The importance of vegetation to understand terrestrial water storage variations

Tina Trautmann[1,2], Sujan Koirala[1], Nuno Carvalhais[1,3], Andreas Güntner[4,5], Martin Jung[1]

[1]Max-Planck Institute for Biogeochemistry, Jena, 07745, Germany
[2]International Max Planck Research School for Global Biogeochemical Cycles, Jena, 07745, Germany
[3]Universidade Nova de Lisboa, Caparica, 2829-516, Portugal
[4]German Research Centre for Geoscience, Potsdam, 14473, Germany
[5]University of Potsdam, Potsdam, 14476, Germany

Correspondence to: Tina Trautmann (ttraut@bgc-jena.mpg.de)

Abstract. So far, various studies aimed at decomposing the integrated terrestrial water storage variations observed by satellite gravimetry (GRACE, GRACE-FO) with the help of large-scale hydrological models. While the results of the storage decomposition depend on model structure, little attention has been given to the impact of the way how vegetation is represented in these models. Although vegetation structure and activity represent the crucial link between water, carbon and energy cycles, their representation in large-scale hydrological models remains a major source of uncertainty. At the same time, the increasing availability and quality of Earth observation-based vegetation data provide valuable information with good prospects for improving model simulations and gaining better insights into the role of vegetation within the global water cycle.

In this study, we use observation-based vegetation information such as vegetation indices and rooting depths for spatializing the parameters of a simple global hydrological model to define infiltration, root water uptake and transpiration processes. The parameters are further constrained by considering observations of terrestrial water storage anomalies (TWS), soil moisture, evapotranspiration (ET) and gridded runoff (Q) estimates in a multi-criteria calibration approach. We assess the implications of including varying vegetation characteristics on the simulation results, with a particular focus on the partitioning between water storage components. To isolate the effect of vegetation, we compare a model experiment ~~with~~in which vegetation parameters ~~varying~~vary in space and time to a baseline experiment in which all parameters are calibrated as static, globally uniform values.

Both experiments show good overall performance, but explicitly including varying vegetation data ~~led~~leads to even better performance and more physically plausible parameter values. Largest improvements regarding TWS and ET ~~were~~are seen in supply-limited (semi-arid) regions and in the tropics, whereas Q simulations improve mainly in northern latitudes. While the total fluxes and storages are similar, accounting for vegetation substantially changes the contributions of ~~snow and~~ different soil water storage components to the TWS variations~~, with the dominance of~~. This suggests an ~~intermediate water pool that interacts with the fast plant accessible soil moisture and the delayed water storage. The findings indicate the~~ important role of the representation of vegetation in hydrological models for interpreting TWS variations. Our simulations further indicate a

major effect of deeper moisture storages ~~as well as~~and groundwater-soil moisture-vegetation interactions as a key to understanding TWS variations. We highlight the need for further observations to identify the adequate model structure rather than only model parameters for a reasonable representation and interpretation of vegetation-water interactions.

1 Introduction

Since 2002 the Gravity Recovery and Climate Experiment (GRACE) mission has facilitated global monitoring of terrestrial water storage (TWS) variations from space – a milestone of global hydrology (Rodell 2004, Famiglietti and Rodell 2013). Observed TWS variations from GRACE have since become a cornerstone for diagnosing trends in water resources due to climate change or anthropogenic activities (Rodell et al. 2018, Reager et al. 2015, Scanlon et al. 2018, Syed et al. 2009, Tapley et al. 2019), as well as for benchmarking and improving global hydrological models (GHMs) (Scanlon et al. 2016, Döll et al. 2014, Werth et al. 2009, Zhang et al. 2017, Kumar et al. 2016, Eicker et al. 2014). Significant co-variations between GRACE-TWS and the global land carbon sink (Humphrey et al. 2018) and surface temperatures (Humphrey et al. 2021) highlight the importance of the water-cycle as nexus in the Earth System.

However, GRACE TWS estimates represent a vertically integrated signal of all water stored in snow, ice, soil, surface and groundwater. Thus, understanding processes and mechanisms of TWS variations requires attribution of TWS variations to individual storage components. Despite advancements in remote sensing, large-scale quantification of these components based on observations remains challenging. For example, remote sensing-based estimates of soil moisture only capture depths up to 5 cm and do not necessarily reflect the moisture availability in the deeper soil column (Dorigo et al. 2015). While these observations can be extrapolated to derive estimates of root zone moisture, either by using statistical relationships (Zhuang et al. 2020) or by data assimilation into land surface models (Reichle et al. 2017, Martens et al. 2017), such products rely on many assumptions. Therefore, GHMs have been necessary to interpret TWS variations in terms of contributions by snow, soil moisture, ground or surface water. However, several studies suggested that current state-of-the-art GHMs cannot reproduce key patterns of observed TWS variations and show partly diverging TWS partitioning (Scanlon et al. 2018, Schellekens et al. 2017, Zhang et al. 2017, Kraft et al. 2021). This uncertainty of the available tools to interpret TWS variations is clearly a major obstacle for diagnosing and understanding global changes of the water cycle~~, which is increased by differing model structures and grown complexity of existing GHMs~~.

~~Among previous studies, Trautmann et al. (2018) showed that a simple large-scale hydrological model that is calibrated in a multi-criteria fashion against multiple global hydrological data streams simultaneously yields very good model performance compared to state-of-the art GHMs. This study contributed important~~ To improve model performance and reliability, GHMs are traditionally calibrated against measured discharge time series at the outlet of catchments (Müller Schmied et al. 2021, Telteu et al. 2021). However, discharge provides an integrated response of an entire catchment with very limited information on the interplay of different processes and spatial heterogeneities. In fact, the use of spatio-temporal data, e.g., from remote sensing, for model calibration has been suggested (Su et al, 2020). While using spatio-temporal vegetation data, e.g., NDVI, seemed promising for this at the catchment scale (Ruiz-Perez et al. 2017), many GHMs still have a limited usage of such data in their modelling approach. Some large-scale studies have shown clear improvements in model performance when a larger number of observational constraints are used to constrain the model parameters, especially when using TWS variations from GRACE (e.g., Lo et al. 2010, Rakovec et al. 2016, Bai et al. 2018, Mostafaie et al. 2018, Trautmann, 2018). Among them, Trautmann et al. 2018 contributed insights in the drivers of TWS variations across spatial and temporal scales in northern high latitudes, in particular with respect to contributions by snow vs liquid water storages. In this study, we follow a similar framework of using multiple observational data streams to constrain a simple hydrological model to understand the ~~partitioning~~role of varying vegetation characteristics for the partitioning of TWS components~~, yet expand to~~ at global scale~~ and with further partitioning of liquid water storages~~.

Among liquid water storages, especially the differentiation between soil moisture and groundwater poses a challenge. Reflecting on the determinants of rather shallow soil moisture vs deeper groundwater storage variations, it is apparent that under most conditions the soil moisture state itself is the first order control valve. In particular, it determines the amount of water that is available for soil water uptake for evapotranspiration but also for percolation into deeper soil layers and consequently recharge into the groundwater storage. The two key processes that shape soil moisture dynamics, infiltration and evapotranspiration (ET), are strongly mediated by the presence and properties of vegetation (Wang et al. 2018). For example, vegetation promotes infiltration over surface runoff due to larger surface roughness, dampened precipitation intensities, and more soil macro pores due to rooting and biological activity. In fact, such roles of vegetation in a global climate model were already envisioned and evaluated almost 4 decades ago (Rind, 1984). Besides, vegetation alters soil properties like soil texture and organic matter content. Such soil properties together with rooting depth control the size of the soil moisture reservoir that is available for ET, and how plants respond to drought stress conditions (Baldocchi et al. 2021, Yang et al. 2020). Furthermore, deep roots may connect to the groundwater and provide access to the deeper moisture storages, and thus have wider implications on the hydrological cycle. Rooting depth is species-specific and, in addition, determined by the infiltration depth and groundwater table depth, and thus has a high spatial heterogeneity both across the globe and at the local scale (Fan et al. 2017). The significance of interactions between vegetation and soil moisture are at the heart of ecohydrology ~~(Rodriguez-Iturbe et al., 2001) and have become evident in many theoretical and experimental studies. While many studies analysed effects of water availability on vegetation functioning (Porporato et al., 2004;Reyer et al., 2013;Wang et al., 2001;Yang et al., 2014), the inverse pathway of how vegetation properties influence dynamics of water pools and the partitioning of TWS in large scale models has received surprisingly little attention~~(Rodriguez-Iturbe et al. 2001) and have become evident in many theoretical and experimental studies. Many studies analyzed effects of water availability on vegetation functioning (Porporato et al. 2004, Reyer et al. 2013, Wang et al. 2001, Yang et al. 2014), and the effect of changing vegetation cover on ecosystem water consumption (Du et al. 2021). While large-scale hydrologic models usually apply simplified and static vegetation characteristics (Quevedo et al. 2008, Weiss et al. 2012, Telteu et al. 2021), spatio-temporal variations of vegetation pattern are vital for good predictions of available water resources (Andersen et al. 2010). On ecosystem scale, Xu et al. 2016 showed the advantage of accounting for different plant hydraulic traits in an ecosystem model. And on a global scale, Weiss et al. 2012, for instance, showed the positive influence on modelled evaporation when static vegetation characteristics are replaced by monthly LAI estimates in a climate model. However, how the representation of vegetation affects global water storages and in particular the partitioning of TWS in large-scale hydrological models has received surprisingly little attention so far.

Therefore, the objective of this study is to investigate the effect of vegetation-dependent parameterizations of key hydrological processes on TWS partitioning ~~in~~at the global scale using a multi-criteria model data fusion approach. The model, an expanded version of Trautmann et al. (2018), is a simple conceptual 4-pool water balance model. Model parameters are calibrated against TWS variations from GRACE (Wiese, 2015), ET from FLUXCOM (Jung et al. 2019), runoff from GRUN (Ghiggi et al. 2019) and ESA CCI soil moisture (Dorigo et al. 2017). We contrast two experiments which differ only with respect to how vegetation-related parameters are defined: 1) a baseline experiment with global uniform parameters, 2) a vegetation experiment where vegetation parameters vary in space and partly in time. In contrast to the traditional approach of spatializing vegetation parameters by plant functional types or land cover classes and keeping this a-priori parameterization fixed during model application, we take advantage of continuous information on few key properties that link vegetation and hydrological processes: 1) spatially distributed and time-varying active vegetation cover that influences transpiration demand and interception storage, 2) spatial pattern of soil water supply for transpiration via roots, and 3) spatially distributed and time-varying ~~the~~ influence of vegetation cover on infiltration and runoff generation. Specifically, we are addressing the following questions:

1) Where, when, and by how much are global hydrological simulations improved by spatially distributed and time varying vegetation parameters?
2) To what extent does the attribution and interpretation of TWS variations for individual storage components change when introducing spatial and temporal variation of vegetation parameters?

~~1.1~~2 Methods

In the first section we give a general overview on the design of this study ~~and it's spatial and temporal coverage~~. Subsequently, the used model and data streams as well as the calibration and evaluation approach are explained in more detail.

2.1~~.2~~ Overview

To assess the potential effect of including continuous information on vegetation, we compare two model variants that are based on the same conceptual structure: 1) a base model with static, globally uniform parameter values (**B**), and 2) a model variant that includes spatially (and temporally) varying vegetation characteristics by defining vegetation parameters as function of global data products (**VEG**). We additionally performed an experiment that discretizes vegetation parameters for distinct classes of plant functional types, similar to some other GHMs. This **PFT** experiment is explained and shown in S9. Forced with global climate-data, the parameters of each variant are calibrated for a spatial subset against multiple Earth -observation-based data. In the **B** experiment, the parameters themselves are calibrated and globally constant parameter values are obtained. While the optimized parameters implicitly account for the effect of the nearly ubiquitous presence of vegetation, they cannot represent effects of spatially or and temporally varying vegetation properties. In the **VEG** experiment, we describe vegetation related parameters as ~~the~~a linear ~~product of a calibrated scalar and~~ function of spatio-temporal varying vegetation variables~~.~~, i.e., we calibrate scalars representing the slopes of these functions. By calibrating the ~~scalar~~slope, we include the continuous pattern from the data, but ~~weight~~scale it to best fit ~~with~~the observational constraints. Hence, vegetation related parameters vary explicitly spatially and partly temporally.

Once the parameters are calibrated, the simulations for the whole domain (global) are used to evaluate the model performance at different spatial and temporal ~~scale~~scales. To finally delineate the effect of including varying vegetation ~~data~~characteristics on the composition of simulated ~~total water storage~~TWS across temporal (mean seasonal, inter-annual) and spatial (local grid scale, spatially aggregated) scales, we use the Impact Index as defined by Getirana et al. (2017).

The model is run on daily time steps at a 1°x1° latitude/longitude resolution, focusing on vegetated regions under primarily natural conditions. To avoid biases of the calibrated model parameters due to processes that are not represented in the model structure, we exclude grid cells with > 10% permanent snow and ice cover, > 50% water fraction, > 20% bare land surface and > 10% artificial land cover fraction. These grid cells are masked out using the Globland20 fractional landcover v2 (Chen et al. 2014). Additionally, we exclude regions with a large human influence, mainly related to groundwater extraction, on the trend in GRACE TWS variations (Rodell et al. 2018) (see Fig. 2). The final study area comprises 74% of global land area. All other data sets used in this study were resampled to the 1°x1° grid and subset to the same grid cells.

Due to the temporal coverage of forcing data and observational constraints, we calibrate the model for the period 01/2002-12/2014, while the global-scale model runs and analyses are performed for the period 03/2000-12/2014. Prior to each model run, all states are initialized by a 8-year spin-up period. The forcing for the spin-up period is assembled by randomly rearranging complete years of the forcing data.

2.2 Model Description

The conceptual hydrological model is forced by daily precipitation, air temperature and net radiation (Table 1). It includes a snow component (see Trautmann et al. (2018)), a 2-layer soil water storage (*wSoil*), a deep soil water storage (*wDeep*) and a delayed, slow water storage (*wSlow*). The schematic structure of the model is shown in Fig. 1 and calibration parameters are explained in Table 2.

Depending on air temperature (*Tair*), precipitation (*Precip*) is partitioned into snow fall (*Snow*), that accumulates in the snow storage (*wSnow*), and rain fall (*Rain*), that partly is retained in an interception storage. Interception throughfall together with snow melt are ~~distributed among soil through~~partitioned into infiltration and infiltration excess depending on the ratio of actual soil moisture and maximum soil water capacity following Bergström 1995:

$$I_{exc} = I_{in} \cdot \left[\frac{\sum_{l=1}^{2} wSoil(l)}{\sum_{l=1}^{2} wSoil_{\max(l)}} \right]^{p_{berg}} \quad (1)$$

where, I_{exc} is the infiltration excess, I_{In} is the incoming water from throughfall and snow melt, $wSoil(l)$ is the soil moisture and $wSoil_{max(l)}$ the maximum soil water capacity of each soil layer l, and p_{berg} is a ~~global~~ calibration parameter. While p_{berg} <1 allocates a small fraction of the incoming water to the soil water pool even if it is nearly empty, p_{berg} >1 allows a large fraction of incoming water to infiltrate into the soil when soil saturation is already high, and p_{berg} = 1 describes a linear relationship between soil water saturation and the amount of incoming water that infiltrates.

