AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2016JD025209

Key Points:
- Large-scale water storage change in 25 catchments around the world
- Relative temporal variability of soil moisture in 10 U.S. catchments
- Simple large-scale soil moisture model reproduces relative temporal variability

Supporting Information:
- Supporting Information S1

Correspondence to:
L. Verrot,
lucile.verrot@natgeo.su.se

Citation:
Verrot, L., and G. Destouni (2016), Data-model comparison of temporal variability in long-term time series of large-scale soil moisture, *J. Geophys. Res. Atmos.*, *121*, 10,056–10,073, doi:10.1002/2016JD025209.

Received 11 APR 2016
Accepted 26 JUL 2016
Accepted article online 29 JUL 2016
Published online 15 SEP 2016

Data-model comparison of temporal variability in long-term time series of large-scale soil moisture

Lucile Verrot[1] and Georgia Destouni[1]

[1]Department of Physical Geography, Stockholm University, Stockholm, Sweden

Abstract Soil moisture is at the heart of many processes connected to water cycle, climate, ecosystem, and societal conditions. This paper investigates the ability of a relatively simple analytical soil moisture model to reproduce temporal variability dynamics in long-term data series for (i) remotely sensed large-scale water storage change in 25 large catchments around the world and (ii) measured soil water content and groundwater level in individual stations within 10 smaller catchments across the United States. The model-data comparison for large-scale water storage change (i) shows good model ability to reproduce the observed temporal variability around long-term average conditions in most of the large study catchments. Also, the model comparison with locally measured data for soil water content and groundwater level in the smaller U.S. catchments (ii) shows good representation of relative seasonal and longer-term fluctuations and their timings and frequencies. Overall, the model results tend to underestimate rather than exaggerate the range of temporal soil moisture fluctuations and storage changes. The model synthesis of large-scale hydroclimatic data is based on fundamental catchment-scale water balance and is as such useful for identifying flux imbalance biases in the hydroclimatic data series that are used as model inputs.

1. Introduction

Soil moisture is a dynamic variable of great importance in the hydrologic cycle [*Corradini*, 2014] as well as for climate, environmental, and societal conditions [*Seneviratne et al.*, 2010]. It refers to the amount of water stored in a given volume of soil, which varies temporally depending on fluctuations of both hydroclimate at the surface [*Rodriguez-Iturbe et al.*, 1991] and groundwater table level in the subsurface [*Destouni and Verrot*, 2014]. It also varies spatially depending on several factors, including soil type, vegetation, topography, and spatial variation of hydroclimatic conditions [*Destouni and Cvetkovic*, 1989; *Russo*, 1998].

Models describing soil moisture variation focus on selected aspects of its full complexity, depending on the field of application. Temporal near-surface fluctuations on relatively large spatial scales are of main interest for energy balance questions related to climate-land surface interactions. Water resource questions focus on more local interactions on time scales that can vary greatly depending on specific study questions.

While being at the center of complex processes of interest across disciplines and over a wide range of scales, soil moisture cannot be directly measured over large areas, because of both cost and disruptive soil effects of direct measurements. Soil moisture upscaling is thus a major challenge in hydrological, geophysical, and climate change science [*Beven*, 2001; *Lahoz and De Lannoy*, 2014]. More generally, scaling issues lie at the heart of current research in atmospheric science and hydrology, questioning our ability to use powerful models and data retrieval techniques for the understanding of global and regional hydroclimatic processes [*Bengtsson et al.*, 2014].

In a recent study, *Destouni and Verrot* [2014] addressed the problem of soil moisture upscaling by proposing an analytical modeling framework for screening long-term soil moisture variability, linking large-scale hydroclimatic variables to local characteristics, processes, and interactions in both the unsaturated and the saturated soil zone. The framework incorporates large-scale average soil water content and its temporal fluctuations under changing climate and groundwater table conditions. *Verrot and Destouni* [2015] further extended this framework to also account for effects of snow storage and melt dynamics and thus be able to account for more wide-ranging climatic conditions in the screening of long-term soil moisture variability.

The modeling framework of *Destouni and Verrot* [2014] and *Verrot and Destouni* [2015] approximates area- and depth-averaged soil moisture (volume of water in a given volume of bulk soil) over the entire unsaturated zone and some depth into the saturated zone of a whole catchment. To do this, the framework models both the average soil water content over the whole (time-variable) unsaturated zone and the depth of the groundwater

table (also time variable) within a fixed total considered depth of average soil moisture quantification over a whole catchment. The present study aims at comparing relevant results of this large-scale modeling framework with observed large-scale changes in total subsurface water storage (study part I) and observed relative temporal variability around the respective long-term average values of locally measured data of soil water content and groundwater level (study part II). Of main interest in the modeling and its comparison with these different types of data (I and II) is the relative temporal variability around each respective long-term average value and the possible changes in the latter exhibited in long-term time series of the investigated variables.

The two-part model-data comparison is done for two sets of study catchments: (i) one set of 25 large catchments spread over the world (Figure 1a) regarding large-scale changes in water storage and (ii) one set of smaller catchments spread over the United States (U.S., Figure 1b) regarding locally measured soil water content and groundwater level. The selection of study catchments is based on the availability of necessary data for testing the relative temporal variability and possible changes in long-term average conditions in these model output variables and for characterizing the catchments in terms of the hydroclimactic and soil data that are needed as model inputs.

In study part (II), large-scale model results are compared with local observations in single or few individual measurement stations existing within each catchment. This cross-scale comparison is made in terms of relative temporal variability around the respective long-term average value of soil moisture at each scale. Such cross-scale comparison is interesting as a complement to the consistently large-scale comparison with satellite data (study part I) because local data from individual measurement stations are still the most commonly available type of data sets for soil water content and groundwater level. The comparison is also relevant because previous results have indicated that calibration of one or two unknown model parameters in order to obtain a relevant absolute long-term average value (which likely differs between different scales) does not much affect the modeled relative temporal variability of soil moisture around this value [*Destouni and Verrot*, 2014; *Verrot and Destouni*, 2015].

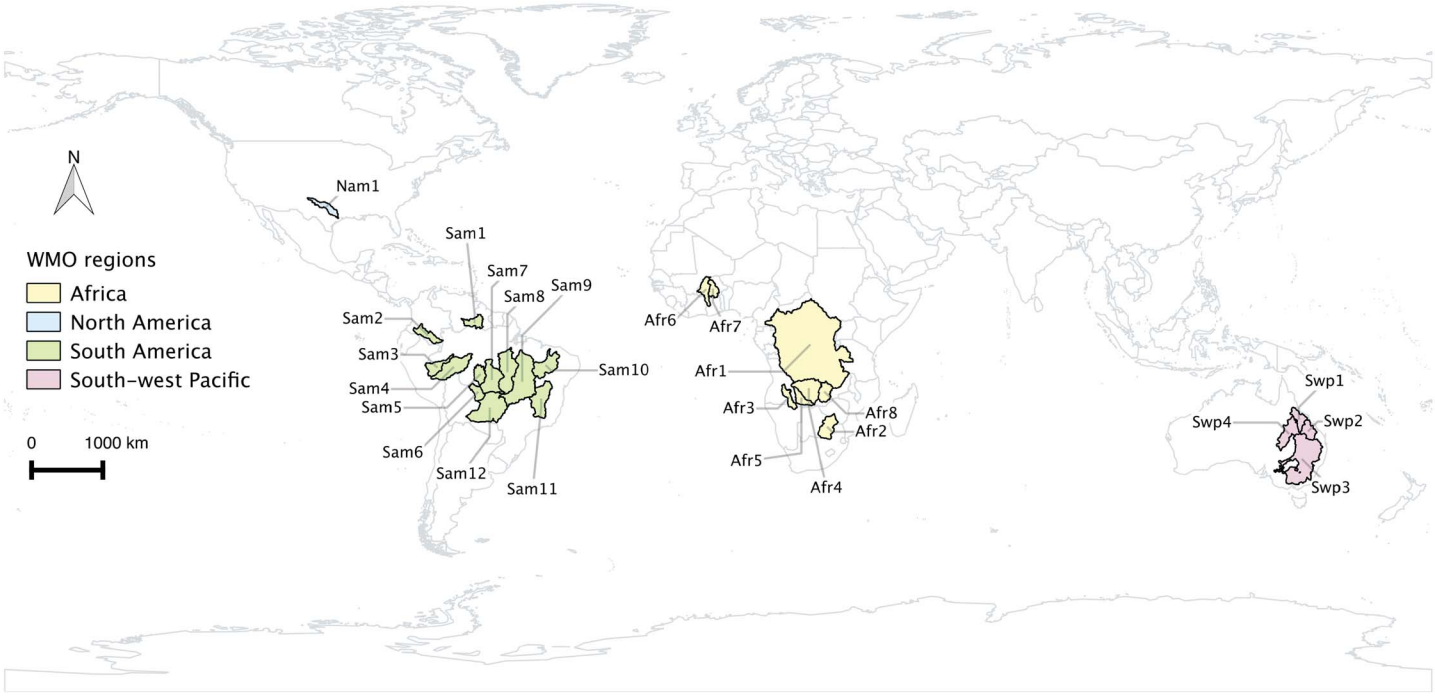
2. Materials and Methods

2.1. General Approach to the Model-Data Comparison

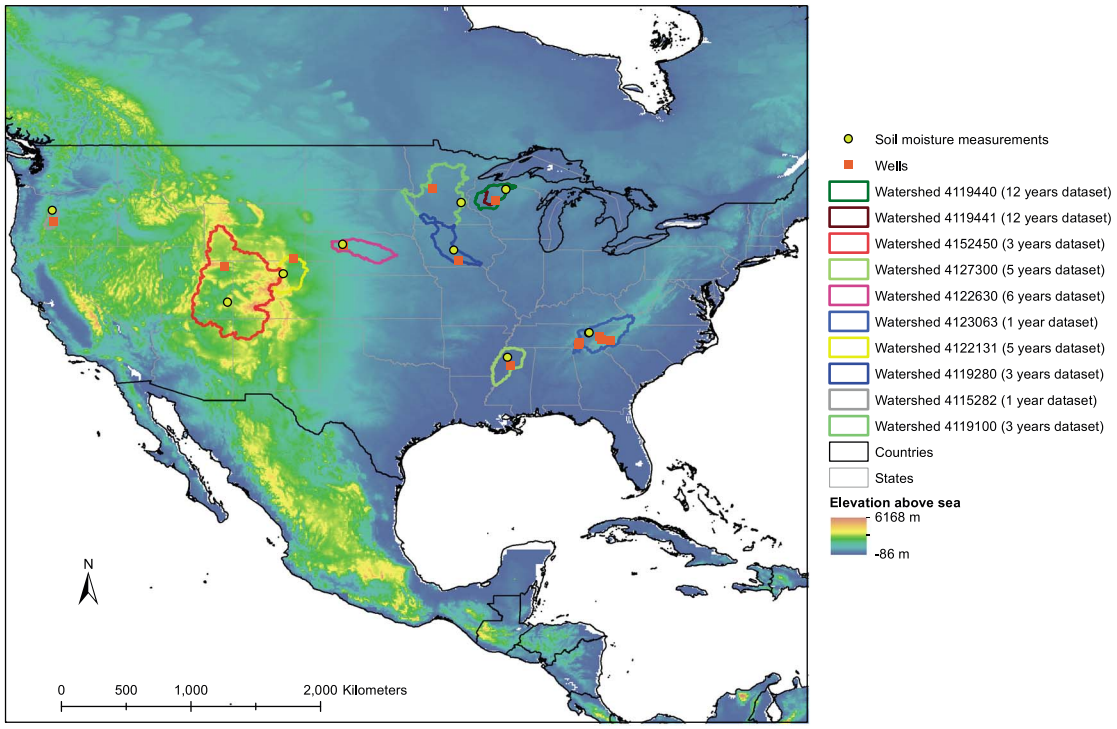
As described by *Destouni and Verrot* [2014] and *Verrot and Destouni* [2015] concerning their modeling framework, and as noted more generally by *Zhu and Lin* [2011] and *Mittelbach and Seneviratne* [2012], both local conditions in the landscape and large-scale hydroclimatic drivers can influence soil moisture and its variability. In this investigation, we select the two sets of study catchments for the two-part model-data comparison (I and II) based on the general criterion of consistent availability to as long-term time series as possible (and at least one full year of data) for all relevant model output and input data. Besides data for the model output variables to be tested, the model-data comparisons also require soil texture and hydroclimatic data as inputs to the modeling; the modeling approach combines these input variables to enable screening of long-term hydroclimatic time series for determination of associated soil moisture variability and change. The choice of catchments for the model-data comparison is thus determined by the availability of data sets with consistent spatiotemporal coverage for all relevant output and input variables, which is overall limited [*Fekete et al*., 2002; *Bukovsky and Karoly*, 2007].

