Toward continental hydrologic–hydrodynamic modeling in South America

Vinícius A. Siqueira[1], Rodrigo C. D. Paiva[1], Ayan S. Fleischmann[1], Fernando M. Fan[1], Anderson L. Ruhoff[1], Paulo R. M. Pontes[2], Adrien Paris[3,4,a], Stéphane Calmant[3], Walter Collischonn[1]

[1]Instituto de Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, 91501-970, Brazil
[2]Instituto Tecnológico Vale (ITV), Belém, 66055-090, Brazil
[3]LEGOS, Université de Toulouse, CNRS, CNES, IRD, UPS, Toulouse, France
[4]GET, Toulouse, France
[a]now at : Collecte Localisation Satellite (CLS), Ramonville-Saint-Agne, 31520, France

Correspondence to: Vinícius A. Siqueira (vinisiquera@gmail.com)

Abstract. Providing reliable estimates of streamflow and hydrological fluxes is a major challenge for water resources management over national and transnational basins in South America. Global hydrological models and land surface models are a possible solution to simulate the terrestrial water cycle at the continental scale, but issues on parameterization and limitations in representing lowland river systems can place constraints on these models to meet local needs. In an attempt to overcome such limitations, we extended a regional, fully coupled hydrologic–hydrodynamic model (MGB) to the continental domain of South America and assessed its performance using daily river discharges, water levels from independent sources (in situ, satellite altimetry), estimates of terrestrial water storage (TWS) and evapotranspiration (ET) from remote sensing and other available global datasets. In addition, river discharges were compared with outputs from global models acquired through the eartH2Observe project (HTESSEL/CaMa-Flood, LISFLOOD and WaterGAP3), providing the first cross-scale assessment (regional/continental × global models) that makes use of spatially distributed, daily discharge data. A satisfactory representation of discharges and water levels was obtained (NSE > 0.6 in 55 % of the cases) and the continental model was able to capture patterns of seasonality and magnitude of TWS and ET especially over the largest basins of South America. After the comparison with global models, we found that it is possible to get considerable improvement on daily river discharges even by using current global forcing data, just by combining parameterization and better routing physics based on regional experience. Issues on potential sources of errors related to both global- and continental-scale modeling are discussed, as well as future directions for improving large-scale model applications in this continent. We hope that our study provides important insights to reduce the gap between global and regional hydrological modeling communities.

Keywords: Model validation; Regional hydrological models; Global hydrological models; Cross-scale assessment; Model comparison.

1 Introduction

Reliable simulations of streamflow dynamics and related processes are vital to support water resources management regarding water security, natural hazards, navigation, agriculture and energy production. Therefore, improved predictions of the hydrological system can aid policymakers and stakeholders in making better decisions, also fostering actions to reduce risk and impacts on water resources under current and future conditions. In South America, recent important floods (e.g., Marengo et al., 2012; Hoyos et al., 2013; Ovando et al., 2016) and droughts (Melo et al., 2016; Erfanian et al., 2017), together with uncertainties about the potential effects of climate change (Marengo et al., 2009) are encouraging new strategies for meeting social, economic and environmental water demands in large river basins all over the continent, some of them extending beyond political borders.

In this context, large-scale hydrological models arise as important tools for simulating the terrestrial phase of the water cycle. Despite limitations related to observed (in situ) data, especially in developing countries, advances in computational resources and remote sensing technologies are pushing such models toward continental and global scales with increasing resolution (Wood et al., 2011; Bierkens, 2015; Bierkens et al., 2015; Sood and Smathkin, 2015). Currently, estimates of water fluxes at these scales are usually obtained using two modeling frameworks, namely global hydrological models (GHMs) and land surface models (LSMs) (Haddeland et al., 2011). While GHMs are more concerned with water resources assessment and lateral transfer of water, thus enabling quantification of human impacts and water scarcity/stress (e.g., Döll et al., 2009; Wada et al., 2011), LSMs were primarily developed to provide lower boundary conditions for atmospheric circulation models, i.e., having a particular focus on vertical fluxes of heat and water (e.g., Balsamo et al., 2009). The latter are often coupled (i.e, in offline mode) to global river routing models designed for transporting water along drainage networks (e.g., Decharme et al., 2008; Zaitchik et al., 2010; Yamazaki et al., 2011; Getirana et al., 2012), which also enables the conversion from surface and groundwater runoff into river discharge and other surface water variables (e.g., flood extent, water level).

Although global-scale models can provide valuable spatiotemporal estimates of water fluxes and projections of those estimates (Sood and Smathkin, 2015), their ability to reproduce discharge observations at basin scale and to address practical water management issues is still limited (Archfield et al., 2015; Hattermann et al., 2018). Inaccuracies in runoff estimation from GHMs and LSMs may be first attributed to the uncertainty in global satellite precipitation products (Tian and Peters-Lidard, 2010; Sperna Weiland et al., 2015), but several studies have shown considerable differences between model outputs even when using the same meteorological forcing, given the lack of knowledge about runoff generation processes and deficiencies in parameter estimation (e.g., Haddeland et al., 2011; Gudmundsson et al., 2012; Zhou et al., 2012; Beck et al., 2017a). In particular, calibration has been found to have the largest impact on storage fluxes, evapotranspiration and discharge in comparison to variations in model structure and forcing data (Müller Schmied et al., 2014), which is a reason to call for efforts on this exercise as many of the GHMs and LSMs are not calibrated (Sood and Smathkin, 2015; Zhang et al., 2016; Beck et al., 2017a).

An alternative to overcome some limitations of GHMs and LSMs is to expand the spatial domain of hydrological models that were initially developed for catchment to regional scales. Applying these models at national (e.g., Crooks et al., 2014) to continental domains (e.g., Abbaspour et al., 2015; Pechlivanidis and Arheimer, 2015; Donelly et al., 2016) translates into a better use of local expert knowledge and country-specific datasets that may be difficult to reach globally. At the same time, it

5 is possible to focus on regionally relevant processes that are usually not included or not well resolved in global models. In South America, for example, several previous studies suggested that lateral water fluxes in large lowland rivers should be resolved using hydrodynamic routing (e.g., Paiva et al., 2011, 2013; Paz et al., 2011, 2014; Yamazaki et al., 2011; Pontes et al., 2017; Zhao et al, 2017), while GHMs generally apply methods based on constant/variable velocity or a kinematic simplification of the St. Venant equations (see the overview by Kauffeldt et al., 2016 and Bierkens, 2015). Even if LSMs can

10 be offline coupled to more physically based global river routing models (e.g., Yamazaki et al., 2011; Getirana et al., 2017b), calibration in the latter is likely to compensate for errors in runoff generation (Pappenberger et al., 2010; Getirana et al., 2013; Hodges, 2013) and lack of relevant vertical hydrological processes linked to river–floodplain dynamics (e.g., Pedinotti et al., 2012; Paz et al., 2014; Fleischmann et al., 2018). In turn, fully coupled large-scale hydrologic–hydrodynamic models (e.g., Paiva et al., 2013) can handle the above interactions while using one single modeling framework, and are now feasible

15 for using in continental domains because recent routing schemes (e.g., Bates et al., 2010) have proved to be computationally efficient for both regional (Getirana et al., 2017b; Pontes et al., 2017; Fleischmann et al., 2018) and global simulations (Yamazaki et al., 2013).

Over the past decades, skill in streamflow prediction has been emphasized in catchment to regional-scale modeling (Archfield et al., 2015), but there is a growing opportunity to perform further spatial analyses rather than relying just on

20 point measurements. Currently, a wide range of remote sensing products can be used to assess other variables than discharge, such as terrestrial water storage (e.g., Tapley et al., 2004; Watkins et al., 2015), evapotranspiration (e.g., Miralles et al., 2011; Zhang et al., 2018), soil moisture (e.g. Kerr et al., 2012) or water surface elevation derived from satellite altimetry (e.g., Santos da Silva et al., 2010). Previous studies have shown the utility of the aforementioned datasets not only to validate hydrological/routing models (e.g., Alkama et al., 2010; Getirana et al., 2012; Paiva et al., 2013) but also as an interesting tool

25 to constrain and estimate model parameters (Getirana, 2010; Werth and Guntner, 2010; Lopez et al., 2017). Therefore, remote sensing products can be helpful for continental-scale modeling in assessing regions where streamflow data are scarce, as well as to outline areas in which future model improvements are potentially needed.

In parallel, the interest in building catchment/regional models up to continental domains, together with global models trying to be locally relevant through hyper-resolution (Wood et al., 2011), fosters the need to reduce the gap between these two

30 modeling communities (Bierkens et al., 2015; Archfield et al., 2015). If the primary goal of a continental model is to provide estimates of river discharges to support regional water management practices, the results can also be compared with outputs of global models to assess both the performance and potential shortcomings of these models under a regional perspective. In the last years, there has been an increasing number of studies assessing outputs of LSMs through multimodel intercomparison (e.g., Zaitchik et al., 2010; Gudmundsson et al., 2012; Zhou et al., 2012), sometimes in conjunction with

GHMs (e.g., Haddeland et al., 2011; Beck et al., 2017a). Other recent studies focused on intercomparison of regional models in large basins around the world, usually relating overall performance to a single gauge station and having a particular interest in monthly statistics (e.g., Huang et al., 2016; Eisner et al., 2017; Krysanova et al., 2017). Moreover, little attention has been given to intercomparison of global and regional-scale models, and the existing studies focused only on monthly to annual flows (e.g., Zhang et al., 2016) and projection of climate change impacts using a small number of gauge stations (e.g., Gosling et al., 2011; Hattermann et al., 2018). As streamflow is highly variable over space and at short time scales (i.e., daily), model performance should be assessed with spatially distributed data within large basins and at sub-monthly intervals (e.g., Wu et al., 2014; Beck et al., 2017a; Zhao et al., 2017), but to our knowledge, no cross-scale (i.e., regional/continental × global models) intercomparison with a comprehensive evaluation of daily river discharges has been carried out over South America.

In this paper, we aim to start bridging this gap by (i) extending a regional-scale, fully coupled hydrologic–hydrodynamic model to the entirety of South America, assessing its ability to represent discharge and other hydrological variables across the continent; (ii) exploring how discharge estimates from continental-scale modeling perform when compared to those estimated with state-of-the-art global models and (iii) identifying the issues that should be addressed for improving continental/global-scale modeling in this continent. The next sections provide a brief description of (i) the major river systems of South America, (ii) modeling approaches, (iii) datasets selected for validation, (iv) calibration procedures, (v) global models selected for discharge comparison and (vi) metrics used for assessment of results.

2 Overview of the major South American river systems

South America is one of the most freshwater abundant regions on Earth, contributing around 30 % of the global runoff to the oceans (Clark et al., 2015) despite having only 12 % of the total land area. Because of a combination of wide latitudinal extent (10º N–55º S), major orographic features and strong oceanic influences (Garcia and Mechoso, 2005; Vera et al., 2006; Garreaud et al., 2009), the continent is subject to a diverse climate that feeds six out of the 10 largest basins in the world in terms of mean annual discharge, four of them only within the Amazon (Latrubesse et al., 2005). In particular, the Amazon is probably the most relevant hydrological system of the world, which draws great scientific attention due to its ecological importance and role in local to global climate (Werth and Avissar, 2002; Vera et al., 2006).

Figure 1 presents general information about South America, including hydroclimatic characteristics, major wetlands and hydrological regions. According to classifications of the Food Agriculture Organization (FAO) and Brazilian National Water Agency (ANA), the continent is partitioned into 27 major hydrological divisions, with the largest ones (Amazon, La Plata and Orinoco basins) sharing much of their water between different countries. Because almost 80 % of the territory lies between the tropics, the continent is dominated by an equatorial (tropical) climate that drives high amounts of rainfall especially near the equator, but extreme conditions also exist and range from very arid (such as the Atacama desert in the northern Chile) to polar areas over the Andes Cordillera and south of Patagonia (southern Argentina and Chile). As a result,

water availability is subject to be highly variable with large discrepancies in mean annual flows, for example, 8.7 mm yr^{-1} at Desaguadero River in the Colorado basin (Canalejas gauge station, 1.8 10^5 km²) compared with 2100 mm yr^{-1} at Japura River (Puerto Cordoba gauge station, 1.5 10^5 km²), the latter located in the Amazon basin.

5   (Figure 1)

Regarding river dynamics, flows in the Amazon are largely affected by floodplains over extensive flat terrains, causing significant flood peak delay and attenuation (Richey et al., 1989; Alsdorf et al., 2007; Yamazaki et al., 2011; Paiva et al., 2013). Because flood waves have travel times in the order of a few months, sometimes out of phase because of the seasonal

10 differences in precipitation (Richey et al., 1989), rivers are subject to strong backwater effects that extend for several hundred kilometers upstream of the river mouth or the confluence of its tributaries (Meade et al., 1991; Getirana and Paiva, 2013; Paiva et al., 2013), existing at both high- and low-water periods (Trigg et al., 2009). To the north, the Orinoco basin shares some characteristics of the Amazon, such as unimodal flood pulse and low interannual variability of floodplain inundation, especially in the Llanos region (Hamilton et al., 2002). In addition, the high amplitude of mainstem water level

15 — 14 to 16 m — produces backwater effects in its tributaries with strong hydrological gradients (Rosales et al., 2002).

At the heart of South America, the La Plata basin plays a major role in terms of agriculture, hydroelectricity and the economy in general, corresponding for almost 70 % of the combined Gross Domestic Product in countries such as Argentina, Paraguay, Bolivia and Brazil (Vera et al., 2006; Barros et al., 2006). La Plata is mainly composed of the Parana and Paraguay Rivers and, to a lesser extent, the Uruguay River, which are completely different with respect to flow conditions.

20 For example, the Upper Parana is largely regulated by reservoirs (about 50 % of mean annual flow, according to Su and Lettenmeier, 2009) and provides around 75 % of discharge up to the confluence with Paraguay river, despite the similar drainage area of both basins (~10^6 km²) (Barros et al., 2006). The Paraguay basin, on the other hand, is largely influenced by one of the largest wetlands in the world — the Pantanal —, a complex anabranching river–floodplain system characterized by very gentle slopes that can be lesser than 1.5 cm.km^{-1} (Tucci and Clarke, 1998; Berbery and Barros, 2002; Paz et al.,

25 2011; Bravo et al., 2012). Much of the water stored in the floodplain at high water does not return to the channels during the drying phase and become available for evaporation and infiltration (Paz et al., 2014), so that flood waves are lagged by about 4–6 months (Tucci and Clarke, 1998; Hamilton et al., 2002), up to 60 % water volume can be lost (Gonçalves et al., 2011) and the shape of downstream hydrographs are strongly modified (Bravo et al., 2012). Moreover, huge areas on the right overbank of the Paraguay River (Chaco Plain) have poorly defined drainage networks associated with alluvial megafans

30 (Latrubesse, 2015), which makes this basin one of the most challenging regions for hydrological modeling in South America. The headwaters of La Plata also border important hydrological systems such as the Tocantins–Araguaia and Sao Francisco in a tropical wet–dry biome called Brazilian Cerrado (i.e., Brazilian savanna). The former is composed of Tocantins and Araguaia Rivers, flowing parallel northwards until joining approximately 550 km upstream of the basin mouth near the Amazon Delta. While the Tocantins is marked by a cascade of large dams, the Araguaia River is much less altered and hosts

the huge Bananal plain, which contributes up to 30 % of reduction in peak discharge due to floodplain inundation (Lininger and Latrubesse, 2016). Regarding Sao Francisco, two-thirds of the runoff is generated at the upper part of the basin (Allasia et al., 2006) and the mainstem crosses a semiarid region known as the "drought polygon", which affects several parts of Northeast Brazil including the Parnaiba basin. At the south of the continent, rivers flowing to the Atlantic Ocean correspond

5 to less than 40 % of the area (i.e., > 60 % is related to endorheic basins) (Pasquini and Depetris, 2007) and their annual cycles usually show two maxima, one associated with the winter rainfalls and another to snowmelt during spring and early summer (Rivera et al. 2018).

3 Methods

3.1 MGB model

10 #### 3.1.1 Model description

The MGB, *Modelo hidrológico de Grandes Bacias* (Large-Scale Hydrological Model) is a conceptual, semi-distributed, large-scale hydrological model first presented by Collischonn et al., (2007). The choice of MGB for this study was motivated by several past applications in South America (Allasia et al., 2006), which encompassed rapid response (e.g., Collischonn et al., 2005, Siqueira et al., 2016a) to markedly seasonal and often slow response basins (e.g., Bravo et al., 2012; Paiva et al.,

15 2013; Fan et al., 2016; Pontes et al., 2017). In its most recent version, basins are divided into unit-catchments (Paiva et al., 2011; Pontes et al., 2017), each one containing a single river reach with associated floodplain and hydrological vertical water balance. Combinations of soil type and land use within each unit-catchment are categorized as Hydrological Response Units (HRUs). Water and energy budgets are computed independently for each HRU of each unit-catchment, and soil is depicted as a bucket model with a single layer. Canopy interception is represented in terms of leaf area index (LAI) and

20 evapotranspiration (soil, plant transpiration and open water evaporation) is calculated using the Penman–Monteith equation. Surface runoff is produced using the variable contribution area concept following the Arno model (Todini, 1996), while groundwater and subsurface flows are computed, respectively, with linear and nonlinear functions according to water availability in the soil layer. Runoff from each one of the components (surface, subsurface and groundwater) is propagated to the stream network using linear reservoirs (i.e., hillslope routing). As MGB was primarily developed for tropical regions,

25 snow processes are not represented in the current model version.

Flow routing in river channels can be computed using the Muskingum–Cunge method (Collischonn et al., 2007), one-dimensional full hydrodynamic (Paiva et al., 2013) or the local inertial method (Pontes et al., 2017). In this work, the MGB was applied with the inertial routing as described by Pontes et al., (2017), which uses the 1D version of the explicit local inertial approximation proposed by Bates et al., (2010). The routing structure of MGB is similar to that one described by

30 Yamazaki et al., (2011, 2013), i.e., the volume of water stored in a given unit-catchment is the only prognostic variable, while other variables such as flow depth and flooded area are diagnosed from the stored volume using floodplain profiles

derived with sub-grid topography. The floodplain is treated as a simple storage model and the water level for a given time step is assumed to be constant along the entire unit-catchment. In addition, the model accounts for evaporation in floodplains and infiltration from flooded areas to the unsaturated soil (Fleischmann et al., 2018), thus feedbacks between hydrological and hydrodynamic modules can be also represented (i.e., a two-way coupling approach). Further details on model water

5 balance and flow routing equations are presented in Supplementary Material S1.

3.1.2 GIS processing

All geoprocessing steps were conducted with an adapted version of IPH-Hydro Tools package (Siqueira et al., 2016b), using the 15 arcsec HydroSHEDS flow direction map (Lehner et al., 2008) as the main input. We chose the latter because it has received extensive corrections to address topological problems in flat areas and endorheic regions (Lehner et al. 2008) and

10 has been successfully applied for river routing in other studies (e.g., Yamazaki et al., 2013; Zhao et al., 2017). An upstream area threshold of 1000 km² was adopted for the onset of drainage networks, while unit-catchments and river reaches were delineated using a fixed-length vector-based discretization of $\Delta x = 15$ km (see Supplementary Material S1.3 for details). This length threshold was selected to ensure a balance between model stability and efficiency, resulting in an improved resolution when compared with configurations used by Yamazaki et al., (2013) and Getirana et al., (2017b) for river routing with the

15 local inertial method at global and Amazon domains (grids with 0.25° resolution), respectively.