~~Part~~A fraction of the infiltration excess (defined by the global calibration parameter rf_{Slow}) then replenishes a delayed water storage (*wSlow*~~),~~) that acts as a linear reservoir and generates slow runoff (Q_{slow}). The remaining infiltration excess represents fast direct runoff (Q_{fast}). Q_{fast} and Q_{slow} together represent total runoff Q, that flows out of the system, i.e., grid cell. Infiltrated water is distributed among 2 soil layers following a top-to-bottom approach, where the maximum capacity of the first soil layer is prescribed as 4 mm, in order to match the tentative depth of satellite soil moisture observations, while the storage capacity of the 2nd soil layer is a calibration parameter ($wSoil_{max(2)}$). The 2nd soil layer is connected with a deeper water storage (*wDeep*). The size of *wDeep* is defined as a multiple of $wSoil_{max(2)}$ by the calibrated scaling parameter s_{deep} ~~(Eq. 2).~~. Depending on the moisture gradient between the two storages, water either percolates from the 2nd soil layer to the deeper soil, or it rises from the deeper storage into the 2nd soil layer, ~~limited~~by scaling to a maximum flux rate~~.~~

$$\cancel{f_{pot} = f_{max} \cdot w_{grad}} \quad \cancel{(2)}$$

~~Where, f_{pot} is~~ (defined by the ~~potential flux between both layers, f_{max} is the maximum flux rate (~~global calibration parameter~~)~~ ~~and w_{grad} is the gradient of moisture calculated as~~

$$\cancel{w_{grad} = \frac{wDeep}{s_{deep} \cdot wSoil_{\max(2)}} - \frac{wSoil(2)}{wSoil_{\max(2)}}} \quad \cancel{(3)}$$

~~where s_{deep} is the scaling parameter to derive the maximum capacity of *wDeep* as multiple of $wSoil_{max(2)}$ (calibration parameter).~~

, f_{max}). The deeper storage therefore acts as a storage buffer~~,~~ that linearly discharges further to the delayed water storage (*wSlow*)~~. The *wSlow*~~, which also receives part of the infiltration excess~~, is thus representative of all delayed storage components~~.

Evapotranspiration (ET) is represented by a demand-supply approach that is driven by a potential ET demand following Priestley-Taylor, and is limited by the available soil moisture supply. ~~The~~ ET is partitioned into interception evaporation (E_{Int}), bare soil evaporation from the first soil layer (E_{Soil}) and plant transpiration from the two soil layers (E_{Transp}). Interception and plant transpiration are only calculated for the vegetated fraction of each grid cell, while bare soil evaporation is limited to the non-vegetated fraction of each grid cell.

$$\cancel{E_{Soil} = min(E_{Soil-sup}, E_{Soil-dem})} \qquad \cancel{(4)}$$

$$\cancel{E_{Soil-dem} = PET.(1 - p_{veg})} \qquad \cancel{(5)}$$

$$\cancel{E_{Soil-sup} = wSoil(1).k_{Soil}} \qquad \cancel{(6)}$$

~~where, *PET* is the potential ET,~~ ~~p_{veg} is the vegetation fraction of each grid cell (calibration parameter), and~~ ~~k_{Soil} is the proportion of the first soil layer available for evaporation (calibration parameter).~~

~~Similarly, transpiration is calculated as~~

$$\cancel{E_{Transp} = min(E_{Transp-sup}, E_{Transp-dem})} \qquad \cancel{(7)}$$

$$\cancel{E_{Transp-dem} = PET.\alpha_{veg}.p_{veg}} \qquad \cancel{(8)}$$

$$\cancel{E_{Transp-sup} = \sum_{l=1}^{2} wSoil(l).k_{Transp}} \qquad \cancel{(9)}$$

~~where,~~ ~~α_{veg} is the alpha coefficient of the Priestley-Taylor formula (calibration parameter) and~~ ~~k_{Transp} is the proportion of soil water available for transpiration (calibration parameter). In supply limited conditions,~~ ~~k_{Transp} effectively acts as a decay parameter that defines the depletion of the soil water storage via transpiration.~~

While water in *wSoil* is directly available for ET, *wDeep* is only indirectly accessible by capillary rise, and the water stored in *wSlow* is not plant-accessible. Total water storage is the sum of all water storages, including *wSnow*, *wSoil*, *wDeep* and *wSlow*. Although groundwater and surface water storages are not implemented explicitly, they are effectively included in *wDeep* and *wSlow*, especially after calibration of associated storage parameters against GRACE TWS.

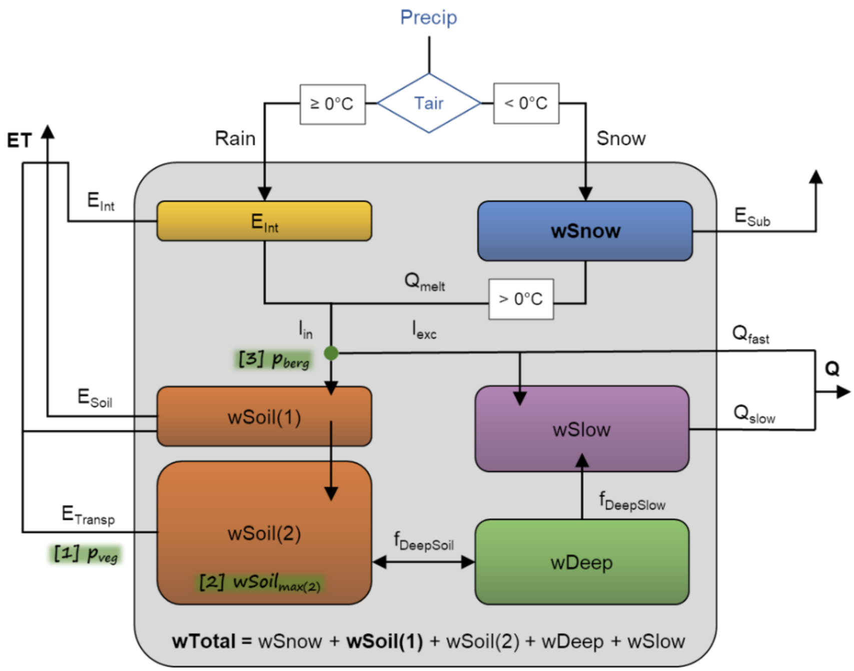


Figure 1 Schematic of the underlying model structure, with blue font denoting forcing data: *Precip* = precipitation, *Tair* = air temperature. Boxes represent states: E_{int} = interception storage, *wSnow* = snow water storage, *wSoil(1)* = upper soil layer, *wSoil(2)* = second soil layer, *wDeep* = deep water storage, and *wSlow* = slowly varying water storage. Arrows denote fluxes: *Rain* = rain fall, *Snow* = snow fall, E_{Sub} = sublimation, Q_{melt} = snow melt, I_{in} = incoming water from throughfall and snow melt, I_{exc} = infiltration excess, Q_{fast} = fast direct runoff, Q_{slow} = slow runoff, *Q* = total runoff, E_{Int} = evaporation from interception storage, E_{Soil} = soil evaporation, E_{Transp} = plant transpiration, *ET* = total evpotranspiration, $f_{DeepSoil}$ = flux between *wSoil* and *wDeep* (percolation resp. capillary rise), $f_{DeepSlow}$ = flux from *wDeep* to *wSlow*. Bold print highlights model variables that are constrained in the calibration. Green highlights show where vegetation influence is included explicitly: [1] the parameter p_{veg} to define each grid cell's vegetation fraction, [2] the parameter $wSoil_{max(2)}$ that defines the maximum plant available soil water, and [3] the parameter p_{berg} to define the infiltration and runoff generation partitioning.

Table 1 Data used for model forcing, for description of vegetation characteristics and for model calibration.

	Product	Space	Time	Data Uncertainty	Reference
Forcing					
Precip	GPCP 1dd v1.2	global	daily		Huffmann et al. 2000
Tair	CRUNCEP v6	global	daily		Vivoy et al. 2015
Rn	CERES Ed4A	global	daily		Wielicki et al. 1996
Vegetation Characteristics					
EVI	based on MCD43C1 v6 (MODIS daily BDRF), calculated via MODIS standard EVI formula		daily climatology		Schaaf & Wang 2015
RD1	maximum rooting depth		static		Fan et al. 2017
RD2	effective rooting depth		static		Yang et al. 2016
RD3	maximum soil water storage capacity		static		Wang-Erlandson et al. 2016
RD4	maximum plant available water		static		Tian et al. 2019

	capacity				
Calibration					
TWS	GRACE mascon RL06	global	monthly	with product	Wiese et al. 2018
wSoil	ESA CCI SM v4.04 (combined product)	~global	daily	with product	Dorigo et al. 2017
ET	FLUXCOM RS ensemble	global	daily	with product	Jung et al. 2018
Q	GRUN v1	global	monthly	~ 50%	Ghiggi et al. 2019

2.~~2.1 Including~~3 Vegetation Characteristics

We include three aspects of vegetation influence on hydrological processes: 1) the specific transpiration demand by vegetation, 2) the soil water supply for transpiration via roots, and 3) the influence of vegetation on infiltration and runoff generation. These three aspects are controlled by three corresponding model parameters, namely the grid cell's ~~(active)~~ vegetation fraction (p_{veg}), the maximum plant available soil water (~~$wSoil_{max}$2),~~$wSoil_{max(2)}$), and the runoff generation/infiltration coefficient (p_{berg}). In the **VEG** experiment, scalar parameters are used as linear multipliers of observation-based spatio-temporal patterns~~. This step is considered necessary~~ to ~~handle the differences among different datasets, while still harvesting~~harvest the information of spatial and temporal patterns from the ~~observations. Alternatively, the scalars can be interpreted as scaling factors or weights of different observational~~continuous data ~~streams. By calibrating the scalars, the weight given to observational pattern is constrained by the data streams used for model calibration.~~products.

2.~~2.1~~3.1 Vegetation Fraction

The parameter p_{veg} reflects the~~ active~~ vegetation cover of each grid cell that influences the grid's interception storage, transpiration demand, and partitioning of evapotranspiration components. To describe its spatial and seasonal variations, we include the mean seasonal cycle (MSC) of the Enhanced Vegetation Index (EVI). Therefore, p_{veg} at time step is defined as linear function of EVI, where s_{EVI} is the calibrated scaling parameter:

$$p_{veg} = s_{EVI} \cdot EVI$$ (~~10~~2)

with $0 \leq p_{veg} \leq 1$.

EVI data is calculated via the MODIS standard formula (Didan & Barreto-Munoz) using the daily BRDF, nadir BRDF adjusted reflectance values MCD43C1 v6 (Schaaf &Wang 2015) for the period 01.2001 – 12.2014:

$$EVI = 2.5 \frac{NIR - Red}{NIR + 6 \cdot Red - 7.5 \cdot Blue + 1}$$ ~~(11~~ (3)

Since the daily EVI time series are not continuous due to noise and missing values during cloudy conditions, snow and darkness, the data was preprocessed to be used in the model. For each grid cell, we calculate the median seasonal cycle, fill long gaps during winter time with a low value, interpolate missing values, and smooth the time series. ~~Therefor~~Therefore, winter is defined as days with negative net radiation and gaps are considered long when 10 consecutive days of EVI data ~~is~~are missing. The winter time gaps are filled with the 5th percentile of available winter time data. The remaining missing values are linearly interpolated and finally the resulting seasonal cycle is smoothed by a local regression with weighted linear least squares and a 1st order polynomial model.

2.~~2.1~~3.2 Plant available Soil Water

In order to determine the soil water supply for transpiration as a function of vegetation, we define the maximum soil water capacity of the 2nd soil layer $wSoil_{max(2)}$ based on ~~either~~ rooting depth ~~or~~and soil water storage capacity data. We include the maximum rooting depth by Fan et al. (2017) (RD1), effective rooting depth by Yang et al. (2016) (RD2), maximum soil water capacity by Wang-Erlandsson et al. (2016) (RD3) and maximum plant accessible water capacity by Tian et al. (2019) (RD4). Due to our definition of $wSoil_{max(2)}$ as maximum plant accessible water, all four data are, theoretically, suitable when focusing on spatial patterns. Practically, though, they vary in their definition, underlying approaches, spatial coverage and derived spatial pattern. The RD1 and RD2 are based on principles of vegetation optimality and plant ~~adaption~~adaptation, and RD3 and RD4 are based on a water-balance perspective but using Earth ~~-~~observations and/or data assimilation techniques. Therefore, we employ an approach in which ~~the weight~~we obtain a linear combination of the four products where the weights of each ~~data is~~product are calibrated~~, and their calibrated values are necessarily constrained by either the ET or TWS data. The maximum soil water capacity is therefore calculated as~~ during the multi-criteria parameter optimization:

$$\cancel{wSoil_{max\,(2)} = \sum_{d=1}^{4} s_{RD(d)} \cdot RD(d)} \qquad \cancel{(12)}$$

~~where $RD(d)$ is the data from each data stream d and $s_{RD(d)}$ are the corresponding scaling factors that are calibrated specifically for each RD data. As RD4 from Tian et al. (2019) is only available for arid to moderately humid vegetated land area and excludes tropical forests (Tian et al., 2019), resulting gaps in the study area are filled by the calibration parameter $wSoil_{max(RD4)}$ prior to scaling RD4.~~

$$wSoil_{max\,(2)} = \sum_{d=1}^{4} s_{RD(d)} \cdot RD(d) \qquad (4)$$

where $RD(d)$ is the data from each data stream d and $s_{RD(d)}$ are the corresponding scaling factors that are calibrated. As RD4 from Tian et al. (2019) is only available for arid to moderately humid vegetated land area and excludes tropical forests (Tian et al. 2019), resulting gaps in the study area are filled by the calibration parameter $wSoil_{max(RD4)}$ prior to scaling RD4.

2.~~2.1~~3.3 Runoff/Infiltration Coefficient

Finally, vegetation structure also affects the infiltration and runoff generation process as it alters the surface and sub-surface characteristics. To reflect this influence, we describe the infiltration/runoff parameter p_{berg} (Eq. 1) as linear function of vegetation fraction p_{veg}:

$$p_{berg} = s_{berg} \cancel{\cdot p_{veg}} \cdot p_{veg} \qquad (\cancel{13}5)$$

where s_{berg} is the calibrated scaling parameter.

2.4 Model Calibration

In order to keep computational costs low and to avoid overfitting of model parameters, we perform model calibration for a ~~spatial~~ subset of ~~selected~~904 (8%) grid cells~~ that include only 8%~~. Since model parameters are expected to vary much more in space than in time (between years), and due to the rather short time period of ~~the total study area. The~~available constraints, we build two subsets of data for calibration and validation data in the spatial domain rather than in time (spatial split sample approach). Calibration grid cells are chosen by a stratified random sampling method that maintains the overall proportion of different climate and hydrological regimes defined by Köppen-Geiger climate regions (Kottek et al. 2006).

Since this study focuses on the impact of vegetation and in order to keep the number of calibration parameters low, we do not optimize snow related parameters and ~~focus on the vegetation related parameters. Instead,~~use the optimized snow parameters from Trautmann et al. 2018~~ are used~~. This results in a total of 11 calibration parameters for the **B** model and a total of 16 parameters for the **VEG** model (Table 2).

In order to constrain different aspects of the water cycle, we use a multi-criteria calibration approach similar to Trautmann et al. 2018. The parameters of each model variant are simultaneously optimized against multiple observational constraints, including monthly TWS anomalies from GRACE (Wiese et al. 2018), ESA CCI Soil Moisture (Dorigo et al. 2017), evapotranspiration estimates from FLUXCOM-RS ensemble (Jung et al. 2019) and gridded runoff from GRUN (Ghiggi et al. 2019) (Table 1).