The 25 large catchments in the first study part (II) are spread over different parts of the world (Figure 1a) where possible effects of snow storage-melting dynamics may be expected to be relatively small. Even though the snow dynamics effects can be handled in the model [*Verrot and Destouni*, 2015], as explained further below and done in the model-data comparison for the U.S. set of catchments, site-specific availability to the required input data for including this dynamics is particularly limited. The first, large-scale model-data testing step (I) is therefore focused on the basic soil moisture model structure [*Destouni and Verrot*, 2014], without addition of the data-limited/data-limiting snow dynamics component. The model's performance in the large-scale model-data comparison is here investigated with regard to remote sensing data for large-scale water storage changes.

For the second study part (II) with a set of 10 smaller catchments, the model's performance is compared with local measurements of soil water content and groundwater level and an approximative approach to snow dynamics account is used and analyzed under the site-specific data limitations. The 10 catchments in this



(a)



(b)

Figure 1. Map of the location of study catchments. (a) The 25 large catchments around the world used for model-data comparison of large-scale changes in subsurface water storage. (b) The 10 study catchments in the United States used for model-data comparison of locally measured unsaturated water content and groundwater level. In Figure 1b, the legend lists the number of years with water content and groundwater table data in each catchment and the background is the elevation map provided by *Amante and Eakins* [2009].

study part are spread over the U.S. (Figure 1b) based on the open availability to directly measured data for soil water content and groundwater level in these catchments.

For the model comparison with large-scale water storage changes (I), remote sensing data for this variable are derived from the Gravity Recovery and Climate Experiment (GRACE) satellite observations (CSR-RL05, from *Swenson* [2012], *Landerer and Swenson* [2012], and *Swenson and Wahr* [2006]) and, where applicable (catchment Afr1), also from the DAHITI database for lake level data [*Schwatke et al*., 2015]. Regarding input variables, soil texture is determined from the Harmonized World Soil Database v1.1 [*Nachtergaele et al*., 2008; *FAO*, 2012], and hydroclimatic data are obtained from the Global Precipitation Climatology Center (GPCC) [*Schneider et al*., 2011], the Global Historical Climatology Network (GHCN)-Climate Monitoring System [*Fan and van den Dool*, 2008], the satellite product of Moderate Resolution Imaging Spectroradiometer (MODIS) [*Oak Ridge National Laboratory Distributed Active Archive Center* (*ORNL DAAC*), 2011], and the *Global Runoff Data Center* (*GRDC*) [2015]. For the 25 catchments included in this large-scale comparison (Figure 1a), the spatial resolution of the gridded data sets allows for meaningful spatial averaging over the whole catchment scale. Furthermore, all the catchments have matching data sets for all variables over multiple years.

For the model comparison with locally measured soil water content and groundwater level (II), the data for these output variables are obtained for the 10 U.S. study catchments from the Ameriflux platform of the *United States Geological Survey* (*USGS*) [2015]. Regarding input variables, soil texture is determined from the First Most Predominant Surface Soil Classes Over NLDAS Domain issued by the North American Land Data Assimilation System [*Xia et al*., 2012], the *Global Runoff Data Center* (*GRDC*) [2015], the reanalysis product of the National Center for Atmospheric Prediction [*National Centers for Environmental Prediction* (*NCEP*), 2004], the Climate Prediction Center (CPC) Unified Gauge-Based Analysis of Daily Precipitation over CONUS [*Climate Prediction Center* (*CPC*) *U.S*., 2015], and the satellite product MODIS [*ORNL DAAC*, 2011]. Among the data sets available from *USGS* [2015] for these 10 U.S. catchments, data for a continuous period of more than 10 years is available for groundwater level in two nested catchments, where also the longest consistent data series is available for soil water content (78 months of data over the 12 years of data available for the groundwater level). Multisite comparison is enabled by also considering the eight additional catchments with shorter data series length down to the limit of at least one full year of data.

2.2. Modeling Approach

2.2.1. Basic Model Structure

The analytical modeling framework of *Destouni and Verrot* [2014] evaluates the average volumetric water content θ_z [−] over the whole extent of some selected soil depth z [L] from the surface, considering also the groundwater table position z_{gw} along z; the vertical z axis is then positive upward, and the land surface position along z is set to zero for simplicity and without loss of generality. The time-dependent groundwater table level z_{gw} (negative or zero, as it is below or at the land surface) further determines the variable depth extent of the unsaturated zone, $-z_{gw}$, and that of the groundwater zone, z_{gw}-z, within the considered soil depth extent $-z$.

The dynamics of large-scale average water content θ_z over the whole soil depth z depends on the unsaturated zone depth z_{gw}, the depth-averaged volumetric water content θ_{uz} [−] over z_{gw} (we refer to θ_{uz} as unsaturated water content in the following), and the saturated water content θ_s [−] in the groundwater zone below z_{gw} as $\theta_z = (\theta_s(z - z_{gw}) + \theta_{uz} z_{gw})/z$. In the framework of *Destouni and Verrot* [2014], the large-scale depth-averaged θ_{uz} magnitude is approximated based on the *Brooks and Corey* [1964] relation to unsaturated hydraulic conductivity K [LT^{-1}]:

$$K(\theta_{uz}) = K_s \left(\frac{\theta_{uz} - \theta_{ir}}{\theta_s - \theta_{ir}} \right) \tag{1}$$

where K_s [LT^{-1}] is saturated hydraulic conductivity and θ_s is saturated soil water content, which can be assumed equal to porosity [*Kumar*, 1999; *Entekhabi et al*., 2010]. Furthermore, θ_{ir} [−] is the residual soil water content, and $\beta = \alpha/(3\alpha + 2)$ [−] and α [−] are characteristic soil texture parameters, linked to the pore size distribution of different soil types [*Rawls et al*., 1982; *Saxton et al*., 1986]. These parameters are also related to corresponding ones in the alternative constitutive relations of *van Genuchten* [1980] and *Morel-Seytoux et al*. [1996].

A first-order approximation of unit hydraulic gradient, based on gravity being a main, even though not the only driver of large-scale flow through the unsaturated zone, has previously been used for quantification of large-scale water content statistics, originally by *Dagan and Bresler* [1979] and *Bresler and Dagan* [1981] and in multiple

other studies thereafter [e.g., *Destouni and Cvetkovic*, 1989, 1991; *Destouni*, 1993; *Destouni and Graham*, 1995]. With this approximation, the unsaturated hydraulic conductivity K in equation (1) can be equated with the average vertical soil water flux through the unsaturated zone, q [LT^{-1}], yielding the following approximate estimate of large-scale and depth-averaged θ_{uz} over the (variable) depth extent (z_{gw}) of the whole unsaturated zone as

$$\theta_{uz} = \left(\frac{q}{K_s}\right)^{\beta}(\theta_s - \theta_{ir}) \approx \left(\frac{R_{\text{eff}}}{K_s}\right)^{\beta}(\theta_s - \theta_{ir}) + \theta_{ir} \tag{2}$$

The second part of equation (2) is introduced for data-based quantification of the temporal variability of large-scale depth-averaged unsaturated water content θ_{uz} around its long-term average value. This equation part expresses the assumption of *Destouni and Verrot* [2014] and *Verrot and Destouni* [2015] that on the scale of a whole catchment, both the long-term average unsaturated soil flux and the temporal q variability around it can be estimated from and constrained by available observation data for runoff R through the catchment. More specifically, *Verrot and Destouni* [2015] approximate q by the effective subsurface runoff component $R_{\text{eff}} = \gamma R$ [LT^{-1}] (with $0 \le \gamma \le 1$), which feeds into the total R of the catchment through the unsaturated and the groundwater zone over a considered time period of temporal averaging; this subsurface runoff component complements the runoff component $(1-\gamma)R$ of overland and pure (not fed by subsurface water into the) surface water flow, which also adds to the total R over the same period. On an annual averaging basis, simulations have shown $\gamma = R_{\text{eff}}/R$ to be typically above 0.5 and in many cases close to 1 for a wide range of investigated temperate, through cold, to permafrost region conditions [*Bosson et al*., 2012]. Relevant γ values may also be variable in time, which can be readily accounted for, along with the observed temporal variability of R, through R_{eff} in equation (2) [*Verrot and Destouni*, 2015].

For a long climatic period of 20 or more years, the long-term average R_{eff} should relatively well approximate the long-term average q because the subsurface water storage change is close to zero when averaged over such long time periods [*Jaramillo et al*., 2013; *Destouni et al*., 2013; *Jaramillo and Destouni*, 2015]. Over shorter time scales, such as a month, a possible difference between q and R_{eff} may be associated with a nonzero water storage change: the water flux q through the soil may be transiently partitioned between feeding water into R_{eff} and increasing water storage in the soil, and conversely R_{eff} may be fed by both q and a transient decrease in soil water storage. However, the relative variability of R_{eff} around its long-term average value may still be relevant and sufficient for estimating the corresponding relative variability of large-scale depth-averaged unsaturated water content θ_{uz} around its long-term average value through equation (2).

The results of this approach to approximating the long-term average value and temporal variability of θ_{uz} will here be directly compared with and thus tested against various relevant observation data. Overall, the tested assumption combination underlying the above-described θ_{uz} estimate can be summarized as follows: (i) gravity is a first order, even though not the sole driver of large-scale water flux q through the unsaturated zone; (ii) the component R_{eff} of runoff through the soil-groundwater system of a catchment provides a realistic data-based constraint for the temporal variability of q in equation (2); and (iii) the combined conditions (i) and (ii) imply that equation (2) may be a sufficient first-order approximation of large-scale depth-averaged unsaturated water content θ_{uz} and its temporal variability.