To estimate sub-grid floodplain topography, we first computed the Height Above Nearest Drainage (HAND) map (Rennó et al., 2008) using flow directions and drainage networks derived from HydroSHEDS together with the Bare-Earth SRTM v.1 Digital Elevation Model (DEM) (O'Loughlin et al., 2016), and further a floodplain profile was created at each unit-catchment relating HAND value, flooded area and water volume similar as done by Yamazaki et al., (2013). The Bare-Earth

20 SRTM, resampled from 3 to 15 arcsec to match the HydroSHEDS resolution, was adopted to account for vegetation biases in floodplains since the C-band radar used by the original SRTM is not able to penetrate fully through the canopy (Carabajal and Hardling, 2005; Berry et al., 2007). Channel bed elevation at a given unit-catchment was estimated subtracting channel bankfull depth from river bank height (i.e., the elevation at bankfull depth) (Yamazaki et al., 2011; Paiva et al., 2013), the latter also derived from Bare-Earth SRTM. However, one of the drawbacks of using an unconditioned DEM such as the

25 Bare-Earth SRTM is the high level of noise affecting channel bank elevations, which need to be attenuated to avoid excessive inundation in low-relief areas. Instead of applying smoothing algorithms that modify the original DEM values (e.g., Paiva et al., 2011), a simple linear regression was fitted to DEM pixels located over drainage networks within each unit-catchment (river reach). Channel bank heights were set as the smoothed elevation associated with the center pixel of each river reach, while the original DEM values remained unchanged, for example, when computing the HAND model and

30 associated floodplain profiles (see Supplementary Material S1.5 and S1.6, for more details).

3.1.3 River hydraulic geometry

Because flow routing is very sensitive to river geometry (Yamazaki et al., 2011; Getirana et al., 2012; Paiva et al., 2013), channel parameters such as bankfull width and depth must be properly defined. However, detailed information about channel geometry is usually not available for large-scale basins and a very common approach is to adopt classic hydraulic geometry relationships (HGs) (Leopold and Maddock, 1953) for specific sites according to drainage area or discharge (Decharme et al., 2008; Yamazaki et al., 2011; Getirana et al., 2012; Paiva et al., 2013; Luo et al., 2017; Pontes et al., 2017). Here, the global database of Andreadis et al., (2013) was used to set initial values of bankfull widths and depths, which were derived from two-year return period flows using the Global Runoff Database Center (GRDC) data and universal HGs obtained from several rivers around the world. In addition, regional HGs (Beighley and Gummaldi, 2011; Paiva et al., 2011, 2013, Pontes, 2016) and width estimates based on satellite imagery from Pontes (2016) were included to improve the global channel geometries of Andreadis et al., (2013) for Amazon and La Plata basins.

3.1.4 Model forcing

The Multi-Source Weighted Ensemble Precipitation — MSWEP v1.1 (Beck et al., 2017b) — was used as precipitation input to the rainfall-runoff module of the MGB model. This is a 3-hourly, global-scale dataset (0.25º resolution) that optimally combines satellite, reanalysis and daily gauge data, and it has been evaluated with satisfactory results in a recent comparison of several precipitation datasets (Beck et al., 2017c). Regarding climate variables used to compute evapotranspiration, mean monthly data for the period 1961–1990 were retrieved from the Climate Research Unit — CRU Global Climate v.2 (New et al., 2002), which provides long-term climatologies of temperature, pressure, radiation and wind speed for all land areas at 10′ resolution.

3.1.5 Land use and soil data

Herein, we used the 400-m resolution HRU merged product (soil + land use) for South America (Fan et al., 2015), which is available at https://www.ufrgs.br/lsh. Basically, the soil map is a combination of the Brazilian database RADAMBrasil and the FAO Digitized Soil Map of the World and Derived Soil Properties, the latter included to account for areas lying outside Brazil. Land use classification was retrieved from the Global Land Cover map, which was generated using Envisat MERIS fine-resolution (300 m) satellite imagery over the year of 2009. Regional land use maps of some Brazilian states were further included in the HRU merged product to improve level of detail.

3.2 Validation datasets

3.2.1 Discharge and water level data

Daily records of discharge were collected from several national hydrological services including: Agência Nacional de Águas (ANA/Brazil: http://www.snirh.gov.br/hidroweb/), Operador Nacional do Serviço Elétrico (ONS/Brazil, Reservoir

naturalized flows: http://ons.org.br/), Instituto Nacional del Agua (INA/Argentina: http://bdhi.hidricosargentina.gob.ar/), Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM/Colombia: http://www.ideam.gov.co/solicitud-de-informacion), Servicio Nacional de Meteorología y Hidrología (SENAMHI/Peru and Bolivia), Dirección General de Aguas (DGA/Chile: http://snia.dga.cl/BNAConsultas/) and other databases such as the Environmental Research Observatory for
5 geodynamical, Hydrological and Biogeochemical control of erosion/alteration and material transport in the Amazon (ORE-HyBam: http://www.ore-hybam.org) and the GRDC (http://www.bafg.de/GRDC/, for Ecuador and NE South America). Based on expert knowledge, we masked out gauges heavily influenced by upstream reservoirs and included only gauges with more than 10 000 km² of drainage area. In a few cases, however, this threshold was lowered to include, at least, a small number of gauges due to the lack of available data. Short time series with less than five years of records were also excluded
10 from analysis. The complete list of discharge gauge stations can be found on Supplementary Table S.4.1.

Satellite altimetry data were obtained from the THEIA/Hydroweb website (http://hydroweb.theia-land.fr/). Within this database, time series of water surface elevation (WSE) were extracted manually using the methodology presented in Santos da Silva et al., (2010) and are provided at virtual stations (VSs) where the satellite ground track forms a cross-over with the river network (~10–40 cm of water level accuracy). A total of 841 VS were found over the Amazon basin, 10 over the
15 Orinoco basin and 29 over the La Plata basin. Data are derived from the observations of Envisat and Jason-2 for the period of 2002–2010 (35-day repeat orbit) and 2008–2010 (10-day repeat orbit), respectively. In situ stage data from ANA gauge stations were also obtained for the Brazilian territory and were filtered using the same criteria as discharge data.

3.2.2 Terrestrial water storage (TWS)

Launched in 2002, the Gravity Recovery and Climate Experiment (GRACE) measures temporal changes in the Earth's
20 gravity field (Tapley et al., 2004). Several studies have shown the ability of GRACE to detect continental water storage variations at large spatial scales (e.g., Wahr et al., 1998; Ramillien et al., 2004; Tapley et al., 2004), which can provide insights into hydrological modeling about potential deficiencies in process description, parameters and input data (Schmidt et al., 2008). Here, we used the Release 05 JPL RL05M v2 mass concentration (mascon) estimates available at GRACE Tellus website (https://grace.jpl.nasa.gov). The JPL RL05M mascon solution solves for monthly gravity anomalies in terms of 3º ×
25 3º equal-area spherical cap mascons, while using a Coastline Resolution Improvement (CRI) filter to discriminate between land and ocean mass portions of each mascon that spans coastlines (Wiese et al., 2016). Regarding the traditional spherical harmonic (SH) approach, which has been widely used in the last decade for global studies (Wahr et al., 1998; Landerer and Swenson, 2012), mascon solutions can also be applied to regional scales (Scanlon et al., 2016) and do not require the user to apply any postprocessing filters to the data, lowering the dependence on using scale factors to recover signal loss (Watkins et
30 al., 2015). Uncertainties in GRACE mascon solutions (3º × 3º) over South America are around 10–15 mm of equivalent water thickness, and have been found to be similar or slightly lower in relation to SH solutions (Scanlon et al., 2016).

Despite of the native resolution (3º × 3º), mascon grids are provided with a spatial sampling of 0.5º × 0.5º. We kept the original resolution computing a simple average of 0.5º grid pixels located inside 3º × 3º mascon locations, as signals at sub-mascon resolution cannot be considered independent of each other. Time series of simulated TWS were first derived by summing water stored in all hydrological compartments, including rivers, floodplains, soil, groundwater and vegetation canopy, at each time step. Similar to Paiva et al., (2013), the modeled TWS was then resampled as the weighted mean of TWS of all unit-catchments within each 3º × 3º equal-area mascon cell, using the former drainage area as weight. To ensure agreement with GRACE data, anomalies of simulated TWS were obtained by subtracting the long-term mean computed for the period between 2004 and 2009.

3.2.3 Evapotranspiration (ET)

Reference values of ET were extracted from the Climate Data Record (CDR) (Zhang et al., 2018), which is available at http://stream.princeton.edu:8080/opendap/MEaSUREs/WC_MULTISOURCES_WB_050/. Within this dataset, 10 gridded global ET products estimated from satellite (five), reanalysis (two) and LSMs (three) were optimally combined at 0.5º resolution using weighted averaging and a Bayesian merging technique. The weight of each product is related to the inverse of the ensemble spread and the deviation from the ensemble mean is assumed as a proxy of the uncertainty/error in individual products. Together with other optimally merged variables provided by the CDR dataset (Precipitation, Runoff and TWS), estimates of ET were further adjusted with a Constrained Kalman Filter to ensure terrestrial water budget closure at each 0.5º grid cell (Zhang et al., 2018). For comparison purposes, the modeled ET was spatially aggregated into cells of 0.5º resolution using the unit-catchment drainage area as weight.

3.3 Model adjustment

Model calibration is commonly performed to improve agreement between observations and model results. However, the traditional gauge-by-gauge calibration used in regional hydrological modeling is not very common in continental to global domains (Archfield et al., 2015; Bierkens, 2015, Samaniego et al., 2017; Mizukami et al. 2017) because it can lead to spatial discontinuities of parameters (i.e., patchwork patterns) and overfitting to account for limitations in data and model structure. In other words, good results of discharge may not reflect a suitable depiction of the underlying hydrological processes, so that modelers are more subject to "get the right answers for the wrong reasons" (Kirchner, 2006)

In an attempt to reduce the only dependency of river gauges during calibration, regions of parameter sets were derived by intersecting the global map of lithology/geology of Durr et al., (2005) with large South American basins/hydrological regions, the latter shown in Fig 1a. The Amazon and La Plata basins were further divided into their main tributaries prior to intersection due to their large spatial extent. As parameter sets do not correspond to a single gauge station, but rather to regions defined by geological characteristics, multiple gauges were calibrated at the same time using the same parameter set. It is worth mentioning that calibration still remains a challenge for hydrological modeling with respect to large-scale

domains (Mizukami et al., 2017) and assessing the suitability of emerging parameter regionalization techniques (e.g., Samaniego et al., 2017) may be investigated in the future because it is beyond the scope of this paper.

The set of parameters used for model calibration is listed in Table 1, including their respective typical ranges of values. Other parameters used to compute energy balance and evapotranspiration (e.g., LAI, superficial resistance, albedo and canopy height) were defined *a priori* for each HRU vegetation type according to Collischonn (2001, and references within). Automatic calibration was not used herein to keep coherent values according to soil type and land cover, thus aiding to reduce model overparameterization. Therefore, sensitivity analysis of rainfall-runoff parameters was continuously performed as part of the manual calibration process. Regarding the hydrodynamic module, downstream boundary condition at oceans and lakes (endorheic basins) was set to a constant water level for simplification. Manning coefficient values were globally set to 0.03, with adjustments in specific rivers of the Amazon basin according to Paiva et al., (2013). Infiltration from floodplains to the soil column was considered and calibrated only for Pantanal region ($K_{inf} = 10$ mm day^{-1}) since previous studies showed that vertical hydrological processes largely influence model results in this area (e.g., Paz et al., 2014). It is worth mentioning that model sensitivity to river geometry and infiltration parameters were previously assessed by Paiva et al., (2013) and Fleischmann et al., (2018), respectively.

(Table 1)

As a result of the calibration procedure, several model parameter sets were manually adjusted and can be summarized into the following median values and percentile ranges (p5–p95): W_m = 500 (50–1500) mm; b = 0.2 (0.02–1.5); K_{bas} = 0.2 (0.01–3.0) mm day^{-1}, K_{int} = 2 (0.1–50) mm day^{-1}, XL = 0.67 (0.1–0.67), C_s = 15 (5–35), C_i = 120 (20–200) and C_b = 1200 (800–6000) h.

3.4 River discharge from GHMs and LSMs

Discharge outputs from state-of-the-art global models were acquired through the eartH2Observe Water Cycle Integrator (WCI, ftp://wci.earth2observe.eu). The WCI hosts multidecadal global water resources reanalysis datasets produced by 10 GHMs and LSMs, providing multi-scale (regional, continental and global) estimates of meteorological and hydrological water balance variables. We selected outputs from the 0.25º resolution Water Resources Re-analysis run 2 (WRR-2) baseline, which is an improved dataset over the WRR-1 (0.5º) produced by the initial project run (Schellekens et al., 2017). Models in the WRR-2 baseline are forced with MSWEP precipitation (1979–2014) and bias-corrected ERA-Interim data using the WFDEI correction methodology (see Martinez-de la Torre et al., 2018). Among the global models in WRR-2, river discharge at 0.25º resolution was available only for one LSM, the HTESSEL offline coupled to CaMa-Flood (Balsamo et al., 2009; Yamazaki et al., 2011), and two GHMs, namely, LISFLOOD (van der Knijf et al., 2010) and WaterGAP3 (Döll et al., 2009). The latter two have some degree of calibration and performed relatively well in terms of runoff in a recent model intercomparison (Beck et al., 2017a), while the former is uncalibrated but uses a state-of-the-art hydrodynamic routing

model. Within the eartH2Observe project, LISFLOOD and WaterGAP3 were run at, respectively, 0.1° and 0.08333° resolutions, and discharges were then resampled to 0.25° for WRR-2 (Martinez-de la Torre et al., 2018). A brief overview of the structure of these models is shown in Supplementary Material Table S2.1.

Because these models are grid-based, we followed a similar procedure to that in Zhao et al., (2017) to match grid cells to corresponding river gauge stations. First, we applied an automatic routine to find the cell coordinates nearest to the gauge locations. Cells were selected when the difference in the upstream area was within 5 %; otherwise, the surrounding cell with minimum upstream area difference was selected. Gauges associated with cells whose drainage area differed more than 15 % were excluded from the analysis. This procedure was performed separately for each global model to deal with differences between their respective drainage networks. Moreover, due to the spatial resolution mismatch of LISFLOOD and WaterGAP3, flow accumulation grids were recomputed using their respective flow direction maps (at 0.1° and 0.08333°) and were resampled to the same resolution of discharge grids (0.25°). The corresponding cells were then extensively validated with a thorough, GIS-assisted visual inspection, supported by long-term mean annual discharge grids (derived from each global model) to minimize errors of gauge mislocation.

3.5 Metrics for assessment of results

MGB simulation was carried out between 01–Jan–1990 and 31–Dec–2009 using a daily time step and a warmup period of two years to eliminate the influence of initial conditions. Model results were assessed in terms of discharge, water levels, ET and TWS, while simulated river discharges were further compared with the output of global models. **Table 2** lists all efficiency metrics used for assessment of model results. Statistics such as the Kling–Gupta Efficiency and Delay Index were the same as used in Kling et al., (2012) and Paiva et al., (2013), respectively.

(Table 2)

4 Results and discussion

4.1 Model validation

4.1.1 River discharges

Simulated daily discharges were compared with in situ observations and results were presented in maps of performance metrics (r, KGE and NSE) at each gauge station (Fig. 2). In addition, the runoff coefficient (RC = Q_{mean}/P_{mean}) was calculated for each gauge station and was plotted against its respective KGE and drainage area (Fig. 3). There is a good agreement between simulated and observed flows in several regions of South America, as NSE and KGE values are larger than 0.6 in 55 % and 70 % of the cases, respectively. Model performance is clearly higher in the southern and southeastern regions of Brazil, including central Amazon. On the other hand, performance decreases in regions marked by semiarid to arid climates, such as in Northeast Brazil, west and southwest of the La Plata basin, most parts of Argentina and northern Chile. For

instance, a poor correlation ($r < 0.2$) is observed in a semiarid region covered by the Colorado basin, where snow/glacier melt has a large contribution to total runoff and corresponds to the main source of water for human activities (Rivera et al., 2017). Other locations with lower performance (NSE < 0.2) usually refer to regions strongly influenced by orography (around Andes Cordillera), which are expected due to larger uncertainties of satellite-derived precipitation in these areas

5   (Tian and Peters-Lidard, 2010; Paiva et al., 2013). River discharges at gauge stations with RC ranging between 0.3 and 0.6 are generally well represented in all spatial scales, while performance tends to be lower for RC < 0.3 and highly variable for rivers with lower drainage areas.

(Figure 2)

(Figure 3)

Figure 4 shows daily simulated discharges for some of the large South American rivers. The agreement between simulated and observed discharges is notable for both high and low flows in most of the cases, which indicates the model's ability to

15   simulate regional to continental-scale rivers (10^5 km² to $4.7 \ 10^6$ km²) with different flow regimes. Results in the Amazon basin (e.g., Obidos, $NSE_{HD} = 0.89$) are comparable to other regional studies (e.g., Getirana et al., 2012; Paiva et al., 2013; Luo et al., 2017) while better performance is found over several of its tributaries (e.g., Purus, Madeira and Japura Rivers). Figure 4 also highlights the improvements of MGB using hydrodynamic (HD) over a non-hydrodynamic (noHD) routing method. In the Paraguay River, peak flows are dramatically reduced at the Amolar gauge station when using the HD routing

20   (up to –75 %), and a similar behavior can be seen at Puerto Bermejo ($NSE_{noHD} = -5.8$ to $NSE_{HD} = 0.42$) located about 1600 km downstream near the confluence with the Parana River. Previous attempts of regional hydrological modeling in this basin that did not account for the floodplain inundation in the Pantanal (e.g., using the calibrated VIC model; see Su and Lettenmeier, 2009) reported negative NSE values for Puerto Bermejo, even at the monthly time scale. Differences in performance between noHD and HD are also quite remarkable (especially in terms of NSE) at gauge stations of Conceicao

25   do Araguaia and Calamar in the lower Magdalena, where a pronounced attenuation effect is observed. On the other hand, in some rivers such as the Uruguay at Garruchos ($NSE_{noHD} = 0.85$; $NSE_{HD} = 0.82$), Parana at Itaipu ($NSE_{noHD} = 0.91$; $NSE_{HD} = 0.87$) and Tocantins at Descarreto ($NSE_{noHD} = 0.72$; $NSE_{HD} = 0.70$) the routing method has a minor impact. In the case of Orinoco at Ciudad Bolivar, both hydrographs look similar, but NSE suggests that results are improved when HD routing is used ($NSE_{noHD} = 0.83$; $NSE_{HD} = 0.9$).