~~For each of the observational constraints, we calculate a cost term that considers the data's specific strengths and uncertainties, as well as each grid cell's area. Thereby~~When using observational data sets from several sources, it is essential to consider possible inconsistencies between them that arise from their respective characteristics and uncertainties (Zeng et al. 2015, Zeng et al. 2020). Therefore, we derived the monthly water (im)balance of the observations following a similar approach as Rodell et al. 2015 (see S10). Although we did not find major systematic inconsistencies at the global scale, we take into account each data set's characteristics and uncertainties in model calibration via the cost term at the grid cell level. To this end, we only use grid cells and time steps with available observations, which vary for the different data streams. To retrieve one cost term per observational constraint, we concatenate the ~~timeseries~~time series of all grid cells into a single vector for which costs are calculated. The individual cost terms are considered to have the full weight of 1, resulting in a total cost value ($cost_{total}$) as ~~a~~the sum of individual ~~cost~~costs. The total cost is then minimized during the optimization process using a global search algorithm, the Covariance Matrix Evolutionary Strategy (CMAES) algorithm (Hansen and Kern, 2004).

$$cost_{total} = \cancel{\sum_{ds=1}^{nds} cost(ds)}$$

$$\cancel{(14}\sum_{ds=1}^{nds} cost(ds) \qquad (6)$$

where, *cost(ds)* is the cost for each data stream *ds*. For ~~wTWS~~TWS, ET and Q, the cost terms are based on the weighted Nash ~~Sutcliff~~Sutcliffe Efficiency (Nash and Sutcliffe, 1970), which explicitly considers the observational uncertainty σ:

$$\cancel{cost = \frac{\sum_{i=1}^{n} \frac{(x_{obs,i}-x_{mod,i})^2}{\sigma i}}{\sum_{i=1}^{n} \frac{(x_{obs,i}-\bar{x}_{obs})^2}{\sigma_i}}} \qquad \cancel{(15)}$$

$$cost = \frac{\sum_{i=1}^{n} \frac{(x_{obs,i}-x_{mod,i})^2}{\sigma i}}{\sum_{i=1}^{n} \frac{(x_{obs,i}-\bar{x}_{obs})^2}{\sigma_i}} \qquad (7)$$

where $x_{mod,i}$ is the modelled variable, $x_{obs,i}$ is the observed variable, $\bar{x}_{obs}$ is the average of x_{obs}, and σ_i is the uncertainty of x_{obs} of each data point *i*. The cost criterion reflects the overall fit in terms of variances and biases, with an optimal value of 0 and a range from 0-∞.

Owing to the larger uncertainties of Q_{obs} on inter-annual scales (Ghiggi et al. 2019), we only use the monthly mean seasonal cycle, while for the other variables, full monthly time series were used.

To define σ of ET_{obs}, we utilize the median absolute deviation of the FLUXCOM-RS ensemble. For Q_{obs}, we assume an average uncertainty of 50% based on values reported in Ghiggi et al. (2019). For TWS_{obs}, the spatially and temporally varying uncertainty information provided with the GRACE data is used. Besides, the largest monthly values of ~~wTWS~~$_{\text{~~obs~~}}$ TWS_{obs} (< -500 mm and > 500 mm) were masked out to avoid the effect of outliers on optimization results. Note that these outliers represent less than 0.5~~%,~~% of the data, and are mainly located in coastal arctic regions, and are, thus, potentially related to land and sea-ice and/or leakage from neighboring grid cells over ocean. Before calculating $cos\cancel{t_{wTWS}}t_{TWS}$, the monthly means of observed and modelled ~~wTWS~~TWS are respectively removed to calculate anomalies over a common time period 01.01.2002–31.12.2012.

Since remote sensing-based soil moisture only captures the top few centimeters of soil depth, usually about 5 cm, $cost_{wSoil}$ is calculated based on the modelled soil moisture in the first soil layer. As the combined ESA CCI soil moisture imposes absolute values and ranges from GLDAS-Noah (Dorigo et al. 2015), we use Pearson's correlation coefficient as $cost_{wSoil,}$ and focus on soil moisture dynamics that is most reflective of the original remote sensing observation. Only estimates from 01.01.2007 onwards are considered, as data before that period ~~is~~are sparse. Further, $cost_{wSoil}$ is calculated from the monthly ~~average~~averaged values to ~~avoid large variations and potentially~~circumvent the large noise in the daily data. Thereby, only months with observations available for at least 10 days are considered. Due to snow cover, the temporal coverage of the product decreases with increasing latitude. Therefore, to prevent a bias towards northern summer months, we also exclude grid cells that lack more than 40% of monthly estimates. After ~~these~~ filtering for missing data, monthly surface soil moisture time series for 56% of the total study area and 51% of the calibration ~~grids~~grid cells are available.

2.5 Model Evaluation and Analysis

For model evaluation, we contrast the optimized parameter values and their uncertainties. The relative uncertainty in the optimized parameter vector is estimated by quantifying each parameter's standard error according to Omlin and Reichert (1999) and Draper and Smith (1981), similar to Trautmann et al. (2018).

For each experiment, the optimized parameter sets are used to produce model simulations for the ~~entire~~global study area. Their performances are then evaluated using Pearson's correlation coefficient and the uncertainty weighted Nash-Sutcliff efficiency (*wNSE*) for ~~wTWS~~TWS, ET and Q observations (Eq. ~~16~~8). The performances are evaluated on local (for each grid cell individually), regional and on global scales.

$$wNSE = 1 - \frac{\sum_{i=1}^{n} \frac{(x_{obs,i} - x_{mod,i})^2}{\sigma i}}{\sum_{i=1}^{n} \frac{(x_{obs,i} - \bar{x}_{obs})^2}{\sigma_i}} \quad (16)$$

$$wNSE = 1 - \frac{\sum_{i=1}^{n} \frac{(x_{obs,i} - x_{mod,i})^2}{\sigma i}}{\sum_{i=1}^{n} \frac{(x_{obs,i} - \bar{x}_{obs})^2}{\sigma_i}} \quad (8)$$

For the regional analysis, we derive 5 hydroclimatic regions by performing a cluster analysis using the spatiotemporal characteristics of ~~wTWS~~TWS, ET and Q observations, as well as each grid cell's latitude. By that, each zone is characterized by similar seasonal dynamics and amplitudes of the ~~constraints~~water cycle variables, allowing for a better comparison ~~when evaluating~~of regional averages than e.g., the commonly used Köppen-Geiger regions which lump regions with very different amplitudes and phasing of the water cycle variables. The resulting regions are shown in Fig. 2. Region 1 comprises the snow dominated northern latitudes (*Cold*), while region 2 includes the moderate mid latitudes (*Temperate*). Very humid and mostly tropical regions are combined in region 3 (*Humid*). Region 4 is characterized by a distinct rain season (*Sub-humid*), while region 5 includes semi-arid areas in low latitudes (*Semi-arid*). Although we hereafter use these hydroclimatic cluster regions for model evaluation, the same analysis for Köppen-Geiger climate zones is presented in S11 to facilitate comparison with other studies.

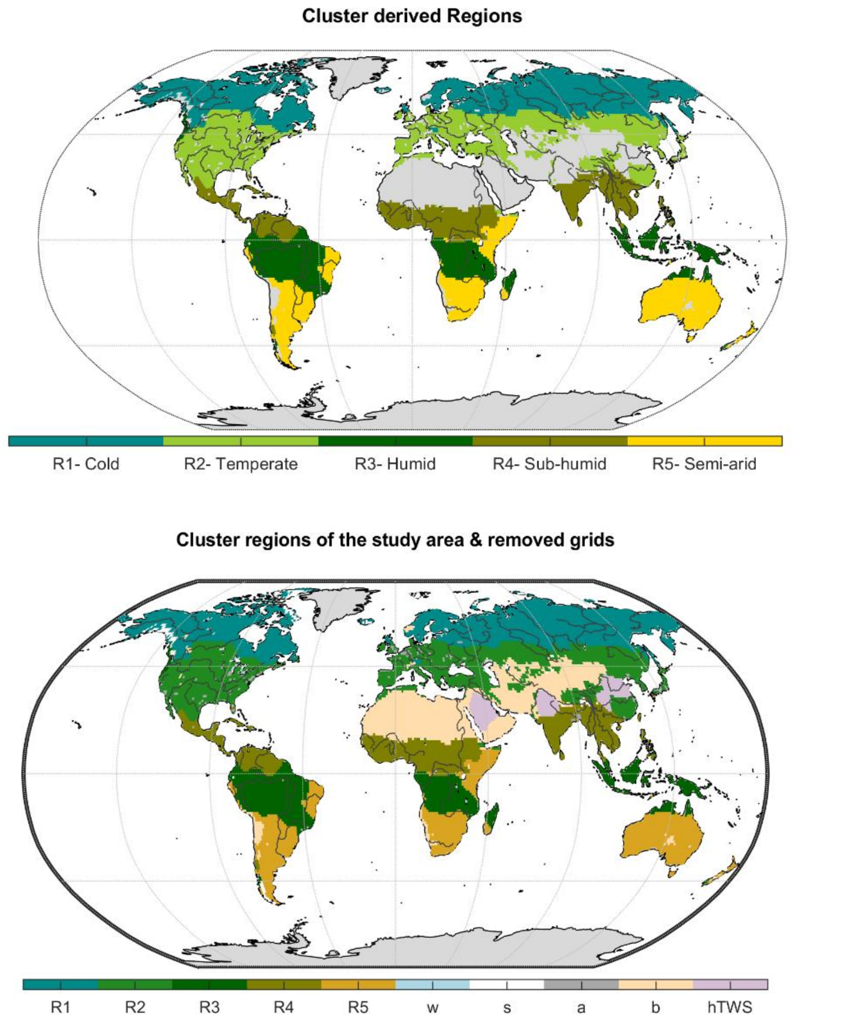


Figure 2 Hydroclimatic cluster regions of the study area (R1 - Cold, R2 - Temperate, R3 - Humid, R4 - Sub-humid, R5 - Semi-arid), and grid cells that have been excluded from this study (w = water fraction >50%, s = permanent snow and ice cover > 10%, a = artificial land cover fraction > 10%, b = bare land surface > 20%, hTWS = direct human impact on the trend in GRACE TWS).

Finally, we assess the contributions of the four water storage components, *wSnow*, *wSoil*, *wDeep* and *wSlow*, to seasonal and inter-annual variations of the total water storage across spatial scales, i.e.~~.~~, the local grid cell, the regional and the global average. To do so, we apply the Impact Index I following Getirana et al. (2017). The metric describes the contribution C of each water storage s as the sum of its absolute monthly anomaly:

$$C_s = \cancel{\sum_{t=1}^{nt} |s_t - \bar{s}|} \qquad \cancel{(17}\sum_{t=1}^{nt} |s_t - \bar{s}| \qquad (9)$$

Where, $\bar{s}$ is the average storage of the timesteps t-nt, with nt = 12 for mean seasonal and nt = 178 for inter-annual dynamics. The Impact Index I_s is then defined as the ratio of each water storage component contribution C_s to the total contributions from all storage components:

$$I_s = \frac{C_s}{\sum_{s=1}^{n} C_s} \qquad (\cancel{18}10)$$

The value of I_s range from 0-1, with 0 indicating no impact and 1 indicating full control of all variations.

3 Results

In the following section we first evaluate ~~the performance of~~ both calibrated model variants by comparing ~~the~~their calibrated model parameters and ~~validate~~by comparing modelled ~~wTWS~~TWS, ET and Q against observations at global, regional and local scale. Subsequently, we show the contribution of individual storage components to TWS variability for **B** and **VEG** on different spatial and temporal scales.

3.1 Model Evaluation

3.1.1 Calibrated Parameters

Table 2 summarizes the calibrated parameters and their uncertainties for the **B** and **VEG** model experiments. Overall, including varying vegetation characteristics leads to more plausible parameter values after calibration, while in **B** several parameters hit their ~~expected~~prescribed bounds. Furthermore, very high parameter uncertainties present in **B**, ~~indicating unconstrained~~that indicate poorly constrained values, could be ~~drastically~~strongly reduced in **VEG** ~~and the interactions among different parameters are reduced with fewer parameters showing correlation~~ (S3).

For **B**, p_{veg} suggests that on average only 37% of each grid cell are covered with vegetation globally. ~~Despite having a good performance of modelled ET, the~~This low vegetation fraction is ~~relatively low, but transpiration is increased~~counteracted by a high α_{veg} value (2.25), which is much higher than commonly used alpha coefficient of the Priestley-Taylor equation of ~~1.2 or 1.26 (Lu et al., 2005). At the same time, a very low fraction of the first layer soil layer is available for evaporation, as~~ ~~k_{Soil}~~

~~hits the~~around 1.2 (Lu et al. 2005) to yield good performance of modelled ET (Fig. 3). At the same time, a very low fraction of the first soil layer is available for soil evaporation, as k_{Soil} hits its lower bound of 10%. Besides, the parameters controlling the drainage from deep and slow water storage (d_{Deep}, d_{Slow}) are high resulting in a fast drainage, and effectively ~~discarding the hydrological influences~~discard any influence of these water pools.

For **VEG**, the median vegetation fraction is 73%, leading to a more realistic fraction of soil moisture being available for evaporation (k_{Soil} = 0.4), which is similar to the ~~mode~~modal value of 0.33 reported by McColl et al. (2017), and a more realistic α_{veg} value of 0.92, that effectively leads to the median Priestley-Taylor alpha coefficient of 0.81 (S2). In comparison to **B**, the resulting $wSoil_{max(2)}$ of **VEG** ~~is~~ with a median value of 52 mm is considerably lower. Its spatial pattern mainly originates from RD3 (Wang-Erlandsson et al. 2016) and RD4 (Tian et al. 2019) data, while RD1 (Fan et al. 2017) contributes only little and RD2 (Yang et al. 2016) data is negligible. The resulting spatial patterns of the maximum soil water capacity from the combination of all datasets (S2) are yet consistent with those from other estimates and patterns of rooting depth (e.g., Schenk and Jackson (2005)). We note here that the soil water capacity data are favoured over the rooting depth data. This agrees with Küçük et al. (2020), who suggest that estimating plant storage capacity based on Earth observation data may be more suitable than those using optimality principles. ~~Besides, $wSoil_{max(2)}$ is defined as maximum plant accessible soil storage, but also used to describe the maximum soil water storage. Due to their retrieval, the rooting depth data RD1 and RD2 suggest deep roots in warm/dry regions, while shallow roots would be sufficient in wet conditions. Therefore, RD1 and RD2's spatial pattern with deep roots in dry/warm regions and shallow roots in wet conditions behaves anti-proportional to the actual (maximum) soil water storage aimed for in this study.~~ Related to the limited size of *wSoil*, calibration enforces a deeper and a slow water storage with reasonable depletion parameters (d_{Deep}, d_{Slow}~~). Overall, in **VEG**, only the k_{Transp}~~) in order to match observed TWS variations. By that, the considerable low $wSoil_{max(2)}$ parameter is counteracted by refilling *wDeep*, which indirectly provides plant accessible water via capillary rise. Likewise, k_{Transp}, which describes the fraction of the 2nd soil layer that is available for transpiration, is relatively high, as a larger fraction of the small soil water storage needs to transpire to match observed ET. Hence, calibrated k_{Transp} is higher than ~~reported~~empirical values of ET decay between 0.02- 0.08 ~~(Teuling et al., 2006).~~, that are based on assuming 1 soil water pool (Teuling et al. 2006).