Both numerical experimentation [*Destouni*, 1991] and field experimentation [*Graham et al*., 1998] over different soil depths and time scales of averaging $q \approx R_{\text{eff}}$ have previously indicated the approximate expression (2) of area-depth-averaged θ_{uz} as practically useful, in those cases for quantifying large-scale solute transport through the unsaturated zone. For quantification of the soil hydraulic properties K_s, θ_s, θ_{ir}, and β in equation (2), previous studies have used both statistical [*Destouni*, 1993] and deterministic [*Destouni*, 1991] approaches to averaging their values over large scales and various depths of interest, as discussed in more detail by *Destouni and Verrot* [2014]. The selection procedure for identifying representative soil data in the present study catchments is discussed in the following section and further detailed in supporting information section S1.

Furthermore, the framework of *Destouni and Verrot* [2014] also estimates the change in subsurface water storage from an initial time t_0 to time t and the associated temporal variability of groundwater table position z_{gw}. The water storage change is estimated from fundamental water balance as

$$S(t; t_0) = \int_{t_0}^{t} \{\gamma[P_{\text{eff}}(\tau) - \text{ET}(\tau)] - R_{\text{eff}}(\tau)\} d_\tau \tag{3}$$

where τ is a dummy time variable for the integration, ET [LT^{-1}] is actual evapotranspiration, and P_{eff} [LT^{-1}] is effective precipitation of liquid water. The latter relates to the partitioning of the measured total precipitation P [LT^{-1}] between liquid and frozen water and to the dynamics of snow storage and melting, which are accounted for in the model as described by *Verrot and Destouni* [2015] and also detailed here in supporting information section S2. Furthermore, the runoff fraction $\gamma = R_{eff}/R$ is used to also quantify a runoff-consistent partitioning of the total net water input $P_{eff} - ET$ between the water fraction γ that is available for percolation through the soil and the complementary water fraction $1-\gamma$ of overland and only (without first flowing through the subsurface and then into) surface water flow. On average over long time periods, the expectation is that $\gamma(P_{eff} - ET) \approx q \approx R_{eff}$. On shorter time scales, however, the available water flux for percolation $\gamma(P_{eff} - ET)$ may be transiently partitioned between the subsurface runoff R_{eff} and a soil water storage change $\Delta S = \gamma(P_{eff} - ET) - R_{eff}$ at each time step of the integration (3). A change in the depth of the groundwater table at each integration time step is then calculated by distributing the subsurface storage change ΔS over the available unsaturated pore space per unit area (θ_s-θ_{uz}), with the total z_{gw} change from initial time t_0 to time t thus being

$$z_{gw}(t;t_0) = z_{gw-0}(t_0) + \int_{t_0}^{t} \frac{\gamma[P_{eff}(\tau) - ET(\tau)] - R_{eff}(\tau)}{\theta_s - \theta_{uz}(\tau)} d\tau \quad (4)$$

where $z_{gw\text{-}0}$ is the initial groundwater level position at time t_0.

2.2.2. Facilitating Model Comparison With GRACE Data

For the model-data comparison in the first set of large catchments regarding large-scale subsurface water storage change, we note that the remote sensing data from GRACE [*Swenson*, 2012; *Landerer and Swenson*, 2012; *Swenson and Wahr*, 2006] (a) quantify total water storage anomalies (TWS), i.e., including also water storage in lakes (LWS) and (b) do not differentiate between storage in the unsaturated zone and that in the saturated zone as the model does. Where applicable based on other available data for lake storage LWS in the large study catchments (Figure 1a), we distinguish then the subsurface water storage change (Δ) as

$$\Delta SWS = \Delta TWS - \Delta LWS \quad (5)$$

where ΔTWS is the total water storage change derived from GRACE [*Swenson*, 2012; *Landerer and Swenson*, 2012; *Swenson and Wahr*, 2006] and ΔLWS is lake water storage change derived from applicable lake level time series in the DAHITI database [*Schwatke et al.*, 2015], with both as well as ΔSWS being normalized by the whole catchment area.

Furthermore, from the basic model described above, we calculate the modeled subsurface water storage change as

$$\Delta SWS = \Delta SWS_{UZ} - \Delta SWS_{SZ} \quad (6)$$

Here ΔSWS_{UZ} is the water storage change in the unsaturated zone, calculated from the soil water content θ_{uz} (equation (2)) as

$$\Delta SWS_{UZ} = \Delta\theta_{UZ} \cdot |z_{gw-0}| \quad (7)$$

In consistency, ΔSWS_{SZ} is the water storage change in the saturated zone, calculated from the relevant hydroclimatic data series as

$$\Delta SWS_{SZ} = \gamma.(P_{eff} - ET - R) \quad (8)$$

2.3. Sites and Data

The aim of this study is to compare key model outputs from the framework of *Destouni and Verrot* [2014] and *Verrot and Destouni* [2015] with relevant long-term observation data. This is done for large-scale subsurface water storage change (ΔSWS) in the first set of 25 large study catchments (Figure 1a), and for locally measured unsaturated water content (θ_{uz}) and groundwater level (z_{gw}) in the second set of the 10 U.S. catchments (Figure 1b). The primary criterion for relevant such comparison is availability of continuous data for these output variables over a relatively long time period (with the lower limit set to at least one full year of data); other relevant criteria for study catchment selection are outlined below.

2.3.1. The World Set of 25 Large Study Catchments—Study Part I

Out of 608 catchments around the world for which the *Global Runoff Data Center* (*GRDC*) [2015] provides runoff data time series, 25 are found that fulfill the following selection conditions: matching time windows for the necessary data sets (total water storage TWS, lake level change LWS, precipitation P, evapotranspiration ET, and runoff R), catchment area greater than 100,000 km^2 in order to avoid downscaling of the gridded data sets (TWS, P, and ET) so that only the considered model performance is assessed here, and no negative surface temperatures in order to be able to neglect snow dynamics effects in the modeling ($P \approx P_{\text{eff}}$ in equations (3), (4), and (8)). The selected catchments based on these conditions are spread over different parts of the world (see Figure 1a), which we cluster in four of the six hydrological regions defined by the *World Meteorological Organization* [2014] as 1 in North America (Nam 1), 12 in South America (Sam1 to Sam12), 8 in Africa (Afr1 to Afr8), and 4 in Southwest Pacific (only in Australia, Swp1 to Swp4). The runoff stations defining these catchments are listed in Table A2.

Most of these catchments do not have large lakes (defined here as lakes with greater relative area than 0.5% of the total catchment area) that can considerably affect the values of subsurface water storage changes, as obtained from equation (5) with data for ΔTWS derived from GRACE [*Swenson*, 2012; *Landerer and Swenson*, 2012; *Swenson and Wahr*, 2006]. Table A2 lists the lakes that were taken into account and the smaller ones that were neglected for each applicable catchment.

As there is no consensus global data set of groundwater level throughout the world, the long-term average depth of initial groundwater level $z_{\text{gw-}0}$ (equation (4); with the absolute value of $z_{\text{gw-}0}$ also representing the thickness of the unsaturated zone in equation (7)) is calibrated for each study catchment. Similarly, the γ factor for effective runoff R_{eff} and in equations (3), (4), and (8) is calibrated for this set of catchments. The calibration procedure for both $z_{\text{gw-}0}$ and γ is described in the model-data comparison section 2.4.1.

As described in supporting information section S1, the most representative U.S. Department of Agriculture (USDA) soil texture [*Baldwin et al*., 1938] in each catchment is found from the Harmonized World Soil Database v1.1 [*Nachtergaele et al*., 2008; *FAO*, 2012] and corresponding soil hydraulic parameters used to evaluate θ_{uz} from equation (2) are obtained from *Xia et al*. [2012], as listed in Table A2.

The model comparison with the data for the large-scale subsurface water storage change is based on monthly hydroclimatic data values; it is only the account of snow dynamics, which is not included here (but is included for the set of 10 U.S. catchments, see the following subsection), that requires use of daily hydroclimatic data values. For the monthly hydroclimatic data (shown for all catchments in supporting information Figure S1), P is given by the Global Precipitation Climatology Center (GPCC) [*Schneider et al*., 2011], a worldwide 1° × 1° data sets that extends from 1901 to near present; ET is extracted from the MODIS product [*ORNL DAAC*, 2011], a 0.05° × 0.05° global grid available for the period 2000–2015; and time series for R are derived from the discharge values given by the *Global Runoff Data Center* (*GRDC*) [2015] (see stations in Table A2).

2.3.2. The U.S. Set of 10 Study Catchments—Study Part II

Among available data sets for locally measured unsaturated water content θ_{uz} and groundwater level z_{gw} from the Ameriflux platform [*USGS*, 2015], two nested catchments have z_{gw} data over a continuous period of more than 10 years (12 years) along with the longest associated data set for θ_{uz} (over 78 months in the larger 4119440 catchment and the smaller 4119441 catchment within it, Figure 1b). In the following, we consider the largest catchment with the longest data records, 4119440, as a base case catchment, with the smaller nested catchment with equally long-term data, 4119441, as duplicate observation for essentially the same site. To enable some multisite comparison, we also consider the eight additional comparative catchments with consistent, even though shorter, data availability down to at least one full year of data (Figure 1b shows all 10 study catchments and their number of years with relevant θ_{uz} and z_{gw} data). From the *Global Runoff Data Center* (*GRDC*) [2015], corresponding runoff R data is obtained for all of these study catchments, with Table A3 summarizing the station details for the available daily θ_{uz}, z_{gw}, and R data sets.

For this set of catchments, effective runoff R_{eff} is estimated from the time series of observed R with the associated γ factor based on simulations results of *Bosson et al*. [2012] for the ratio of groundwater recharge and subsequent runoff (R_{gw} in their notation) to total runoff R under various hydroclimatic conditions. On average, this ratio was found to be 0.73 for temperate and 0.53 for cold climate conditions (the latter without any prevailing permafrost). For consideration of γ variability effects in the various climate conditions of the present study catchments, we use $\gamma = 0.53$ for months with negative average temperature (when both R and γ are relatively small due to snow storage

and frozen ground conditions) and $\gamma = 0.73$ for months with positive average temperature (when both R and γ are relatively large due to snow melt and unfrozen ground conditions); these choices are consistent with those of a previous study accounting for such γ variability effects [*Verrot and Destouni*, 2015]. The required data for air temperature T_A are extracted from the daily reanalysis product of the National Center for Atmospheric Prediction [*NCEP*, 2004], a national gridded data set of approximately 0.3° resolution.

The additional hydroclimatic data needed for the modeling of z_{gw} (cf. equations (3) and (4)) include P data, which are extracted for each of the 10 study catchments from NARR [*NCEP*, 2004], and ET data, which are extracted from the MODIS product [*ORNL DAAC*, 2011].