(Figure 4)

4.1.2 Water levels

Results of performance metrics regarding water levels are presented in Fig. 5. For a suitable comparison, observed data and modeled WSE were first converted into anomalies (i.e., by subtracting their respective long-term mean: $h_{new} = h - \bar{h}$) to keep values with the same reference. In addition, Fig. 6 shows time series of simulated water level anomalies (hereafter referred to as water levels) for some of the large rivers of South America, which were plotted against in situ water levels and satellite altimetry. In general, the results obtained for the assessed gauge and VS stations are considered satisfactory in terms of correlation ($r \geq 0.8$ in 80 % of cases) and Nash–Sutcliffe (NSE ≥ 0.6 in 60 % of cases), with a reasonable performance for amplitudes (–30 % $< \sigma$BIAS < 30 % in 50 % of the cases). Like prior studies in the same region (e.g., Getirana et al., 2012, 2017b; Paiva et al., 2013), water levels are well represented in central Amazon, where a good performance is observed at tributaries and along the main river down to its mouth in the Atlantic Ocean. Amplitudes are overestimated (σBIAS > 30 %) in southeast tributaries such as Madeira, Xingu and Tapajos, as well as in headwaters located in the northwest part of the basin. Outside the Amazon, there are acceptable results in the Orinoco (e.g., lower Meta), Uruguay and Tocantins–Araguaia basins, where the model generally performs well in all assessed metrics. Large overestimations in the standard deviation (σBIAS $> +50$ %) are systematically found over the Sao Francisco main stem, which are reflected by very low values of NSE (< 0.2). On the Paraguay River, a reasonable agreement between observed and simulated water levels is observed at Amolar, but performance significantly reduces in both correlation and σBIAS for downstream regions (e.g., at Porto Murtinho). In the latter case, model results are clearly advanced in time and are not capturing rapid variations of water level originating from lateral contributions of tributaries.

(Figure 5)

(Figure 6)

Performance in water levels is directly related to the agreement between simulated and observed discharges. On the other hand, even if discharges are well represented, there are uncertainties related to Manning values and also to river widths and depths derived from HGs, which do not reflect singularities of cross sections such as narrowing or widening of rivers at both gauge and VS locations. Previous studies have demonstrated the large influence of channel geometry and roughness on both amplitude and timing of water levels, especially over the Amazon basin (e.g., Yamazaki et al., 2011; Paiva et al., 2013; Paris et al., 2016; Luo et al., 2017). Moreover, river bed profiles are subject to DEM errors that can hardly be reduced through simple profile-smoothing procedures. For example, datasets used to remove the vegetation bias in Bare-Earth SRTM (IceSAT, vegetation height maps, uncorrected SRTM) have different spatial resolutions that lead to artifacts around the edges of vegetation patches (O'Loughlin et al., 2016), producing additional noise on river bed elevations. In addition to vegetation, other SRTM error sources such as stripe noise (Rodriguez et al., 2006) significantly affect large flat areas on the

La Plata basin (Yamazaki et al., 2017), which can ultimately impact model results. Model resolution and the ability to route discharge in downstream multi-directions (e.g., rivers with bifurcations and anabranching networks) can affect simulated water levels and flooded areas (e.g., Mateo et al., 2017), which has been taken into account in recent studies with MGB (e.g., Pontes et al., 2017; Fleischmann et al., 2018) but not in this continental model application. Other simplifications in the model
5   may also affect results, like the assumptions of rectangular cross-sections and zero velocity of water stored on floodplains.

4.1.3 Evapotranspiration (ET)

Figure 7 shows the magnitude and seasonality of ET averaged for major basins in South America, as well as the magnitude of errors (RMSE) comparing modeled values to the optimal estimate of the CDR dataset. The ratio of the RMSE to the CDR uncertainty ($RMSE_{unc}$) was calculated to outline regions where simulated ET tends to deviate from the optimal CDR value,
10   i.e., values above unity indicate that the model error is larger than the mean deviation of all datasets (used in CDR) from their ensemble mean.

Results show that MGB can capture patterns of ET over the South America region. Simulated ET values are within the CDR uncertainty range in most of the continent ($RMSE_{unc} < 1$), with errors varying between 10 and 30 mm month^{-1}. A good agreement in terms of magnitude and seasonality of ET is observed for the Amazon, Tocantins–Araguaia and mainly for La
15   Plata basin, but the model also performs reasonably well in Sao Francisco. Conversely, larger deviations ($RMSE_{unc} > 1$) are found at low latitudes (20º S–10º N) where RMSE values reach up to 50 mm month^{-1}. ET is underestimated during the dry season in basins such as Orinoco (DJF), Amazon and Tocantins–Araguaia (JJA), while it is largely overestimated from the onset to the end of the dry (wet) season in Sao Francisco and Parnaiba (Orinoco). The latter two are clearly affected by a temporal lag in ET seasonality, where simulated values are delayed by approximately one month with respect to CDR
20   estimations. At mid latitudes (> 20º S), large RMSE values (above 50 mm month^{-1} and $RMSE_{unc} > 2$) are observed only in a narrow N–S range over the southern Andes.

Regarding issues on timing and magnitude of simulated ET, meteorological forcing probably has an influence on model performance because long-term mean climate data are used for ET computations. Another possible reason is the lack of spatial variability of moisture in the MGB soil column. Guswa et al., (2002) compared a simple bucket (one layer) to a
25   physically based Richards model, showing that large discrepancies can occur with respect to the relationship between ET and average root-zone saturation, as well as in timing and intensity of transpiration, especially for water-limited conditions. Moreover, Wang et al., (2006) found that time scales of evapotranspiration can differ significantly between a single-layer and multi-layer soil scheme due to nonlinear interactions that occur in the latter. Indeed, ET is expected to respond quickly at the beginning of the rainy season due to an increase of water availability at the soil surface layer, which cannot be well
30   represented with a single-layer, bucket-type model like MGB. In contrast to some of the datasets used in CDR (e.g., reanalysis and LSMs), the MGB does not account for snow processes, which may explain the large RMSE values over the southern Andean region.

Although it is beyond the scope of this study for a full assessment of ET estimations derived from hydrological models and
5   other sources, it is important to note that errors presented here correspond to the difference between both estimates (MGB and CDR). ET is one of the most uncertain water balance variables due to its high spatial and temporal variability, thus it is difficult to validate given the lack of ground observations (Miralles et al., 2011; Zhang et al., 2018). Even accounting for ~70 % of the weight in CDR dataset in comparison to LSM and reanalysis (Zhang et al., 2018), remote sensing products of ET are based on Penman–Monteith or Priestley–Taylor equations that depend on vegetation indices and meteorological forcing
10   derived from satellite/reanalysis data, which are associated to many uncertainties (Miralles et al., 2011; Vinukollu et al., 2011). Christoffersen et al., (2014) and Maeda et al., (2017) showed that most remote sensing and land surface models are unable consistently to reproduce ET seasonal cycles in tropical areas (across the Amazon basin) when compared with eddy covariance measurements and ET estimates from water balance. In the Amazon, for example, ET seasonality is regulated by radiation, rainfall and how vegetation assimilates water and energy (Restrepo-Coupe et al., 2013; Maeda et al., 2017). Other
15   limitations are also associated with the vegetation cover fraction and how ET is partitioned between transpiration, soil and canopy evaporation (Miralles et al., 2011).

4.1.4 Terrestrial water storage (TWS)

Figure 8 shows the performance of simulated TWS anomalies in comparison to observations from GRACE mascon solutions. To evaluate the ability of MGB to reproduce monthly variations of TWS, both simulated and observed time series
20   of TWS were averaged to the scale of large basins in South America and are presented in Fig. 9.
In general, the results show that MGB has the ability to represent TWS anomalies over the continent. There is a good temporal correlation in most part of tropical South America ($r > 0.75$), as well as in temperate regions with dry summer between latitudes of 30º S and 40º S. Amplitudes of TWS are reasonably well simulated (–20 % < σBIAS < 20 %) mainly in central Brazil, parts of the northeast, south of La Plata and areas of South Chile. On the other hand, performance typically
25   decreases in semiarid to arid climates, as can be seen in regions such as North Chile, Colorado basin, west of La Plata and South Argentina. High negative bias with large RMSE (> 150 mm) is observed in the northeast Amazon and west of the Orinoco, whereas large overestimations are found mainly over coastlines at low latitudes in the southern hemisphere (0–20º S). Moreover, in some regions characterized by polar climate, at the extreme south or in areas over the Andes, modeled TWS anomalies are markedly underestimated (σBIAS < –80 %).
30   Modeled TWS is in good agreement with GRACE observations over the Amazon, Tocantins–Araguaia, Sao Francisco and Parnaiba basins ($r > 0.9$, |σBIAS| < 15 % and RMSE < 45 mm), capturing both the interannual variability and amplitude of TWS anomalies for the analyzed period. MGB is also successful in representing annual changes of TWS in the La Plata basin, but with an overestimation (σBIAS = 22 %) probably caused by high positive σBIAS in the Paraguay and Chaco

regions. Errors in the La Plata basin (RMSE = 24 mm) are in the same order as those in the Amazon (RMSE = 26 mm). In addition, larger amplitude differences clearly occur in the Orinoco (σBIAS = –32 %), but with a pronounced RMSE (> 60 mm) that mainly originates from the eastern part of the basin.

(Figure 8)

(Figure 9)

The good agreement found in the Amazon basin can be attributed to the explicit representation of surface water reservoir

10 (channels and floodplains), which has been demonstrated to play an important role on both magnitude and timing of TWS (Alkama et al., 2010; Paiva et al., 2013; Getirana et al., 2017a). Other authors have pointed out that the contribution of surface storage to TWS is also potentially high in the Orinoco (~45 %) (e.g., Frappart et al., 2014), suggesting a large underestimation of the soil storage (in the eastern part of the basin) because anomalies of water level were reasonably well simulated. Indeed, surface storage has been understood as a major component of TWS variability over tropical regions of

15 South America, and may also be relevant for large rivers crossing semiarid areas such as the Sao Francisco (Getirana et al., 2017a). In the case of La Plata, the TWS amplitude is likely to be amplified if surface water is anticipated in time (Getirana et al., 2017a), which is probably occurring due to the low correlation of water levels previously simulated for the Paraguay River. In addition, the absence of a well-defined river system due to very flat terrains (e.g., Chaco Region, in the west part of La Plata) potentially favors the dominance of the groundwater dynamics over TWS, as already reported by Kuppel et al.,

20 (2015) in the Western Pampas more in the south.

It is worth mentioning that, for regions of South America located in mid latitudes, TWS is dominated by interannual variability rather than the seasonal cycle (Humphrey et al., 2016), where TWS amplitudes are generally lower and errors more apparent. Previous studies showed a strong negative trend in GRACE mascon solutions in the Colorado basin (e.g. Scanlon et al., 2016) that can be associated to a decrease in snow water equivalent over the dry Central Andes (Rivera et al.,

25 2017). Moreover, negative variations in glacier mass have been reported in southern Argentina/Chile over the Patagonia Icefields (Chen et al., 2007), which is probably the main responsible for the large RMSE observed at the extreme south of the continent. TWS in nearby semiarid areas is potentially affected by snow/glacier melting because the latter one is an important water source of Patagonian rivers flowing to the Atlantic Ocean (Pasquini and Depetris, 2007; Rivera et al. 2018). Inconsistencies along coastlines are also expected because of the smaller size of land mascons that increase uncertainty in

30 GRACE estimates (Wiese et al., 2016). Finally, in addition to issues related to model parameterization and depiction of hydrological processes (e.g., snowmelt), artificial reservoirs (dams) and lakes are not included in the current version of the South America MGB model, leading to additional uncertainties in TWS estimation.

4.2 Cross-scale comparison of river discharges from continental × global models

This section presents an assessment of MGB simulated discharges in comparison to the outputs from HTESSEL/CaMa-Flood, LISFLOOD and WaterGAP3 models, extracted from WRR-2 in the context of the eartH2Observe project (Schellekens et al., 2017; Martinez-de la Torre et al., 2018). This offers an interesting opportunity to evaluate to what extent discharge estimates can be improved at the continental scale of South America, as well as to identify the major shortcomings that should be addressed. To provide a concise spatial analysis, discharges from global models were reduced to their ensemble mean (Ensemble GM) and results are presented in terms of the difference of each metric (indicated by "d_metric"), i.e., by subtracting the performance of MGB from the performance of the Ensemble GM (Fig. 10). Bias and DI values are given in terms of absolute differences (d_Abs(metric)) to make both under- and overestimations comparable. Therefore, positive values indicate that MGB outperforms the ensemble mean of global models and vice-versa. Detailed performance metrics of each model can be found in Supplementary Material S3.

The continental model presents improvements for all metrics over most of the South America regions when compared with the global ensemble mean. In relative terms, a better agreement of simulated and observed discharges (d_KGE > 0.8) is observed over semiarid regions (e.g., East/Northeast Brazil and most part of Argentina), which are strongly impacted by bias in the Ensemble GM (d_Abs (BIAS) > 60 %). In tropical regions with marked seasonality and dry winter (e.g., upper Parana headwaters), differences in bias are lower (–10 % < d_Abs (BIAS) < 20 %), which indicates that KGE performance depends mainly on the variability of flows that is not captured by the Ensemble GM. Correlation is considerably higher over the Paraguay River (d_r > 0.4), highlighting the strong influence of hydrodynamic effects and complex processes in the Pantanal and Chaco regions, as documented by regional studies (e.g., Paz et al., 2011, 2014; Bravo et al., 2012, Pontes, 2016). There is also a clear correlation improvement in rivers such as the Araguaia, Amazonas and lower Magdalena, which are also affected by river–floodplain interactions with consequent flood peak attenuation (e.g., Paiva et al., 2013; Lininger and Latrubesse, 2016; Angarita et al., 2017; Pontes et al., 2017). A similar performance is observed for timing (d_Abs (DI)) with absolute differences being larger than 20 days, which also occur in the main stem of the Orinoco basin. In terms of NSE, the largest differences in performance previously observed for the KGE now extend to the main Amazon River, to its tributaries in the eastern region (i.e., Tapajos and Xingu) and also to both Magdalena and Tocantins–Araguaia basins, with values of d_NSE ≥ 0.8. With respect to low to medium flows (d_NSElog), there is a similar pattern to d_KGE (except for East Brazil), although with more pronounced differences in the Amazon and Magdalena regions.

The Ensemble GM performs relatively well in all statistics over temperate regions with the absence of lowland rivers (e.g., southern Brazil and southern Chile) and outlines specific locations where the continental approach may be somewhat limited. For example, correlation is slightly reduced for the continental model (–0.1 < d_r < –0.2) in areas over the Parnaiba basin and Chile, while a marked decrease in timing performance (d_Abs (DI) < –20 days) is observed in dry Argentinian rivers like Salado (southwest of La Plata basin) and Desaguadero (Colorado basin). Poor estimates of river geometry and large overestimation of flows in these regions may be causing excessive flooded areas and consequent peak attenuation.

Regarding intermediate to low flows, considerable differences in model performance are observed mainly over specific rivers in East Brazil and parts of Amazon basin near the Andes Cordillera (d_NSElog < –0.8), as well as in regions over southern Chile (–0.2 < d_NSElog < –0.6) that are potentially affected by snowmelt.

(Figure 10)

Table 3 shows differences in median discharge statistics for each global model and also for the Ensemble GM in comparison to the MGB continental model. Because LISFLOOD and WaterGAP3 account for reservoir impacts in their model structure, gauge stations with naturalized discharge data ($n = 98$) were excluded from the analysis to provide a fair assessment. Except

10 for the Ensemble GM, differences in performance regarding each pair of models (i.e., difference between continental and global) are quite similar for KGE (~0.45) and NSElog (~0.5), while being highly variable for both NSE (~1 to ~1.8) and bias (~4 % to ~30 %). Differences in median DI are between 1 and 2 days, which can be important for cases where flood timing is around this order of magnitude. Among the estimates from global models only, the Ensemble GM outperforms four out of the six metrics analyzed (KGE, NSE, NSElog and DI) with correlation (d_r = 0.03) equivalent to the best of global models

15 for this metric (LISFLOOD, d_r = 0.02). A reduction in performance occurs only when bias is evaluated, where 50 % of the gauge stations have an absolute difference equal to or greater than 11 % compared with differences in HTESSEL/CaMa-Flood (d_Abs (BIAS) ≈ 8 %) and WaterGAP3 (d_Abs (BIAS) ≈ 4 %). In the assessment by Beck et al., (2017a) for basins < 10 000 km² around the world, LISFLOOD also had an advantage in correlation when compared with other global models, while WaterGAP3 demonstrated problems related to baseflow index, which may be indicated here by the largest difference

20 of NSElog (d_NSElog ≈ 0.6).

(Table 3)

The set of boxplots shown in Fig. 11 summarizes the individual performance of continental and global models. Results are

25 presented for some of the representative South America basins and for the entire continental region, using a subset of metrics (KGE, NSE, BIAS and DI). In addition, a further analysis of the continental model performance was carried out using a few degraded configurations: hydrodynamic routing with calibrated rainfall-runoff parameters, i.e., the reference simulation (MGB_HD_calib), hydrodynamic routing with uncalibrated rainfall-runoff parameters (MGB_HD_noCalib), non-hydrodynamic routing with calibrated rainfall-runoff parameters (MGB_noHD_calib) and non-hydrodynamic routing with

30 uncalibrated rainfall-runoff parameters (MGB_noHD_noCalib). For the uncalibrated MGB versions (noCalib), a single set of parameters was adopted corresponding to the median values resulting from model adjustment (as shown in sect. 3.3). It is important to note that only rainfall-runoff parameters were reduced to their median values, while river routing parameters (Manning coefficient and river geometries) remained unchanged.

(Figure 11)

Results indicate that global models have important limitations in representing daily discharges in South America. In absolute terms, more than 40 % (60 %) of the gauge stations show negative or close to zero KGE (NSE) values. These models tend to overestimate discharges in the continent, with median bias ranging between +10 % and +50 %. In general, the performance among global models is variable according to the analyzed region and metric, which is supported by the large boxplot ranges. None of the models has a clear advantage with respect to all statistics, and this is especially valid for NSE and KGE. In the Amazon, KGE values present a more uniform pattern than in other regions with a median value close to 0.5, while models agree in a reasonable number of positive KGE and NSE values. Performance in the La Plata basin is highly variable between models, and this is the only region in which both systematic underestimation (HTESSEL/CaMa-Flood with median BIAS ≈ –20 %) and overestimation (LISFLOOD with median BIAS ≈ +20 %) is observed. Moreover, performance of the global models in basins with semiarid regions (e.g., Sao Francisco and Parnaiba) is extremely poor for KGE and NSE (median < –1 and < –2 respectively), which is probably associated to a dramatic overestimation of flows in these regions. WaterGAP3 shows a lower bias for all basins, but simulated peak flows occur too early according to DI for South America (percentile 25 % of DI ≈ –10 days). On the other hand, LISFLOOD appears to have a systematic delay in flow timing with more pronounced values over the Amazon (median DI ≈ +10 days), and a strong wet bias. For instance, median values of bias in LISFLOOD are larger than 40 % for the entire continent and exceed 100 % in basins such as Sao Francisco and Parnaiba. Absolute DI values are generally lower for HTESSEL/CaMa-Flood (median DI closer to 0) and this model usually shows an intermediate performance with respect to other metrics in comparison to LISFLOOD and WaterGAP3. Furthermore, the Ensemble GM shows a better overall performance when compared with each of these models alone, but still producing a similar number of negative KGE and NSE values (33 % and 60 % of the gauges, respectively).