Table 2 Calibrated model parameters, their description, range and calibrated values for experiments B and VEG. Red fonts highlight calibrated values at the predefined parameter bounds.

Parameter	Description	Units	Default Value	Range	Calibrated Values ± Uncertainty (%) B	VEG
Vegetation Fraction						
p_{veg}	active vegetation fraction of the grid cell		0.5	0.3 - 1	0.37 ± 0.05	
s_{EVI}	scaling parameter to derive active vegetation fraction from EVI data		1	0 - 5		3.89 ± 0.05
Evapotranspiration						

p_{Int}	interception storage	mm	1	0 - 10	1.0	± 0.08	0.6	± 0.02
k_{Soil}	fraction of 1st soil layer available for evaporation		0.5	0.1 - 0.95	0.1	± 0.01	0.4	± 0.08
α_{veg}	alpha parameter of the Priestley-Taylor equation		1	0.2 - 3	2.25	± 0.15	0.92	± 0.00
k_{Transp}	fraction of soil water available for transpiration		0.02	0 - 1	0.12	± 0.32	0.48	± 1.76
Infiltration/Runoff								
p_{berg}	runoff-infiltration coefficient		1.1	0.1 - 5	1.32	± 0.02		
s_{berg}	scaling parameter to derive the runoff-infiltration coefficient from p_{veg}		3	0.1 - 10			3.08	± 0.02
Soil Moisture								
$wSoil_{max\,(2)}$	maximum (available) water capacity of the 2nd soil layer	mm	300	10 - 1000	752	± 0.02		
$s_{RD(1)}$	weight to include maximum rooting depth by Fan et al. 2017		0.05	0 - 5			0.01	± 0.00
$s_{RD(2)}$	weight to include effective rooting depth by Yang et al. 2016		0.05	0 - 5			0.00	± 0.00
$s_{RD(3)}$	weight to include maximum soil water storage capacity by Wang-Erlandson et al. 2016		0.05	0 - 5			0.15	± 0.06
$s_{RD(4)}$	weight to include plant available water capacity by Tian et al. 2019		0.05	0 - 5			0.15	± 0.07
$wSoil_{max\,(RD4)}$	maximum (available) water capacity of the 2nd soil layer for grids with missing estimates in Tian et al. 2019	mm	50	0 - 1000			145	± 0.08
Deep Soil								
s_{deep}	scaling parameter to derive the maximum deep soil storage from $wSoil_{max\,(2)}$		0.5	0 - 50	9.1	± 461317	5.6	± 0.21
f_{max}	maximum flux rate between deep soil and the 2nd soil layer	mm d^{-1}	10	0 - 20	1.5	± 0.00	5.1	± 0.01
d_{Deep}	depletion coefficient from deep soil to delayed water storage		0.5	0 - 1	1.0	± 5.61	0.01	± 0.00
Delayed Water Storage								
rf_{Slow}	recharge fraction of infiltration excess into delayed water storage		0.5	0 - 1	0.78	± 1.72	0.68	± 0.01
d_{Slow}	depletion coefficient from delayed water storage to slow runoff		0.01	0 - 1	1.0	± 2329	0.02	± 0.03

3.1.2 Model Performance

Table 3 contrasts the overall model performance metrics for ~~wTWS~~TWS, ET and Q for the two experiments for the calibration subset of 8% grid cells (*opti*) and the entire study domain (*global*). The metrics are calculated in the same way as during optimization, i.e., by concatenation of the ~~timeseries~~time series of all ~~grids~~grid cells into a single vector for which statistics are calculated. In general, the differences between *opti* and *global*, as well as between **B** and **VEG** are marginal. For **VEG**, results mainly improve for ~~wTWS~~TWS, and slightly for ET. Although the models were only calibrated for the spatial subset in *opti*, equally good or even better performances are obtained when the calibrated parameters are applied over the entire study domain. This suggests that the calibration subset was representative of the entire study domain and the calibration did not overfit ~~the~~ model parameters.

Among the variables, the best model performances in terms of *wNSE* and *corr* is obtained for ET. While the correlation between observed and simulated ~~wTWS~~TWS is high, the overall *wNSE* is relatively low, which mainly results from higher uncertainties in TWS_{obs} and a larger variance error, likely originating from grid cells with low observed TWS variance.

Table 3 Overall model performance metrics in terms of weighted Nash-Sutcliff efficiency (wNSE) and Pearson's correlation coefficient (corr) of total water storage (~~wTWS~~TWS), evapotranspiration (ET) and runoff (Q) in B and VEG experiments for the calibration subset (opti) and the entire study domain (global).

	~~wTWS~~TWS				**ET**				**Q**			
	wNSE		***corr***		***wNSE***		***corr***		***wNSE (MSC)***		***corr (MSC)***	
	opti	global	opti	global	opti	global	opti	global	opti	global	opti	global
B	0.33	0.33	0.69	0.69	0.97	0.97	0.90	0.90	0.63	0.63	0.86	0.86
VEG	0.38	0.41	0.71	0.72	0.98	0.98	0.90	0.91	0.60	0.57	0.85	0.85

Similar to the global metrics, the average mean seasonal cycle of different regions shows an equally good or slightly better performance of **VEG** compared to **B** regarding all variables (Fig. ~~2~~3). At regional scale (Fig. 4), the general pattern of grid-wise Pearson correlation is similar for both experiments. However, the difference between the correlation coefficients highlights an improvement using **VEG** for a large proportion of grid cells, and regarding all ~~variables~~TWS, ET, and Q (indicated by brown color). ~~In the next section, regional and local performances are explained for wTWS, ET and finally Q.~~

For ~~wTWS~~TWS, the amplitude at the global scale is well-captured, yet with a phase difference of ~1 month in both model variants, ~~which~~where both ~~lead the~~model variants show an earlier timing of peak storage (Fig. 3). The phase shift is also apparent in the *Temperate* and *Cold* regions, while the seasonal dynamics in *Sub-humid* and *Humid* region is captured well, yet with an underestimation of the amplitude. Though differences are small, **VEG** obtains higher correlation ~~and a smaller bias~~ except for the *Semi-arid* region. At local scale, correlation with GRACE TWS is lowest in rather semi-arid grid cells (Fig. 4), where ~~wTWS~~TWS variation is low~~ and it has lower impact on the global *wNSE*, which determines the cost in model calibration.~~. However, including spatial ~~pattern~~patterns of vegetation improves ~~wTWS~~TWS mainly in these (semi-)arid regions.

Regarding ET, both experiments reproduce seasonal dynamics in all regions quite well, yet tend to underestimate ET in the *Semi-arid*, *Sub-humid* and *Humid* regions, especially in months with low ET (Fig. 3). At grid-scale (Fig. 4), correlation of

ET is very high, except for tropical regions~~.~~ due to low seasonality. Compared to **B**, **VEG** improves correlation here, as well as in some (semi-)arid regions such as the Sahel zone and the Western US.

In contrast to ET, performance for Q is generally the best in regions with poorer model performance in terms of ET (*Semi-arid*, *Sub-humid* and *Humid* regions) (Fig. 3), suggesting a trade-off between the two different observation data streams~~.~~, i.e. the inability of matching both observed fluxes simultaneously. Nonetheless, including varying vegetation characteristics improves peak runoff in all regions and reduces the underestimation of Q especially in the *Cold* region. While the improvement of Q simulations in Northern latitudes gets even more obvious at grid-scale, **B** shows higher correlation with observations in Africa and the Mediterranean (Fig. 4).

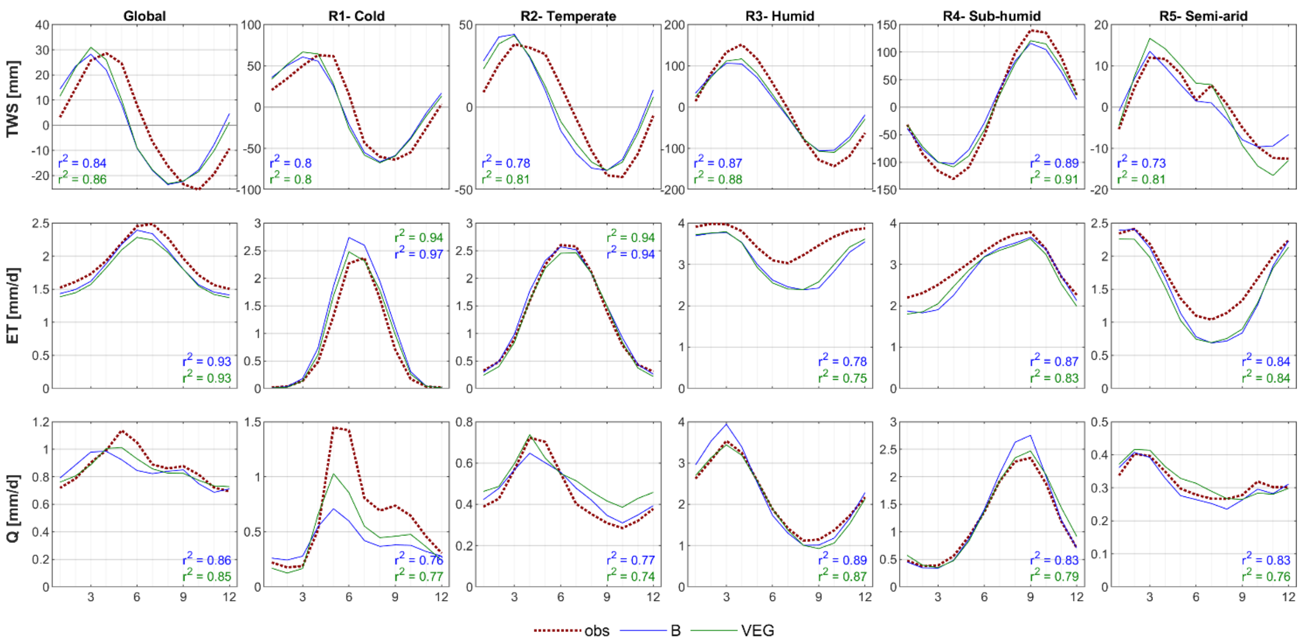
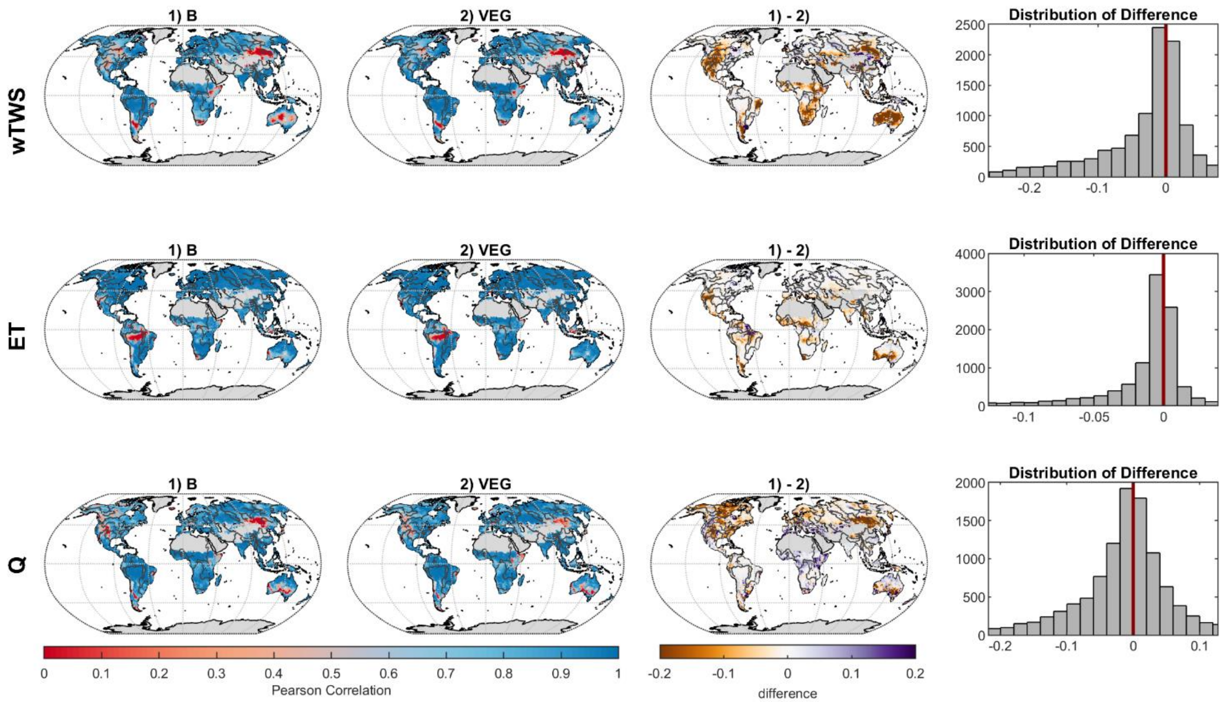


Figure 3 Global and regional mean seasonal cycles of total water storage (~~wTWS~~TWS), evapotranspiration (ET) and runoff (Q) for the B and VEG experiments compared to the observational constraints by GRACE (~~wTWS~~TWS), FLUXCOM (ET) and GRUN (Q).


1) B
2) VEG
1) - 2)
Distribution of Difference
wTWS
ET
Q
Pearson Correlation
difference
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
-0.2 -0.1 0 0.1 0.2

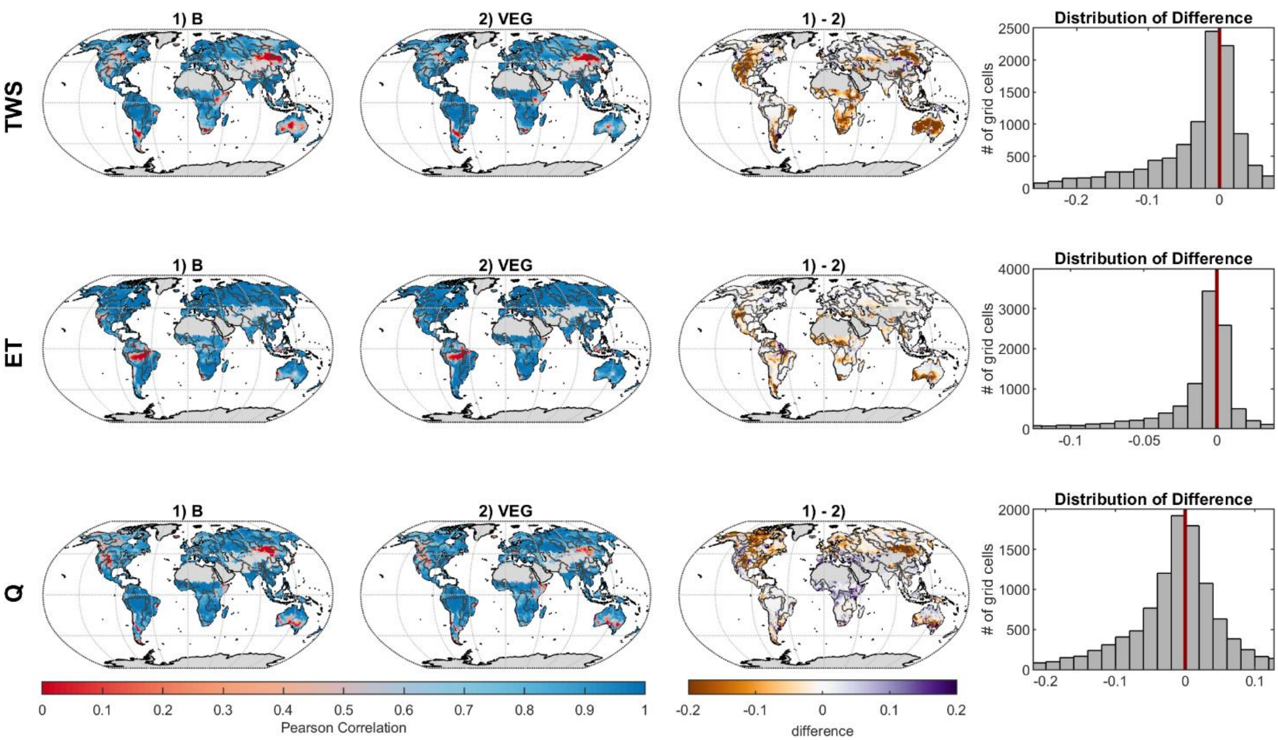

1) B
2) VEG
1) - 2)
Distribution of Difference
# of grid cells
TWS
ET
Q
Pearson Correlation
difference
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
-0.2 -0.1 0 0.1 0.2

Figure 4 Grid-wise Pearson's correlation coefficient for total water storage (~~wTWS~~TWS), evapotranspiration (ET) and runoff (Q) between 1) observations and B, and 2) observations and VEG, as well as differences between 1) and 2) (brown color, i.e., negative values indicate higher correlations for VEG, while purple color, i.e., positive values indicate better correlation values for B).