Regarding the handling of snow storage and melting dynamics, which are accounted for here through the estimation of P_{eff} from observed P, the underlying degree-day model [*Rankinen et al.*, 2004a; *Verrot and Destouni*, 2015] (cf. equations (S1)–(S3) of supporting information section S2) requires estimation of four temperature-related parameters. The values of these are not specifically known for the present study catchments but assigned as in previous studies for sites of different climate conditions [*Rankinen et al.*, 2004b; *Verrot and Destouni*, 2015]. The influence of the degree-day parameter choices on model results is further assessed by sensitivity analysis, with methodology and detailed results presented in supporting information section S3. Overall, in the 10 investigated U.S. catchments, the model exhibits small sensitivity to the precise snow parameter values.

With regard to groundwater level z_{gw}, we consider here wells with average groundwater level shallower than or around −3 m over the time period of investigation, in consistency with previous numerical testing of θ_{uz} estimation based on equation (2) [*Destouni*, 1991], and in order to focus on the variability of relatively shallow groundwater that can affect the near-surface variability of soil moisture. Soil parameter values used to evaluate θ_{uz} from equation (2) are further extracted from *Rawls et al.* [1982]. As described in supporting information section S1, parameters corresponding to different soil types present within each catchment are considered in the modeling of θ_{uz} and compared to available θ_{uz} data. For each catchment, the soil data providing the most relevant average level of modeled θ_{uz} is selected for catchment-scale soil representation; *Destouni and Verrot* [2014] and *Verrot and Destouni* [2015] have previously shown that the choice of parameter values for soil texture does not impact the model results for relative variability of soil water content and groundwater level around their respective long-term average values. The selected soil data for each study catchment are listed in Table A2.

2.4. Model-Data Comparison

2.4.1. The World Set of 25 Large Catchments—Study Part I

For this set of catchments, we compare model results for subsurface water storage change (ΔSWS, equations 6–8) with corresponding large-scale data (equation (5)). Out of the 25 catchments, Afr4 has the longest time series period with data (12.5 years) and is used for exemplification of results in the World Meteorological Organization (WMO) region of Africa in the main paper, with corresponding results for the other catchments presented in supporting information figures. Similarly, Sam11, Nam1, and Swp1, which have time series with data over 8.3 years, 12.5 years, and 10 years respectively, are used for exemplfication of results in the WMO regions of South America, North America, and Southwest Pacific.

Single values of $z_{gw\text{-}0}$ and γ are calibrated over the entire time series of each catchment through a simple Monte Carlo algorithm, assuming a normal probability distribution within the intervals [−0.5 m; −20 m] for $z_{gw\text{-}0}$ and [0; 1] for γ. The calibrated values optimize the Nash-Sutcliffe model efficiency coefficient (NS) defined as [*Nash and Sutcliffe*, 1970]

$$\mathrm{NS} = 1 - \frac{\sum_{t=0}^{T}\left(Q_{d,t} - Q_{m,t}\right)^2}{\sum_{t=0}^{T}\left(Q_{d,t} - \mu_d\right)^2} \tag{9}$$

where $Q_{d,t}$ and $Q_{m,t}$ represent the data and the modeled variable value at time t, respectively, and μ_d is the mean variable value over the whole time series period T. Using the single $z_{gw\text{-}0}$ and γ values obtained from this calibration process, the model performance is assessed with regard to the monthly values of water storage change in the average year over T for each catchment.

2.4.2. The U.S. Set of 10 Study Catchments—Study Part II

For this set of catchments, we compare model results with locally measured field data for groundwater level z_{gw} within 3 m depth below the surface and unsaturated water content θ_{uz}, averaged down to the deepest point measured at each catchment site. For the longest possible temporal comparison between data and

model results, we consider and present in the main text results for the base case catchment 4119440. In supporting information figures we present corresponding results for the smaller nested catchment 4119441 with the same data series length and for the other eight study catchments with shorter data availability (Figure 1b).

Original data for water content include point values measured at different depths of a soil profile within each catchment (Table 3). As the modeled θ_{uz} represents a depth-averaged value, the point data values are averaged over depth for relevant model-data comparison. Also, z_{gw} measurements represent mostly point values within each catchment; where several measurement wells exist within a catchment, however, it is their spatial average that is compared with the modeled z_{gw}.

In general, the long-term average values of θ_{uz} and z_{gw} in a single soil profile may not represent the long-term average values of θ_{uz} and z_{gw} over a whole catchment. However, wherever the necessary field data are available, the procedure for selecting representative soil parameter values (section 2.3.2 and supporting information section S1) can equally well be used to reproduce catchment-averaged θ_{uz} and z_{gw} conditions as single-profile conditions. In any case, this study addresses the relative temporal variability around long-term average values, which may be similar for catchment-averaged and single-profile data because the hydroclimatic conditions that drive this temporal variability are relatively homogeneous over relatively large spatial scales.

With regard to temporal variability around long-term averages, we quantify and compare detrended and normalized temporal signals and their spectra of modeled and data time series of θ_{uz} and z_{gw}. For the spectral analysis, we use a standard procedure: a fast Fourier transform (FFT) applied to the detrended and normalized original signals. We also investigate cleaned signals with focus on seasonal (intra-annual) variability by applying a filter with cutoff frequency of 1.5 year^{-1}. The signal and spectral analysis is performed for the five study catchments with data time series containing more than 50 monthly values: 4119440, 4119441, 4122131, 4122630, and 4127300 (the latter of which has 50 months data only for z_{gw}).

With regard to detrending, the available *P* time series from NARR [*NCEP*, 2004] exhibits a decreasing trend, with no corresponding trend shown by the independent *R* or ET data sets (supporting information Figure S2). Furthermore, also the alternative independent (monthly) *P* data set of the CPC Unified Gauge-Based Analysis of Daily Precipitation over CONUS [*CPC US*, 2015] does not show any similar decreasing *P* trend as the NARR data set for *P* [*NCEP*, 2004] (supporting information Figure S2). The three independent data sets for ET, *R*, and *P* from *CPC U.S.* [2015] thus combine in indicating a trend bias in the *P* data set from NARR [*NCEP*, 2004].

The daily *P* data from NARR [*NCEP*, 2004] is nevertheless needed to account for effects of snow storage and melting dynamics in the degree-day model component of *Verrot and Destouni* [2015] (supporting information section S2). This account may be important for a majority of study catchments that have cold winters with monthly average temperatures below 0°C. To bias correct the daily *P* data from NARR [*NCEP*, 2004], we use the monthly *P* data ($P_{m,cpc}$ for each month *m*) from *CPC US* [2015] as

$$P_{d-m,corr} = f_{m,corr} \cdot P_{d-m} \qquad (10)$$
$$f_{m,corr} = P_{m,cpc}/P_m$$

where $P_{d\text{-}m}$ is original and $P_{d\text{-}m,corr}$ is corrected daily *P* for each day *d* in month *m*, and P_m is aggregated monthly *P* for month *m* from the NARR [*NCEP*, 2004] data set.

The bias correction of $P = P_{d\text{-}m,corr}$ does not remove the entire flux imbalance implied by the *P* data in relation to the ET and *R* data (supporting information Figure S2), and the remaining imbalance is also not reflected as any long-term change trend in the observed z_{gw} and θ_{uz} data (see section 3). Various additional biases may thus affect any or all of the data from different sources, e.g., related to *P* undercatch, *ET* misinterpretation from satellite data, and/or limited catchment-scale representativity of z_{gw} and θ_{uz} data. To focus on the model ability to capture the relative temporal variability around any data-given long-term average of z_{gw}, expressed as μ, all modeled time series are centered at the mean of their respective long-term data series. Equation (4) is then calculated as follows:

$$z_{gw}(t;t_0) = \overline{z_{gw-obs}} + \int_{t_0}^{t_{0-a}} \left(\frac{\gamma[P_{eff}(\tau) - ET(\tau)] - R_{eff}(\tau)}{\theta_s - \theta_{uz}(\tau)} \right) d\tau - \mu \qquad (11)$$

where $\overline{z_{gw-obs}}$ is the long-term average groundwater level, as given by the actual z_{gw} data over the entire time period of each data series.

3. Results

3.1. The World Set of 25 Large Catchments—Study Part I

The model-data comparison for large-scale water storage change shows that the calibrated model reproduces well the overall relative temporal variation of ΔSWS (top panels in Figures 2a–2d) and its monthly variation through the average year (bottom panels in Figures 2a–2d); Figure 2 exemplifies these results for catchment Afr4 (Figure 2a), Sam11 (Figure 2b), Nam1 (Figure 2c), and Swp1 (Figure 2d) and corresponding results for all catchments are shown in supporting information Figures S3 and S4.

For both the whole time series (Figure S3) and the average year (Figure S4), the lowest NS values are obtained for three Southwest Pacific catchments Swp2, Swp3, and Swp4, and for the North American catchment (Nam1). Overall, the fluctuation patterns are well reproduced by the model, but the minimum and maximum fluctuation levels, i.e., the fluctuation range, may not be fully captured.

The resulting Nash-Sutcliffe model efficiency coefficient NS is mostly somewhat higher in the model-data comparison for the intra-annual variability over the average year (bottom panels in Figures 2a–2d and Figure S4) than for the total temporal variability over the entire time series (top panels in Figures 2a–2d and Figure S3). The standard deviation of the intra-annual temporal variability in the average year tends to be smaller for the model results than for the data (bottom panels in Figures 2a–2d and Figure S4). Overall, there is a pattern of the model efficiency being higher (NS values closer to 1) for catchments with larger temporal variability in their data; this applies to results for the whole time series (Figure S3) and for the average year (Figure S4); i.e., for both the total and the intra-annual temporal variability, the model captures larger fluctuations better than smaller ones. This result reflects a general model bias toward capturing the main processes and patterns of relatively large temporal fluctuations but missing those of relatively small fluctuations; this bias implies a model tendency to somewhat underestimate rather than overestimate the temporal changes in subsurface water storage around its long-term average value.

The calibrated model values of γ and $z_{gw\text{-}0}$ are shown in supporting information Figure S5 and presented and discussed in supporting information section S4. While some absolute values of these two calibrated parameters are relatively low ($\gamma \approx 0.1$ in Swp1 and $|z_{gw\text{-}0}| \approx 0.05$ in four catchments, Afr5, Sam9, Sam12, and Swp4), overall the calibration results show γ values closer to 1 and groundwater level depth $|z_{gw\text{-}0}|$ on the order of several decimeters or meters rather than centimeters. Moreover, previous results have shown that calibration of soil texture and long-term groundwater level parameters affects mainly the long-term average value of soil moisture but does not impact much the modeled relative temporal variability around that value for each catchment [*Destouni and Verrot*, 2014; *Verrot and Destouni*, 2015]. In analogy, the present calibration of γ and $z_{gw\text{-}0}$ aims to reproduce a data-consistent long-term average water storage in each catchment but does not much affect the temporal storage changes around that value, which is the main result addressed and discussed in this study.

