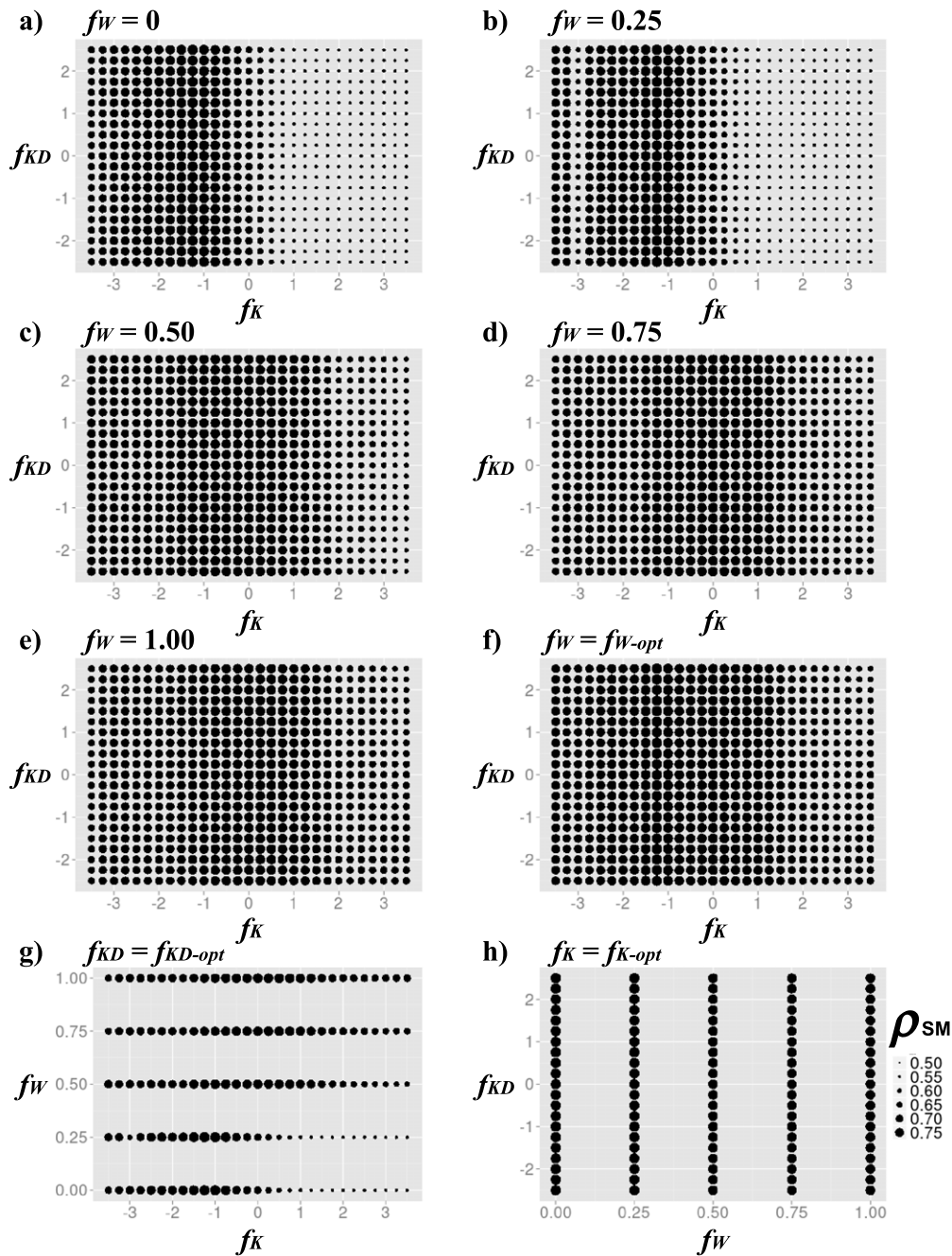
Figure 9. Prefactor spaces corresponding to the soil moisture correlation objective function ρ_{SM}—indicated by areas of circles—for pixel 112 (see Figure 1e, for the pixel location). The subscript *opt* in Figure 9f–9h indicates several optimal values of a certain prefactor that give the highest ρ_{SM} in a combination with the other two prefactors. These results shows that soil moisture data are less suitable for determining optimal prefactors for this pixel, which is located in an intermediate site between low-lying and mountainous regions.

discharge calibration. For pixels having $\rho_{SM} \geq 0.75$—most of which are located in the Dutch and Flemish lowlands, the recalibration of f_K was constrained such that f_K must be higher or equal to the reference value (i.e., $f_K \geq 0$). This constraint was added because we found that too low f_K values may lead to good soil moisture reproduction but at that same time lead to bad discharge reproduction. This is consistent to the analysis during the discharge calibration in section 4.1, in which we found that high saturated hydraulic conductivities are required to yield good discharge reproduction.

[51] The run identified from this stepwise calibration procedure results in good discharge performance with $NS_{eff} = 0.62$ and $NS_{ano} = 0.62$ for its daily discharge simulation at Lobith (Rhine), and $NS_{eff} = 0.66$ and $NS_{ano} = 0.77$ for its daily simulated discharge at Borgharen (Meuse).