Simulated discharges after setting MGB with a single set of rainfall-runoff parameters (MGB_noHD_noCalib) results in positive values for both KGE and NSE median values, varying between 0.3 and 0.6 for the entire continent. The uncalibrated version of the continental model outperforms global models in South America except for basins with semiarid regions (e.g., Sao Francisco and Parnaiba), where performances seem to be very dependent on parameter adjustment. The introduction of hydrodynamic routing (MGB_HD_noCalib) causes a slight improvement in NSE and KGE but this effect is more evident in the Amazon and especially over the La Plata basin (percentile 25 % of NSE changes from –1.5 to 0). Improvements in flow timing (DI) for both Amazon and La Plata are also observed after including the HD routing method, although excessive delays occur in Sao Francisco and Parnaiba because of the large bias that leads to an excess of floodplain attenuation (see MGB_HD_noCalib × MGB_HD_Calib). Furthermore, boxplot ranges are considerably smaller for KGE, NSE and bias in the default MGB simulation (MGB_HD_Calib) with respect to global models, and this reduction can be mostly seen in both MGB calibrated versions.

Our results agree with other studies from the literature, highlighting the large influence of model structure and parameterization in addition to meteorological forcing (e.g., Haddeland et al., 2011; Gudmunsson et al., 2012; Zhou et al., 2012; Beck et al., 2017a). Regarding global models, other studies also stress the large number of negative NSE values resulting from LSMs and GHMs in many basins around the world (e.g., Zhang et al., 2016; Beck et al., 2017a), including

5   South America (e.g., Wu et al., 2014). In particular, WaterGAP3 is expected to produce a lower bias because it is calibrated in terms of mean annual flow (Döll et al., 2009; Müller Schmied et al., 2014), and the systematic advance in timing is probably caused by the simple variable velocity equation (based on Manning) used for computing flow routing. In the case of LISFLOOD, large overestimation of flows in comparison to other global models has already been reported in the context of the eartH2Observe project (Beck et al., 2017a), where it showed the lowest estimate of potential evaporation (Schellekens

10   et al., 2017). Indeed, this excessive wet bias is one of the possible reasons for the observed delay in flows, but this is not the only factor because large overestimations are concomitantly found in regions where DI is negative. This is the case of eastern Amazon tributaries located downstream of Obidos, such as the Tapajos and Xingu (see Supplementary Material S3). The interplay between Manning coefficient and groundwater parameters and their influence on flow timing of LISFLOOD has been shown in recent studies (e.g., Revilla-Romero et al., 2015; Zajac et al., 2017), and here may compensate for

15   limitations of the double-kinematic wave (channel + floodplain) used for river routing, especially in the Amazon. This suggests that calibrating large basins with lowland river systems using few downstream stations (such as Obidos) should be taken with care if hydrodynamic routing is not accounted for in the model structure.

Although some authors pointed to a clear (and general) underestimation of HTESSEL (e.g., Haddeland et al., 2011; Beck et al., 2017a), our results showed that it occurs only in the La Plata Basin but not in other regions, which may be related to the

20   precipitation forcing used in WRR-2 (MSWEP). With respect to model performances, the relatively better flow timing of HTESSEL can be attributed to CaMa-Flood routing, but the advantages of this coupling were below the expected ones when looking at other statistics. It is worth mentioning that default parameters for river routing were used within WRR-2 simulations, i.e., the CaMa-Flood model was not calibrated (Martinez-de la Torre et al., 2018). This could be one of the reasons that low values of NSE are found over the Amazon main stem for this model (see Supplementary Material S3).

25   Characteristics such as timing and magnitude of flood waves in the hydrodynamic routing are very sensitive to channel geometry and roughness (Yamazaki et al., 2011; Paiva et al., 2013), but also to DEM vegetation effects (Baugh et al., 2013) that can impact the subgrid floodplain profiles (Paiva et al., 2011; Yamazaki et al., 2011). Furthermore, Zhao et al., (2017) emphasize that the benefit of CaMa-Flood highly depends on the runoff fields simulated by the coupled LSM. Our results showed that discharge estimates of an uncalibrated model are improved over Amazon and (mainly) La Plata basins after

30   inclusion of hydrodynamic routing (MGB_HD_noCalib), provided that channel geometry and floodplain topography are reasonably well estimated. Apart from the particular issues of WaterGAP3, LISFLOOD and HTESSEL/CaMa-Flood over South America, our findings reinforce the conclusion of other authors who recommend the ensemble mean of global models as the most reliable estimate (e.g., Haddeland et al., 2011; Gudmunsson et al., 2012; Schellekens et al., 2017; Hattermann et al., 2018), and it occurs even when discharges of a small number of models are averaged (three in the present case).

As outlined by many studies (e.g., Haddeland et al., 2011; Gudmundsson et al., 2012; Pechlivanidis and Arheimer, 2015; Zhang et al., 2016; Krysanova et al., 2017), performance of both regional and global models generally reduces when there is a transition from wet to dry conditions. In semiarid regions, satellites have several limitations in capturing rainfall intensities due to the local, convective nature of the precipitation, and they often overestimate the occurrence of rainfall because raindrops are likely to evaporate (i.e., sub-cloud evaporation) before reaching the surface (Dinku et al., 2010; Sunilkumar et al., 2015; Beck et al., 2017c). In addition, runoff generation mechanisms are strongly nonlinear and depend too much on storage processes, which are parameterized with large uncertainty (Gudmundsson et al., 2012). For instance, there is little knowledge about the influence of transmission losses, their partitioning between its main components (e.g., infiltration/evaporation from channels or floodplains) (Jarihani et al., 2015) and the dominant mechanisms of losing/gaining water according to different periods of the wet season (Costa et al., 2013). Processes such as reinfiltration of surface runoff, lateral redistribution of subsurface runoff and hydraulic-connected stream-aquifer interactions have been shown to be necessary for hydrological modeling in Northeast Brazil (Güntner and Bronstert, 2004; Costa et al., 2012, 2013), but are not explicitly accounted for in any structure of the assessed models. Therefore, a systematic underestimation of continental ET and consequent overestimation of flows is expected in dry regions (e.g., Alkama et al., 2010; Haddeland et al., 2011; Zhang et al., 2016). Uncertainties about human interferences in water resources (i.e., small ponds and reservoirs, water abstractions) may also play an important role (Hanasaki et al., 2018), especially in regions where data are scarce. Nevertheless, the complexity of the global models assessed herein makes it difficult to explain the real factors that impact discharge estimates. Model resolution can partially explain performances in headwater catchments or in areas with complex orography. In addition to issues related to the quality of satellite products (Tian and Peters-Lidard, 2010), aspects that potentially affect these regions are the shape (grid or unit-catchment) and size of computational elements, as well as the downscaling method of rainfall fields to force the hydrological models (Rahman et al., 2009), which is in our case a simple inverse distance weighting interpolation. It is worth noting that we have used a drainage area threshold of 1000 km² for headwater catchments (i.e., the onset of drainage networks), while this same area for global models varies between ~100 and ~625 km² according to their respective grid resolutions (between 0.08333º to 0.25º). In addition, results of discharge were evaluated at gauges monitoring at least 10 000 km², meaning that at these points the continental model has at least ~20 unit-catchments forced by ~16 MSWEP pixels. On the other hand, response to precipitation at these smaller basins may occur at hourly scale, while model forcing and analyses were performed at daily scale.

6 Summary and conclusions

For the first time, a regional-scale, fully coupled hydrologic–hydrodynamic model (MGB) was applied to a continental domain (South America). Model results were assessed using observed discharges and water levels from both in situ/satellite altimetry at an unprecedented gauge network over the continent, together with estimates of TWS and ET from remote sensing and other data sources. In addition, a cross-scale assessment (i.e., regional/continental × global models, the latter

acquired from the eartH2Observe project) was conducted with the novelty of using spatially distributed, daily discharge data for model comparison.

Regarding continental modeling, analyses showed a satisfactory agreement between simulated and observed discharges, with NSE (KGE) > 0.6 in more than 55 % (70 %) of the gauges. The performance was generally better in large rivers and humid regions, but worse in areas with semiarid to arid climates, influence of snowmelt or draining complex orography such as the Andes Cordillera. Similar results were found for water levels (both in situ and satellite altimetry), despite having 50 % of the gauges with large under- and overestimation of amplitudes (> |30 %|). The model was able to capture patterns of seasonality and magnitude of ET and TWS in many parts of the continent, especially when results were averaged to the scale of large South American basins (e.g., Amazon, La Plata, Orinoco, Tocantins–Araguaia, Sao Francisco, Parnaiba). In addition, model errors in simulating discharges were also found in other hydrological variables, which demonstrate the importance of assessing model results using multiple data sources. Uncertainties were attributed to deficiencies in process representation and simplifications in parameterization, as well as to limitations of the datasets used as model input and validation.

The cross-scale comparison shed light on what extent it is possible to improve discharge estimates in South America. Global models (HTESSEL/CaMa-Flood, LISFLOOD and WaterGAP3) presented negative NSE values in a large number of gauges (> 60 %) and showed highly variable performances when evaluated over multiple gauges within large basins. A considerable improvement in performance was found when the continental model was compared with individual global models, reaching median differences of around 0.45 for both KGE and NSElog, being larger than unity for NSE. By using the ensemble mean of global models as their best estimate, large differences in absolute bias (> 60 %) were detected mainly in East/Northeast Brazil and regions over Argentina, as well as in Sao Francisco, Parnaiba and Magdalena basins. Differences in timing of more than 20 days were found in rivers with floodplain effects, such as the Amazon, La Plata, Tocantins–Araguaia, Orinoco and lower Magdalena. Nevertheless, global models demonstrated a good ability to predict daily discharges over temperate, humid regions with the absence of lowland rivers (e.g., southern Brazil and south Chile), while performing reasonably in the Amazon basin.

The analyses also showed that model calibration and hydrodynamic routing cannot be neglected if simulation of daily discharges in this continent is desired. Calibration was found to be a key factor to model performance in most regions but mainly in drier basins (e.g., Parnaiba and Sao Francisco), where models generally fail to represent the underlying hydrological processes. In addition, a hydrodynamic routing module was essential to achieve a suitable representation of both magnitude and timing in major river systems, especially in cases where flows are dramatically attenuated by floodplains (e.g., the Paraguay River). However, the expected benefit of coupling hydrodynamic and hydrological models occurs when river geometries are reasonably well represented and calibration of rainfall-runoff parameters is performed together. This must be conducted by looking to different flow signatures (low and high flows, bias, timing of hydrographs) in a spatially distributed way, i.e., not considering only a single downstream gauge of large basins to reduce potential issues related to parameter compensation. As many of the approaches used in this study are applicable to global models, our findings suggest

that large improvements on estimated discharges can be achieved by the latter even without a significant increase in the number of computational elements.

Regardless of the scale (global or continental), limitations still remain in some regions of South America and can be explored as "stress tests" in model evaluation studies. For instance, characteristics of the La Plata basin such as complex floodplains, extensive rivers with mild slopes, significant reservoir regulation and existence of several climatic zones make it a unique test bed for model assessments. In addition, large basins with semiarid conditions located at Northeast Brazil (e.g., Parnaiba) or also with snowmelt-driven regimes at the South of the continent (e.g., Colorado) are interesting examples to stress model performance. Despite the cross-scale assessments conducted herein encompassed only river discharges, we recommend that other studies should also include in situ water levels and observation inferred variables (e.g., TWS from GRACE, ET from remote sensing / global datasets, and satellite altimetry) when possible. To facilitate access to in situ data, we call for cooperation among South American countries to produce a continental dataset that can be used by a broader audience, thus contributing to reduce the gap between regional and global modeling communities.

Finally, this study shows that extending a regional, fully coupled hydrologic-hydrodynamic model to the whole South American continent is feasible. This underscores the importance of regional expert knowledge, which can indicate relevant hydrological processes and datasets to be included in continental/global model simulations. We hope that moving from regional toward continental hydrologic–hydrodynamic modeling will bring new opportunities for operational practices such as real-time hydrological forecasting, which is the topic of an ongoing research. Nevertheless, several improvements should be carried out in the model structure not only to achieve a better understanding of the underlying processes but also to provide further insights about human impacts on South American water resources. This includes the representation of reservoirs, lakes and water abstractions. Uncertainties in model parameters are also important to be addressed and should be further investigated.

Data availability. Results from the MGB model are available to the public at http://www.ufrgs.br/lsh. All other datasets used in the present study can be accessed using the websites cited in this manuscript.

Author Contributions. V. A. S. worked on the MGB model for South America, designed the study, performed model calibration/simulations, managed all datasets (in situ, remote sensing and global models) and produced the manuscript; R. C. D. P. helped on study design, MGB code optimization, interpretation of results, contributed to writing of the manuscript and conducted paper review; A. S. F. and F. M. F. assisted on MGB calibration and conducted paper review; A. L. R. helped with the interpretation of results. P. R. M. P. provided support with GIS routines and MGB simulations over La Plata basin; A. P. and S. C. provided support with satellite altimetry data and associated analyses; W. C. was the research supervisor, contributed to writing of the manuscript and conducted paper review.

Competing interests. The authors declare that they have no conflict of interests.

Acknowledgments. The first author would like to acknowledge the Brazilian National Council for Scientific Research (CNPq) for the financial support (Project: "South America Flood Awareness System – SAFAS", under grant number
5   422422/2016–9); the institutions ANA, SENAMHIs, IDEAM, DGA, INA, HyBAM, GRDC and ONS for providing in situ discharge and water stage data; the LEGOS–HydroWeb for both Envisat and Jason-2 satellite altimetry data; the Jet Propulsion Laboratory (JPL) for GRACE mascon data; Princeton University for the CDR–ET data; the eartH2Observe project for WRR-2 discharges from global models (WaterGAP3, LISFLOOD and HTESSEL/CaMa-Flood) and, in particular, Martina Floerke, Gabriel Fink, Ad de Roo, Emanuel Dutra and Gianpaolo Balsamo for their kind support with the
10   global models used herein.

References

Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H., and Kløve, B.: A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model, Journal of Hydrology, 524, 733-752, 10.1016/j.jhydrol.2015.03.027, 2015.

15   Alkama, R., Decharme, B., Douville, H., Becker, M., Cazenave, A., Sheffield, J., Voldoire, A., Tyteca, S., and Le Moigne, P.: Global evaluation of the ISBA-TRIP continental hydrological system. Part I: Comparison to GRACE terrestrial water storage estimates and in situ river discharges, Journal of Hydrometeorology, 11, 583-600, 10.1175/2010JHM1211.1, 2010.

Allasia, D. G., Da Silva, B. C., Collischonn, W., and Tucci, C. E. M.: Large basin simulation experience in South America,
20   IAHS-AISH Publication, 2006, 360-370,

Alsdorf, D., Bates, P., Melack, J., Wilson, M., and Dunne, T.: Spatial and temporal complexity of the Amazon flood measured from space, Geophysical Research Letters, 34, 10.1029/2007GL029447, 2007.

Andreadis, K. M., Schumann, G. J. P., and Pavelsky, T.: A simple global river bankfull width and depth database, Water Resources Research, 49, 7164-7168, 10.1002/wrcr.20440, 2013.

25   Angarita, H., Wickel, A. J., Sieber, J., Chavarro, J., Maldonado-Ocampo, J. A., Herrera-R, G. A., Delgado, J., and Purkey, D.: Basin-scale impacts of hydropower development on the Mompós Depression wetlands, Colombia, Hydrology and Earth System Sciences, 22, 2839-2865, 10.5194/hess-22-2839-2018, 2018.

Archfield, S. A., Clark, M., Arheimer, B., Hay, L. E., McMillan, H., Kiang, J. E., Seibert, J., Hakala, K., Bock, A., Wagener, T., Farmer, W. H., Andréassian, V., Attinger, S., Viglione, A., Knight, R., Markstrom, S., and Over, T.: Accelerating
30   advances in continental domain hydrologic modeling, Water Resources Research, 51, 10078-10091, 10.1002/2015wr017498, 2015.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., and Betts, A. K.: A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System, Journal of Hydrometeorology, 10, 623-643, 10.1175/2008jhm1068.1, 2009.

Barros, V., Clarke, R., and Dias, P. S.: Climate change in the La Plata basin, Publication of the Inter-American Institute for Global Change Research (IAI), São José dos Campos, Brazil, 2006.

Bates, P. D., Horritt, M. S., and Fewtrell, T. J.: A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling, Journal of Hydrology, 387, 33-45, 10.1016/j.jhydrol.2010.03.027, 2010.

Baugh, C. A., Bates, P. D., Schumann, G., and Trigg, M. A.: SRTM vegetation removal and hydrodynamic modeling accuracy, Water Resources Research, 49, 5276-5289, 10.1002/wrcr.20412, 2013.

Beck, H. E., van Dijk, A. I. J. M., de Roo, A., Dutra, E., Fink, G., Orth, R., and Schellekens, J.: Global evaluation of runoff from 10 state-of-the-art hydrological models, Hydrology and Earth System Sciences, 21, 2881-2903, 10.5194/hess-21-2881-2017, 2017a.

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo, A.: MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data, Hydrology and Earth System Sciences, 21, 589-615, 10.5194/hess-21-589-2017, 2017b.

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P., Brocca, L., Pappenberger, F., Huffman, G. J., and Wood, E. F.: Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling, Hydrology and Earth System Sciences, 21, 6201-6217, 10.5194/hess-21-6201-2017, 2017c.

Beighley, R. E., and Gummadi, V.: Developing channel and floodplain dimensions with limited data: A case study in the Amazon Basin, Earth Surface Processes and Landforms, 36, 1059-1071, 10.1002/esp.2132, 2011.

Berbery, E. H., and Barros, V. R.: The hydrologic cycle of the La Plata basin in South America, Journal of Hydrometeorology, 3, 630-645, 10.1175/1525-7541(2002)003<0630:THCOTL>2.0.CO;2, 2002.

Berry, P. A. M., Garlick, J. D., and Smith, R. G.: Near-global validation of the SRTM DEM using satellite radar altimetry, Remote Sensing of Environment, 106, 17-27, 10.1016/j.rse.2006.07.011, 2007.

Bierkens, M. F. P.: Global hydrology 2015: State, trends, and directions, Water Resources Research, 51, 4923-4947, 10.1002/2015wr017173, 2015.

Bierkens, M. F. P., Bell, V. A., Burek, P., Chaney, N., Condon, L. E., David, C. H., de Roo, A., Döll, P., Drost, N., Famiglietti, J. S., Flörke, M., Gochis, D. J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R. M., Reager, J. T., Samaniego, L., Sudicky, E., Sutanudjaja, E. H., van de Giesen, N., Winsemius, H., and Wood, E. F.: Hyper-resolution global hydrological modelling: what is next?, Hydrological Processes, 29, 310-320, 10.1002/hyp.10391, 2015.

Bravo, J. M., Allasia, D., Paz, A. R., Collischonn, W., and Tucci, C. E. M.: Coupled Hydrologic-Hydraulic Modeling of the Upper Paraguay River Basin, Journal of Hydrologic Engineering, 17, 635-646, 10.1061/(ASCE)HE.1943-5584.0000494, 2012.