3.2. The U.S. Set of 10 Study Catchments—Study Part II

The model-data comparison with local measurements in these catchments regards the relative temporal variability of soil water content θ_{uz} (Figure S6, showing also the model calibration of soil parameters according to supporting information section S1) and groundwater level z_{gw} (Figure S7, showing also the model sensitivity analysis for snow dynamics according to supporting information section S3). The model comparison with the relative variability and spectral signals of θ_{uz} in catchment 4119440 with the longest data series (Figure 3) shows that the model captures most of the temporal fluctuation pattern around the long-term temporal mean, and the occurrence frequencies of the main fluctuations, but not fully the fluctuation magnitudes. The model captures particularly well the power, occurrence frequency, timing, and often also the actual magnitude of the main annual peak (frequency of around 1 $year^{-1}$). Also for the other catchments with relatively long data records the main annual peak in θ_{uz} (frequency of around 1 yr^{-1}) is mostly reproduced relatively accurately, even with regard to its magnitude, by the model (supporting information Figure S8).

Concerning the groundwater level z_{gw} for the catchment 4119440 with the longest data time series (Figure 4 and supporting information Figure S9 for the other relevant catchments), conclusions from the relative variability and spectral signal analysis are mostly the same as for the unsaturated soil water content θ_{uz} (Figure 3 and supporting information Figure S8). A main annual peak (frequency of around 1 $year^{-1}$) is also evident in the z_{gw} data and relatively well reproduced by the model for most catchments. The annual peak is less well reproduced

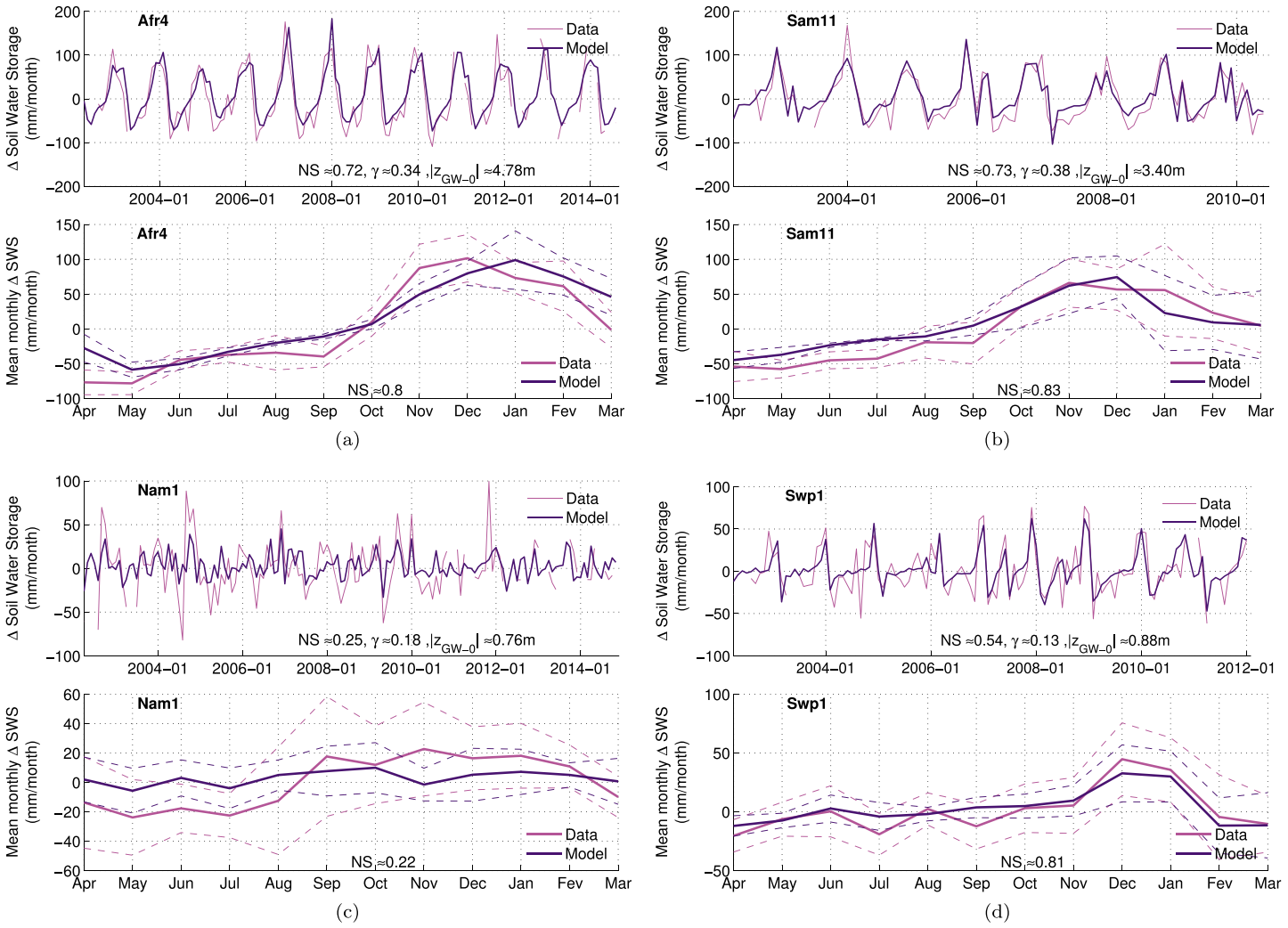


Figure 2. Subsurface water storage change ΔSWS in catchments (a) Afr4, (b) Sam11, (c) Nam1, and (d) Swp1 from Figure 1a. (top panels) Time series of ΔSWS obtained from GRACE data (pink line) and from the calibrated model (purple line). (bottom panels) Intra-annual variation in ΔSWS over the average year, as obtained from the calibrated model (purple line) and GRACE data (pink line). In each top panel, the Nash-Sutcliffe model efficiency coefficient NS (see equation (9)), the optimized temporal shift, and the values of the calibrated factor γ and thickness of the unsaturated zone $|z_{gw\text{-}0}|$ (with $z_{gw\text{-}0}$ being initial groundwater level) are indicated at the bottom of the panel. In each bottom panel, the variation in the mean monthly ΔSWS is represented by the solid lines, the value of one temporal standard deviation around the mean is represented by dotted lines, and the *NS* value is indicated at the bottom of the panel for the model-data comparison for the average year (and not for the entire time series). Corresponding results for all catchments in Figure 1a are presented in supporting information Figures S3 and S4.

for the longest time series catchment 4119440 (Figure 4) than for its duplicate nested catchment 4119441 and the other relevant catchments (supporting information Figure S9). However, the main longer-term peaks with smaller frequency than 1 year^{-1} are instead better reproduced for catchment 4119440 than for most other catchments, while also these are relatively well reproduced for the nested catchment 4119441. The seasonal (intra-annual) dynamics of the cleaned low-frequency z_{gw} signals are reasonably well reproduced by the model for most catchments, even though mostly underestimating the fluctuation magnitudes.

4. Discussion

Soil moisture is an important variable in large-scale hydroclimatic and environmental models. However, model-data comparisons of long-term soil water content and groundwater level dynamics over large spatial scales are hampered by several data limitations and scaling problems. Continuous, long-term and large-scale subsurface monitoring, in particular, of soil water content and also of groundwater level is rare. Also, for the

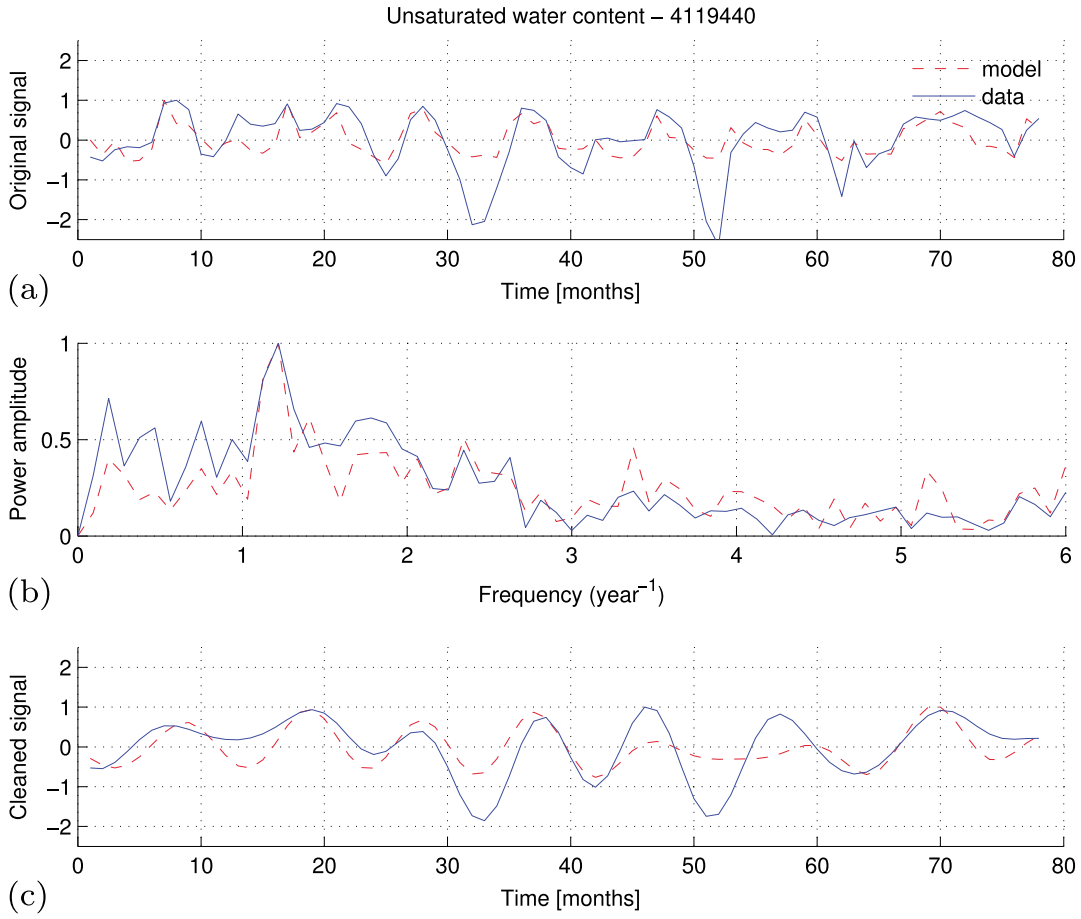


Figure 3. Model-data comparison for relative temporal variability of depth-averaged unsaturated soil water content. Results are shown for catchment 4419440 among the U.S. catchments in Figure 1b, with red dashed lines showing model results and blue solid lines showing data-based results for (a) the detrended and normalized signal, (b) the power spectrum of the detrended and normalized signal, and (c) a cleaned detrended and normalized signal by application of a low-pass filter with cutoff frequency of 1.5 years^{-1}. The detrending procedure consists of removing from the absolute data values the associated long-term mean value (if essentially constant) or best fit regression line (in the least square sense). Normalization by division of the time series values by the maximum value found in the detrended signal was done for illustrative purposes (improved clarity) in Figure 3a. The fast Fourier transform (FFT) of the signal for Figure 3b is not applied to the normalized time series. For all panel results, gaps in the data time series and corresponding model results are not included. Corresponding results for all catchments in Figure 1b with sufficient time series length for FFT application are presented in supporting information Figure S8.

present model-data comparison the scarcity of long-term soil moisture monitoring has been a limiting factor in finding relevant study catchments.