Table 3. Parameter Identifiabilities and Optimal Values According Each Observation Type

Observation Type	f_W	f_K	f_{KD}	Performance[a]
Discharge at Lobith (Rhine)—Figure 3	0.75	1.5	−1.5	$NS_{eff} \leq 0.62$ (0.43)
Discharge at Borgharen (Meuse)—Figure 3	0.75	3.0	−0.75	$NS_{eff} \leq 0.66$ (0.50)
SWI for pixel 21 (lowland)—Figure 6	NI[b]	−2 and 2.75[c]	NI	$\rho_{SM} \leq 0.86$ (0.67)
SWI for pixel 25 (lowland)	NI	−2.25 and 2.75	NI	$\rho_{SM} \leq 0.82$ (0.61)
SWI for pixel 38 (lowland)	NI	−2 and 1.75	NI	$\rho_{SM} \leq 0.84$ (0.56)
SWI for pixel 39 (lowland)	NI	−2.25 and 2.5	NI	$\rho_{SM} \leq 0.83$ (0.57)
SWI for pixel 112 (intermediate)—Figure 7a	NI	NI	NI	$\rho_{SM} \leq 0.80$ (0.69)
SWI for pixel 117 (mountainous)	NI	NI	NI	$\rho_{SM} \leq 0.64$ (0.54)
SWI for pixel 197 (mountainous)—Figure 7b	NI	NI	NI	$\rho_{SM} \leq 0.29$ (0.19)
SWI study area-wide average values—Figure 7c	NI	−1	NI	$\rho_{SM} \leq 0.65$ (0.58)

[a]Values within parentheses are from the REFERENCE run, which has $f_W = 0$, $f_K = 0$, and $f_{KD} = 0$.
[b]NI indicates that a parameter is nonidentifiable.
[c]Two optimal values of f_K are found for the pixels located in low-lying areas.

These efficiency coefficients are similar to the ones obtained from calibrating on discharge alone (see Figure 4 and Table 4).

4.4. Verification to Observed Groundwater Head Time Series

[52] The performances of all calibration scenarios defined in sections 4.1–4.3 were verified by evaluating their simulated groundwater head time series to observations at 4250 stations. These groundwater head performances are summarized as violin plots in Figure 10. A violin plot is a modified box plot with its sides showing the shape of its density trace (or "smoothed histogram") and indicates the distributional characteristics of data (for more detailed explanation, see *Hintze and Nelson* [1998]). In each violin plot, the median is shown as a circle, and 25th and 75th percentiles are indicated as the low and high ends of bold lines. The whisker (or thin line) extends to the most extreme data which are no more than 1.5 times the interquantile range of the data [see also *Tukey*, 1977]. Figures 10a and 10b present the groundwater head performance indicators of all calibration scenarios—measured in ρ_{HEAD} and MAE_{ano}—for head stations located in pixels 21 and 112—arbitrarily chosen in order to illustrate overall phenomena found. Figure 10c gives ρ_{HEAD} and MAE_{ano} for all head stations used. In Figure 10, there are six groups described as follows.

[53] The first group, REFERENCE, is the uncalibrated model run, having $f_W = 0$, $f_K = 0$, and $f_{KD} = 0$. The results from this group are already quite good. In terms of ρ_{HEAD}, there are 3034 stations (71%) having $\rho_{HEAD} > 0.5$ and there

are 1897 stations (45%) having $\rho_{HEAD} > 0.7$. In terms of MAE_{ano}, there are 2044 stations (48%) having $MAE_{ano} < 0.25$ m. This groundwater head performance is summarized in Table 4 that also lists the performance of the other groups, including their discharge performance.

[54] The last group in Figure 10, HEAD-BEST, ("by proxy") represents the "best results"—based on groundwater head observations. For this group, in each groundwater head station (total: 4250 stations), we identified the best run (from 3045 runs that we simulated), i.e., the highest ρ_{HEAD} and the lowest MAE_{ano}. HEAD-BEST is a collection of "best runs" and none of the runs performed in this study would have the equal performance as HEAD-BEST, as indicated by its superiority in Table 4.

[55] The second group, DISCHARGE, represents the calibrated model run based on discharge data alone, as explained in section 4.1. Figure 10 shows that, in terms of model fit to heads, the performance of DISCHARGE is almost similar to the performance of REFERENCE. Yet, DISCHARGE gives a better discharge simulation (see Table 4).

[56] The SM-BASIN and SM-PIXEL groups represent the calibration scenarios using (only) SWI data, based on the objective function ρ_{SM}. Given their highest ρ_{SM}, SM-BASIN and SM-PIXEL represent the "best runs" in simulating soil moisture dynamics. The SM-BASIN is the run with the highest study area-average value of ρ_{SM}, as discussed in section 4.2. Figure 10 shows that the head performances of REFERENCE, DISCHARGE, and SM-BASIN are similar. However, only DISCHARGE provides good discharge simulation (see Table 4). The SM-PIXEL

Table 4. Groundwater Head and Discharge Performance for Each Calibration Scenario

	Head Performance			Discharge Performance			
	Number of Head Stations			Rhine (Lobith)		Meuse (Borgharen)	
Calibration Scenarios	$\rho_{HEAD} > 0.5$	$\rho_{HEAD} > 0.7$	$MAE_{ano} < 0.25$ m	NS_{eff}	NS_{ano}	NS_{eff}	NS_{ano}
REFERENCE	3034 (71%)	1897 (45%)	2044 (48%)	0.43	0.43	0.50	0.67
DISCHARGE	2733 (65%)	1770 (42%)	1744 (41%)	0.62	0.63	0.66	0.77
SM-BASIN	2743 (65%)	1611 (38%)	2170 (51%)	0.01	0.04	0.13	0.25
SM-PIXEL[a]	1765 (42%)	870 (20%)	452 (11%)	NA	NA	NA	NA
SM-DISCHARGE	2709 (64%)	1748 (41%)	1772 (42%)	0.62	0.62	0.66	0.77
HEAD-BEST[a]	4060 (96%)	3706 (87%)	3400 (91%)	NA	NA	NA	NA