3.2 ~~Contribution~~ Importance of varying Vegetation Properties to TWS Variability

In this section, we present the influences of vegetation on TWS partitioning into snow (*wSnow*), plant-accessible soil moisture (*wSoil*), not directly plant-accessible deep soil water (*wDeep*) and non-plant-accessible slow water storages (*wSlow*) at different spatial and temporal scales. We first focus on mean seasonal dynamics and continue with the contribution of each component to inter-annual TWS variability at local grid-cell and regional scales, respectively, before presenting the analysis at the global scale.

3.2.1 Local & Regional Scale

Figure 5 shows the contribution of individual water storages to mean seasonal TWS variations at local grid-scale. For both **B** and **VEG**, *wSnow* has the highest impact in Northern latitudes and high altitudes where snow fall occurs regularly. Locally, the contribution of liquid water increases gradually with decreasing latitude and, finally, causes all TWS variations south of ~45° N. Within the liquid water storages, **B** attributes nearly all variations to directly plant accessible soil moisture *wSoil*, with an average of 76% over all grid cells. While showing a similar pattern of increasing contribution towards lower latitudes, the **VEG** experiment only has an average of 17% contribution from *wSoil*. Instead, most variations (40%) are due to variability in the deeper soil storage, *wDeep*. Besides, the average impact of slow water storages *wSlow* (20%) is comparable to that of *wSnow* (22%) in **VEG**, though it is spatially much more limited to tropical regions, such as the Amazon basin.

Mean seasonal dynamics averaged globally and for different regions are shown in Fig. 6. As indicated by the grid-scale results, *wSnow* dominates TWS variations in the northern *Cold* region (73% in **B**, resp. 69% in **VEG**), and plays a considerable role in the *Temperate* region (28% resp. 26%). For the other regions, **B** attributes nearly all remaining variability to *wSoil*, while in **VEG** *wDeep* has the highest Impact Index (59% in *Semi-arid*, 50% in *Sub-humid* and 43% in *Humid*).

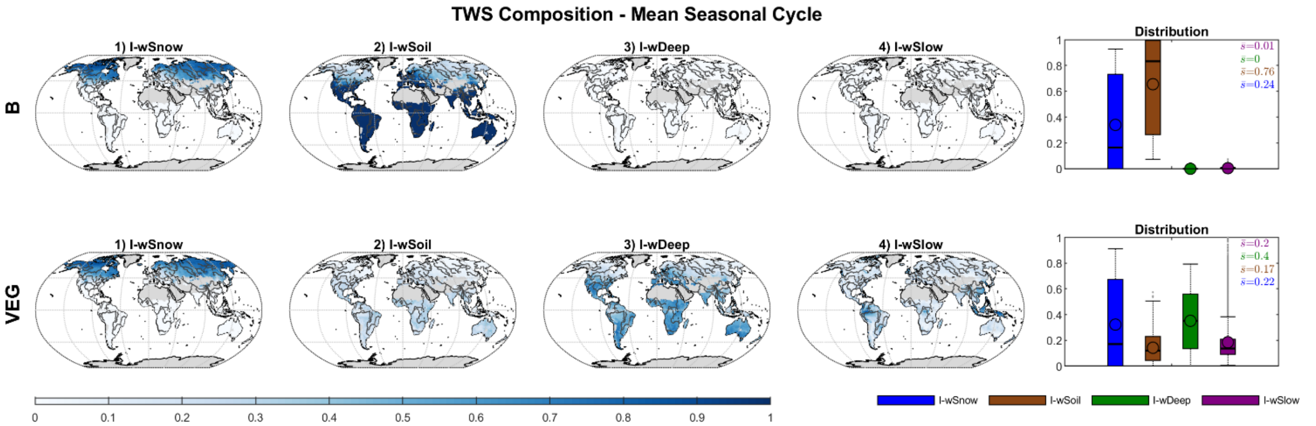


Figure 5 Global distribution of the Impact Index *I* for the contribution of simulated snow (*wSnow*), soil (*wSoil*), deep water storage (*wDeep*) and delayed water storage (*wSlow*) to the mean seasonal cycle of total water storage, for B and VEG.

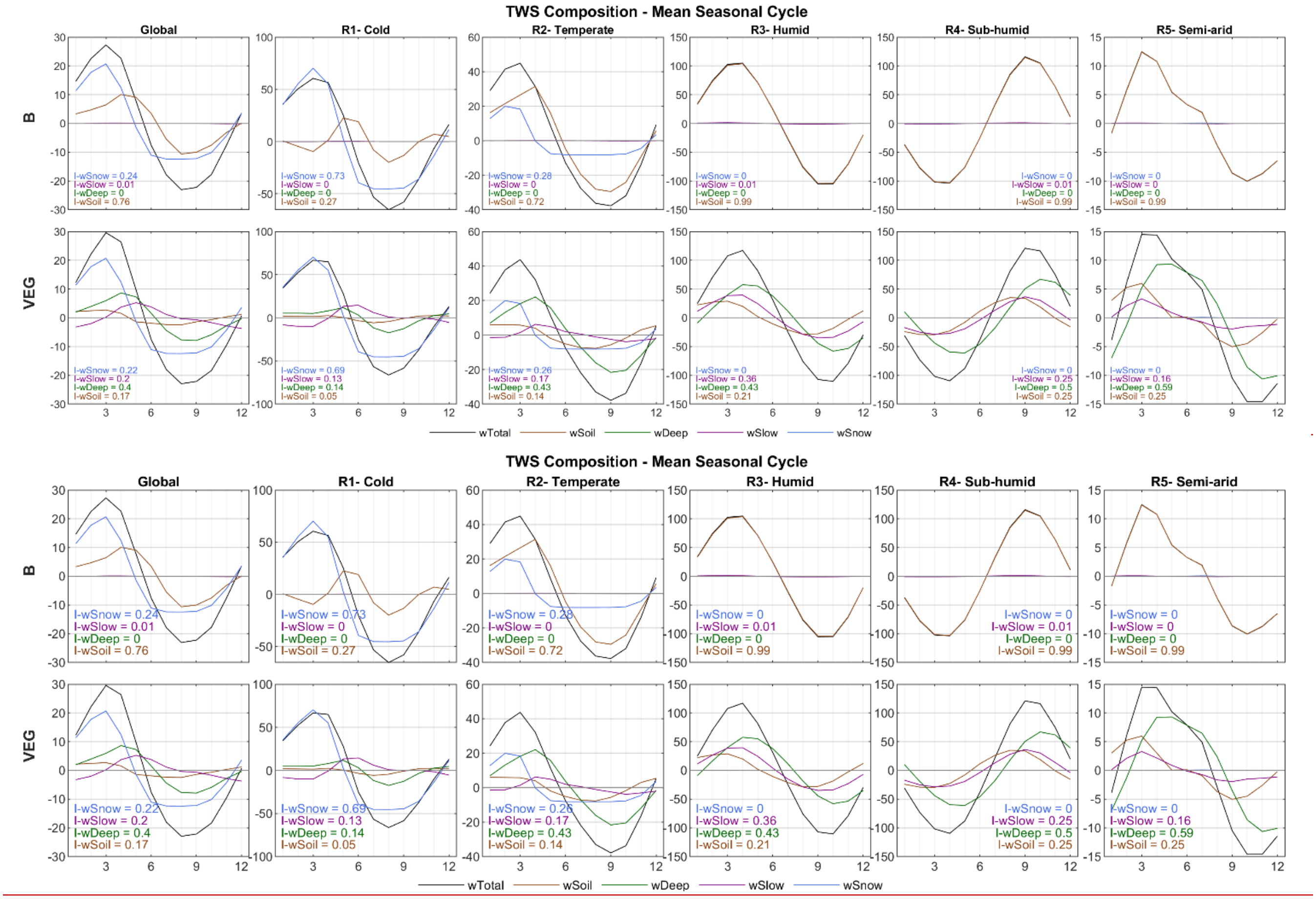


Figure 6 Global and regional average mean seasonal cycles of simulated total water storage and its components for B and VEG, including the regional Impact Index *I* for each storage.

At the inter-annual scales, the impact of *wSnow* decreases to 10% (**B**) respectively 12% (**VEG**) locally (Fig. 7). For most of the grid cells, all inter-annual TWS variations are caused by *wSoil* in **B**. In **VEG** however, the deeper soil layer *wDeep* is again the most important storage, with an average Impact Index of 53% for all grid cells. The contribution of *wSoil* and *wSlow* remain more or less the same as those for seasonal TWS variations.

Average contributions for different regions and globally (S4) show again that, in **B,** nearly all inter-annual ~~wTWS~~TWS variability is caused by *wSoil* (87-99%). Only in the *Cold* region, the impact of *wSoil* decreases to 69% in the favor of *wSnow* (31%). Similar to the local scale, in **VEG,** *wDeep* explains > 50% of ~~wTWS~~TWS variability in most regions~~,~~. Only in the *Cold* region, the contribution of *wDeep* is similar to *wSnow* (39% vs. 38%). The contribution of *wSoil* ranges from 9% (*Cold*) to 19% (*Semi-arid*), while the impact of *wSlow* is between 16-18% in most regions and increases in *Sub-humid* (24%) and *Humid* (34%) regions.

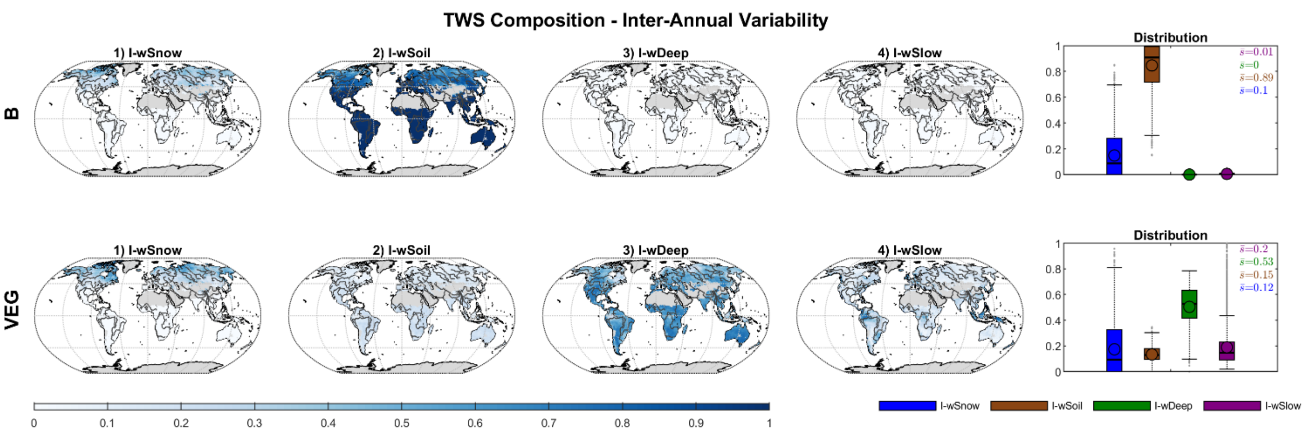


Figure 7 Global distribution of the Impact Index *I* for the contribution of simulated snow (*wSnow*), soil (*wSoil*), deep water storage (*wDeep*) and delayed water storage (*wSlow*) to the inter-annual variability of total water storage, for B and VEG.

3.2.2 Global Scale

Finally, Fig. 8 contrasts the impact of water storage components to the total storage, in **B** and **VEG**, at the global scale. As with the local and regional scales, including varying vegetation characteristics differentiates the composition of global TWS variations drastically. In both experiments, *wSnow* clearly dominates the spatially aggregated mean seasonal cycle with an Impact Index of 71% (**B**) and ~~69~~61% (**VEG**). These contributions are considerably higher than the average local Impact Index over all grid cells (**B** 24%, **VEG** 22%; Fig. 5). As already seen at local scale, liquid water storages dominate the inter-annual TWS variability, whereby **B** and **VEG** differ in the attribution to different components of the liquid water storage. In **B,** all variations other than *wSnow* ~~originates~~originate from *wSoil*, but *wDeep* dominates in **VEG**. Especially at inter-annual scales, *wDeep* accounts for half of all TWS variations. In contrast to **B**, in **VEG**, *wSoil* only has a minor impact of 7% at seasonal and 13% at inter-annual scale. Instead, *wSlow* has a moderate contribution of 11% (mean seasonal) and. 17% (inter-annual). In contrast to the mean seasonal dynamics in which the dominating ~~storage~~storages are different at local and global scales, the inter-annual dynamics are consistent across scales with the same storage component dominating at both local and global scale (Fig. 5,7,8).

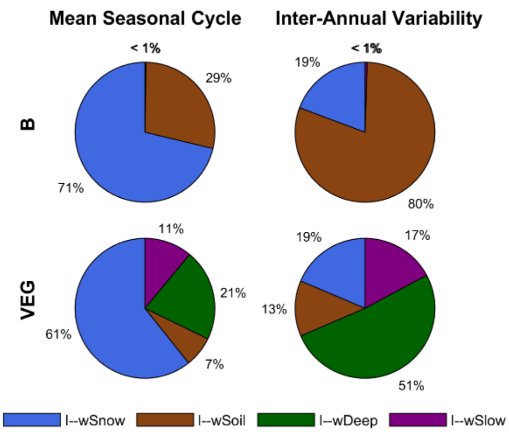


Figure 8 Impact Index *I* for the contribution of simulated snow (*wSnow*), soil (*wSoil*), deep water storage (*wDeep*) and delayed water storage (*wSlow*) to the global average mean seasonal cycle and inter-annual variability of total water storage, for B and VEG.

4 Discussion

In order to address the two main research questions of this study, the following section discusses the above shown differences between **B** and **VEG,** first regarding model performance and finally regarding the modelled partitioning of ~~wTWS~~TWS.

4.1 Model Performance

Both experiments ~~have~~show good performance against the observational constraints, and the differences between **B** and **VEG** are ~~marginal over~~relatively small at the global scale. However, there are ~~not only~~ systematic ~~spatial differences~~improvements for **VEG** at the regional and local scale, ~~but also~~and calibrated parameter values for **VEG** are more realistic and ~~with smaller interactions among each other (less equifinal) in the **VEG** experiment. The latter~~better constrained. This suggests a more realistic representation of fluxes and states in **VEG** ~~even though they are not directly constrained during the calibration process~~overall. Remaining discrepancies compared to observations can be associated with shortcomings and uncertainties in the observational data, as well as to the processes that are not represented in the rather simple model structure.