The question of how to compare large-scale model results with available local measurements is both important and interesting [*Western et al.*, 1999 and *Teuling and Troch*, 2005]. Not least the scale transition required for reconciliation of local data and large-scale modeling of soil moisture has been a hot topic from the beginning of the modeling era [*Corwin et al.*, 2006]. One way to resolve such hydrological scaling problems is to statistically upscale point measurements to the catchment scale based on available variability information and assumptions that can bridge the information and data gaps [see *Hopmans et al.*, 2002]. Along this line, many authors have addressed the problem of soil heterogeneity using stochastic models [*Destouni*, 1993] and statistical tools such as Monte Carlo analysis [*Foussereau et al.*, 2000]. Among the limitations of this approach, however, is the need for multiple point measurements in order to realistically quantify relevant statistical distributions of point variables for large spatial scales. As in the present set of the 10 U.S. study catchments (Figure 1b), locally measured soil moisture data may often only be available for a single soil profile within a whole catchment. The present comparison of large-scale model with such local, single station data is therefore interesting; it is also relevant by considering the

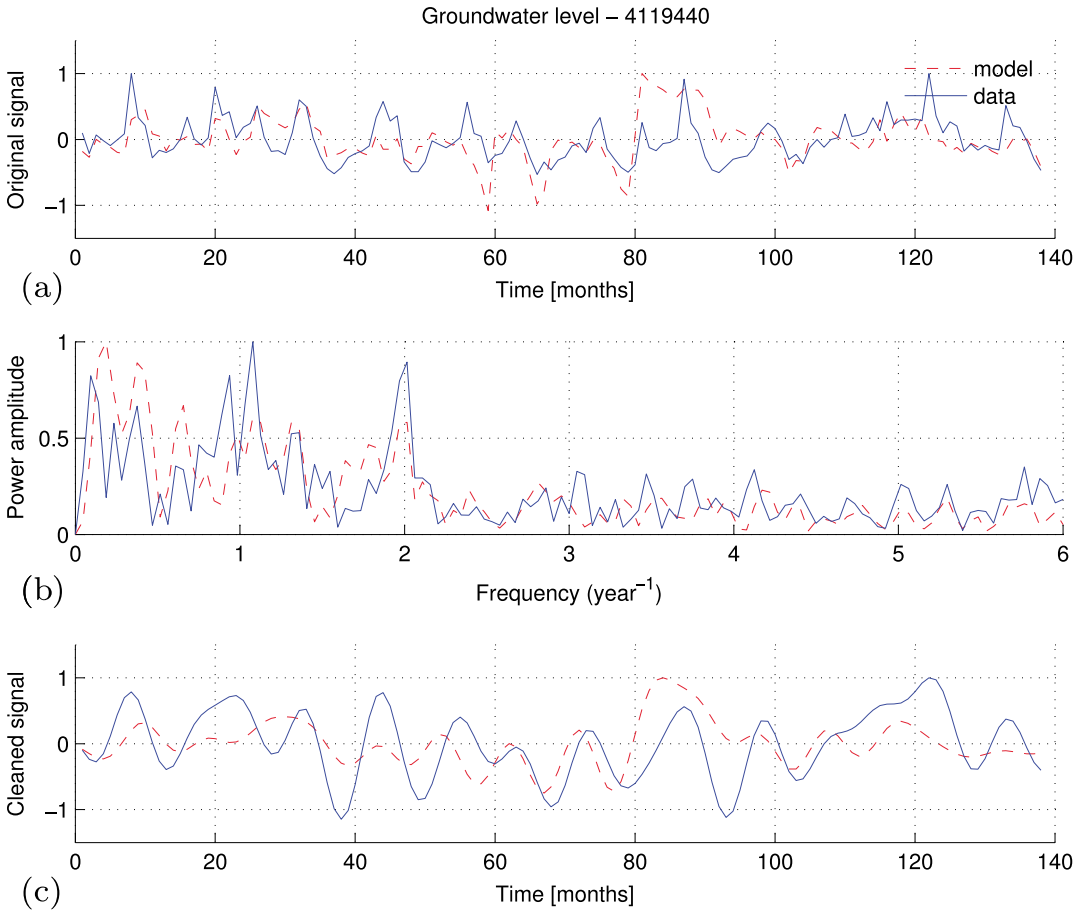


Figure 4. Model-data comparison for relative temporal variability of groundwater level. Results are for catchment 4419440 among the U.S. catchments in Figure 1b, with red dashed line showing model results and blue solid line showing data-based results for (a) the detrended and normalized signal, (b) the power spectrum of the detrended and normalized signal, and (c) a cleaned detrended and normalized signal by application of a low-pass filter with cutoff frequency of 1.5 years^{-1}. The detrending procedure consists of removing from the absolute data values the associated long-term mean value (if essentially constant) or best fit regression line (in the least square sense). Normalization by division of the time series values by the maximum value found in the detrended signal was done for illustrative purposes (improved clarity) in Figure 4a. The fast Fourier transform (FFT) of the signal for Figure 4b is not applied to the normalized time series. For all panel results, gaps in the data time series and corresponding model results are not included. Corresponding results for all catchments in Figure 1b with sufficient time series length for FFT application are presented in supporting information Figure S9.

relative temporal variability around the respective long-term average value at each scale, even though latter is likely to differ between the scales in absolute terms.

Another way to resolve the scaling problem may be to use relevant satellite products, such as the GRACE data used in the present study, and/or passive and active microwave sensors of relevant data. The latter may, for instance, provide soil moisture values for the top soil layer or proxies for soil moisture over the root or vadose zones [*Wang et al.*, 2007; *Qin et al.*, 2013], which can be further used in and/or compared with soil moisture modeling [*Das and Mohanty*, 2006; *Scott et al.*, 2003]. Although results are promising, they too have limitations, including that soil moisture information from satellite/airborne products may only be available for relatively short and recent time periods and that the usefulness of such products for retrieving soil moisture data may also depend on other variables, such as vegetation type, climate, and soil properties. The present study also clarifies that GRACE data do not differentiate the water in the unsaturated zone from that in the saturated zone and can be used to evaluate relative water storage changes but not readily absolute soil moisture values; product adjustments may be required to overcome such limitations in the future [*Schmidt et al.*, 2006; *Awange et al.*, 2009]. Due to the orbital speed and period of revolution of the satellites (including GRACE satellites), data derived from their observations represent snapshots in time and do not continuously capture the change evolution between two measurement points in time.

Large-scale hydroclimatic input data may also have biases and limitations. Data for atmospheric climate (T_A and P), runoff through the landscape (R), and the landscape-atmosphere flux link of evapotranspiration (ET) are provided by different sources for different scales and are not often consistently combined to check their joint water balance implications on catchment scales. The tested modeling approach in this study combines such large-scale hydroclimatic data, including the catchment-integrating runoff R, with remotely sensed large-scale soil data and/or locally measured soil data. This data combination showed considerable bias and flux imbalance, implying water storage change that was not supported as realistic by concurrent soil moisture (θ_{uz} and z_{gw}) data. Other recent studies have shown similar biases and flux imbalances also in climate model projections of large-scale water fluxes, with unrealistic implications for long-term water storage changes in the landscape [*Bring et al*., 2015].

Locally measured data can facilitate resolution of soil moisture variability and change on smaller catchment scales and for other time periods than those relevant for large-scale GRACE data. Such local data can be spatially upscaled if/where they exist for multiple locations within a catchment, such that relevant spatial distributions can be derived over the whole catchment scale. Such data existence is possible but was not the case for the present study catchments. The periods of available subsurface data time series are also often relatively short (between 1 and 12.5 years in the present study catchments), whereas the modeling framework aims at screening and synthesizing long-term times series of hydroclimatic and soil data for assessing the associated soil moisture variability and change.

Nevertheless, with the data synthesis facilitated by the present modeling approach, relevant model-data comparison can still be carried out, providing physically reasonable and useful results. The results from the model comparison with GRACE data for the set of 25 large catchments (Figure 1a) show good model representation of large-scale subsurface water storage changes across different parts of the world. The results from the model comparison for the set of 10 U.S. catchments show that the model captures most of the relative fluctuation pattern around the long-term temporal mean of locally measured soil water content and groundwater level. This provides support for the underlying assumption that relative soil moisture variability patterns may be similar locally as for larger scales, even though representative (average) soil moisture values may differ between scales in absolute terms.

The model also captures well the occurrence and frequency of main soil moisture fluctuations but does not fully capture the whole fluctuation range; i.e., it tends to underestimate rather than exaggerate the latter. This is physically reasonable if, for instance, water content data are measured over relatively shallow depth and/or locally, because the temporal variation of water content close to the land surface and in a single point is generally expected to be greater than for the model-quantified depth- and area-averaged water content all the way down to the groundwater table and over a whole catchment [*Destouni*, 1991, 1992; *Li and Islam*, 1999]. Previously, *Destouni and Verrot* [2014] have investigated temporal variability changes for θ_{uz} and z_{gw} by use of the tested modeling approach, finding increased hydrological and agricultural drought risk due to hydroclimatic shifts driven by twentieth century agricultural expansion-intensification [*Destouni et al*., 2013]. Not least for such risk studies and results, it is important that the present results show that the model tends to underestimate rather than exaggerating the temporal variability and thereby also the associated changes in drought and flood risk under hydroclimatic change.

5. Conclusion

We have investigated the ability of a relatively simple analytical modeling framework [*Destouni and Verrot*, 2014; *Verrot and Destouni*, 2015] to reproduce main temporal variability dynamics of soil moisture around relevant long-term average values of various key measures and components. Large-scale subsurface change in water storage is one of these measures, with model-data comparison results for 25 large study catchments around the world showing that its relative temporal variability around long-term average conditions is well reproduced by the model in most of the catchments.

Unsaturated water content and groundwater table level are two other key measures and components of soil moisture over some depth from the land surface. Model-data comparison for soil water content and relatively shallow groundwater level in the 10 study catchments across the U.S. shows that the tested modeling approach consistently reproduces relative variability dynamics in terms of seasonal and longer-term fluctuation timings and frequencies.

Overall, the model tends to underestimate rather than exaggerate the temporal variability range and temporal changes of soil moisture around and from its long-term average conditions. This is particularly so for the fluctuation magnitudes of local measurement data because local quantities generally tend to fluctuate more than corresponding large-scale average quantities. At any rate, the present results show that the tested large-scale modeling framework fulfills reasonably well its main aim of relatively simple and unexaggerated determination of temporal variability and change around long-term average soil moisture conditions.

The model represents a synthesis of large-scale time-variable hydroclimatic data and relatively time-invariant soil data that enables screening of long-term soil moisture variability and change on catchment scales. By such a catchment synthesis, which is based on fundamental catchment-scale water balance, the model also facilitates identification of flux imbalance biases in the data time series that are used as its large-scale hydroclimatic drivers.

The model ability to screen available historic long-term data series of catchment-scale precipitation, evapotranspiration, and runoff in order to determine associated soil moisture variability and change has here been tested against different types and scales of available observation data and for various soil moisture aspects (quantities). This ability is expected to also be useful for screening model-projected long-term hydroclimatic time series to determine scenarios of future variability and change in large-scale soil moisture.

Appendix A

Tables A1, A2, and A3 present the list of the catchments and their respective discharge station, soil properties, USGS well identification number, and Ameriflux soil water content data. Table A4 lists the snow parameters used in the degree-day model presented in the supporting information section S2.