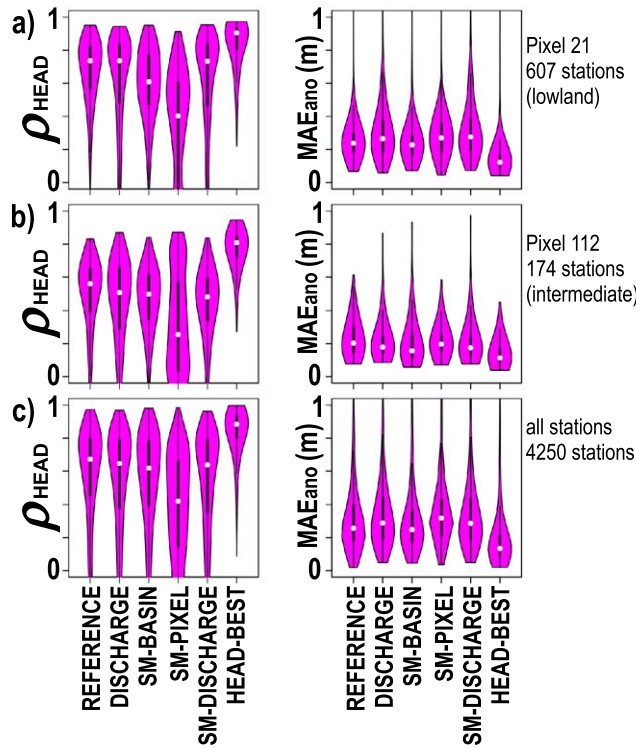[a]SM-PIXEL and HEAD-BEST are collections of runs (not a single run).

Figure 10. Violin plots summarizing groundwater head performance in the measurement stations located in pixels (a) 21, (b) 112, and (c) in the entire study area.

group represents the "best runs" that give the highest ρ_{SM}, locally, at any half arc-degree pixel (e.g., as identified from the highest ρ_{SM} in Figure 6 for the case of pixel 21). Here, we find that the reproduction of groundwater head observations worsens, as indicated in Figure 10 and Table 4.

[57] The SM-DISCHARGE represents the results from the stepwise calibration procedure, using the combination of discharge data and remotely sensed soil moisture data of SWI, as discussed in section 4.3. From Figure 10 and Table 4, we find that the performance of this group is similar to REFERENCE, DISCHARGE, and SM-BASIN. However, SM-DISCHARGE is the only calibration scenario that can reproduce both soil moisture and discharge reasonably well, while at the same time producing acceptable groundwater heads.

[58] Figure 11 compares the simulated head dynamics of SM-DISCHARGE to observation data, as well as those of some other calibration scenarios, specifically DISCHARGE and SM-PIXEL. The time series of REFERENCE and SM-BASIN are not given for the sake of clarity of plots and due to the fact that it is clear that simulated discharge time series for these cases perform poorly (see Table 4). Figures 11a and 11b show time series from points a and b in Figure 1f. Figure 11c plots the average time series from all measurement stations in pixel 21. These figures suggest that the simulated time series of DISCHARGE and SM-DISCHARGE provide equally good groundwater head performance. Both of them also provide similar discharge performance (see Table 4). Yet, as discussed previously in section 4.3, SM-DISCHARGE gives a better soil moisture simulation as the upper soil saturated hydraulic conductiv-

ities (f_K) were recalibrated to improve the soil moisture performance (section 4.3).

5. Discussion

[59] The results of this study show that SWI time series can be used to calibrate the upper soil saturated hydraulic conductivities (f_K), determining the recharge to the deeper groundwater zone. However, calibration using only SWI data is difficult for two major reasons. First, remotely sensed soil moisture data of SWI for mountainous regions often contain artifacts. This makes SWI more suitable to calibrate soil moisture dynamics in low-lying areas. Second, there is the problem of equifinality as there is more than one optimal parameter set capable of reproducing the soil moisture dynamics of SWI. To overcome these problems, discharge observations are needed, particularly to determine the optimal upper soil saturated hydraulic conductivities (f_K) in nonlow-lying areas and to constrain aquifer transmissivities (f_{KD}) and runoff-infiltration partitioning (f_W) parameters, which are hardly identifiable from the evaluation of soil moisture dynamics alone. A combined and stepwise calibration approach using both discharge observations and remotely sensed soil moisture data yields a model that can fit both discharge and soil moisture, as well as predicting the dynamics of groundwater head with acceptable accuracy.