Carabajal, C. C., and Harding, D. J.: ICESat validation of SRTM C-band digital elevation models, Geophysical Research Letters, 32, 5, 10.1029/2005gl023957, 2005.

Chen, J. L., Wilson, C. R., Tapley, B. D., Blankenship, D. D., and Ivins, E. R.: Patagonia icefield melting observed by gravity recovery and climate experiment (GRACE), Geophysical Research Letters, 34, 10.1029/2007gl031871, 2007.

Christoffersen, B. O., Restrepo-Coupe, N., Arain, M. A., Baker, I. T., Cestaro, B. P., Ciais, P., Fisher, J. B., Galbraith, D., Guan, X. D., Gulden, L., van den Hurk, B., Ichii, K., Imbuzeiro, H., Jain, A., Levine, N., Miguez-Machor, G., Poulter, B., Roberti, D. R., Sakaguchi, K., Sahoo, A., Schaefer, K., Shi, M. J., Verbeeck, H., Yang, Z. L., Araujo, A. C., Kruijt, B., Manzi, A. O., da Rocha, H. R., von Randow, C., Muza, M. N., Borak, J., Costa, M. H., de Goncalves, L. G. G., Zeng, X. B., and Saleska, S. R.: Mechanisms of water supply and vegetation demand govern the seasonality and magnitude of evapotranspiration in Amazonia and Cerrado, Agricultural and Forest Meteorology, 191, 33-50, 10.1016/j.agrformet.2014.02.008, 2014.

Clark, E. A., Sheffield, J., van Vliet, M. T. H., Nijssen, B., and Lettenmaier, D. P.: Continental Runoff into the Oceans (1950-2008), Journal of Hydrometeorology, 16, 1502-1520, 10.1175/jhm-d-14-0183.1, 2015.

Collischonn, W.: Simulação Hidrológica de Grandes Bacias, Phd Dissertation (in portuguese), Doutorado em Recursos Hídricos e Saneamento Ambiental, Instituto de Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, 270 pp., 2001.

Collischonn, W., Haas, R., Andreolli, I., and Tucci, C. E. M.: Forecasting River Uruguay flow using rainfall forecasts from a regional weather-prediction model, Journal of Hydrology, 305, 87-98, 10.1016/j.jhydrol.2004.08.028, 2005.

Collischonn, W., Allasia, D., Da Silva, B. C., and Tucci, C. E. M.: The MGB-IPH model for large-scale rainfall-runoff modelling, Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 52, 878-895, 10.1623/hysj.52.5.878, 2007.

Costa, A. C., Bronstert, A., and De Araújo, J. C.: A channel transmission losses model for different dryland rivers, Hydrology and Earth System Sciences, 16, 1111-1135, 10.5194/hess-16-1111-2012, 2012.

Costa, A. C., Foerster, S., de Araújo, J. C., and Bronstert, A.: Analysis of channel transmission losses in a dryland river reach in north-eastern Brazil using streamflow series, groundwater level series and multi-temporal satellite data, Hydrological Processes, 27, 1046-1060, 10.1002/hyp.9243, 2013.

Crooks, S. M., Kay, A. L., Davies, H. N., and Bell, V. A.: From Catchment to National Scale Rainfall-Runoff Modelling: Demonstration of a Hydrological Modelling Framework, Hydrology, 1, 63-88, 10.3390/hydrology1010063, 2014.

Decharme, B., Douville, H., Prigent, C., Papa, F., and Aires, F.: A new river flooding scheme for global climate applications: Off-line evaluation over South America, Journal of Geophysical Research-Atmospheres, 113, 11, 10.1029/2007jd009376, 2008.

Dinku, T., Ceccato, P., Cressman, K., and Connor, S. J.: Evaluating detection skills of satellite rainfall estimates over desert locust recession regions, Journal of Applied Meteorology and Climatology, 49, 1322-1332, 10.1175/2010JAMC2281.1, 2010.

Doll, P., Fiedler, K., and Zhang, J.: Global-scale analysis of river flow alterations due to water withdrawals and reservoirs, Hydrology and Earth System Sciences, 13, 2413-2432, 2009.

Donnelly, C., Andersson, J. C. M., and Arheimer, B.: Using flow signatures and catchment similarities to evaluate the E-HYPE multi-basin model across Europe, Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 61, 255-273, 10.1080/02626667.2015.1027710, 2016.

Durr, H. H., Meybeck, M., and Durr, S. H.: Lithologic composition of the Earth's continental surfaces derived from a new digital map emphasizing riverine material transfer, Global Biogeochemical Cycles, 19, 10.1029/2005gb002515, 2005.

Eisner, S., Florke, M., Chamorro, A., Daggupati, P., Donnelly, C., Huang, J., Hundecha, Y., Koch, H., Kalugin, A., Krylenko, I., Mishra, V., Piniewski, M., Samaniego, L., Seidou, O., Wallner, M., and Krysanova, V.: An ensemble analysis of climate change impacts on streamflow seasonality across 11 large river basins, Climatic Change, 141, 401-417, 10.1007/s10584-016-1844-5, 2017.

Erfanian, A., Wang, G., and Fomenko, L.: Unprecedented drought over tropical South America in 2016: Significantly under-predicted by tropical SST, Scientific Reports, 7, 10.1038/s41598-017-05373-2, 2017.

Fan, F. M., Buarque, D. C., Pontes, P. R. M., and Collischonn, W.: Um mapa de unidades de resposta hidrológica para a América do Sul, XXI Simpósio Brasileiro de Recursos Hídricos, Brasilia, 2015, PAP019919,

Fan, F. M., Collischonn, W., Quiroz, K. J., Sorribas, M. V., Buarque, D. C., and Siqueira, V. A.: Flood forecasting on the Tocantins River using ensemble rainfall forecasts and real-time satellite rainfall estimates, Journal of Flood Risk Management, 9, 278-288, 10.1111/jfr3.12177, 2016.

Fleischmann, A. S., Siqueira, V. A., Paris, A., Collischonn, W., Paiva, R. C. D., Pontes, P. R. M., Crétaux, J.-F., Bergé-Nguyen, M., Biancamaria, S., Gosset, M., Calmant, S., and Tanimoun, B. A.: Modeling hydrologic and hydrodynamic processes in basins with semi-arid wetlands, Journal of Hydrology, 10.1016/j.jhydrol.2018.04.041, in press, 2018.

Frappart, F., Papa, F., Malbeteau, Y., León, J., Ramillien, G., Prigent, C., Seoane, L., Seyler, F., and Calmant, S.: Surface Freshwater Storage Variations in the Orinoco Floodplains Using Multi-Satellite Observations, Remote Sensing, 7, 89-110, 10.3390/rs70100089, 2014.

Garcia, N. O., and Mechoso, C. R.: Variability in the discharge of South American rivers and in climate, Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 50, 459-478, 10.1623/hysj.50.3.459.65030, 2005.

Garreaud, R. D., Vuille, M., Compagnucci, R., and Marengo, J.: Present-day South American climate, Palaeogeography Palaeoclimatology Palaeoecology, 281, 180-195, 10.1016/j.palaeo.2007.10.032, 2009.

Getirana, A. C. V.: Integrating spatial altimetry data into the automatic calibration of hydrological models, Journal of Hydrology, 387, 244-255, 10.1016/j.jhydrol.2010.04.013, 2010.

Getirana, A. C. V., Boone, A., Yamazaki, D., Decharme, B., Papa, F., and Mognard, N.: The Hydrological Modeling and Analysis Platform (HyMAP): Evaluation in the Amazon Basin, Journal of Hydrometeorology, 13, 1641-1665, 10.1175/jhm-d-12-021.1, 2012.

Getirana, A. C. V., Boone, A., Yamazaki, D., and Mognard, N.: Automatic parameterization of a flow routing scheme driven by radar altimetry data: Evaluation in the Amazon basin, Water Resources Research, 49, 614-629, 10.1002/wrcr.20077, 2013.

Getirana, A. C. V., and Paiva, R. C. D.: Mapping large-scale river flow hydraulics in the Amazon Basin, Water Resources Research, 49, 2437-2445, 10.1002/wrcr.20212, 2013.

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, 2017a.

Getirana, A., Peters-Lidard, C., Rodell, M., and Bates, P. D.: Trade-off between cost and accuracy in large-scale surface water dynamic modeling, Water Resources Research, 53, 4942-4955, 10.1002/2017wr020519, 2017b.

Goncalves, H. C., Mercante, M. A., and Santos, E. T.: Hydrological cycle, Brazilian Journal of Biology, 71, 241-253, 10.1590/s1519-69842011000200003, 2011.

Gosling, S. N., Taylor, R. G., Arnell, N. W., and Todd, M. C.: A comparative analysis of projected impacts of climate change on river runoff from global and catchment-scale hydrological models, Hydrology and Earth System Sciences, 15, 279-294, 10.5194/hess-15-279-2011, 2011.

Gudmundsson, L., Tallaksen, L. M., Stahl, K., Clark, D. B., Dumont, E., Hagemann, S., Bertrand, N., Gerten, D., Heinke, J., Hanasaki, N., Voss, F., and Koirala, S.: Comparing Large-Scale Hydrological Model Simulations to Observed Runoff Percentiles in Europe, Journal of Hydrometeorology, 13, 604-620, 10.1175/jhm-d-11-083.1, 2012.

Guntner, A., and Bronstert, A.: Representation of landscape variability and lateral redistribution processes for large-scale hydrological modelling in semi-arid areas, Journal of Hydrology, 297, 136-161, 10.1016/j.jhydrol.2004.04.008, 2004.

Guswa, A. J., Celia, M. A., and Rodriguez-Iturbe, I.: Models of soil moisture dynamics in ecohydrology: A comparative study, Water Resources Research, 38, 10.1029/2001wr000826, 2002.

Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G. P., and Yeh, P.: Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results, Journal of Hydrometeorology, 12, 869-884, 10.1175/2011jhm1324.1, 2011.

Hamilton, S. K., Sippel, S. J., and Melack, J. M.: Comparison of inundation patterns among major South American floodplains, Journal of Geophysical Research-Atmospheres, 107, 10.1029/2000jd000306, 2002.

Hanasaki, N., Yoshikawa, S., Pokhrel, Y., and Kanae, S.: A global hydrological simulation to specify the sources of water used by humans, Hydrology and Earth System Sciences, 22, 789-817, 10.5194/hess-22-789-2018, 2018.

Hattermann, F. F., Vetter, T., Breuer, L., Su, B. D., Daggupati, P., Donnelly, C., Fekete, B., Florke, F., Gosling, S. N., Hoffmann, P., Liersch, S., Masaki, Y., Motovilov, Y., Muller, C., Samaniego, L., Stacke, T., Wada, Y., Yang, T., and Krysnaova, V.: Sources of uncertainty in hydrological climate impact assessment: a cross-scale study, Environmental Research Letters, 13, 10.1088/1748-9326/aa9938, 2018.

Hodges, B. R.: Challenges in Continental River Dynamics, Environmental Modelling & Software, 50, 16-20, 10.1016/j.envsoft.2013.08.010, 2013.

Hoyos, N., Escobar, J., Restrepo, J. C., Arango, A. M., and Ortiz, J. C.: Impact of the 2010-2011 La Nina phenomenon in Colombia, South America: The human toll of an extreme weather event, Applied Geography, 39, 16-25, 10.1016/j.apgeog.2012.11.018, 2013.

Huang, S., Kumar, R., Flörke, M., Yang, T., Hundecha, Y., Kraft, P., Gao, C., Gelfan, A., Liersch, S., Lobanova, A., Strauch, M., van Ogtrop, F., Reinhardt, J., Haberlandt, U., and Krysanova, V.: Evaluation of an ensemble of regional hydrological models in 12 large-scale river basins worldwide, Climatic Change, 141, 381-397, 10.1007/s10584-016-1841-8, 2016.

Humphrey, V., Gudmundsson, L., and Seneviratne, S. I.: Assessing Global Water Storage Variability from GRACE: Trends, Seasonal Cycle, Subseasonal Anomalies and Extremes, Surv Geophys, 37, 357-395, 10.1007/s10712-016-9367-1, 2016.

Jarihani, A. A., Larsen, J. R., Callow, J. N., McVicar, T. R., and Johansen, K.: Where does all the water go? Partitioning water transmission losses in a data-sparse, multi-channel and low-gradient dryland river system using modelling and remote sensing, Journal of Hydrology, 529, 1511-1529, 10.1016/j.jhydrol.2015.08.030, 2015.

Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P., and Thielen, J.: Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level, Environmental Modelling & Software, 75, 68-76, 10.1016/j.envsoft.2015.09.009, 2016.

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S. E., Leroux, D., Mialon, A., and Delwart, S.: The SMOS soil moisture retrieval algorithm, IEEE Transactions on Geoscience and Remote Sensing, 50, 1384-1403, 10.1109/TGRS.2012.2184548, 2012.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resources Research, 42, 10.1029/2005wr004362, 2006.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, Journal of Hydrology, 424-425, 264-277, 10.1016/j.jhydrol.2012.01.011, 2012.

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.

Krysanova, V., Vetter, T., Eisner, S., Huang, S., Pechlivanidis, I., Strauch, M., Gelfan, A., Kumar, R., Aich, V., Arheimer, B., Chamorro, A., Van Griensven, A., Kundu, D., Lobanova, A., Mishra, V., Plötner, S., Reinhardt, J., Seidou, O., Wang, X., Wortmann, M., Zeng, X., and Hattermann, F. F.: Intercomparison of regional-scale hydrological models and climate change impacts projected for 12 large river basins worldwide - A synthesis, Environmental Research Letters, 12, 10.1088/1748-9326/aa8359, 2017.

Kuppel, S., Houspanossian, J., Nosetto, M. D., and Jobbagy, E. G.: What does it take to flood the Pampas?: Lessons from a decade of strong hydrological fluctuations, Water Resources Research, 51, 2937-2950, 10.1002/2015wr016966, 2015.

Landerer, F. W., and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resources Research, 48, 11, 10.1029/2011wr011453, 2012.

Latrubesse, E. M., Stevaux, J. C., and Sinha, R.: Tropical rivers, Geomorphology, 70, 187-206, 10.1016/j.geomorph.2005.02.005, 2005.

5   Latrubesse, E. M.: Large rivers, megafans and other Quaternary avulsive fluvial systems: A potential "who's who" in the geological record, Earth-Science Reviews, 146, 1-30, 10.1016/j.earscirev.2015.03.004, 2015.

Lehner, B., and Döll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, Journal of Hydrology, 296, 1-22, 10.1016/j.jhydrol.2004.03.028, 2004.

Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation data, Eos, 89, 93-94,
10   10.1029/2008EO100001, 2008.

Leopold, L. B., and Maddock, T. J.: The hydraulic geometry of stream channels and some physiographic implications, U.S. Geological Survey Professional Paper, 252, 56p, 1953.

Lininger, K. B., and Latrubesse, E. M.: Flooding hydrology and peak discharge attenuation along the middle Araguaia River in central Brazil, Catena, 143, 90-101, 10.1016/j.catena.2016.03.043, 2016.

15   Lopez, P. L., Sutanudjaja, E. H., Schellekens, J., Sterk, G., and Bierkens, M. F. P.: Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products, Hydrology and Earth System Sciences, 21, 3125-3144, 10.5194/hess-21-3125-2017, 2017.

Luo, X. Y., Li, H. Y., Leung, L. R., Tesfa, T. K., Getirana, A., Papa, F., and Hess, L. L.: Modeling surface water dynamics in the Amazon Basin using MOSART-Inundation v1.0: impacts of geomorphological parameters and river flow
20   representation, Geoscientific Model Development, 10, 1233-1259, 10.5194/gmd-10-1233-2017, 2017.

Maeda, E. E., Ma, X. L., Wagner, F. H., Kim, H., Oki, T., Eamus, D., and Huete, A.: Evapotranspiration seasonality across the Amazon Basin, Earth System Dynamics, 8, 439-454, 10.5194/esd-8-439-2017, 2017.

Marengo, J. A., Jones, R., Alves, L. M., and Valverde, M. C.: Future change of temperature and precipitation extremes in south america as derived from the precis regional climate modeling system, International Journal of Climatology, 29,
25   2241-2255, 10.1002/joc.1863, 2009.

Marengo, J. A., Tomasella, J., Soares, W. R., Alves, L. M., and Nobre, C. A.: Extreme climatic events in the Amazon basin, Theoretical and Applied Climatology, 107, 73-85, 10.1007/s00704-011-0465-1, 2012.

Martinez-de la Torre, A. et al. Water Resources Reanalysis 2 ensemble of global land surface and hydrological models: the EartH2Observe Tier-2 dataset. In preparation.

30   Mateo, C. M. R., Yamazaki, D., Kim, H., Champathong, A., Vaze, J., and Oki, T.: Impacts of spatial resolution and representation of flow connectivity on large-scale simulation of floods, Hydrology and Earth System Sciences, 21, 5143-5163, 10.5194/hess-21-5143-2017, 2017.

Meade, R. H., Rayol, J. M., Daconceicao, S. C., and Natividade, J. R. G.: Backwater effects in the Amazon river basin of Brazil, Environmental Geology and Water Sciences, 18, 105-114, 10.1007/bf01704664, 1991.

Melo, D. D. D., Scanlon, B. R., Zhang, Z., Wendland, E., and Yin, L.: Reservoir storage and hydrologic responses to droughts in the Parana River basin, south-eastern Brazil, Hydrology and Earth System Sciences, 20, 4673-4688, 10.5194/hess-20-4673-2016, 2016.

Miralles, D. G., De Jeu, R. A. M., Gash, J. H., Holmes, T. R. H., and Dolman, A. J.: Magnitude and variability of land
5   evaporation and its components at the global scale, Hydrology and Earth System Sciences, 15, 967-981, 10.5194/hess-15-967-2011, 2011.

Mizukami, N., Clark, M. P., Newman, A. J., Wood, A. W., Gutmann, E. D., Nijssen, B., Rakovec, O., and Samaniego, L.: Towards seamless large-domain parameter estimation for hydrologic models, Water Resources Research, 53, 8020-8040, 10.1002/2017WR020401, 2017.

10   Müller Schmied, H., Eisner, S., Franz, D., Wattenbach, M., Portmann, F. T., Florke, M., and Doll, P.: Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration, Hydrology and Earth System Sciences, 18, 3511-3538, 10.5194/hess-18-3511-2014, 2014.

New, M., Lister, D., Hulme, M., and Makin, I.: A high-resolution data set of surface climate over global land areas, Climate Research, 21, 1-25, 2002.

15   O'Loughlin, F. E., Paiva, R. C. D., Durand, M., Alsdorf, D. E., and Bates, P. D.: A multi-sensor approach towards a global vegetation corrected SRTM DEM product, Remote Sensing of Environment, 182, 49-59, 10.1016/j.rse.2016.04.018, 2016.

Ovando, A., Tomasella, J., Rodriguez, D. A., Martinez, J. M., Siqueira-Junior, J. L., Pinto, G. L. N., Passy, P., Vauchel, P., Noriega, L., and von Randow, C.: Extreme flood events in the Bolivian Amazon wetlands, Journal of Hydrology:
20   Regional Studies, 5, 293-308, 10.1016/j.ejrh.2015.11.004, 2016.