Table A1. The Set of 25 Large Study Catchments Across the World (Figure 1a) and the Lakes Accounted for

Catchment Number	GRDC Station	Lake Name (Relative Area in % of the Total Catchment Area)
		Lake With Relative Area < 0.5% (% Given in Parentheses), Not Accounted for in This Study
Afr1	1147010	Lake Tanganyika (0.91), Lake Malawi (0.91) *Lake Kivu (0.07), Lake Bangweulu (0.06), Lake Mweru (0.14)*
Afr2	1196551	
Afr3	1257100	
Afr4	1291100	
Afr5	1291200	
Afr6	1531100	
Afr7	1531450	
Afr8	1591401	
Sam1	3618500	
Sam2	3621400	
Sam3	3624120	
Sam4	3625310	
Sam5	3627650	
Sam6	3627810	
Sam7	3629150	
Sam8	3630120	
Sam9	3649950	*Represa de Tucuruí (0.45)*
Sam10	3650481	
Sam11	3651807	
Sam12	3667060	
Nam1	4150500	
Swp1	5101200	
Swp2	5101301	
Swp3	5404270	
Swp4	5410100	

Table A2. All Considered Soil Parameters for Results in Supporting Information Figure S5 After *Rawls et al.* [1982] and the Selected Ones for the Modeling of Each Study Catchment According to Supporting Information Section S1

USDA Soil Texture	K_s (m/s)	θ_{ir} (−)	θ_s (−)	β (−)	Selected as Representative for Catchment (Worldwide Catchments-*U.S. Catchments*)
Sand	5.83×10^{-4}	0.02	0.44	0.17	Afr3, Afr4, Afr5, *4152450*
Loamy sand	1.70×10^{-4}	0.04	0.44	0.15	*4114282*
Sandy loam	7.19×10^{-5}	0.04	0.45	0.12	Afr7, Sam10, *4119100*, *4122131*
Loam	3.67×10^{-5}	0.03	0.46	0.09	Afr6, Nam1, *4119280*, *4123063*
Silt loam	1.89×10^{-5}	0.02	0.50	0.09	*4119440*, *4119441*, *4122630*
Sandy clay loam	1.19×10^{-5}	0.07	0.40	0.11	Afr1, Afr2, Sam1, Sam2, Sam6, Sam7, Sam8, Sam9, Sam11, Sam12, Swp1, Swp3
Clay loam	6.39×10^{-6}	0.08	0.46	0.09	Afr8, Sam3, Sam4, Swp2, Swp4
Silty clay loam	4.17×10^{-6}	0.04	0.47	0.07	*4127300*
Clay	1.67×10^{-7}	0.09	0.48	0.07	Sam5

Table A3. Catchment and Data Station Information for the Set of U.S. Study Catchments in Figure 1b

Catchment Number/GRDC Station	Water Content Station From Ameriflux	Depth of Water Content Sensors (m)	Groundwater Well (USGS Identification Number)
4152450	Corral Pocket, Utah [*Bowling et al.*, 2010]	0.05, 0.10 (× 2), 0.20, 0.40, 0.60	403158109372201
4127300	Goodwin Creek, Mississippi [*Xiao et al.*, 2010]	0.10, 0.20, 0.30, 0.40, 0.60, 1.00	334215089442701
4123063	Chestnut Ridge, Tennessee	0.05, 0.10, 0.20,	350750085045802
	Walker Branch, Tennessee [*Hollinger et al.*, 2010]	0.50, 1.00	351428085003600
			352315082484401
			352519083272401
			353922083345600
4122630	Nebraska SandHills Dry Valley, Nebraska [*Billesbach et al.*, 2004]	0.05, 0.10, 0.20, 0.50	420204101200502
4122131	Niwot Ridge, Colorado [*Blanken et al.*, 2009]	0.15 (× 2)	410827104501601
4119441	Willow Creek, Wisconsin [*Cook et al.*, 2014]	0.05, 0.10, 0.20, 0.50, 1.00	481713105052804
4119440	Willow Creek, Wisconsin [*Cook et al.*, 2014]	0.05, 0.10, 0.20, 0.50, 1.00	450947090483902
4119280	Brooks Field Site 10, Iowa [*Hernandez-Ramirez et al.*, 2011]	0.05	405820093252501
4119100	KUOM Turfgrass Field, Minnesota [*Peters and McFadden*, 2012]	0.10	455927095123101
4115282	Metolius First Young Pine, Oregon [*Schmidt et al.*, 2011]	0 to 0.30 (× 2)	434206121311701

Table A4. Degree-Day Model Parameters, From *Rankinen et al.* [2004a], for the Snow Dynamics Model Component (Supporting Information Section S2) and Sensitivity Analysis (According to Supporting Information Section S3 and With Results Shown in Supporting Information Figure S7)

	Scenario A	Scenario B	Scenario C	Scenario D	Used for Modeling, All Catchments
T_M (°C)	−0.0036	0.74	−0.0036	0.74	0.30
F_M (mm d^{-1}°C^{-1})	4.88	1.60	4.88	1.60	3.83
T_L (°C)	−2.02	−2.02	−4	−4	−2.02
T_U (°C)	0	0	3.70	3.70	3.70

Acknowledgments

This work has been supported by the Swedish University strategic environmental research program Ekoklim and the Swedish Research Council Formas (project 2012-790). The soil moisture data were downloaded from the Ameriflux website: funding for AmeriFlux data resources was provided by the U.S. Department of Energy's Office of Science. GPCC Precipitation data, GHCN Gridded V2 data, NARR data, and CPC US Unified Precipitation data were obtained from the Web site of NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, at http://www.esrl.noaa.gov/psd/.

References

Amante, C., and B. W. Eakins (2009), ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24, Natl. Geophys. Data Cent., NOAA, doi:10.7289/V5C8276, Mretrieved in March 2015.

Awange, J. L., M. A. Sharifi, O. Baur, W. Keller, W. E. Featherstone, and M. Kuhn (2009), GRACE hydrological monitoring of Australia: Current limitations and future prospects, *J. Spat. Sci.*, *54*(1), 23–36.

Baldwin, M., C. E. Kellogg, and J. Thorp (1938), Soil classification, in *Soils and Men, Yearbook of Agriculture*, pp. 979–1001 , U.S. Gov. Print. Off., Washington, D. C.

Bengtsson, L., R.-M. Bonnet, M. Calisto, G. Destouni, R. Gurney, J. Johannessen, Y. Kerr, W. A. Lahoz, M. Rast (Eds.) (2014), *The Earth's Hydrological Cycle*, vol. 46, Springer, Netherlands, doi:10.1007/978-94-017-8789-5.

Beven, K. J. (2001), Dalton medal lecture: How far can we go in distributed hydrological modelling?, *Hydrol. Earth Syst. Sci.*, *5*(1), 1–12.

Billesbach, D. P., M. L. Fischer, M. S. Torn, and J. A. Berry (2004), A portable eddy covariance system for the measurement of ecosystem-atmosphere exchange of CO_2, water vapor, and energy, *J. Atmos. Oceanic Technol.*, *21*(4), 639–650, doi:10.1175/1520-0426(2004)021<0639:APECSF>2.0.CO;2.

Blanken, P. D., M. W. Williams, S. P. Burns, R. K. Monson, J. Knowles, K. Chowanski, and T. Ackerman (2009), A comparison of water and carbon dioxide exchange at a windy alpine tundra and subalpine forest site near Niwot Ridge, Colorado, *Biogeosciences*, *95*(1), 61–76, doi:10.1007/s10533-009-9325-9.

Bosson, E., U. Sabel, L. G. Gustafsson, M. Sassner, and G. Destouni (2012), Influences of shifts in climate, landscape, and permafrost on terrestrial hydrology, *J. Geophys. Res.*, *117*, D05120, doi:10.1029/2011JD016429.

Bowling, D. R., S. Bethers-Marchetti, C. K. Lunch, E. E. Grote, and J. Belnap (2010), Carbon, water, and energy fluxes in a semi-arid cold desert grassland during and following multi-year drought, *J. Geophys. Res.*, *115*, G04026, doi:10.1029/2010JG001322.

Bresler, E., and G. Dagan (1981), Convective and pore scale dispersive solute transport in unsaturated heterogeneous fields, *Water Resour. Res.*, *17*, 1683–1693, doi:10.1029/WR017i006p01683.

Bring, A., S. M. Asokan, F. Jaramillo, J. Jarsjö, L. Levi, J. Pietroń, C. Prieto, P. Rogberg, and G. Destouni (2015), Implications of freshwater flux data from the CMIP5 multi-model output across a set of Northern Hemisphere drainage basins, *Earth's Future*, *3*, doi:10.1002/2014EF000296.

Brooks, R. H., and A. T. Corey (1964), *Hydraulic Properties of Porous Media*, *Hydrology Papers*, Colorado State Univ, Fort Collins, Colo.

Bukovsky, M. S., and D. J. Karoly (2007), A brief evaluation of precipitation from the North American Regional Reanalysis, *J. Hydrometeorol.*, *8*(4), 837–846.

Cook, B. D., et al. 2014), Carbon exchange and venting anomalies in an upland deciduous forest in northern Wisconsin, USA, *Agric. For. Meteorol.*, *126*(3–4), 271–295, doi:10.1016/j.agrformet.2004.06.008.

Corradini, C. (2014), Soil moisture in the development of hydrological processes and its determination at different spatial scales, *J. Hydrol.*, *516*, 1–5, doi:10.1016/j.jhydrol.2014.02.051.

Corwin, D. L., J. Hopmans, and G. H. de Rooij (2006), From field-to landscape-scale vadose zone processes: Scale issues, modeling, and monitoring, *Vadose Zone J.*, *5*(1), 129–139.

Climate Prediction Center (CPC) U.S. (2015), Unified precipitation data provided by the NOAA/OAR/ESRL PSD, Boulder, Colo. [Available at http://www.esrl.noaa.gov/psd/.]

Dagan, G., and E. Bresler (1979), Solute dispersion in unsaturated heterogeneous soil at field scale: I. Theory, *Soil Sci. Soc. Am. J.*, *43*, 461–467.

Das, N. N., and B. P. Mohanty (2006), Root zone soil moisture assessment using remote sensing and vadose zone modeling, *Vadose Zone J.*, *5*(1), 296–307.

Destouni, G. (1991), Applicability of the steady-state flow assumption for solute advection in field soils, *Water Resour. Res.*, *27*, 2129–2140, doi:10.1029/91WR01115.

Destouni, G. (1992), Prediction uncertainty in solute flux through heterogeneous soil, *Water Resour. Res.*, *28*, 793–801, doi:10.1029/91WR02979.

Destouni, G. (1993), Stochastic modelling of solute flux in the unsaturated zone at the field scale, *J. Hydrol.*, *143*, 45–61.

Destouni, G., and L. Verrot (2014), Screening long-term variability and change of soil moisture in a changing climate, *J. Hydrol.*, *516*, 131–139, doi:10.1016/j.jhydrol.2014.01.059.

Destouni, G., and V. Cvetkovic (1989), The effect of heterogeneity on large scale solute transport in the unsaturated zone, *Nord. Hydrol.*, *20*, 43–52.