[60] Although there is still room for improvement, this study shows the prospect of incorporating soil moisture data in large-extent and high-resolution coupled groundwater-land surface models. As all data used in building and parameterizing the model are globally available [*Sutanudjaja et al.*, 2011], the results of this study suggest that applying the methods in other basins or even globally is in principle possible. In addition discharge data, this study uses and shows the added values of the spaceborne soil moisture data of Soil Water Index (SWI) for model calibration. It suggests the potential of using other satellite soil moisture products for groundwater modeling studies, such as those from the Advanced Microwave Scanning Radiometer-Earth Observation System (AMSR-E) [*Njoku et al.*, 2003; *de Jeu and Owe*, 2003] and Soil Moisture and Ocean Salinity (SMOS) [*Kerr et al.*, 2001] missions. Soil moisture data sets from these missions should be explored in future studies. Future studies may also aim for a combined evaluation or calibration approach using discharge, soil moisture, and GRACE data. The latter has been used in many large-scale or even global-extent modeling evaluation [e.g., *Döll et al.*, 2012; *Kim et al.*, 2009; *Pokhrel et al.*, 2011]. As GRACE is typically used to evaluate total water storage variations (i.e., soil moisture and groundwater) at lower resolutions (400 km), remotely sensed soil moisture data, such as SWI, can be used to constrain soil moisture dynamics at higher resolutions (50 km), as shown here, specifically in lowland or shallow groundwater areas, i.e., the areas where shallow groundwater matters for the surface energy balance, agriculture, and ecosystem biodiversity [*Fan et al.*, 2013]. This complementary approach may be further encouraged for other available remote sensing products, such as the recently issued 1° scaled (100 km) GRACE data [*Landerer and Swenson*, 2012] and an improved SWI from *de Lange et al.* [2008], derived by
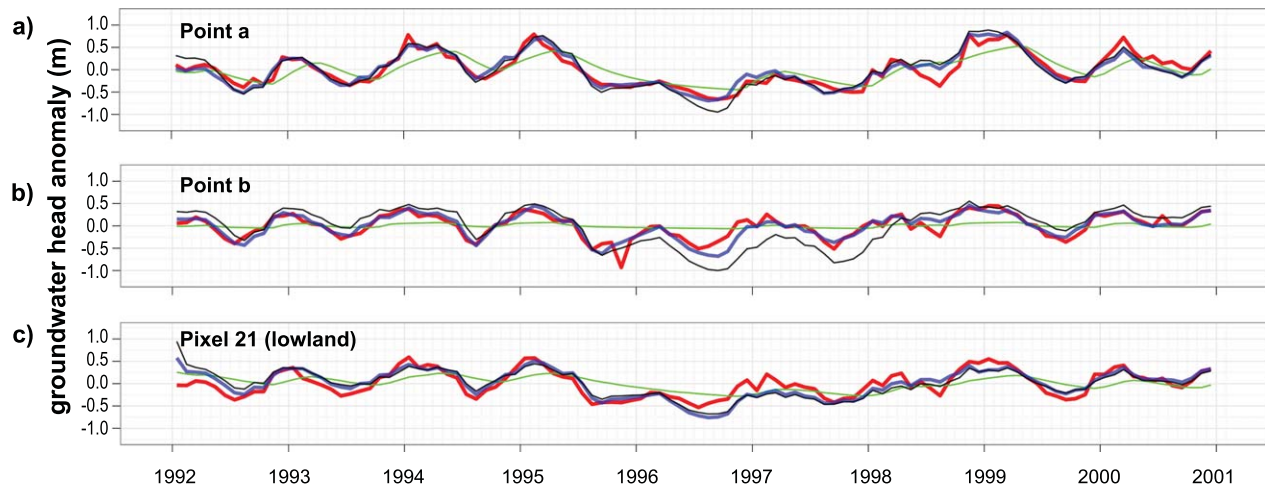
Figure 11. Comparison of observed heads (red) to model simulation results from the calibration scenarios DISCHARGE (black), SM-PIXEL (green), and SM-DISCHARGE (blue) for two locations (a) and (b), as indicated in Figure 1f, and (c) using the average value from all simulated and measured time series belonging to all stations in pixel 21 (see Figure 1e, for the pixel location). All time series are presented in anomalies to the average values.

incorporating the influence of soil texture. By combining several data sets, various objective functions can be used to evaluate different parts or components of the model. Alternatively, an error-based weighting approach may be used to integrate different types of measurement data into a single objective function [*Hill and Tiedeman*, 2007; *Foglia et al.*, 2009, 2013], specifically considering low flow conditions in a more balanced way. Low flow metrics are not investigated in the present study, but should be the concern for future research, specifically in areas with high groundwater demand problems and large base flow contribution to streamflow.

[61] In this paper, we did not consider human interventions in the water cycle whereas in many parts of the world groundwater and surface water is abstracted [*Rodell et al.*, 2009; *Wada et al.*, 2011; *Gleeson et al.*, 2012] in great amounts such that it will affect both discharge and groundwater heads. Also, we simplified the present study by excluding the downstream areas in Netherlands where water management practices are dominant. In future studies, such human interventions should be part of the model if parameters are to be estimated through calibration more accurately.

[62] Nevertheless, the successful inclusion of the well-known MODFLOW [*McDonald and Harbaugh*, 1988] code groundwater model in PCR-GLOBWB-MOD shown in this study should suggest an opportunity to improve any existing global or large-extent hydrological models, such as the original version of PCR-GLOBWB [*van Beek et al.*, 2011], WASMOD-M [*Widén-Nilsson et al.*, 2007], and variable infiltration capacity (VIC) [*Liang et al.*, 1994], which still do not have the ability to calculate spatio-temporal groundwater heads, and hence do not simulate any lateral flows in their groundwater compartments. Incorporating such lateral flows are important for future global hydrological models that may have 1 km or even finer resolutions [*Wood et al.*, 2011]. The inclusion of MODFLOW in the model is particularly relevant for areas with large

sedimentary basins or pockets, such as the basins of Danube, Mekong, Yellow, and Ganges-Brahmaputra Rivers. The ability to simulate groundwater head dynamics in such large-extent or global-extent high-resolution models will greatly enhance the applicability of such models as groundwater head is an important variable in many land-surface processes and real-world applications, including land subsidence, irrigation, nutrient transport, salt water intrusion, drought, and forest fire hazard.