The differences in the seasonal phase of global ~~wTWS~~TWS in both model experiments mainly originate from the *Temperate* and *Cold* regions, and such model simulation differences have been reported previously (Döll et al. 2014, Schellekens et al. 2017, Trautmann et al. 2018). One of the potential reasons is the ~~intermediate~~temporary storage of melt water during spring in rivers and other surface water bodies, which occurs coherently over large areas in mid-to-high latitudes (Döll et al. 2014, Schellekens et al. 2017, Schmidt et al. 2008, Kim et al. 2009), and which delays the storage decay. In this context, also lateral water transport may additionally affect the ~~wTWS~~TWS variations in downstream grid cells. Yet, such processes and conditions are neither represented in **B** nor **VEG**.

Weaker performance of ~~wTWS~~TWS in (semi-) arid regions~~, on the one hand, originates from a low influence of these regions on global *wNSE*, because of~~ is likely mainly due to low observed TWS variations and a ~~high~~low signal-to-noise ratio (Scanlon et al. 2016). ~~Therefore~~Hence, less weight is also given to those grid cells in the cost component during calibration~~.~~ ~~On the other hand~~ due to their small variations. In addition, alteration by human activities like groundwater withdrawal, dams and irrigation to overcome the natural water shortage in such regions as North-East China and the American (Mid-)West can be regionally large in relative terms. While we aimed to exclude grid cells with large human impact a priori, we cannot completely exclude the influence of the aforementioned anthropogenic processes~~,~~ that are not explicitly represented in our model experiments. It should, however, be noted that the observational EVI data used in the **VEG** experiment do have an imprint ~~(~~of ~~the effects) of~~ e.g. irrigated agriculture ~~in terms of increased~~as the measured surface reflectance includes the higher vegetation activity~~. This may be associated with an~~ due to irrigation. The better representation of ET in semi-arid regions due to the EVI constraint contributes to the improved simulation of ~~wTWS~~TWS variations~~ in such regions~~ in the **VEG** experiment.

While overall ET performance is good, tropical regions ~~peak out with~~show low correlation. These areas are associated with higher uncertainties in the FLUXCOM ET estimates ~~due to issues with~~(Jung et al. 2019) due to underlying data uncertainties of the eddy covariance observations. Those uncertainties are related to poor station coverage and energy balance closure gap~~.~~, but also to issues of the satellite data inputs caused by cloud coverage. Nonetheless, including varying vegetation characteristics data improves simulated ET here, suggesting a better representation of the characteristic highly active vegetation compared to other regions and to global averages. Besides, **VEG** improves ET mainly in water supply-limited regions for the reasons already presented above for improved TWS performance in (semi-) arid regions.

The trade-off between the performances, ~~more specifically~~ in particular in terms of the bias~~,~~ of Q and ET~~ simulations suggest possible~~, suggests either larger uncertainties in one of the data streams for these regions, inconsistencies between the ET and Q constraints from independent sources~~. Additionally, the bias regarding either ET or Q, may relate~~, and/or model structure deficits. A small tendency to ~~shortcomings in~~ a negative water balance in the consistency checks of the observational data for these regions (S10) implies either underestimation of the precipitation forcing or overestimation of FLUXCOM ET or GRUN Q. Global precipitation datasets tend to underestimate precipitation (Trenberth et al. 2007, Contractor et al 2020) due to limitations of the satellite retrieval, gauge measurements and, if combined, the combination method (Fekete et al. 2004). Validation of the GPCD 1DD data used in this study showed an underestimation of precipitation in complex terrain and regionally during spring and autumn, while precipitation in winter time tends to be overestimated (Huffman et al. 2001). While we accounted for the latter by reducing snowfall (via a scaling parameter that ~~doesn't~~was calibrated in Trautmann et al. 2018), we don't consider potential underestimation in the rainfall forcing. Therefore, precipitation forcing may not provide sufficient water input ~~to support both outgoing water fluxes.~~for ET and Q in the model to achieve the magnitudes given by the observation-based products. Lastly, some ~~remaining~~ deterioration of performance of Q in **VEG** may originate from deficiencies in the GRUN product itself which was generated with climatic drivers only, disregarding information on spatio-temporal variations in vegetation (Ghiggi et al. 2019).

The improvement of Q in Northern latitudes is associated with the activation of the slow and delayed storage in the **VEG** experiment with spatial varying parameterization of soil water storage capacity. The relatively low storage capacity in these regions facilitates more fast saturation excess runoff. In addition, the slow storage represents better the runoff delay in surface water and rivers in these regions that results in improvements of low flow during winter as well as the increase of runoff during spring (Fig. 3). Such delayed runoff also improves the simulation of peak runoff in the *Sub-humid* and *Humid* regions.

The remaining deficiencies in model performance, especially in the *Cold* region, indicate missing processes in the simple model structure. ~~Such processes include freeze/thaw dynamics, permafrost and ice jam in river channels that would increase surface water storage and allow high spring flood. Besides, snow parameters have been calibrated against remote sensing-based GlobSnow Snow Water Equivalent that is known to saturate for deep snow conditions (Luojus et al., 2014) (see Trautmann et al. (2018)).~~Such processes include freeze/thaw dynamics and permafrost (Yu et al. 2020) as well as ice jam in river channels that would increase surface water storage and allow high spring flood (Kim et al. 2009). Besides, snow parameters have been calibrated against remote sensing-based GlobSnow Snow Water Equivalent that is known to saturate for deep snow conditions (Luojus et al. 2014) (see Trautmann et al. (2018)). Although the calibration process considered this shortcoming, an underestimation of modelled snow accumulation is possible – leading to an underestimation of peak snow pack in winter that would result in an underestimation of runoff due to lower snowmelt in spring.

While the **VEG** experiment presented here considers all 3 aspects of vegetation influences on hydrological processes explicitly (see section 2.2.1), we also run experiments that include ~~with including~~ these aspects separately into model calibration (not shown). These analyses found that the largest improvement was obtained when including soil water storage capacity as a function of rooting depth and storage capacity data, and a rather low impact when considering the runoff/infiltration partitioning as a function of vegetation fraction. This highlights the central role of soil water storages and the importance of adequately describing soil moisture pattern and dynamics in hydrological models.

4.2 Contribution to TWS Variability

~~Both model experiments agree with previous studies (Trautmann et al., 2018) in showing a dominating role of snow accumulation and depletion on global seasonal TWS variability, and the same of liquid water storage on inter-annual~~

~~scales.~~Albeit their global coverage, the above presented results agree with the previous regional study that focused on Northern mid-to-high latitudes (Trautmann et al. 2018). Similarly, both model experiments show a dominating role of snow accumulation and depletion on global seasonal TWS variability, whereas liquid water storages determine inter-annual TWS variations. At the same time, the contribution of individual storages to TWS variations differ at the local grid-scale compared to when they are averaged over a region or globally. ~~In **B**, all variations other than *wSnow* originates from directly plant accessible soil moisture, whereas, in **VEG**, the deeper soil storage *wDeep* becomes the most important. Therefore, including observation-based information on vegetation changes the attribution of TWS variations drastically, while the variations of total TWS themselves do not change significantly.~~

~~In the following, we briefly focus on the varying TWS composition among spatial scales, and finally discuss the systematic differences between **B** and **VEG**.~~

~~Differences among scales~~

~~Albeit their global coverage, the presented results agree with previous regional studies focusing on Northern mid-to-high latitudes (Trautmann et al., 2018). Snow dominates seasonal TWS variations locally and regionally in higher northern latitudes and altitudes (Güntner et al., 2007). Its stronger contribution~~The stronger contribution of snow on spatially aggregated signals can be explained by the spatial coherence of snow accumulation over larger areas. Liquid water storages, on the other hand, are more spatially heterogeneous, with increasing and decreasing dynamics across regions that cancel out and compensate each other when spatially aggregated (Trautmann et al. 2018, Jung et al. 2017). In contrast to the mean seasonal dynamics, the inter-annual Impact Indices of the storage components at the global scale are similar to the average local Impact Indices (Fig. 7 and Fig. 8). This suggests that at inter-annual time scales, there is no spatially coherent pattern of one single storage component that leads to higher accumulated Impact Indices than the local averages. However, while both experiments agree in the general pattern of the impact of snow versus liquid water storages, they systematically differ in the allocation of water among liquid storage compartments. In **B**, all variations other than *wSnow* originate from directly plant accessible soil moisture, whereas, in **VEG**, the deeper soil storage *wDeep* becomes the most important. Therefore, including observation-based information on vegetation changes the attribution of TWS variations drastically, while the variations of total TWS themselves do not change significantly.

~~Differences among model experiments~~

Differences in the composition of TWS variability between **B** and **VEG** are effectively reflected in the differences of calibrated parameters. In **B**, the directly plant accessible soil water storage is larger, due to a higher effective $wSoil_{max(2)}$, while delayed water storages are 'turned off' because of increased drainage (d_{Deep}, d_{Slow}), reducing the variations in *wDeep* and *wSlow*. Although **VEG** has been calibrated in the same way with the same observational constraints, calibrated model parameters differ as the included data on vegetation characteristics provides complementary information on spatial and temporal patterns. Therefore, the resulting calibrated parameters can be assumed to be more realistic. For example, they enable (delayed) longer-term water storage as well as capillary rise from the deeper soil water storage when the directly plant accessible storage dries out. Due to this process, TWS variations are mainly controlled by *wDeep* in **VEG**.

In detail, the ~~increasing~~increased importance of the indirect plant-accessible storage *wDeep* in **VEG** can be related to the limited maximum soil water capacity $wSoil_{max(2)}$ that is constrained by rooting depth/soil water capacity data, and to a higher k_{Transp} parameter. The smaller *wSoil* storage ~~forces~~increases percolation to *wDeep*, but the water is still available when needed due to the capillary rise from *wDeep* to *wSoil*.

Removing capillary flux from *wDeep* to *wSoil* in fact increases the contribution of *wSoil* to seasonal variability, while the impact of *wDeep* remains high on inter-annual scales (S7). ~~Thus, the question is whether the derived contributions to TWS variability are robust or an artefact of the model formulation.~~ While the contribution of capillary rise to total ET is < 20% for most grid cells, it becomes more important in arid-to-wet transition regions, e.g., sub-Saharan Sahel, Savannas, northern Australia and the Indian subcontinent (Fig. 9). These are regions with high precipitation seasonality, where vegetation often grows deep roots to access deep unsaturated zone storage and groundwater during the dry season. ~~The spatial patterns do not only agree with the findings of Koirala et al. (2014),~~The spatial patterns of ET supported by capillary rise agree with the findings of Koirala et al. (2014), who applied the physically-based model MATSIRO to investigate the effect of capillary flux to hydrological variables~~, but also with observations. For example, the spatial patterns are in line with the predicted probability of deep rooting by Schenk and Jackson (2005), and are supported by Tian et al.~~. The spatial patterns are also in line with the predicted probability of deep rooting by Schenk and Jackson (2005), and are supported by Tian et al. 2019 who found that vegetation remains active long into the dry season in Africa, suggesting that soil-deep soil/groundwater interaction plays a considerable role. Therefore, the spatial pattern of ~~where~~the interactions of *wDeep* ~~interacts~~ with *wSoil* ~~is~~in **VEG** seems reasonable~~, indirectly validating the~~ and our results ~~from **VEG**.~~indicate that capillary rise appears to be a process of large scale relevance.

While defined as 'fraction of soil water available for transpiration', k_{Transp} is ~~also~~ an effective decay parameter for the depletion of *wSoil* via transpiration processes~~.~~ under water limited conditions. Plausible values derived from eddy covariance observations of ET are in the order of 10^{-3} – $10^{-\cancel{2},1}$ (Teuling et al. 2006), similar in magnitude to delay coefficients for baseflow. By calibrating a model against GRACE TWS, it is difficult to decide whether water leaves the system slowly via ET or by Q~~. Additional including ET and Q data in model calibration should have ideally reduced equifinality. However, our results suggest that the model is still not constrained well enough:~~ especially during drydown periods. In **B,** k_{Transp} is much smaller than in **VEG** and more ~~plausible~~consistent with expected magnitudes, yet other slow depleting storages are effectively 'turned off'. In contrast, **VEG** with additional ~~vegetational~~vegetation data, simulates ~~a reasonable~~an important slow storage that contributes to Q~~, but~~ and also to soil moisture via capillary rise, and has a rather high calibrated k_{Transp}. ~~Fixing~~To better understand the implications of parameterising supply limited ET decay in the model we conducted another experiment where we fixed k_{Transp} in **VEG** to 0.05 ~~and optimizing~~(about the median value of empirically derived k_{Transp} from Teuling et al. 2006) and optimized all other parameters ~~would result~~ again~~ in~~. This caused that most TWS variations ~~being~~

~~caused by~~originate from *wSoil*, but with less improvement in model performance compared to **B** (S8). Therefore, TWS decomposition is very sensitive to ~~(certain) parameter values.~~parameters controlling ET under water limited conditions. However, **VEG** and **VEG** with fixed k_{Transp} qualitatively agree in the importance of the slow water storage in *Humid* regions, which was also shown by Getirana et al. (2017). Overall, our results imply that the representation of ET under water limited conditions in the models plays a decisive role on the simulated partitioning of TWS in soil moisture and slow water pools.

~~We showed, that even with the same calibration procedure and underlying model structure, TWS composition differs drastically between model experiments, and that is indicative of~~ The large impact of the role of vegetation and of transpiration water supply within the model.~~–~~ is also supported by a complementary experiment, in which vegetation parameters were discretized for plant functional type classes and calibrated with the same multi-criteria approach (S9).

As with the presented model variants, TWS composition simulated with existing large-scale hydrological models differs widely (Scanlon et al. 2018, Schellekens et al. 2017, Zhang et al. 2017). For example, PCR-GLOBWB and W3RA attribute seasonal TWS variations in the tropics to groundwater, while other models suggest it is mainly caused by soil moisture. Those results are largely dependent on model structure and parametrization, which is potentially a challenge when models are used to decompose the integrated GRACE TWS signal, and when implications of different processes and interactions are drawn. ~~For example, Humphrey et al. (2018) analysed how the CO_2 growth rate is affected by inter-annual fluctuations~~For example, Humphrey et al. (2018) analysed how the CO_2 growth rate, a proxy for the land carbon balance fluctuations, is affected by inter-annual variations in GRACE TWS, assuming that these represent fluctuations in plant accessible water that influence the carbon uptake of land ecosystems. In contrast, our study, along with previous reports, show that a significant proportion of the GRACE TWS signal in tropics is not directly plant accessible soil moisture, but deeper soil water and slow storage ~~component~~components. The latter comprises surface water storage, whose importance for TWS variations in tropical regions has been shown by several studies (e.g., Güntner et al. 2007, Getirana et al. (2017)).