Destouni, G., and V. Cvetkovic (1991), Field scale mass arrival of sorptive solute into the groundwater, *Water Resour. Res.*, *27*(6), 1315–1325, doi:10.1029/91WR00182.

Destouni, G., and W. Graham (1995), Solute transport through an integrated heterogeneous soil-groundwater system, *Water Resour. Res.*, *31*(8), 1935–1944, doi:10.1029/95WR01330.

Destouni, G., F. Jaramillo, and C. Prieto (2013), Hydroclimatic shifts driven by human water use for food and energy production, *Nat. Clim. Change*, *3*, 213–217.

Entekhabi, D., R. H. Reichle, R. D. Koster, and W. T. Crow (2010), Performance metrics for soil moisture retrievals and application requirements, *J. Hydrometeorol.*, *11*, 832–840.

Fan, Y., and H. Van den Dool (2008), A global monthly land surface air temperature analysis for 1948–present, *J. Geophys. Res.*, *113*, D01103, doi:10.1029/2007JD008470.

FAO (2012), *Harmonized World Soil Database (Version 1.2)*, FAO, Rome, Italy and IIASA, Luxemburg, Austria.

Fekete, B. M., C. J. Vörösmarty, and W. Grabs (2002), High-resolution fields of global runoff combining observed river discharge and simulated water balances, *Global Biogeochem. Cycles*, *16*(3), 1042, doi:10.1029/1999GB001254.

Foussereau, X., W. Graham, A. Aakpoji, G. Destouni, and P. S. C. Rao (2000), Stochastic analysis of transport in unsaturated heterogeneous soils under transient flow regimes, *Water Resour. Res.*, *36*, 911–921, doi:10.1029/1999WR900343.

Graham, W., G. Destouni, G. Demmy, and X. Foussereau (1998), Prediction of local concentration statistics in variably saturated soils: Influence of observation scale and comparison with field data, *J. Contam. Hydrol.*, *32*, 177–199.

Global Runoff Data Center (GRDC) (2015), The Global Runoff Data Centre, D-56002 Koblenz, Germany, retrieved in February 2015.

Hernandez-Ramirez, G., J. L. Hatfield, T. B. Parkin, T. J. Sauer, and J. H. Prueger (2011), Carbon dioxide fluxes in corn-soybean rotation in the midwestern U.S.: Inter- and intra-annual variations, and biophysical controls, *Agric. For. Meteorol.*, *151*(12), 1831–1842, doi:10.1016/j.agrformet.2011.07.017.

Hollinger, D. Y., et al. (2010), Albedo estimates for land surface models and support for a new paradigm based on foliage nitrogen concentration, *Global Change Biol.*, *16*(2), 696–710, doi:10.1111/j.1365-2486.2009.02028.x.

Hopmans, J. W., D. R. Nielsen, and K. L. Bristow (2002), How useful are small-scale soil hydraulic property measurements for large-scale vadose zone modeling? in *Environmental Mechanics: Water, Mass and Energy Transfer in the Biosphere: The Philip Volume*, vol. 129, edited by P. A. C. Raats, D. Smiles, and A. W. Warrick, pp. 247–258, AGU, Washington, D. C.

Jaramillo, F., and G. Destouni (2015), Local flow regulation and irrigation raise global human water consumption and footprint, *Science*, *350*(6265), 1248–1251.

Jaramillo, F., C. Prieto, S. W. Lyon, and G. Destouni (2013), Multimethod assessment of evapotranspiration shifts due to non-irrigated agricultural development in Sweden, *J. Hydrol.*, *484*, 55–62.

Kumar, P. (1999), A multiple scale state-space model for characterizing subgrid scale variability of near-surface soil moisture, *IEEE Trans. Geosci. Remote Sens.*, *37*, 182–197.

Lahoz, W. A., and G. J. De Lannoy (2014), Closing the gaps in our knowledge of the hydrological cycle over land: Conceptual problems, *Surv. Geophys.*, *35*(3), 623–660.

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.

Li, J., and S. Islam (1999), On the estimation of soil moisture profile and surface fluxes partitioning from sequential assimilation of surface layer soil moisture, *J. Hydrol.*, *220*(1), 86–103.

Mittelbach, H., and S. I. Seneviratne (2012), A new perspective on the spatio-temporal variability of soil moisture: Temporal dynamics versus time-invariant contributions, *Hydrol. Earth Syst. Sci.*, *16*(7), 2169–2179.

Morel-Seytoux, H. J., P. D. Meyer, M. Nachabe, J. Tourna, M. V. Genuchten, and R. J. Lenhard (1996), Parameter equivalence for the Brooks-Corey and van Genuchten soil characteristics: Preserving the effective capillary drive, *Water Resour. Res.*, *32*, 1251–1258, doi:10.1029/96WR00069.

Nachtergaele, F., et al. (2008), Harmonized world soil database. Food and Agriculture Organization of the United Nations.

Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I—A discussion of principles, *J. Hydrol.*, *10*(3), 282–290.

National Centers for Environmental Prediction (NCEP) (2004), North American Regional Reanalysis: A long-term, consistent, high-resolution climate dataset for the North American domain, as a major improvement upon the earlier global reanalysis datasets in both resolution and accuracy, Fedor Mesinger et. al, submitted to BAMS (2004), NARR data provided by the NOAA/OAR/ESRL PSD, Boulder, Colo., [Available at http://www.esrl.noaa.gov/psd/.]

Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC) (2011), MODIS subsetted land products, collection 5. Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC), Oak Ridge, Tenn. [Avaialble at http://daac.ornl.gov/MODIS/modis.html> (accessed 5.10.2015).]

Peters, E. B., and J. P. McFadden (2012), Continuous measurements of net CO_2 exchange by vegetation and soils in a suburban landscape, *J. Geophys. Res.*, *117*, G03005, doi:10.1029/2011JG001933.

Qin, J., K. Yang, N. Lu, Y. Chen, L. Zhao, and M. Han (2013), Spatial upscaling of in-situ soil moisture measurements based on MODIS-derived apparent thermal inertia, *Remote Sens. Environ.*, *138*, 1–9.

Rankinen, K., Ø. Kaste, and D. Butterfield (2004a), Adaptation of the integrated nitrogen model for catchments (INCA) to seasonally snow-covered catchments, *Hydrol. Earth Syst. Sci.*, *8*, 695–705.

Rankinen, K., T. Karvone, and D. Butterfield (2004b), A simple model for predicting soil temperature in snow-covered and seasonally frozen soil: Model description and testing, *Hydrol. Earth Syst. Sci.*, *8*, 706–716.

Rawls, W. J., D. L. Brakensiek, and K. E. Saxton (1982), Estimation of soil water properties, *Trans. Am. Soc. Agric. Eng.*, *25*, 1316–1320.

Rodriguez-Iturbe, I., D. Entekhabi, and R. L. Bras (1991), Nonlinear dynamics of soil moisture at climate scales 1. Stochastic analysis, *Water Resour. Res.*, *27*, 1899–1906, doi:10.1029/91WR01035.

Russo, D. (1998), Stochastic analysis of flow and transport in unsaturated heterogeneous porous formation: Effects of variability in water saturation, *Water Resour. Res.*, *34*, 569–581, doi:10.1029/97WR03619.

Saxton, K. E., W. Rawls, J. S. Romberger, and R. I. Papendick (1986), Estimating generalized soil-water characteristics from texture, *Soil Sci. Soc. Am. J.*, *50*, 1031–1036.

Schmidt, A., C. Hanson, J. Kathilankal, and B. E. Law (2011), Classification and assessment of turbulent fluxes above ecosystems in North-America with self-organizing feature map networks, *Agric. For. Meteorol.*, *151*(4), 508–520, doi:10.1016/j.agrformet.2010.12.009.

Schmidt, R., et al. (2006), GRACE observations of changes in continental water storage, *Global Planet. Change*, *50*(1), 112–126.

Schneider, U., A. Becker P. Finger, A. Meyer-Christoffer, B. Rudolf, and M. Ziese (2011), *GPCC Full Data Reanalysis Version 6.0 at 0.5: Monthly Land-Surface Precipitation From Rain-Gauges Built on GTS-Based and Historic Data*, Global Precipitation Climatology Centre, doi:10.5676/DWD_GPCC.FD_M_V6_050.

Schwatke, C., D. Dettmering, W. Bosch, and F. Seitz (2015), DAHITI—An innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry, *Hydrol. Earth Syst. Sci.*, *19*(10), 4345–4364.

Scott, C. A., W. G. Bastiaanssen, and M. U. D. Ahmad (2003), Mapping root zone soil moisture using remotely sensed optical imagery, *J. Irrig. Drain. Eng.*, *129*(5), 326–335.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010), Investigating soil moisture–climate interactions in a changing climate: A review, *Earth Sci. Rev.*, *99*, 125–161.

Swenson, S. C. (2012), *GRACE Monthly Land Water Mass Grids NETCDF RELEASE 5.0. Ver. 5.0*, PO.DAAC, CA, USA. doi:10.5067/TELND-NC005.

Swenson, S., and J. Wahr (2006), Post-processing removal of correlated errors in GRACE data, *Geophys. Res. Lett.*, *33*, L08402, doi:10.1029/2005GL025285.

Teuling, A. J., and P. A. Troch (2005), Improved understanding of soil moisture variability dynamics, *Geophys. Res. Lett.*, *32*, L05404, doi:10.1029/2004GL021935.

United States Geological Survey (USGS) (2015), U.S. Geological Survey, Department of the Interior/USGS, Groundwater level data retrieved from. [Available at http://waterdata.usgs.gov/nwis/gw in February 2015.]

van Genuchten, M. T. (1980), A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.*, *44*, 892–898.

Verrot, L., and G. Destouni (2015), Screening variability and change of soil moisture under wide-ranging climate conditions: Snow dynamics effects, *Ambio*, *44*(1), 6–16, doi:10.1007/s13280-014-0583-y.

Wang, X., H. Xie, H. Guan, and X. Zhou (2007), Different responses of MODIS-derived NDVI to root-zone soil moisture in semi-arid and humid regions, *J. Hydrol.*, *340*(1), 12–24.

Western, A. W., R. B. Grayson, G. Blöschl, G. R. Willgoose, and T. A. McMahon (1999), Observed spatial organization of soil moisture and its relation to terrain indices, *Water Resour. Res.*, *35*(3), 797–810, doi:10.1029/1998WR900065.

World Meteorological Organization (2014), *Composition of the WMO*, World Meteorological Organization, Geneva 2, Switzerland, WMO-OMM-No5, CH-1211.

Xia, Y., et al. (2012), Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products, *J. Geophys. Res.*, *117*, D03109, doi:10.1029/2011JD016048.

Xiao, J., et al. (2010), A continuous measure of gross primary production for the conterminous United States derived from MODIS and AmeriFlux data, *Remote Sens. Environ.*, *114*(3), 576–591, doi:10.1016/j.rse.2009.10.013.

Zhu, Q., and H. Lin (2011), Influences of soil, terrain, and crop growth on soil moisture variation from transect to farm scales, *Geoderma*, *163*(1), 45–54.