6. Summary and Conclusions

[63] In areas without ground-based measurements, applications of satellite-based remote sensing for hydrological studies have received increased attention. Yet groundwater studies, which are very important for agricultural areas and ecosystems, rarely utilize the benefits of spaceborne remote sensing. In this study, we explored the possibility of calibrating a high-resolution and large-extent integrated groundwater-land surface model using remotely sensed soil moisture data in the form of Soil Water Index (SWI) and in situ discharge observations. We introduced and used a multiobjective and stepwise calibration approach that exploits the strength of each observation type in order to reproduce the observed dynamics of different parts of the model, i.e., discharge and soil moisture, as well as groundwater head. Results suggest that SWI time series are suitable for calibrating soil moisture dynamics in low-lying areas. In such low-lying areas, SWI data can be used to estimate the upper soil hydraulic conductivities, that determine the recharge to the deeper groundwater zone. Nevertheless, discharge observations are needed to obtain full calibration of the coupled model, particularly to determine the optimal upper soil hydraulic conductivities in mountainous regions and to estimate aquifer transmissivities, as well as to constrain runoff-infiltration partitioning processes. A combined and stepwise calibration approach using both discharge observations and remotely sensed soil moisture data yields a

model that can fit both discharge and soil moisture, as well as predicting the dynamics of groundwater head with acceptable accuracy. Although there is still room for improvement, this study show that with the combination of globally available data sets and remote sensing products, large-extent or even global-extent coupled groundwater-land surface modeling is within the reach.

References

Alkhaier, F., G. N. Flerchinger, and Z. Su (2012a), Shallow groundwater effect on land surface temperature and surface energy balance under bare soil conditions: Modeling and description, *Hydrol. Earth Syst. Sci.*, 16(7), 1817–1831, doi:10.5194/hess-16-1817-2012.

Alkhaier, F., Z. Su, and G. N. Flerchinger (2012b), Reconnoitering the effect of shallow groundwater on land surface temperature and surface energy balance using MODIS and SEBS, *Hydrol. Earth Syst. Sci.*, 16(7), 1833–1844, doi:10.5194/hess-16-1833-2012.

Allen, R. G., L. S. Pereira, D. Raes, and M. Smith (1998), *Crop Evapotranspiration: Guidelines for Computing Crop Requirements*, U. N.-Food and Agric. Organ., Rome. [Available at http://www.fao.org/docrep/X0490E/X0490E00.htm.]

Bastiaanssen, W. G. M., M. Menenti, R. A. Feddes, and A. A. M. Holtslag (1998a), A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation, *J. Hydrol.*, 212–213, 198–212, doi:10.1016/S0022-1694(98)00253-4.

Bastiaanssen, W. G. M., H. Pelgrum, J. Wang, Y. Ma, J. F. Moreno, G. J. Roerink, and T. van der Wal (1998b), A remote sensing surface energy balance algorithm for land (SEBAL): 2. Validation, *J. Hydrol.*, 212–213, 213–229, doi:10.1016/S0022-1694(98)00254-6.

Becker, M. W. (2006), Potential for satellite remote sensing of ground water, *Ground Water*, 44(2), 306–318, doi:10.1111/j.1745-6584.2005.00123.x.

Bergström, S. (1995), The HBV model, in *Computer Models of Watershed Hydrology*, edited by V. Singh, 443–476, Water Resour. Publ., Highlands Ranch, Colo.

Bierkens, M. F. P., and B. J. J. M. van den Hurk (2007), Groundwater convergence as a possible mechanism for multi-year persistence in rainfall, *Geophys. Res. Lett.*, 34, L02402, doi:10.1029/2006GL028396.

Brocca, L., et al. (2011), Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across

Europe, *Remote Sens. Environ.*, 115(12), 3390–3408, doi:10.1016/j.rse.2011.08.003.

Campbell, G. (1974), A simple method for determining unsaturated conductivity from moisture retention data, *Soil Sci.*, 117(6), 311–314.

Clapp, R. B., and G. M. Hornberger (1978), Empirical equations for some soil hydraulic properties, *Water Resour. Res.*, 14(4), 601–604, doi:10.1029/WR014i004p00601.

Crow, W. T., E. F. Wood, and M. Pan (2003), Multiobjective calibration of land surface model evapotranspiration predictions using streamflow observations and spaceborne surface radiometric temperature retrievals, *J. Geophys. Res.*, 108(D23), 4725, doi:10.1029/2002JD003292.

Dankers, R., and S. M. de Jong (2004), Monitoring snow-cover dynamics in Northern Fennoscandia with SPOT VEGETATION images, *Int. J. Remote Sens.*, 25(15), 2933–2949, doi:10.1080/01431160310001618374.

de Jeu, R. A. M., and M. Owe (2003), Further validation of a new methodology for surface moisture and vegetation optical depth retrieval, *Int. J. Remote Sens.*, 24(22), 4559–4578, doi:10.1080/0143116031000095934.

de Lange, R., R. Beck, N. van de Giesen, J. Friesen, A. de Wit, and W. Wagner (2008), Scatterometer-derived soil moisture calibrated for soil texture with a one-dimensional water-flow model, *IEEE Trans. Geosci. Remote Sens.*, 46(12), 4041–4049, doi:10.1109/TGRS.2008.2000796.

de Wit, M. (2001), *Effect of Climate Change on the Hydrology of the River Meuse*, Natl. Inst. for Public Health and the Environ., Bilthoven, Netherlands.

Dee, D. P., and A. M. Da Silva (1998), Data assimilation in the presence of forecast bias, *Q. J. R. Meteorol. Soc.*, 124(545), 269–295, doi:10.1002/qj.49712454512.

Döll, P., F. Kaspar, and B. Lehner (2003), A global hydrological model for deriving water availability indicators: Model tuning and validation, *J. Hydrol.*, 270(1–2), 105–134, doi:10.1016/S0022-1694(02)00283-4.