Paiva, R. C. D., Collischonn, W., and Tucci, C. E. M.: Large scale hydrologic and hydrodynamic modeling using limited data and a GIS based approach, Journal of Hydrology, 406, 170-181, 10.1016/j.jhydrol.2011.06.007, 2011.

Paiva, R. C. D., Buarque, D. C., Collischonn, W., Bonnet, M. P., Frappart, F., Calmant, S., and Bulhões Mendes, C. A.: Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin, Water Resources Research, 49, 1226-
25   1243, 10.1002/wrcr.20067, 2013.

Pappenberger, F., Cloke, H. L., Balsamo, G., Ngo-Duc, T., and Oki, T.: Global runoff routing with the hydrological component of the ECMWF NWP system, International Journal of Climatology, 30, 2155-2174, 10.1002/joc.2028, 2010.

Paris, A., Dias de Paiva, R., Santos da Silva, J., Medeiros Moreira, D., Calmant, S., Garambois, P. A., Collischonn, W., Bonnet, M. P., and Seyler, F.: Stage-discharge rating curves based on satellite altimetry and modeled discharge in the
30   Amazon basin, Water Resources Research, 52, 3787-3814, 10.1002/2014WR016618, 2016.

Pasquini, A. I., and Depetris, P. J.: Discharge trends and flow dynamics of South American rivers draining the southern Atlantic seaboard: An overview, Journal of Hydrology, 333, 385-399, 10.1016/j.jhydrol.2006.09.005, 2007.

Paz, A. R., Collischonn, W., Tucci, C. E. M., and Padovani, C. R.: Large-scale modelling of channel flow and floodplain inundation dynamics and its application to the Pantanal (Brazil), Hydrological Processes, 25, 1498-1516, 10.1002/hyp.7926, 2011.

Paz, A. R., Collischonn, W., Bravo, J. M., Bates, P. D., and Baugh, C.: The influence of vertical water balance on modelling Pantanal (Brazil) spatio-temporal inundation dynamics, Hydrological Processes, 28, 3539-3553, 10.1002/hyp.9897, 2014.

Pechlivanidis, I. G., and Arheimer, B.: Large-scale hydrological modelling by using modified PUB recommendations: The India-HYPE case, Hydrology and Earth System Sciences, 19, 4559-4579, 10.5194/hess-19-4559-2015, 2015.

Pedinotti, V., Boone, A., Decharme, B., Crétaux, J. F., Mognard, N., Panthou, G., Papa, F., and Tanimoun, B. A.: Evaluation of the ISBA-TRIP continental hydrologic system over the Niger basin using in situ and satellite derived datasets, Hydrology and Earth System Sciences, 16, 1745-1773, 10.5194/hess-16-1745-2012, 2012.

Pontes, P. R. M.: Modelagem hidrológica e hidrodinâmica integrada da bacia do Prata, Phd Dissertation (in portuguese), Instituto de Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, 194 pp., 2016.

Pontes, P. R. M., Fan, F. M., Fleischmann, A. S., de Paiva, R. C. D., Buarque, D. C., Siqueira, V. A., Jardim, P. F., Sorribas, M. V., and Collischonn, W.: MGB-IPH model for hydrological and hydraulic simulation of large floodplain river systems coupled with open source GIS, Environmental Modelling & Software, 94, 1-20, 10.1016/j.envsoft.2017.03.029, 2017.

Rahman, S., Bagtzoglou, A. C., Hossain, F., Tang, L., Yarbrough, L. D., Easson, G.: Investigating Spatial Downscaling of Satellite Rainfall Data for Streamflow Simulation in a Medium-Sized Basin, Journal of Hydrometeorology, 10, 1063-1079, 10.1175/2009JHM1072.1, 2009.

Ramillien, G., Cazenave, A., and Brunau, O.: Global time variations of hydrological signals from GRACE satellite gravimetry, Geophysical Journal International, 158, 813-826, 10.1111/j.1365-246X.2004.02328.x, 2004.

Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., Tomasella, J., and Waterloo, M. J.: HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia, Remote Sensing of Environment, 112, 3469-3481, 10.1016/j.rse.2008.03.018, 2008.

Restrepo-Coupe, N., da Rocha, H. R., Hutyra, L. R., da Araujo, A. C., Borma, L. S., Christoffersen, B., Cabral, O. M. R., de Camargo, P. B., Cardoso, F. L., da Costa, A. C. L., Fitzjarrald, D. R., Goulden, M. L., Kruijt, B., Maia, J. M. F., Malhi, Y. S., Manzi, A. O., Miller, S. D., Nobre, A. D., von Randow, C., Sa, L. D. A., Sakai, R. K., Tota, J., Wofsy, S. C., Zanchi, F. B., and Saleska, S. R.: What drives the seasonality of photosynthesis across the Amazon basin? A cross-site analysis of eddy flux tower measurements from the Brasil flux network, Agricultural and Forest Meteorology, 182, 128-144, 10.1016/j.agrformet.2013.04.031, 2013.

Revilla-Romero, B., Beck, H. E., Burek, P., Salamon, P., de Roo, A., and Thielen, J.: Filling the gaps: Calibrating a rainfall-runoff model using satellite-derived surface water extent, Remote Sensing of Environment, 171, 118-131, 10.1016/j.rse.2015.10.022, 2015.

Richey, J. E., Mertes, L. A. K., Dunne, T., Victoria, R. L., Forsberg, B. R., Tancredi, A., and Oliveira, E.: Sources and routing of the Amazon River flood wave, Global Biogeochemical Cycles, 3, 191-204, 10.1029/GB003i003p00191, 1989.

Rivera, J. A., Penalba, O. C., Villalba, R., and Araneo, D. C.: Spatio-temporal patterns of the 2010-2015 extreme hydrological drought across the Central Andes, Argentina, Water (Switzerland), 9, 10.3390/w9090652, 2017.

Rivera, J. A., Araneo, D. C., Penalba, O. C., and Villalba, R.: Regional aspects of streamflow droughts in the Andean rivers of Patagonia, Argentina. Links with large-scale climatic oscillations, Hydrology Research, 49, 134-149, 10.2166/nh.2017.207, 2018.

Rodriguez, E., Morris, C. S., and Belz, J. E.: A global assessment of the SRTM performance, Photogrammetric Engineering and Remote Sensing, 72, 249-260, 10.14358/pers.72.3.249, 2006.

Rosales, J., Vispo, C., Dezzeo, N., Blanco-Belmonte, L., Knab-Vispo, C., González, N., Daza, F., Bradley, C., Gilvear, D., and Escalante, G.: Ecohydrology of riparian forests in the Orinoco River Basin, The ecohydrology of South American rivers and wetlands. Special Publication Nº6. IAHS Press. Wallingford, UK, 93-110, 2002.

Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., Eisner, S., Müller Schmied, H., Sutanudjaja, E., Warrach-Sagi, K., and Attinger, S.: Toward seamless hydrologic predictions across spatial scales, Hydrology and Earth System Sciences, 21, 4323-4346, 10.5194/hess-21-4323-2017, 2017.

Santos da Silva, J., Calmant, S., Seyler, F., Rotunno Filho, O. C., Cochonneau, G., and Mansur, W. J.: Water levels in the Amazon basin derived from the ERS 2 and ENVISAT radar altimetry missions, Remote Sensing of Environment, 114, 2160-2181, 10.1016/j.rse.2010.04.020, 2010.

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

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.

Schmidt, R., Flechtner, F., Meyer, U., Neumayer, K. H., Dahle, C., Konig, R., and Kusche, J.: Hydrological Signals Observed by the GRACE Satellites, Surveys in Geophysics, 29, 319-334, 10.1007/s10712-008-9033-3, 2008.

Siqueira, V. A., Collischonn, W., Fan, F. M., and Chou, S. C.: Ensemble flood forecasting based on operational forecasts of the regional Eta EPS in the Taquari-Antas basin, Rbrh, 21, 587-602, 10.1590/2318-0331.011616004, 2016a.

Siqueira, V. A., Fleischmann, A., Jardim, P. F., Fan, F. M., and Collischonn, W.: IPH-Hydro Tools: A GIS coupled tool for watershed topology acquisition in an open-source environment, Rbrh, 21, 274-287, 10.21168/rbrh.v21n1.p274-287, 2016b.

Sood, A., and Smakhtin, V.: Global hydrological models: a review, Hydrological Sciences Journal, 60, 549-565, 10.1080/02626667.2014.950580, 2015.

Sperna Weiland, F. C., Vrugt, J. A., van Beek, R. P. H., Weerts, A. H., and Bierkens, M. F. P.: Significant uncertainty in global scale hydrological modeling from precipitation data errors, Journal of Hydrology, 529, 1095-1115, 10.1016/j.jhydrol.2015.08.061, 2015.

Su, F. G., and Lettenmaier, D. P.: Estimation of the Surface Water Budget of the La Plata Basin, Journal of Hydrometeorology, 10, 981-998, 10.1175/2009jhm1100.1, 2009.

Sunilkumar, K., Narayana Rao, T., Saikranthi, K., and Purnachandra Rao, M.: Comprehensive evaluation of multisatellite precipitation estimates over India using gridded rainfall data, Journal of Geophysical Research, 120, 8987-9005, 10.1002/2015JD023437, 2015.

Tapley, B. D., Bettadpur, S., Watkins, M., and Reigber, C.: The gravity recovery and climate experiment: Mission overview and early results, Geophysical Research Letters, 31, 4, 10.1029/2004gl019920, 2004.

Tian, Y., and Peters-Lidard, C. D.: A global map of uncertainties in satellite-based precipitation measurements, Geophysical Research Letters, 37, 10.1029/2010GL046008, 2010.

Todini, E.: The ARNO rainfall-runoff model, Journal of Hydrology, 175, 339-382, 1996.

Trigg, M. A., Wilson, M. D., Bates, P. D., Horritt, M. S., Alsdorf, D. E., Forsberg, B. R., and Vega, M. C.: Amazon flood wave hydraulics, Journal of Hydrology, 374, 92-105, 10.1016/j.jhydrol.2009.06.004, 2009.

Tucci, C. E. M., and Clarke, R. T.: Environmental issues in the la Plata Basin, International Journal of Water Resources Development, 14, 157-173, 10.1080/07900629849376, 1998.

Van Der Knijff, J. M., Younis, J., and De Roo, A. P. J.: LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation, International Journal of Geographical Information Science, 24, 189-212, 10.1080/13658810802549154, 2010.

Vera, C., Baez, J., Douglas, M., Emmanuel, C. B., Marengo, J., Meitin, J., Nicolini, M., Nogues-Paegle, J., Paegle, J., Penalba, O., Salio, P., Saulo, C., Dias, M. A. S., Dias, P. S., and Zipser, E.: The South American low-level jet experiment, Bulletin of the American Meteorological Society, 87, 63-77, 10.1175/bams-87-1-63, 2006.

Vinukollu, R. K., Meynadier, R., Sheffield, J., and Wood, E. F.: Multi-model, multi-sensor estimates of global evapotranspiration: climatology, uncertainties and trends, Hydrological Processes, 25, 3993-4010, 10.1002/hyp.8393, 2011.

Wada, Y., Van Beek, L. P. H., and Bierkens, M. F. P.: Modelling global water stress of the recent past: On the relative importance of trends in water demand and climate variability, Hydrology and Earth System Sciences, 15, 3785-3808, 10.5194/hess-15-3785-2011, 2011.

Wahr, J., Molenaar, M., and Bryan, F.: Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, Journal of Geophysical Research: Solid Earth, 103, 30205-30229, 1998.

Wang, A., Zeng, X., Shen, S. S. P., Zeng, Q. C., and Dickinson, R. E.: Time scales of land surface hydrology, Journal of Hydrometeorology, 7, 868-879, 10.1175/JHM527.1, 2006.

Watkins, M. M., Wiese, D. N., Yuan, D. N., Boening, C., and Landerer, F. W.: Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons, Journal of Geophysical Research: Solid Earth, 120, 2648-2671, 10.1002/2014JB011547, 2015.

Werth, D., and Avissar, R.: The local and global effects of Amazon deforestation, Journal of Geophysical Research-Atmospheres, 107, 10.1029/2001jd000717, 2002.

Werth, S., and Guntner, A.: Calibration analysis for water storage variability of the global hydrological model WGHM, Hydrology and Earth System Sciences, 14, 59-78, 2010.

Wiese, D. N., Landerer, F. W., and Watkins, M. M.: Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution, Water Resources Research, 52, 7490-7502, 10.1002/2016wr019344, 2016.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Doll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffe, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P.: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, Water Resources Research, 47, 10.1029/2010wr010090, 2011.

Wu, H., Adler, R. F., Tian, Y. D., Huffman, G. J., Li, H. Y., and Wang, J. J.: Real-time global flood estimation using satellite-based precipitation and a coupled land surface and routing model, Water Resources Research, 50, 2693-2717, 10.1002/2013wr014710, 2014.

Yamazaki, D., Kanae, S., Kim, H., and Oki, T.: A physically based description of floodplain inundation dynamics in a global river routing model, Water Resources Research, 47, 10.1029/2010WR009726, 2011.

Yamazaki, D., De Almeida, G. A. M., and Bates, P. D.: Improving computational efficiency in global river models by implementing the local inertial flow equation and a vector-based river network map, Water Resources Research, 49, 7221-7235, 10.1002/wrcr.20552, 2013.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., and Bates, P. D.: A high-accuracy map of global terrain elevations, Geophysical Research Letters, 44, 5844-5853, 10.1002/2017GL072874, 2017.

Zaitchik, B. F., Rodell, M., and Olivera, F.: Evaluation of the Global Land Data Assimilation System using global river discharge data and a source-to-sink routing scheme, Water Resources Research, 46, 10.1029/2009wr007811, 2010.

Zajac, Z., Revilla-Romero, B., Salamon, P., Burek, P., Hirpa, F., and Beck, H.: The impact of lake and reservoir parameterization on global streamflow simulation, Journal of Hydrology, 548, 552-568, 10.1016/j.jhydrol.2017.03.022, 2017.

Zhang, Y., Zheng, H., Chiew, F. H. S., Arancibia, J. P., and Zhou, X.: Evaluating Regional and Global Hydrological Models against Streamflow and Evapotranspiration Measurements, Journal of Hydrometeorology, 17, 995-1010, 10.1175/jhm-d-15-0107.1, 2016.

Zhang, Y., Pan, M., Sheffield, J., Siemann, A. L., Fisher, C. K., Liang, M., Beck, H. E., Wanders, N., MacCracken, R. F., Houser, P. R., Zhou, T., Lettenmaier, D. P., Pinker, R. T., Bytheway, J., Kummerow, C. D., and Wood, E. F.: A Climate Data Record (CDR) for the global terrestrial water budget: 1984–2010, Hydrology and Earth System Sciences, 22, 241-263, 10.5194/hess-22-241-2018, 2018.

5   Zhao, F., Veldkamp, T. I. E., Frieler, K., Schewe, J., Ostberg, S., Willner, S., Schauberger, B., Gosling, S. N., Schmied, H. M., Portmann, F. T., Leng, G., Huang, M., Liu, X., Tang, Q., Hanasaki, N., Biemans, H., Gerten, D., Satoh, Y., Pokhrel, Y., Stacke, T., Ciais, P., Chang, J., Ducharne, A., Guimberteau, M., Wada, Y., Kim, H., and Yamazaki, D.: The critical role of the routing scheme in simulating peak river discharge in global hydrological models, Environmental Research Letters, 12, 10.1088/1748-9326/aa7250, 2017.

10   Zhou, X., Zhang, Y., Wang, Y., Zhang, H., Vaze, J., Zhang, L., Yang, Y., and Zhou, Y.: Benchmarking global land surface models against the observed mean annual runoff from 150 large basins, Journal of Hydrology, 470-471, 269-279, 10.1016/j.jhydrol.2012.09.002, 2012.

Table 1. Parameters used to calibrate the rainfall-runoff module of MGB, including their respective (typical) range of values.

Parameter	Description	Unit	Min	Max
Wm	Maximum water storage	mm	50	2000
b	Controls the distribution of water storage capacity of the soil	-	0.01	1.6
Kbas	Percolation rate from soil to groundwater	mm Δt^{-1}	0.1	4
Kint	Saturated soil hydraulic conductivity	mm Δt^{-1}	4	40
Kinf	Infiltration rate from floodplains when soil is completely dry	mm Δt^{-1}	-	
XL	Soil porosity index	-	Default = 0.67	
Cb	groundwater reservoir residence time	h	800	8000
Ci	Adjustment factor for subsurface reservoir residence time	-	50	200
Cs	Adjustment factor for superficial reservoir residence time	-	1	30

Table 2. Efficiency metrics used in this study.

Metric	Abbreviation	Assessment	Variables	Equation
Pearson's Correlation Coefficient	r	Linear correlation	Discharge, Water level, TWS	$$\frac{\sum[(x_{sim} - \mu_{sim})(x_{obs} - \mu_{obs})]}{\sqrt{\sum(x_{sim} - \mu_{sim})^2}\sqrt{\sum(x_{obs} - \mu_{obs})^2}}$$
Modified Kling–Gupta Efficiency	KGE	Overall Performance	Discharge	$$1 - \sqrt{(1-r)^2 + \left(1 - \frac{\mu_{sim}}{\mu_{obs}}\right)^2 + \left(1 - \frac{CV_{sim}}{CV_{obs}}\right)^2}$$
Nash–Sutcliffe Efficiency	NSE	High Flows	Discharge, Water level	$$1 - \frac{\sum(x_{sim} - x_{obs})^2}{\sum(x_{obs} - \mu_{obs})^2}$$
Log-transformed Nash–Sutcliffe Efficiency	NSElog	Low Flows	Discharge	$$1 - \frac{\sum[Log(x_{sim}) - Log(x_{obs})]^2}{\sum[Log(x_{obs}) - Log(\mu_{obs})]^2}$$
Overall BIAS (%)	BIAS	Under- and overestimation (volume)	Discharge	$$\left(\frac{\sum x_{sim} - \sum x_{obs}}{\sum x_{obs}}\right).100$$
BIAS in standard deviation (%)	σ BIAS	Under- and overestimation (anomalies)	Water levels, TWS	$$\left(\frac{\sigma_{sim} - \sigma_{obs}}{\sigma_{obs}}\right).100$$
Delay Index (days)	DI	Timing errors	Discharge	$\max r_{xy}[x_{sim}; x_{obs}], \quad \text{lag } [-100; +100]$
Root Mean Square Error	RMSE	Deviation of predicted values	ET, TWS	$$\sqrt{\frac{\sum(x_{sim} - x_{obs})^2}{n}}$$

5 Where: x_{sim} = simulated variable; x_{obs} = observed variable; μ_{sim} = mean of simulated variable; μ_{obs} = mean of observed variable; CV = coefficient of variation, equal to σ/μ; r_{xy} = cross-correlation

Table 3. Median values of discharge metrics for South America, computed as the performance difference between continental and global models. Lower values show better performance for a given global model when benchmarked against the MGB continental hydrologic–hydrodynamic model. Gauge stations with naturalized flows were removed from the analysis to provide a fair comparison.