Although **VEG** can be considered more reliable because of more realistic parameter values and better model performance, the current study still has some shortcomings. Despite using a multi-~~variate~~criteria calibration, individual component fluxes and states may not necessarily be well constrained. To further improve and solidify conclusions, especially on TWS partitioning, more constraints, such as deep soil moisture estimates or high-quality observations of surface water are needed. Furthermore, spatial constraints for defining the depletion of water storages via ET and Q – either with spatial information on the delay parameters (k_{Transp} for ET~~;~~, d_{Slow} *for Q*), or for their sub fluxes (transpiration or evaporation~~;~~, baseflow or direct runoff) would be beneficial. In this context, runoff characteristics as the baseflow index or the baseflow recession coefficient provided by Beck et al. (2015) are potentially useful to define spatial pattern of the slow runoff component. Besides, a GRACE product with daily resolution (Eicker et al. 2020) could enable better decomposition and differentiation of fast and slow storages whose short-term imprints are lumped in the monthly TWS signal.

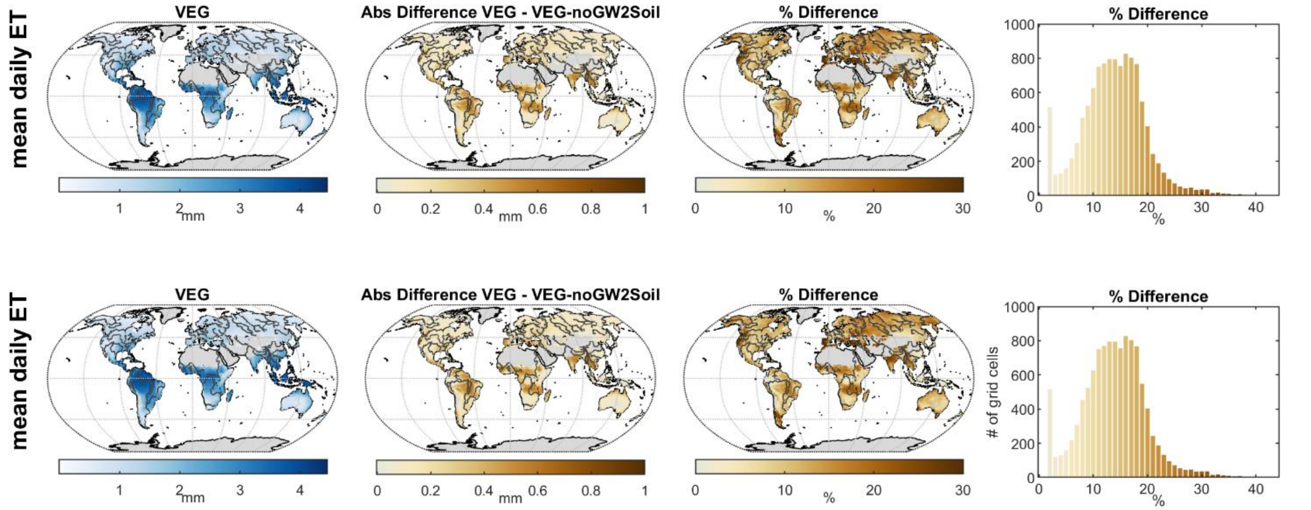


Figure 9 Total evapotranspiration of VEG with capillary flux from the deep soil water storage (left), and difference compared to a model version without capillary flux in mm (center) and as percentage difference (right).

5 Conclusion

In this study, we investigated the effect of varying vegetation characteristics on global hydrological simulations and in particular on the partitioning of TWS variations among snow, plant accessible soil moisture, a deep soil ~~and groundwater~~ water storage, and a slowly varying water pool that ~~represent~~represents groundwater, surface and near-surface water storage. To do so, we included observation-based continuous vegetation information to parameterize the hydrological processes of evapotranspiration, soil water storage and runoff generation in a large-scale hydrological model. With the parsimonious model that was constrained against multiple observations, we highlight the value of observation-based datasets in constraining model parameters of global hydrological models, while maintaining simple model formulations to evaluate the influences of vegetation in the global hydrological cycle.

First, we find that using a multi-criteria calibration approach allows for different model variants to perform relatively well despite major differences in model parameterization among them. In fact, even without accounting for dynamics and patterns of vegetation explicitly, the model performance can be interpreted as reasonable, and more so at the global scale. However, including spatial pattern of vegetation further improved the model performance. For example, large improvements were found in supply-limited regions, i.e., (semi-) arid regions (TWS and ET) and in tropical regions (ET), and Q simulations both globally and regionally in the Northern hemisphere. Undoubtedly, spatio-temporal variations of vegetation characteristics are relevant for ~~the~~ regional and global hydrological simulations.

Interestingly, we find that the calibrated parameter values are also more reasonable when the model is fed with the vegetation information. In particular, parameter interactions and equifinality were reduced even though the same observational constraints were used for calibration.

Lastly, we show how the representation of vegetation can modulate surface and subsurface hydrological process representation in the model, changing the spatial-temporal dynamics of individual storage components while maintaining the same overall response of total hydrological fluxes and storage variations. With or without accounting for varying vegetation characteristics explicitly, seasonal storage variations are dominated by snow at the global scale. However, including varying vegetation characteristics drastically changes the attribution of TWS variations among soil moisture, deep soil water and slow water storages. Without varying vegetation parameters, the soil moisture effectively controls most of the TWS variation, but with varying vegetation characteristics the role of deeper and delayed water storage becomes prominent. In particular the representation of water limited ET by the interplay of its sensitivity to soil moisture, maximum plant accessible water storage capacity, and interactions with deep soil moisture or groundwater seem to play a decisive role for TWS partitioning in the simulations.

In summary, this study highlights the value of including varying vegetation characteristics to further constrain model parameters ~~and to reliably represent hydrological processes despite~~with a parsimonious model structure. The findings further suggest an important role of groundwater-soil moisture-vegetation interactions for TWS variations. ~~In particular,~~Since the representation of vegetation related processes in global hydrological models seems to be a key factor for controlling TWS partitioning ~~of their simulations and associated uncertainties. Besides, it~~we emphasizes the need for further ~~observations and~~studies and improvements of global water cycle models with respect to the role of vegetation by utilizing observational~~-based estimates, as well as~~ constraints on ecohydrological functioning in multi-criteria model ~~experiments to provide and understand identifiability in model structure~~calibration exercises.

~~Besides, this study motivates further multi-model experiments to understand the need and potential of existing and novel observational constraints to increase the identifiability not only regarding model parameters, but also of model structure.~~

Author contribution

TT designed the research in extensive collaboration with SK, NC and MJ. TT and SK programmed the model experiments. NC contributed to parameter estimation and uncertainty analysis. TT performed the analysis and prepared the first draft of the manuscript. All co-authors provided recommendations for the data analysis, participated in discussions about the results, and edited the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Code and Data Availability

The data used for model forcing and calibration is publicly available via the original data provider mentioned in Table 1. The scripts to perform the analysis of this study can be accessed via https://doi.org/10.5281/zenodo.5770238. The processed data and model simulations are available upon request from the corresponding author.

Acknowledgements

All experiments of this study, including the model implementation and calibrations, were performed within the SINDBAD model data integration framework developed at the Max-Planck Institute for Biogeochemistry Jena.

References

Andersen, O. B., Krogh, P. E., Bauer-Gottwein, P., Leiriao, S., Smith, R., and Berry, P.: Terrestrial Water Storage from GRACE and Satellite Altimetry in the Okavango Delta (Botswana), 135, 521-526, 10.1007/978-3-642-10634-7_69, 2010.

Bai, P., Liu, X., and Liu, C.: Improving hydrological simulations by incorporating GRACE data for model calibration, Journal of Hydrology, 557, 291-304, https://doi.org/10.1016/j.jhydrol.2017.12.025, 2018.

Baldocchi, D., Ma, S., and Verfaillie, J.: On the inter- and intra-annual variability of ecosystem evapotranspiration and water use efficiency of an oak savanna and annual grassland subjected to booms and busts in rainfall, Global Change Biology, 27, 359-375, https://doi.org/10.1111/gcb.15414, 2021.

Beck, H. E., de Roo, A., and van Dijk, A. I. J. M.: Global Maps of Streamflow Characteristics Based on Observations from Several Thousand Catchments*, Journal of Hydrometeorology, 16, 1478-1501, 10.1175/jhm-d-14-0155.1, 2015.

Chen, J., Ban, Y., and Li, S.: Open access to Earth land-cover map, Nature, 514, 434-434, 10.1038/514434c, 2014.

Contractor, S., Donat, M. G., Alexander, L. V., Ziese, M., Meyer-Christoffer, A., Schneider, U., Rustemeier, E., Becker, A., Durre, I., and Vose, R. S.: Rainfall Estimates on a Gridded Network (REGEN) – a global land-based gridded dataset of daily precipitation from 1950 to 2016, Hydrol. Earth Syst. Sci., 24, 919-943, 10.5194/hess-24-919-2020, 2020.

Döll, P., Fritsche, M., Eicker, A., and Müller Schmied, H.: Seasonal Water Storage Variations as Impacted by Water Abstractions: Comparing the Output of a Global Hydrological Model with GRACE and GPS Observations, Surveys in Geophysics, 35, 1311-1331, 10.1007/s10712-014-9282-2, 2014.

Dorigo, W. A., Gruber, A., De Jeu, R. A. M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R. M., and Kidd, R.: Evaluation of the ESA CCI soil moisture product using ground-based observations, Remote Sensing of Environment, 162, 380-395, https://doi.org/10.1016/j.rse.2014.07.023, 2015.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions, Remote Sensing of Environment, 203, 185-215, https://doi.org/10.1016/j.rse.2017.07.001, 2017.

Draper, N., and Smith, H.: Applied Regression Analysis 2nd ed Wiley, New York, NY, 1981.

Du, L., Zeng, Y., Ma, L., Qiao, C., Wu, H., Su, Z., and Bao, G.: Effects of anthropogenic revegetation on the water and carbon cycles of a desert steppe ecosystem, Agricultural and Forest Meteorology, 300, 108339, https://doi.org/10.1016/j.agrformet.2021.108339, 2021.

Eicker, A., Jensen, L., Wöhnke, V., Dobslaw, H., Kvas, A., Mayer-Gürr, T., and Dill, R.: Daily GRACE satellite data evaluate short-term hydro-meteorological fluxes from global atmospheric reanalyses, Scientific Reports, 10, 4504, 10.1038/s41598-020-61166-0, 2020.

Eicker, A., Schumacher, M., Kusche, J., Döll, P., and Schmied, H. M.: Calibration/data assimilation approach for integrating GRACE data into the WaterGAP Global Hydrology Model (WGHM) using an ensemble Kalman filter: First results, Surveys in Geophysics, 35, 1285-1309, 2014.

Famiglietti, J. S., and Rodell, M.: Water in the Balance, Science, 340, 1300-1301, 10.1126/science.1236460, 2013.

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., and Otero-Casal, C.: Hydrologic regulation of plant rooting depth, Proceedings of the National Academy of Sciences, 10.1073/pnas.1712381114, 2017.

Fekete, B. M., Vörösmarty, C. J., Roads, J. O., and Willmott, C. J.: Uncertainties in Precipitation and Their Impacts on Runoff Estimates, Journal of Climate, 17, 294-304, 10.1175/1520-0442(2004)017<0294:uipati>2.0.co;2, 2004.

Getirana, A., Kumar, S., Girotto, M., and Rodell, M.: Rivers and Floodplains as Key Components of Global Terrestrial Water Storage Variability, Geophysical Research Letters, 44, 10,359-310,368, 10.1002/2017gl074684, 2017.

Ghiggi, G., Humphrey, V., Seneviratne, S. I., and Gudmundsson, L.: GRUN: an observation-based global gridded runoff dataset from 1902 to 2014, Earth System Science Data, 11, 1655-1674, 2019.

Güntner, A., Stuck, J., Werth, S., Döll, P., Verzano, K., and Merz, B.: A global analysis of temporal and spatial variations in continental water storage, Water Resources Research, 43, 10.1029/2006WR005247, 2007.

Hansen, N., and Kern, S.: Evaluating the CMA Evolution Strategy on Multimodal Test Functions, in: Parallel Problem Solving from Nature - PPSN VIII, edited by: Yao, X., Burke, E., Lozano, J. A., Smith, J., Merelo-Guervós, J. J., Bullinaria, J. A., Rowe, J., Tino, P., Kabán, A., and Schwefel, H.-P., Springer, Berlin, 2004.

Huffman, G. J., Adler, R., Morrissey, M. M., Bolvin, D., Curtis, S., Joyce, R., McGavock, B., and Susskind, J.: Global Precipitation at One-Degree Resolution from Multisatellite Observations, Journal of Hydrometeorology, 2, 36-50, 2000.

Humphrey, V., Berg, A., Ciais, P., Gentine, P., Jung, M., Reichstein, M., Seneviratne, S. I., and Frankenberg, C.: Soil moisture–atmosphere feedback dominates land carbon uptake variability, Nature, 592, 65-69, 10.1038/s41586-021-03325-5, 2021.

Humphrey, V., Zscheischler, J., Ciais, P., Gudmundsson, L., Sitch, S., and Seneviratne, S. I.: Sensitivity of atmospheric CO2 growth rate to observed changes in terrestrial water storage, Nature, 560, 628-631, 10.1038/s41586-018-0424-4, 2018.

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., and Reichstein, M.: The FLUXCOM ensemble of global land-atmosphere energy fluxes, Scientific data, 6, 1-14, 2019.

Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlstrom, A., Arneth, A., Camps-Valls, G., Ciais, P., Friedlingstein, P., Gans, F., Ichii, K., Jain, A. K., Kato, E., Papale, D., Poulter, B., Raduly, B., Rodenbeck, C., Tramontana, G., Viovy, N., Wang, Y. P., Weber, U., Zaehle, S., and Zeng, N.: Compensatory water effects link yearly global land CO2 sink changes to temperature, Nature, 541, 516-520, 10.1038/nature20780, 2017.

Kim, H., Yeh, P. J. F., Oki, T., and Kanae, S.: Role of rivers in the seasonal variations of terrestrial water storage over global basins, Geophysical Research Letters, 36, doi:10.1029/2009GL039006, 2009.

Koirala, S., Yeh, P. J. F., Hirabayashi, Y., Kanae, S., and Oki, T.: Global-scale land surface hydrologic modeling with the representation of water table dynamics, Journal of Geophysical Research: Atmospheres, 119, 75-89, 10.1002/2013JD020398, 2014.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World Map of the Köppen-Geiger climate classification updated, Meteorologische Zeitschrift, 15, 259-263, 10.1127/0941-2948/2006/0130, 2006.

Kraft, B., Jung, M., Körner, M., Koirala, S., and Reichstein, M.: Towards hybrid modeling of the global hydrological cycle, Hydrol. Earth Syst. Sci. Discuss., 2021, 1-40, 10.5194/hess-2021-211, 2021.

Küçük, Ç., Koirala, S., Carvalhais, N., Miralles, D. G., Reichstein, M., and Jung, M.: Characterising the response of vegetation cover to water limitation in Africa using geostationary satellites Earth and Space Science Open Archive, 27, doi:10.1002/essoar.10504964.1, 2020.

Kumar, S. V., Zaitchik, B. F., Peters-Lidard, C. D., Rodell, M., Reichle, R., Li, B., Jasinski, M., Mocko, D., Getirana, A., De Lannoy, G., Cosh, M. H., Hain, C. R., Anderson, M., Arsenault, K. R., Xia, Y., and Ek, M.: Assimilation of Gridded GRACE Terrestrial Water Storage Estimates in the North American Land Data Assimilation System, Journal of Hydrometeorology, 17, 1951-1972, 10.1175/jhm-d-15-0157.1, 2016.

Lo, M.-H., Famiglietti, J. S., Yeh, P. J.-F., and Syed, T. H.: Improving parameter estimation and water table depth simulation in a land surface model using GRACE water storage and estimated base flow data, Water Resources Research, 46, https://doi.org/10.1029/2009WR007855, 2010.