Döll, P., H. Hoffmann-Dobrev, F. Portmann, S. Siebert, A. Eicker, M. Rodell, G. Strassberg, and B. Scanlon (2012), Impact of water withdrawals from groundwater and surface water on continental water storage variations, *J. Geodyn.*, 59–60, 143–156, doi:10.1016/j.jog.2011.05.001.

Drusch, M., E. F. Wood, and H. Gao (2005), Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture, *Geophys. Res. Lett.*, 32, L15403, doi:10.1029/2005GL023623.

Dürr, H. H., M. Meybeck, and S. H. Dürr (2005), Lithologic composition of the Earth's continental surfaces derived from a new digital map emphasizing riverine material transfer, *Global Biogeochem. Cycles*, 19, GB4S10, doi:10.1029/2005GB002515.

Eagleson, P. S. (1978), Climate, soil, and vegetation: 3. A simplified model of soil moisture movement in the liquid phase, *Water Resour. Res.*, 14(5), 722–730, doi:10.1029/WR014i005p00722.

Fan, Y., and G. Miguez-Macho (2010), Potential groundwater contribution to Amazon evapotranspiration, *Hydrol. Earth Syst. Sci.*, 14(10), 2039–2056, doi:10.5194/hess-14-2039-2010.

Fan, Y., and G. Miguez-Macho (2011), A simple hydrologic framework for simulating wetlands in climate and earth system models, *Clim. Dyn.*, 37, 253–278, doi:10.1007/s00382-010-0829-8.

Fan, Y., G. Miguez-Macho, C. P. Weaver, R. Walko, and A. Robock (2007), Incorporating water table dynamics in climate modeling: 1. Water table observations and equilibrium water table simulations, *J. Geophys. Res.*, 112, D10125, doi:10.1029/2006JD008111.

Fan, Y., H. Li, and G. Miguez-Macho (2013), Global patterns of groundwater table depth, *Science*, 339(6122), 940–943, doi:10.1126/science.1229881.

Fenicia, F., H. H. G. Savenije, P. Matgen, and L. Pfister (2007), A comparison of alternative multiobjective calibration strategies for hydrological modeling, *Water Resour. Res.*, 43, W03434, doi:10.1029/2006WR005098.

Foglia, L., M. C. Hill, S. W. Mehl, and P. Burlando (2009), Sensitivity analysis, calibration, and testing of a distributed hydrological model using error-based weighting and one objective function, *Water Resour. Res.*, 45, W06427, doi:10.1029/2008WR007255.

Foglia, L., S. W. Mehl, M. C. Hill, and P. Burlando (2013), Evaluating model structure adequacy: The case of the Maggia valley groundwater system, southern Switzerland, *Water Resour. Res.*, 49, 260–282, doi:10.1029/2011WR011779.

Gardner, W. R. (1958), Some steady-state solutions of the unsaturated moisture flow equation with application to evaporation from a water table, *Soil Sci.*, 85(4), 228–232.

Gleeson, T., L. Smith, N. Moosdorf, J. Hartmann, H. H. Dürr, A. H. Manning, L. P. H. van Beek, and A. M. Jellinek (2011), Mapping

permeability over the surface of the earth, *Geophys. Res. Lett.*, *38*, L02401, doi:10.1029/2010GL045565.

Gleeson, T., Y. Wada, M. F. P. Bierkens, and L. P. H. van Beek (2012), Water balance of global aquifers revealed by groundwater footprint, *Nature*, *488*(7410), 197–200, doi:10.1038/nature11295.

Gupta, H. V., S. Sorooshian, and P. O. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, *34*(4), 751–763, doi:10.1029/97WR03495.

Hagemann, S., and L. D. Gates (2003), Improving a subgrid runoff parameterization scheme for climate models by the use of high resolution data derived from satellite observations, *Clim. Dyn.*, *21*(3), 349–359, doi:10.1007/s00382-003-0349-x.

Hill, M. C., and C. R. Tiedeman (2007), *Effective groundwater model calibration: With analysis of data, sensitivities, predictions, and uncertainty*, John Wiley and Sons, N. Y.

Hintze, J. L., and R. D. Nelson (1998), Violin plots: A box plot-density trace synergism, *Am. Stat.*, *52*, 181–184.

Immerzeel, W. W., P. Droogers, S. M. de Jong, and M. F. P. Bierkens (2009), Large-scale monitoring of snow cover and runoff simulation in Himalayan river basins using remote sensing, *Remote Sens. Environ.*, *113*(1), 40–49, doi:10.1016/j.rse.2008.08.010.

Jackson, T. (2002), Remote sensing of soil moisture: Implications for groundwater recharge, *Hydrogeol. J.*, *10*, 40–51, doi:10.1007/s10040-001-0168-2.

Jung, M., et al. (2010), Recent decline in the global land evapotranspiration trend due to limited moisture supply, *Nature*, *467*(7318), 951–954, doi:10.1038/nature09396.

Kerr, Y., P. Waldteufel, J.-P. Wigneron, J. Martinuzzi, J. Font, and M. Berger (2001), Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission, *IEEE Trans. Geosci. Remote Sens.*, *39*(8), 1729–1735.

Kim, H., P. J.-F. Yeh, T. Oki, and S. Kanae (2009), Role of rivers in the seasonal variations of terrestrial water storage over global basins, *Geophys. Res. Lett.*, *36*, L17402, doi:10.1029/2009GL039006.

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar (2000), A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure, *J. Geophys. Res.*, *105*(D20), 24,809–24,822, doi:10.1029/2000JD900327.

Koster, R. D., et al. (2004), Regions of strong coupling between soil moisture and precipitation, *Science*, *305*(5687), 1138–1140, doi:10.1126/science.1100217.