Model difference	d_r	d_KGE	d_NSE	d_NSElog	d_Abs(BIAS) [%]	d_Abs(DI) [days]
MGB – HTESSEL/CaMa-Flood	0.11	0.48	1.35	0.53	8.3	1
MGB – LISFLOOD	0.02	0.44	1.86	0.48	32.5	1.5
MGB – WaterGAP3	0.16	0.44	1.05	0.57	3.8	2
MGB – Ensemble GM	0.03	0.26	0.91	0.27	11.0	0
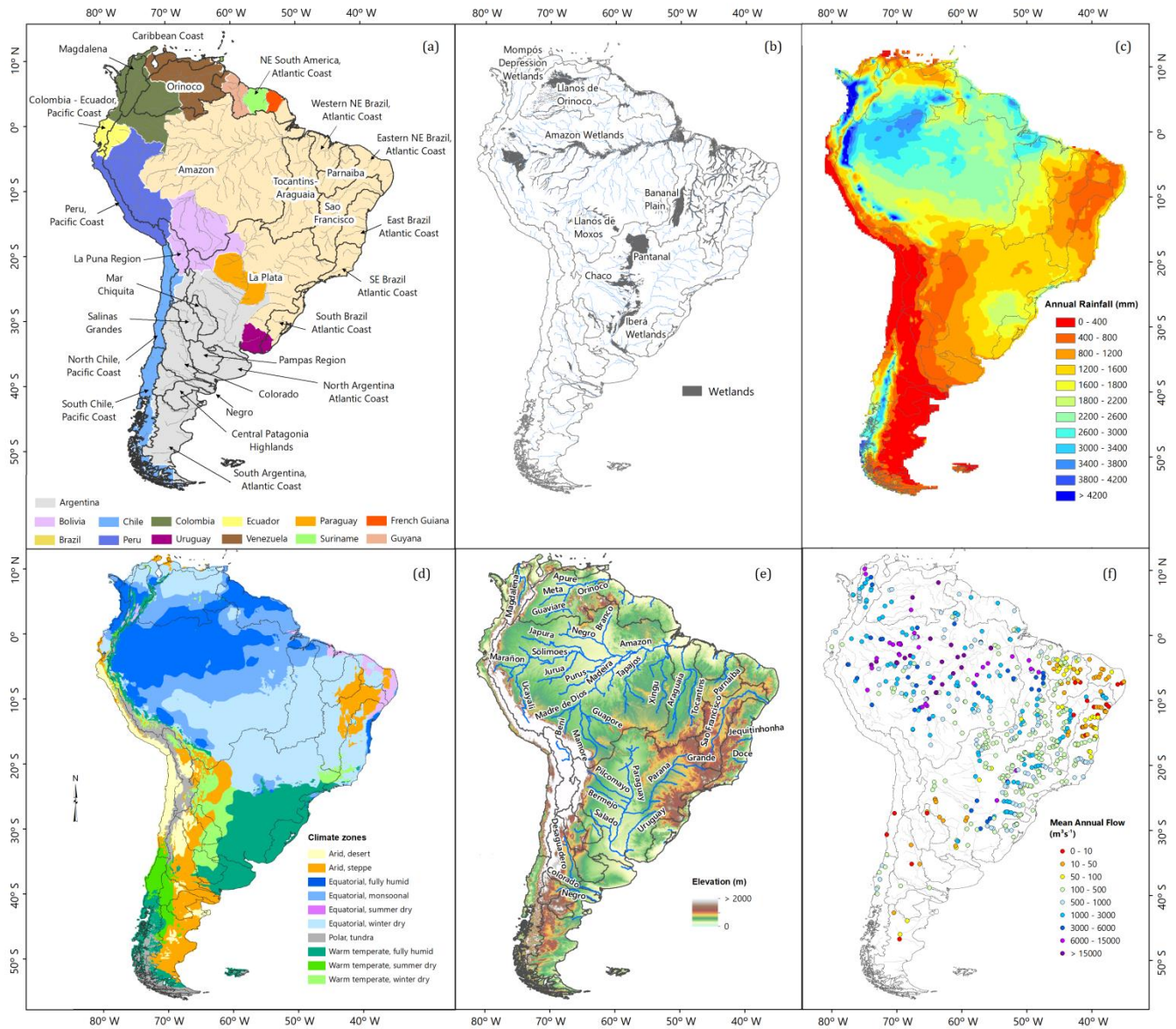
Figure 1: South America maps showing: (a)Countries and major hydrological regions according to FAO and ANA classifications, (b)Major wetlands and lowland regions, adapted from Lehner and Döll (2004), (c)Mean Annual Precipitation derived from MSWEP dataset (Beck et al., 2017b), (d)Köppen–Geiger updated climate classification from Kottek et al., (2006), (e)Relief map based on the Bare-Earth SRTM (O'Loughlin et al., 2016), including main rivers and (f)Mean annual flow at discharge gauge stations used in this study.
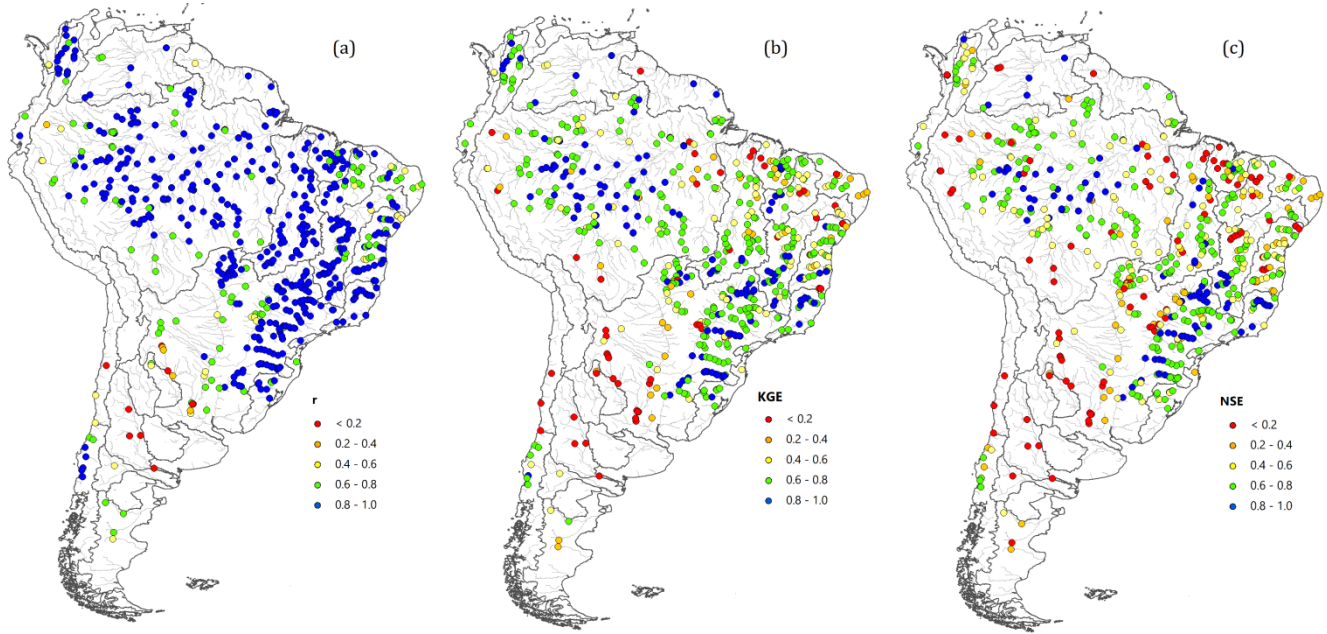
Figure 2: Discharge performance over South America in terms of (a)Correlation (r), (b)Kling–Gupta Efficiency (KGE) and (c)Nash–Sutcliffe (NSE).
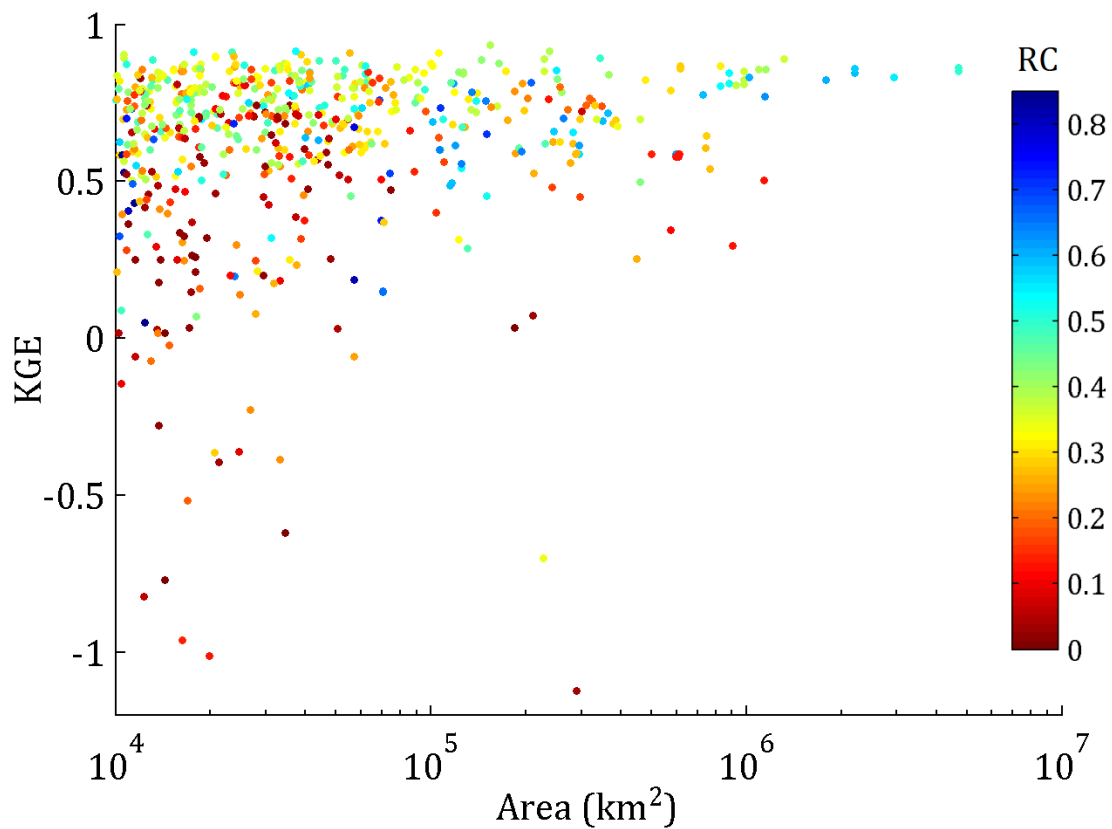
Figure 3: MGB model performance (KGE) versus RC and drainage area over South America.
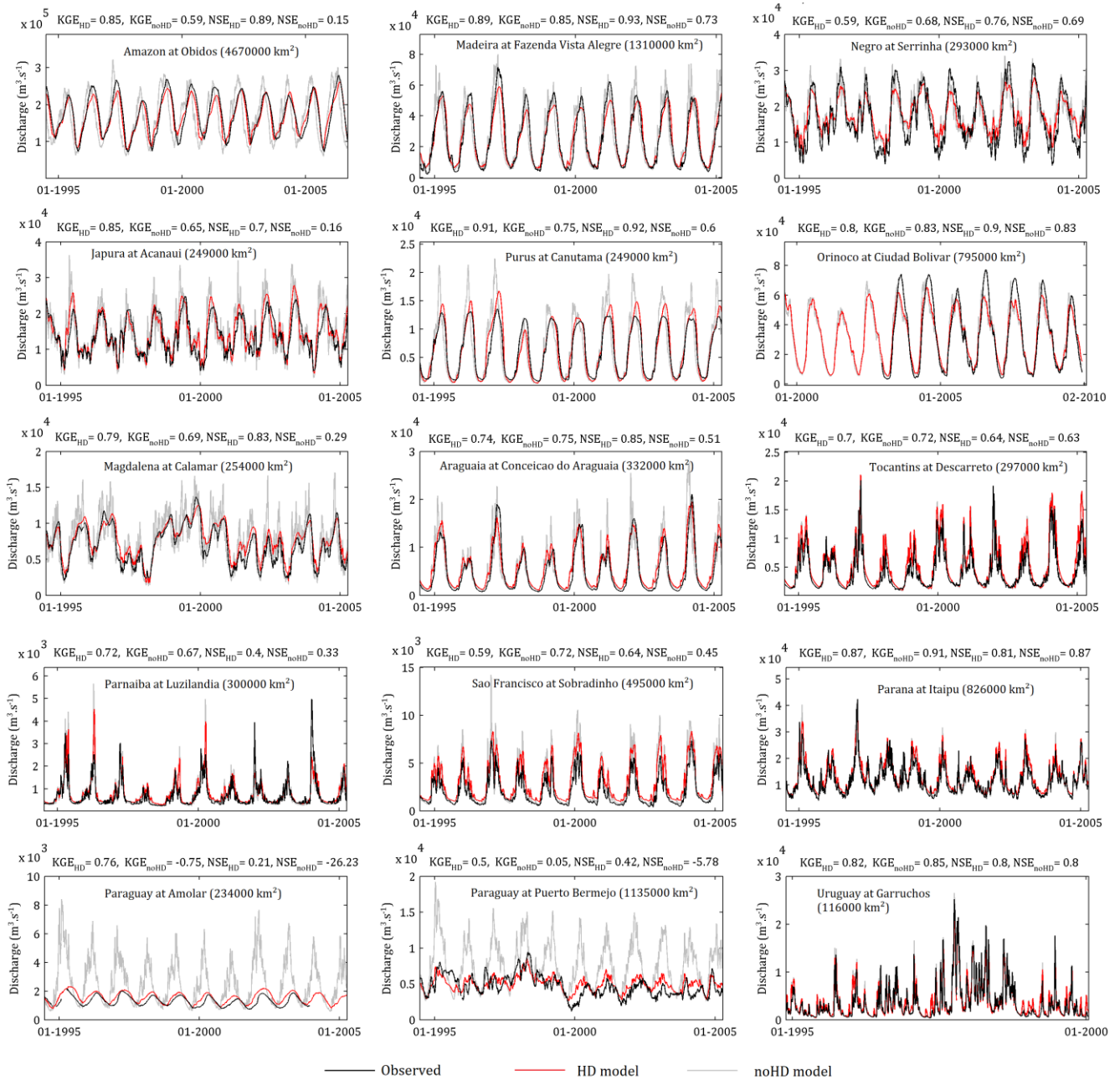
Figure 4: Comparison between observed (black) and simulated discharges for major South American rivers. Model results are shown considering both hydrodynamic (HD; red color) and non-hydrodynamic (noHD; gray color) routing methods.
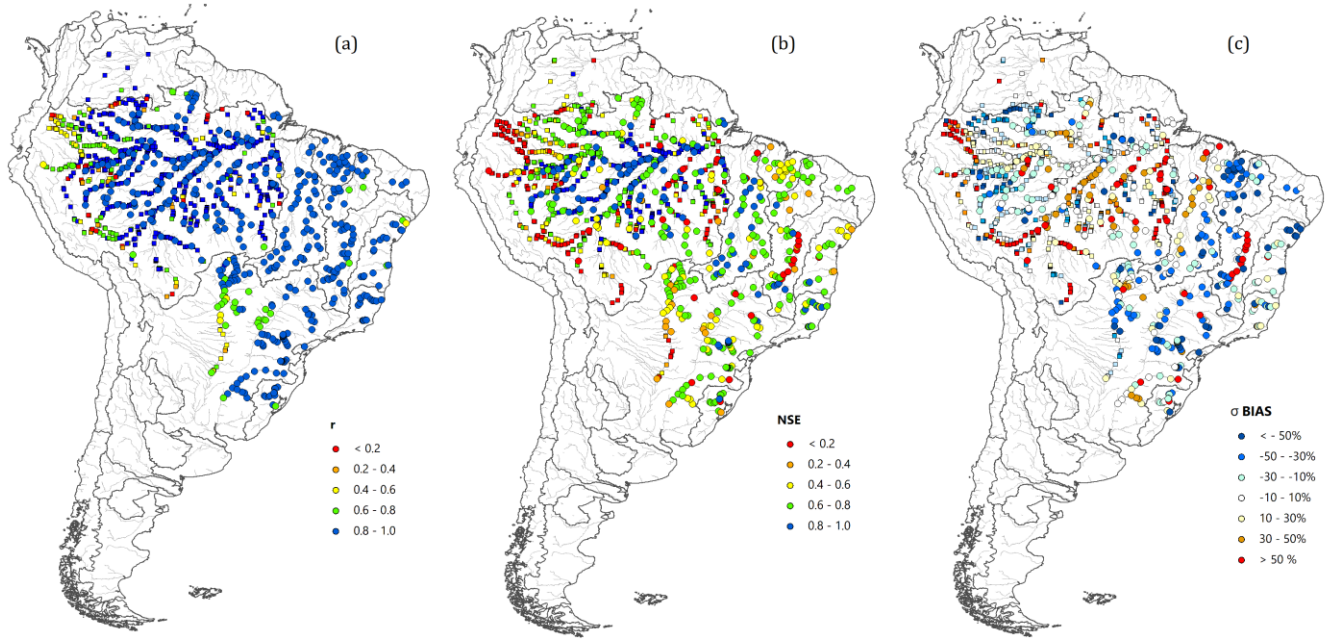
Figure 5: MGB performance for simulated water levels over South America in terms of (a)Correlation (*r*), (b)Nash–Sutcliffe (NSE) and (c)Bias in standard deviation (σBIAS). In situ and satellite altimetry locations are shown in circle and square symbols, respectively.
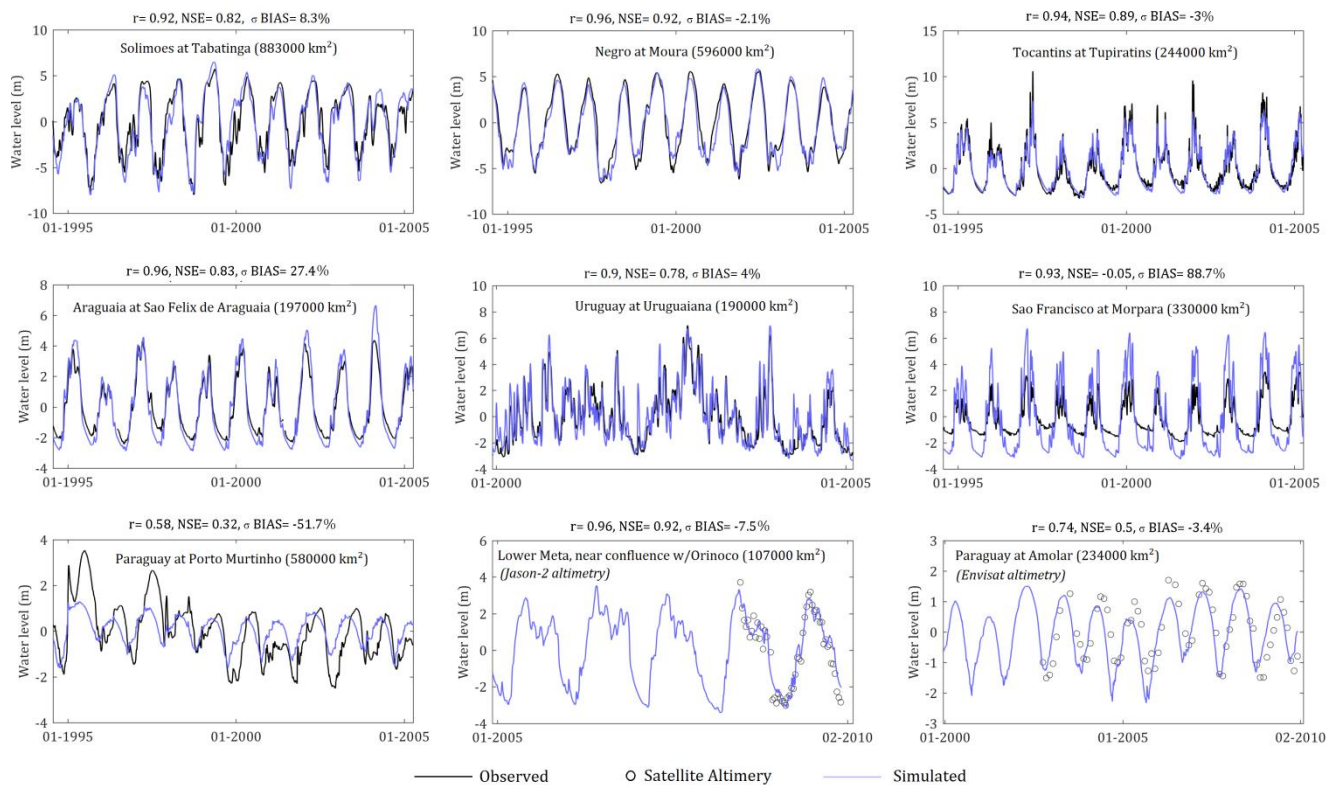