Lu, J., Sun, G., McNulty, S. G., and Amatya, D. M.: A ccomparison of six potential evapotranspiration methods for regional use in the southeastern United States, Journal of the American Water Resources Association, 41, 621-633, 2005.

Luojus, K., Pulliainen, J., Takala, M., Lemmetyinen, J., Kangwa, M., Eskelinen, M., Metsämäki, S., Solberg, R., Salberg, A.-B., Bippus, G., Ripper, E., Nagler, T., Derksen, C., Wiesmann, A., Wunderle, S., Hüsler, F., Fontana, F., and Foppa, N.: GlobSnow2 - Final Report, 2014.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev., 10, 1903-1925, 10.5194/gmd-10-1903-2017, 2017.

McColl, K. A., Wang, W., Peng, B., Akbar, R., Short Gianotti, D. J., Lu, H., Pan, M., and Entekhabi, D.: Global characterization of surface soil moisture drydowns, Geophysical Research Letters, 44, 3682-3690, https://doi.org/10.1002/2017GL072819, 2017.

Mostafaie, A., Forootan, E., Safari, A., and Schumacher, M.: Comparing multi-objective optimization techniques to calibrate a conceptual hydrological model using in situ runoff and daily GRACE data, Computational Geosciences, 22, 789-814, 10.1007/s10596-018-9726-8, 2018.

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., Peiris, T. A., Popat, E., Portmann, F. T., Reinecke, R., Schumacher, M., Shadkam, S., Telteu, C. E., Trautmann, T., and Döll, P.: The global water resources and use model WaterGAP v2.2d: model description and evaluation, Geosci. Model Dev., 14, 1037-1079, 10.5194/gmd-14-1037-2021, 2021.

Nash, J. E., and Sutcliffe, J. V.: River flow forecasting through conceptual models Part I - A discussion of principles, Journal of Hydrology, 10, 282-290, 1970.

Omlin, M., and Reichert, P.: A comparison of techniques for the estimation of model prediction uncertainty, Ecological modelling, 115, 45-59, 1999.

Porporato, A., Daly, E., and Rodriguez-Iturbe, I.: Soil water balance and ecosystem response to climate change, The American Naturalist, 164, 625-632, 2004.

Quevedo, D. I., and Francés, F.: A conceptual dynamic vegetation-soil model for arid and semiarid zones, Hydrol. Earth Syst. Sci., 12, 1175-1187, 10.5194/hess-12-1175-2008, 2008.

Rakovec, O., Kumar, R., Attinger, S., and Samaniego, L.: Improving the realism of hydrologic model functioning through multivariate parameter estimation, Water Resources Research, 52, 7779-7792, https://doi.org/10.1002/2016WR019430, 2016.

Reager, J. T., Thomas, A. C., Sproles, E. A., Rodell, M., Beaudoing, H. K., Li, B., and Famiglietti, J. S.: Assimilation of GRACE Terrestrial Water Storage Observations into a Land Surface Model for the Assessment of Regional Flood Potential, Remote Sensing, 7, 14663-14679, 2015.

Reichle, R. H., Draper, C. S., Liu, Q., Girotto, M., Mahanama, S. P. P., Koster, R. D., and De Lannoy, G. J. M.: Assessment of MERRA-2 Land Surface Hydrology Estimates, Journal of Climate, 30, 2937-2960, 10.1175/jcli-d-16-0720.1, 2017.

Reyer, C. P. O., Leuzinger, S., Rammig, A., Wolf, A., Bartholomeus, R. P., Bonfante, A., de Lorenzi, F., Dury, M., Gloning, P., Abou Jaoudé, R., Klein, T., Kuster, T. M., Martins, M., Niedrist, G., Riccardi, M., Wohlfahrt, G., de Angelis, P., de Dato, G., François, L., Menzel, A., and Pereira, M.: A plant's perspective of extremes: terrestrial plant responses to changing climatic variability, Global Change Biology, 19, 75-89, https://doi.org/10.1111/gcb.12023, 2013.

Rind, D.: The influence of vegetation on the hydrologic cycle in a global climate model, in: Climate Processes and Climate Sensitivity, edited by: Hansen, J. E., and Takahashi, T., AGU Geophysical Monograph 29, Maurice Ewing American Geophysical Union, 73-91, 1984.

Rodell, M., Beaudoing, H. K., L'Ecuyer, T. S., Olson, W. S., Famiglietti, J. S., Houser, P. R., Adler, R., Bosilovich, M. G., Clayson, C. A., Chambers, D., Clark, E., Fetzer, E. J., Gao, X., Gu, G., Hilburn, K., Huffman, G. J., Lettenmaier, D. P., Liu, W. T., Robertson, F. R., Schlosser, C. A., Sheffield, J., and Wood, E. F.: The Observed State of the Water Cycle in the Early Twenty-First Century, Journal of Climate, 28, 8289-8318, 10.1175/jcli-d-14-00555.1, 2015.

Rodell, M., Famiglietti, J., Wiese, D., Reager, J., Beaudoing, H., Landerer, F., and Lo, M.-H.: Emerging trends in global freshwater availability, Nature, 1, 2018.

Rodell, M.: Basin scale estimates of evapotranspiration using GRACE and other observations, Geophysical Research Letters, 31, 10.1029/2004gl020873, 2004.

Rodriguez-Iturbe, I., Porporato, A., Laio, F., and Ridolfi, L.: Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress: I. Scope and general outline, Advances in Water Resources, 24, 695-705, https://doi.org/10.1016/S0309-1708(01)00004-5, 2001.

Ruiz-Pérez, G., Koch, J., Manfreda, S., Caylor, K., and Francés, F.: Calibration of a parsimonious distributed ecohydrological daily model in a data-scarce basin by exclusively using the spatio-temporal variation of NDVI, Hydrol. Earth Syst. Sci., 21, 6235-6251, 10.5194/hess-21-6235-2017, 2017.

Scanlon, B. R., Zhang, Z. Z., Save, H., Wiese, D. N., Landerer, F. W., Long, D., Longuevergne, L., and Chen, J. l.: Global evaluation of new GRACE mascon products for hydrologic applications, Water Resources Research, 52, 9412-9429, 10.1002/2016wr019494, 2016.

Scanlon, B. R., Zhang, Z., Save, H., Sun, A. Y., Müller Schmied, H., van Beek, L. P. H., Wiese, D. N., Wada, Y., Long, D., Reedy, R. C., Longuevergne, L., Döll, P., and Bierkens, M. F. P.: Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data, Proceedings of the National Academy of Sciences, 10.1073/pnas.1704665115, 2018.

Schellekens, J., Dutra, E., Martínez-de la Torre, A., Balsamo, G., van Dijk, A., Sperna Weiland, F., Minvielle, M., Calvet, J.-C., Decharme, B., Eisner, S., Fink, G., Flörke, M., Peßenteiner, S., van Beek, R., Polcher, J., Beck, H., Orth, R., Calton, B., Burke, S., Dorigo, W., and Weedon, G. P.: A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset, Earth System Science Data, 9, 389-413, 10.5194/essd-9-389-2017, 2017.

Schenk, H. J., and Jackson, R. B.: Mapping the global distribution of deep roots in relation to climate and soil characteristics, Geoderma, 126, 129-140, https://doi.org/10.1016/j.geoderma.2004.11.018, 2005.

Schmidt, R., Petrovic, S., Güntner, A., Barthelmes, F., Wünsch, J., and Kusche, J.: Periodic components of water storage changes from GRACE and global hydrology models, Journal of Geophysical Research: Solid Earth (1978–2012), 113, 10.1029/2007JB005363, 2008.

Su, Z., Zeng, Y., Romano, N., Manfreda, S., Francés, F., Ben Dor, E., Szabó, B., Vico, G., Nasta, P., Zhuang, R., Francos, N., Mészáros, J., Dal Sasso, S. F., Bassiouni, M., Zhang, L., Rwasoka, D. T., Retsios, B., Yu, L., Blatchford, M. L., and Mannaerts, C.: An Integrative Information Aqueduct to Close the Gaps between Satellite Observation of Water Cycle and Local Sustainable Management of Water Resources, Water, 12, 1495, 2020.

Syed, T. H., Famiglietti, J. S., and Chambers, D. P.: GRACE-Based Estimates of Terrestrial Freshwater Discharge from Basin to Continental Scales, Journal of Hydrometeorology, 10, 22-40, 10.1175/2008jhm993.1, 2009.

Tapley, B. D., Watkins, M. M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., Sasgen, I., Famiglietti, J. S., Landerer, F. W., Chambers, D. P., Reager, J. T., Gardner, A. S., Save, H., Ivins, E. R., Swenson, S. C., Boening, C., Dahle, C., Wiese, D. N., Dobslaw, H., Tamisiea, M. E., and Velicogna, I.: Contributions of GRACE to understanding climate change, Nature Climate Change, 9, 358-369, 10.1038/s41558-019-0456-2, 2019.

Telteu, C. E., Müller Schmied, H., Thiery, W., Leng, G., Burek, P., Liu, X., Boulange, J. E. S., Andersen, L. S., Grillakis, M., Gosling, S. N., Satoh, Y., Rakovec, O., Stacke, T., Chang, J., Wanders, N., Shah, H. L., Trautmann, T., Mao, G., Hanasaki, N., Koutroulis, A., Pokhrel, Y., Samaniego, L., Wada, Y., Mishra, V., Liu, J., Döll, P., Zhao, F., Gädeke, A., Rabin, S. S., and Herz, F.: Understanding each other's models: an introduction and a standard representation of 16 global water models to support intercomparison, improvement, and communication, Geosci. Model Dev., 14, 3843-3878, 10.5194/gmd-14-3843-2021, 2021.

Teuling, A. J., Seneviratne, S. I., Williams, C., and Troch, P. A.: Observed timescales of evapotranspiration response to soil moisture, Geophysical Research Letters, 33, 10.1029/2006gl028178, 2006.

Tian, S., Van Dijk, A. I. J. M., Tregoning, P., and Renzullo, L. J.: Forecasting dryland vegetation condition months in advance through satellite data assimilation, Nature Communications, 10, 469, 10.1038/s41467-019-08403-x, 2019.

Trautmann, T., Koirala, S., Carvalhais, N., Eicker, A., Fink, M., Niemann, C., and Jung, M.: Understanding terrestrial water storage variations in northern latitudes across scales, Hydrol. Earth Syst. Sci., 22, 4061-4082, 10.5194/hess-22-4061-2018, 2018.

Trenberth, K. E., Smith, L., Qian, T., Dai, A., and Fasullo, J.: Estimates of the Global Water Budget and Its Annual Cycle Using Observational and Model Data, Journal of Hydrometeorology, 8, 758-769, 10.1175/jhm600.1, 2007.

Wang, F., Polcher, J., Peylin, P., and Bastrikov, V.: Assimilation of river discharge in a land surface model to improve estimates of the continental water cycles, Hydrol. Earth Syst. Sci., 22, 3863-3882, 10.5194/hess-22-3863-2018, 2018.

Wang, J., Price, K. P., and Rich, P. M.: Spatial patterns of NDVI in response to precipitation and temperature in the central Great Plains, International Journal of Remote Sensing, 22, 3827-3844, 10.1080/01431160010007033, 2001.

Wang-Erlandsson, L., Bastiaanssen, W. G. M., Gao, H., Jägermeyr, J., Senay, G. B., van Dijk, A. I. J. M., Guerschman, J. P., Keys, P. W., Gordon, L. J., and Savenije, H. H. G.: Global root zone storage capacity from satellite-based evaporation, Hydrol. Earth Syst. Sci., 20, 1459-1481, 10.5194/hess-20-1459-2016, 2016.

Weiss, M., van den Hurk, B., Haarsma, R., and Hazeleger, W.: Impact of vegetation variability on potential predictability and skill of EC-Earth simulations, Climate Dynamics, 39, 2733-2746, 10.1007/s00382-012-1572-0, 2012.

Werth, S., Güntner, A., Petrovic, S., and Schmidt, R.: Integration of GRACE mass variations into a global hydrological model, Earth and Planetary Science Letters, 277, 166-173, 10.1016/j.epsl.2008.10.021, 2009.

Wiese, D. N.: GRACE monthly global water mass grids NETCDF RELEASE 5.0 Ver. 5.0 Mascon Ver. 2, PO.DAAC, CA, USA, 2015.

Xu, X., Medvigy, D., Powers, J. S., Becknell, J. M., and Guan, K.: Diversity in plant hydraulic traits explains seasonal and inter-annual variations of vegetation dynamics in seasonally dry tropical forests, New Phytologist, 212, 80-95, https://doi.org/10.1111/nph.14009, 2016.

Yang, Y. T., Long, D., Guan, H. D., Scanlon, B. R., Simmons, C. T., Jiang, L., and Xu, X.: GRACE satellite observed hydrological controls on interannual and seasonal variability in surface greenness over mainland Australia, J Geophys Res-Biogeo, 119, 2245-2260, 10.1002/2014jg002670, 2014.

Yang, Y., Anderson, M., Gao, F., Hain, C., Noormets, A., Sun, G., Wynne, R., Thomas, V., and Sun, L.: Investigating impacts of drought and disturbance on evapotranspiration over a forested landscape in North Carolina, USA using high spatiotemporal resolution remotely sensed data, Remote Sensing of Environment, 238, 111018, https://doi.org/10.1016/j.rse.2018.12.017, 2020.

Yang, Y., Donohue, R. J., and McVicar, T. R.: Global estimation of effective plant rooting depth: Implications for hydrological modeling, Water Resources Research, 52, 8260-8276, 10.1002/2016WR019392, 2016.

Yu, L., Fatichi, S., Zeng, Y., and Su, Z.: The role of vadose zone physics in the ecohydrological response of a Tibetan meadow to freeze–thaw cycles, The Cryosphere, 14, 4653-4673, 10.5194/tc-14-4653-2020, 2020.

Zeng, Y., Su, Z., Barmpadimos, I., Perrels, A., Poli, P., Boersma, K. F., Frey, A., Ma, X., de Bruin, K., Goosen, H., John, V. O., Roebeling, R., Schulz, J., and Timmermans, W.: Towards a Traceable Climate Service: Assessment of Quality and Usability of Essential Climate Variables, Remote Sensing, 11, 1186, 2019.

Zeng, Y., Su, Z., Calvet, J. C., Manninen, T., Swinnen, E., Schulz, J., Roebeling, R., Poli, P., Tan, D., Riihelä, A., Tanis, C. M., Arslan, A. N., Obregon, A., Kaiser-Weiss, A., John, V. O., Timmermans, W., Timmermans, J., Kaspar, F., Gregow, H., Barbu, A. L., Fairbairn, D., Gelati, E., and Meurey, C.: Analysis of current validation practices in Europe for space-based climate data records of essential climate variables, International Journal of Applied Earth Observation and Geoinformation, 42, 150-161, https://doi.org/10.1016/j.jag.2015.06.006, 2015.

Zhang, L., Dobslaw, H., Stacke, T., Güntner, A., Dill, R., and Thomas, M.: Validation of terrestrial water storage variations as simulated by different global numerical models with GRACE satellite observations, Hydrology and Earth System Sciences, 21, 821-837, 10.5194/hess-21-821-2017, 2017.

Zhuang, R., Zeng, Y., Manfreda, S., and Su, Z.: Quantifying Long-Term Land Surface and Root Zone Soil Moisture over Tibetan Plateau, Remote Sensing, 12, 509, 2020.