Kraaijenhoff van de Leur, D. (1958), A study of non-steady groundwater flow with special reference to a reservoir coefficient, *De Ingénieur*, *70*(19), 87–94.

Kummerow, C., et al. (2000), The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit, *J. Appl. Meteorol.*, *39*(12), 1965–1982, doi:10.1175/1520-0450(2001)040<1965:TSOTTR>2.0.CO;2.

Lacey, G. (1930), Stable channels in alluvium, *Proc. Inst. Civ. Eng.*, *229*, 259–292.

Landerer, F. W., and S. C. Swenson (2012), Accuracy of scaled grace terrestrial water storage estimates, *Water Resour. Res.*, *48*, W04531, doi:10.1029/2011WR011453.

Lehner, B., K. Verdin, and A. Jarvis (2008), New global hydrography derived from spaceborne elevation data, *Eos Trans. AGU*, *89*, 93–94, doi:10.1029/2008EO100001.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *J. Geophys. Res.*, *99*(D7), 14,415–14,428, doi:10.1029/94JD00483.

Liu, Y., W. Dorigo, R. Parinussa, R. de Jeu, W. Wagner, M. McCabe, J. Evans, and A. van Dijk (2012), Trend-preserving blending of passive and active microwave soil moisture retrievals, *Remote Sens. Environ.*, *123*, 280–297, doi:10.1016/j.rse.2012.03.014.

Liu, Y. Y., R. M. Parinussa, W. A. Dorigo, R. A. M. De Jeu, W. Wagner, A. I. J. M. van Dijk, M. F. McCabe, and J. P. Evans (2011), Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, *Hydrol. Earth Syst. Sci.*, *15*(2), 425–436, doi:10.5194/hess-15-425-2011.

Lo, M.-H., J. S. Famiglietti, P. J.-F. Yeh, and T. H. Syed (2010), Improving parameter estimation and water table depth simulation in a land surface model using grace water storage and estimated base flow data, *Water Resour. Res.*, *46*, W05517, doi:10.1029/2009WR007855.

McDonald, M., and A. Harbaugh (1988), *A Modular Three-Dimensional Finite-Difference Ground-Water Flow Model: Techniques of Water-Resources Investigations*, Book 6, U.S. Geol. Surv, Denver, Colo. [Available at http://pubs.water.usgs.gov/twri6a1.]

Mitchell, T. D., and P. D. Jones (2005), An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *Int. J. Climatol.*, *25*(6), 693–712, doi:10.1002/joc.1181.

Mu, Q., F. A. Heinsch, M. Zhao, and S. W. Running (2007), Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, *Remote Sens. Environ.*, *111*(4), 519–536, doi:10.1016/j.rse.2007.04.015.

Nash, J., and J. Sutcliffe (1970), River flow forecasting through conceptual models. Part I: A discussion of principles, *J. Hydrol.*, *10*(3), 282–290, doi:10.1016/0022–1694(70)90255-6.

New, M., M. Hulme, and P. Jones (1999), Representing twentieth-century spacetime climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology, *J. Clim.*, *12*(3), 829–856, doi: 10.1175/1520-0442(1999)012<0829:RTCSTC>2.0.CO;2.

Njoku, E., T. Jackson, V. Lakshmi, T. Chan, and S. Nghiem (2003), Soil moisture retrieval from AMSR-E, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 215–229, doi:10.1109/TGRS.2002.808243.

Pokhrel, Y., N. Hanasaki, S. Koirala, J. Cho, P. J. F. Yeh, H. Kim, S. Kanae, and T. Oki (2011), Incorporating anthropogenic water regulation modules into a land surface model, *J. Hydrometeorol.*, *13*(1), 255–269, doi:10.1175/JHM-D-11-013.1.

Reichle, R. H., and R. D. Koster (2004), Bias reduction in short records of satellite soil moisture, *Geophys. Res. Lett.*, *31*, L19501, doi:10.1029/2004GL020938.

Rodell, M., I. Velicogna, and J. S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in India, *Nature*, *460*(7258), 999–1002, doi:10.1038/nature08238.

Savenije, H. H. G. (2003), The width of a bankfull channel: Lacey's formula explained, *J. Hydrol.*, *276*(1–4), 176–183, doi:10.1016/S0022-1694(03)00069-6.

Schmitz, O., D. Karssenberg, W. van Deursen, and C. Wesseling (2009), Linking external components to a spatio-temporal modelling framework: Coupling MODFLOW and PCRaster, *Environ. Modell. Software*, *24*(9), 1088–1099, doi:10.1016/j.envsoft.2009.02.018.

Shukla, J., and Y. Mintz (1982), Influence of land-surface evapotranspiration on the earth's climate, *Science*, *215*(4539), 1498–1501, doi:10.1126/science.215.4539.1498.

Soil Conservation Service (SCS) (1972), *National Engineering Handbook, Section 4: Hydrology*, U.S. Dept. of Agric., Washington D. C.

Sólyom, P. B., and G. E. Tucker (2004), Effect of limited storm duration on landscape evolution, drainage basin geometry, and hydrograph shapes, *J. Geophys. Res.*, *109*, F03012, doi:10.1029/2003JF000032.

Soylu, M. E., E. Istanbulluoglu, J. D. Lenters, and T. Wang (2011), Quantifying the impact of groundwater depth on evapotranspiration in a semi-arid grassland region, *Hydrol. Earth Syst. Sci.*, *15*(3), 787–806, doi:10.5194/hess-15-787-2011.