Figure 6: Comparison between simulated (blue) and observed (black) water level anomalies in major South American rivers for in situ gauges (continuous lines) and satellite altimetry (circles) at virtual stations (VS).
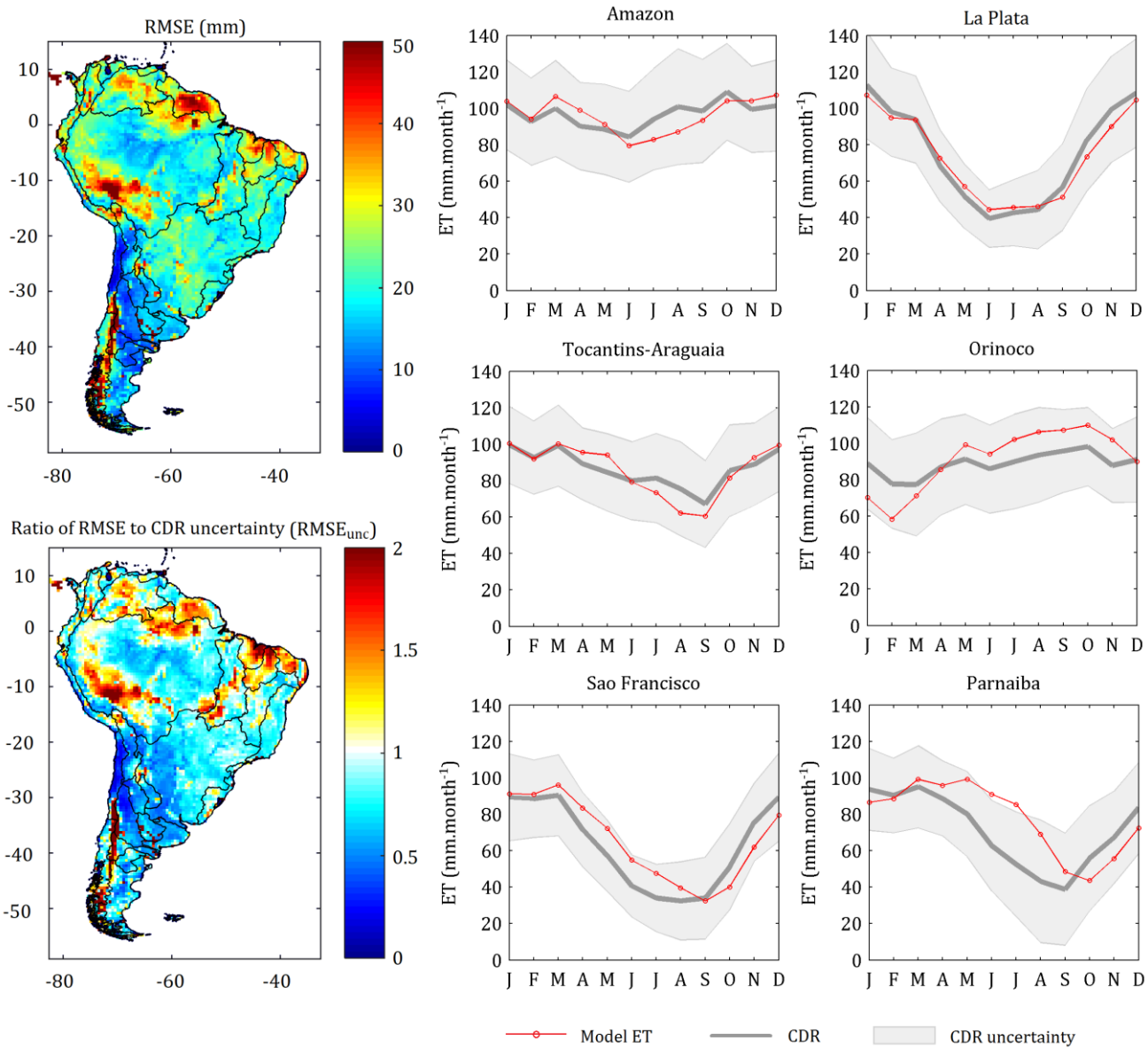
Figure 7: Comparison between MGB and CDR ET estimates in terms of RMSE (left) and seasonality for major South American basins (right). The light gray area represents the proxy of the CDR uncertainty, i.e., the mean deviation of all datasets (within CDR) from the ensemble mean (Zhang et al. 2018).
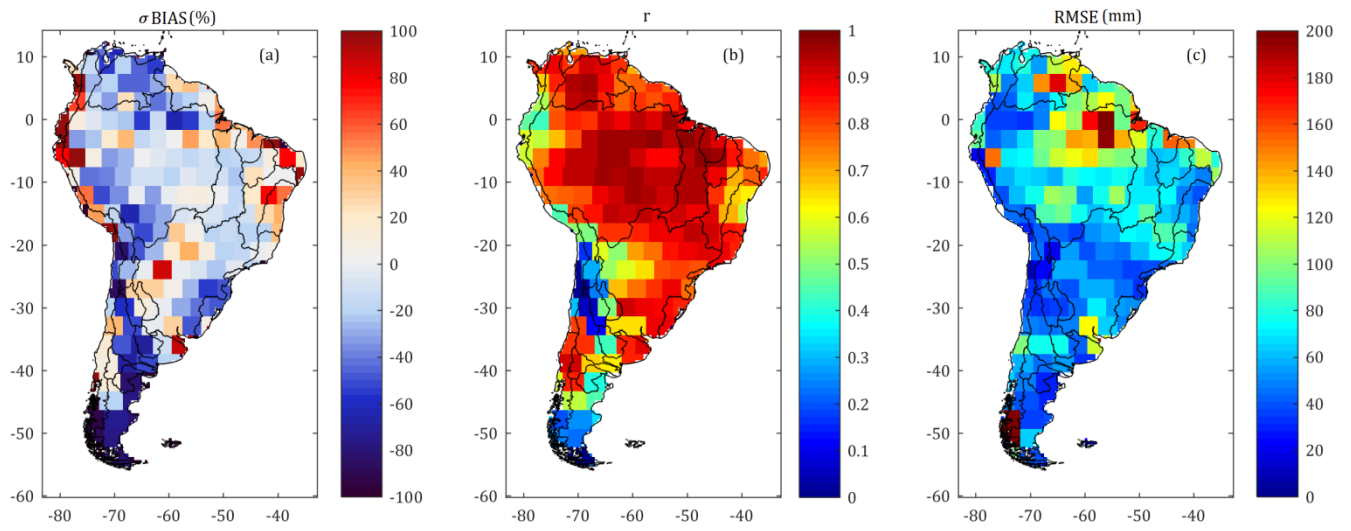
Figure 8: Comparison between MGB and GRACE (JPL RL05 v2 mascon solution) TWS anomalies in terms of (a)Bias in standard deviation, (b)Correlation and (c)RMSE.
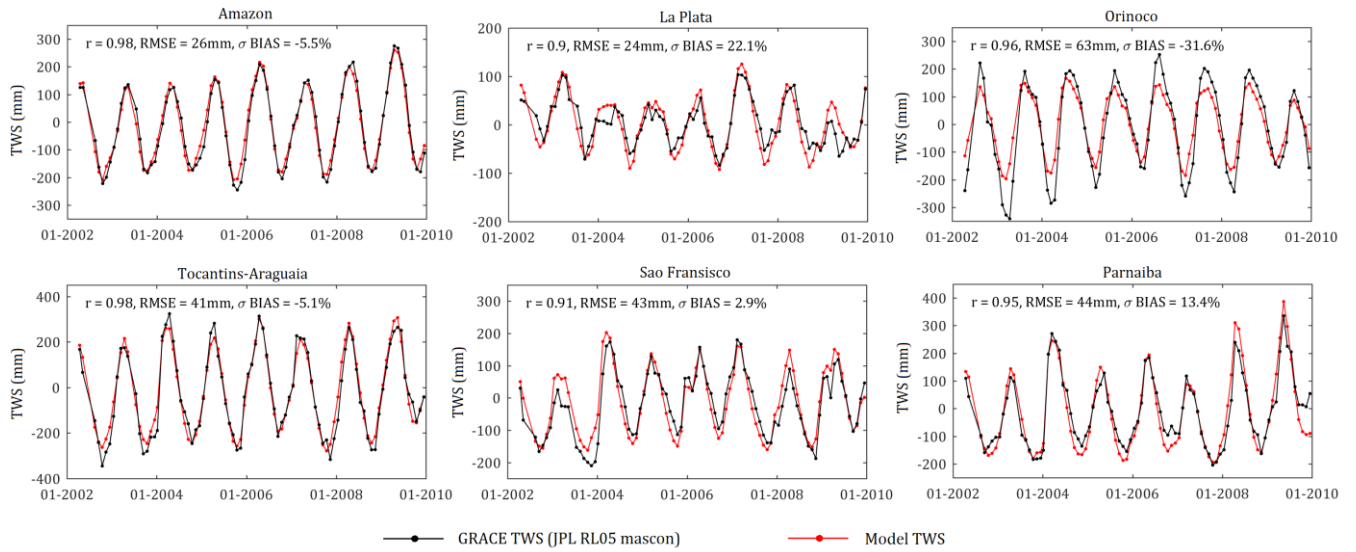
Figure 9: Comparison between MGB and GRACE (JPL RL05 v2 mascon solution) monthly TWS anomalies for major South American basins.
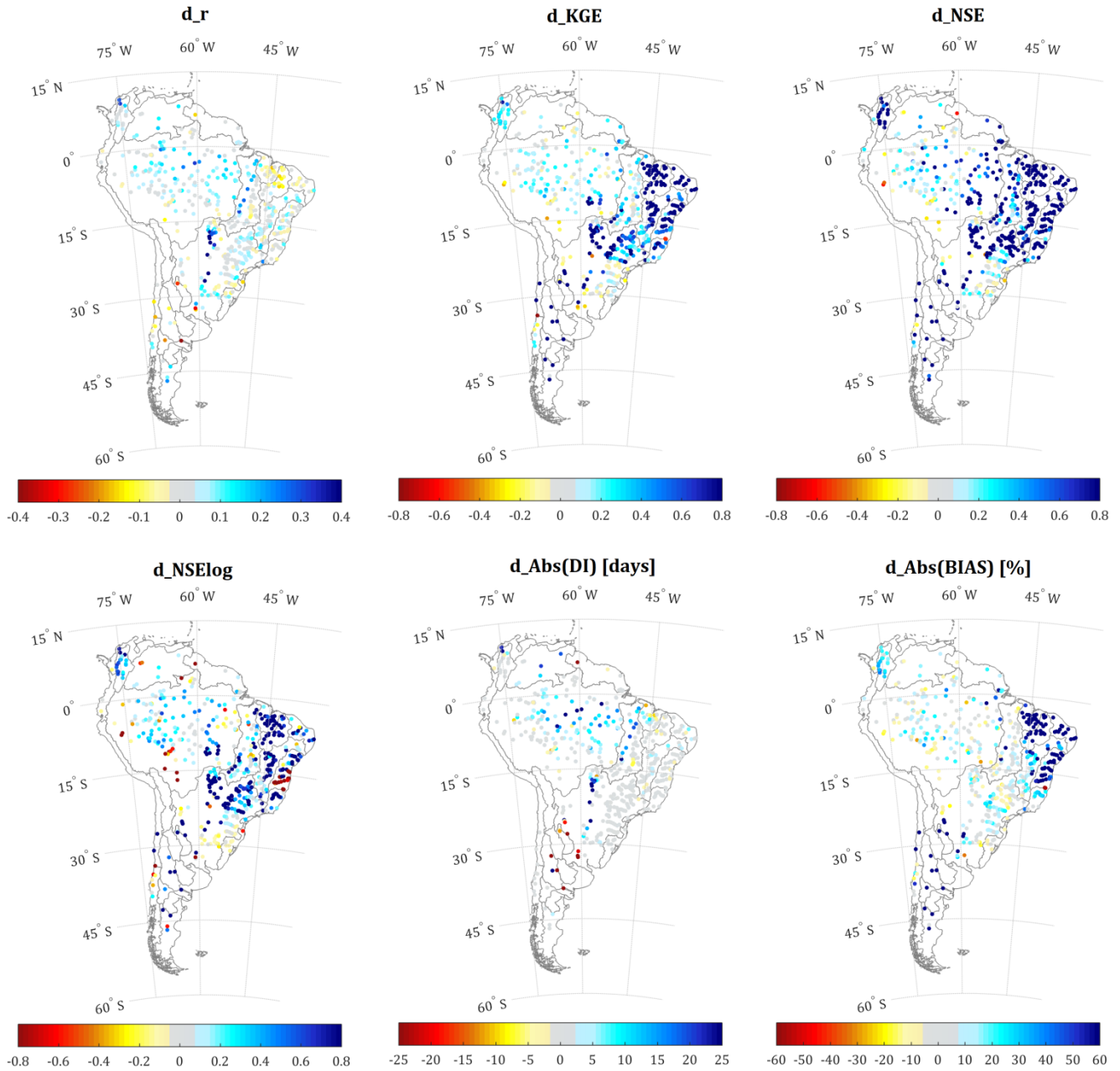
Figure 10: Difference between performances of MGB and the Ensemble GM for discharge metrics. Values considered as not significant (gray) are within the ranges –0.05 to +0.05 (d_r, d_KGE, d_NSE and d_NSElog), –5 % to 5 % (d_Abs(BIAS)) and –2 to +2 days (d_Abs(DI)).
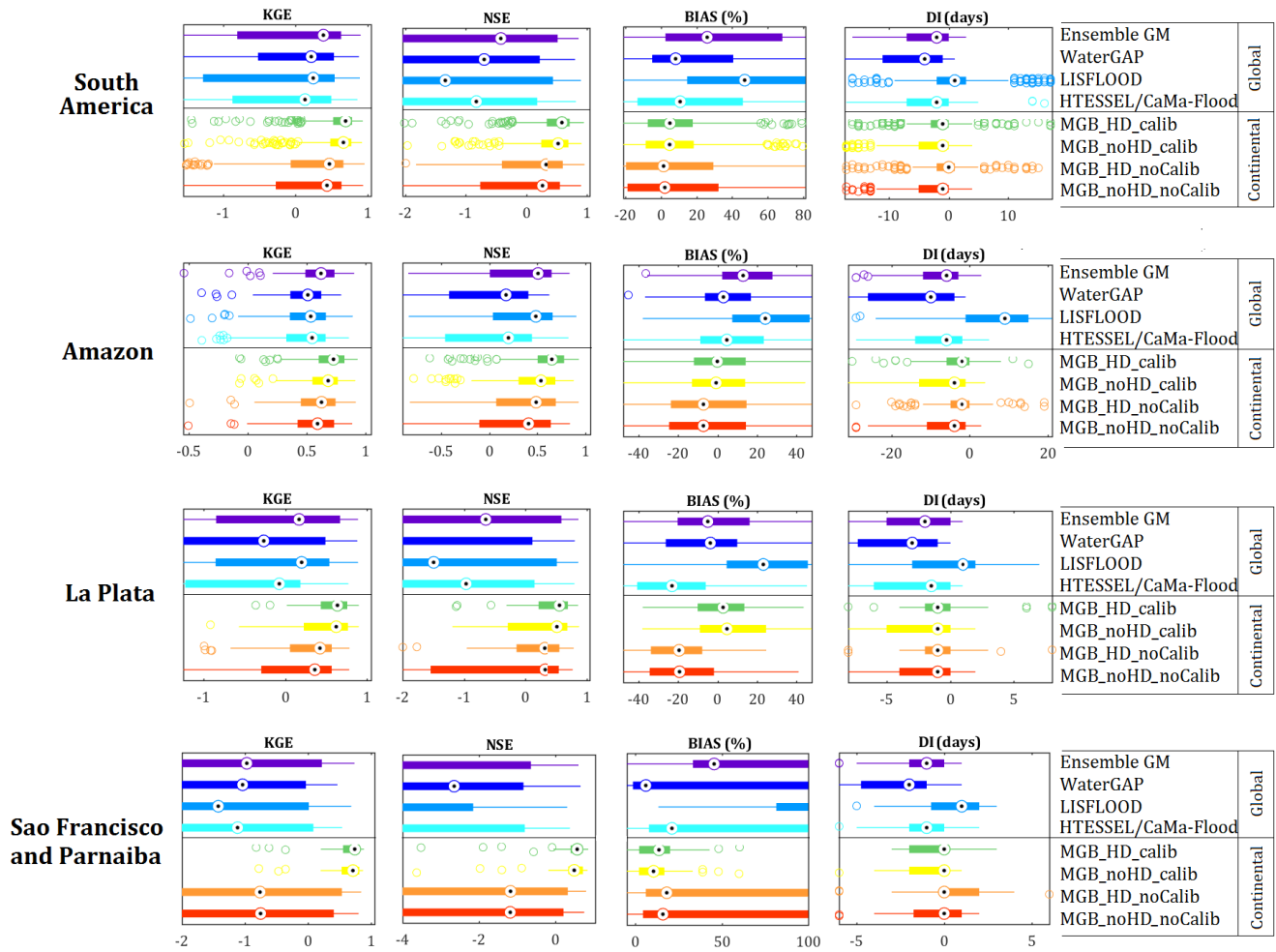
Figure 11: Boxplots of global (above center line) and continental (below center line) model performances for different South American regions. MGB model configurations: hydrodynamic routing with calibrated rainfall-runoff parameters (MGB_HD_calib), hydrodynamic routing with uncalibrated rainfall-runoff parameters (MGB_HD_noCalib), non-hydrodynamic routing with calibrated rainfall-runoff parameters (MGB_noHD_calib) and non-hydrodynamic routing with uncalibrated rainfall-runoff parameters (MGB_noHD_noCalib). Gauge stations with naturalized flows were removed from the analysis to provide a fair comparison.