Su, Z. (2002), The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes, *Hydrol. Earth Syst. Sci.*, *6*(1), 85–100, doi:10.5194/hess-6-85-2002.

Sutanudjaja, E. H. (2012), The use of soil moisture—Remote sensing products for large-scale groundwater modeling and assessment, PhD thesis, Utrecht Univ., Netherlands.

Sutanudjaja, E. H., L. P. H. van Beek, S. M. de Jong, F. C. van Geer, and M. F. P. Bierkens (2011), 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.

Tapley, B. D., S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins (2004), GRACE Measurements of mass variability in the earth system, *Science*, *305*(5683), 503–505, doi:10.1126/science.1099192.

Todini, E. (1996), The ARNO rainfall-runoff model, *J. Hydrol.*, *175*(1–4), 339–382, doi:10.1016/S0022-1694(96)80016-3.

Tukey, J. W. (1977), *Exploratory Data Analysis*, Addison-Wesley, Reading, Mass.

Uppala, S. M., et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012. [Available at http://edoc.mpg.de/256471.]

van Beek, L. P. H. (2008), Forcing PCR-GLOBWB with CRU data, technical report, Dep. of Phys. Geogr., Utrecht Univ., Utrecht, Netherlands. [Available at http://vanbeek.geo.uu.nl/suppinfo/vanbeek2008.pdf.]

van Beek, L. P. H., and M. F. P. Bierkens (2009), The global hydrological model PCR-GLOBWB: conceptualization, parameterization and verification, technical report, Dep. of Phys. Geogr., Utrecht Univ., Utrecht, Netherlands. [Available at http://vanbeek.geo.uu.nl/suppinfo/vanbeek-bierkens2009.pdf.]

van Beek, L. P. H., Y. Wada, and M. F. P. Bierkens (2011), Global monthly water stress: 1. Water balance and water availability, *Water Resour. Res.*, *47*, W07517, doi:10.1029/2010WR009791.

van Loon, A. F., M. H. J. Van Huijgevoort, and H. A. J. Van Lanen (2012), Evaluation of drought propagation in an ensemble mean of large-scale hydrological models, *Hydrol. Earth Syst. Sci.*, *16*(11), 4057–4078, doi: 10.5194/hess-16-4057-2012.

Wada, Y., L. P. H. van Beek, C. M. van Kempen, J. W. T. M. Reckman, S. Vasak, and M. F. P. Bierkens (2010), Global depletion of groundwater resources, *Geophys. Res. Lett.*, *37*, L20402, doi:10.1029/2010GL044 571.

Wada, Y., L. P. H. van Beek, D. Viviroli, H. H. Dürr, R. Weingartner, and M. F. P. Bierkens (2011), Global monthly water stress: 2. Water demand and severity of water stress, *Water Resour. Res.*, *47*, W07518, doi: 10.1029/2010WR009792.

Wagner, W. (1998), Soil moisture retriecval from ERS scatterometer data, PhD thesis, Vienna Univ. of Technol., Vienna.

Wagner, W., G. Lemoine, M. Borgeaud, and H. Rott (1999a), A study of vegetation cover effects on ERS scatterometer data, *IEEE Trans. Geosci. Remote Sens.*, *37*(2), 938–948.

Wagner, W., G. Lemoine, and H. Rott (1999b), A method for estimating soil moisture from ERS Scatterometer and soil data, *Remote Sens. Environ.*, *70*(2), 191–207, doi:10.1016/S0034-4257(99)00036-X.

Wan, Z., and Z.-L. Li (1997), A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data, *IEEE Trans. Geosci. Remote Sens.*, *35*(4), 980–996, doi:10.1109/36.602541.

Wanders, N., D. Karssenberg, M. F. P. Bierkens, R. M. Parinussa, R. A. M. de Jeu, J. van Dam, S. M. de Jong (2012), Observation uncertainty of satellite soil moisture products determined with physically-based modeling, *Remote Sens. Environ.*, *127*, 341–356, doi:10.1016/j.rse.2012.09.004.

Werth, S., A. Gntner, S. Petrovic, and R. Schmidt (2009), Integration of GRACE mass variations into a global hydrological model, *Earth Planet. Sci. Lett.*, *277*(1–2), 166–173, doi: 10.1016/j.epsl.2008.10.021.

Wesseling, C. G., D. Karssenberg, W. P. A. van Deursen, and P. A. Burrough (1996), Integrating dynamic environmental models in GIS: The development of a Dynamic Modelling language, *Trans. GIS*, *1*(1), 40–48, doi:10.1111/j.1467-9671.1996.tb00032.x.

Widén-Nilsson, E., S. Halldin, and C. yu Xu (2007), Global water-balance modelling with WASMOD-M: Parameter estimation and regionalisation, *J. Hydrol.*, *340*(1–2), 105–118.

Wood, E. F., et al. (2011), Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, *Water Resour. Res.*, *47*, W05301, doi:10.1029/2010WR010090.

Yeh, P. J.-F., and E. A. B. Eltahir (2005a), Representation of water table dynamics in a land surface scheme. Part I: Model development, *J. Clim.*, *18*(12), 1861–1880, doi:10.1175/JCLI3330.1.

Yeh, P. J.-F., and E. A. B. Eltahir (2005b), Representation of water table dynamics in a land surface scheme. Part II: Subgrid variability, *J. Clim.*, *18*(12), 1881–1901, doi:10.1175/JCLI3331.1.

Zaitchik, B. F., M. Rodell, and F. Olivera (2010), Evaluation of the global land data assimilation system using global river discharge data and a source-to-sink routing scheme, *Water Resour. Res.*, *46*, W06507, doi: 10.1029/2009WR007811.