https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development
Discussions
Open Access

SYNERGY BETWEEN SATELLITE OBSERVATIONS OF SOIL MOISTURE AND WATER STORAGE ANOMALIES FOR GLOBAL RUNOFF ESTIMATION

Stefania Camici [(1)], Gabriele Giuliani [(1)], Luca Brocca [(1)], Christian Massari [(1)], Angelica Tarpanelli [(1)], Hassan Hashemi Farahani [(2)], Nico Sneeuw [(2)], Marco Restano [(3)], Jérôme Benveniste [(4)]

(1) National Research Council, Research Institute for Geo-Hydrological Protection, Perugia, Italy (s.camici@irpi.cnr.it)

(2) Institute of Geodesy, University of Stuttgart, Geschwister-Scholl-Straße 24D, 70174 Stuttgart, Germany

(3) SERCO c/o ESA-ESRIN, Largo Galileo Galilei, Frascati, 00044, Italy

(4) European Space Agency, ESA-ESRIN, Largo Galileo Galilei, Frascati, 00044, Italy

* Correspondence to: Ph.D. Stefania Camici, Research Institute for Geo-Hydrological Protection, National Research Council, Via della Madonna Alta 126, 06128 Perugia, Italy. Tel: +39 0755014419 Fax: +39 0755014420 E-mail: stefania.camici@irpi.cnr.it.

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

ABSTRACT

This paper presents an innovative approach, STREAM - SaTellite based Runoff Evaluation And Mapping - to derive daily river discharge and runoff estimates from satellite soil moisture, precipitation and terrestrial water storage anomalies observations. Within a very simple model structure, the first two variables (precipitation and soil moisture) are used to estimate the quick-flow river discharge component while the terrestrial water storage anomalies are used for obtaining its complementary part, i.e., the slow-flow river discharge component. The two are then summed up to obtain river discharge and runoff estimates.

The method is tested over the Mississippi river basin for the period 2003-2016 by using Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) rainfall data, European Space Agency Climate Change Initiative (ESA CCI) soil moisture data and Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage data. Despite the model simplicity, relatively high-performance scores are obtained in river discharge simulations, with a Kling-Gupta efficiency index greater than 0.65 both at the outlet and over several inner stations used for model calibration highlighting the high information content of satellite observations on surface processes. Potentially useful for multiple operational and scientific applications (from flood warning systems to the understanding of water cycle), the added-value of the STREAM approach is twofold: 1) a simple modelling framework, potentially suitable for global runoff monitoring, at daily time scale when forced with satellite observations only, 2) increased knowledge on the natural processes, human activities and on their interactions on the land.

Key words: satellite products, soil moisture, water storage variations, data-driven hydrological modelling, rainfall-runoff modelling, Mississippi.

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development Discussions

1. INTRODUCTION

Spatial and temporal continuous river discharge monitoring is paramount for improving the understanding of the hydrological cycle, for planning human activities related to water use as well as to prevent/mitigate the losses due to extreme flood events. To accomplish these tasks, runoff and river discharge data, which represents the aggregated signal of runoff (Fekete et al., 2012), should be available at adequate spatial/temporal resolution, i.e., at basin scale (basin area larger than 10’000 km^2) and at monthly time step for water resources management and drought monitoring up to grid scale (few km)/(sub-) daily time step for flood prediction. The accurate continuous (in space and time) runoff and river discharge estimation at finer spatial/temporal resolution is still a big challenge for hydrologists.

Traditional in situ observations of river discharge, even if generally characterized by high temporal resolution (up to sub-hourly time step), typically offer little information on the spatial distribution of runoff within a watershed. Moreover, river discharge observation networks suffer from many limitations such as low station density and often incomplete temporal coverage, substantial delay in data access and large decline in monitoring capacity (Vörösmarty et al. 2002). Paradoxically, this latter issue is exacerbated in developing nations (Crochemore et al, 2020), where the knowledge of the terrestrial water dynamics deserves greater attention due to huge damages to settlements and especially the loss of human lives that occurs regularly.

This precarious situation has led to growing interest in finding alternative solutions, i.e., model-based or observation-based approaches, for runoff and river discharge monitoring. Model-based approaches, based on the mathematical description of the main hydrological processes (e.g., water balance models, WBMs, global hydrological models, GHMs, e.g., Döll et al., 2003 or, increasing in complexity, land surface models, LSM, e.g., Balsamo et al., 2009; Schellekens et al., 2017), are able to provide comprehensive information on a large number of relevant variables of the hydrological cycle including runoff and river discharge at very high temporal and spatial resolution (up to hourly

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

sampling and 0.05° grid scale). However, the values of simulated water balance components rely on a massive parameterization of the soil, vegetation and land parameters, which is not always realistic, and are strongly dependent on the GHM/ LSM models used, analysis periods (Wisser et al., 2010) and climate forcings selected (e.g Haddeland et al., 2012; Gudmundsson et al., 2012a, b; Prudhomme et al., 2014; Müller Schmied et al., 2016).

Alternatively, the observation-based approaches exploit machine learning techniques and a considerable amount of data to describe the physics of the system (i.e. hydraulic and/or hydrologic phenomena, Solomatine and Ostfeld, 2008) with only a limited number of assumptions. Besides being simpler than model-based approaches, these approaches still present some limitations. At first, as they rely on a considerable amount of data describing the modelled system's physics, the spatial/temporal extent and the uncertainty of the resulting dataset is determined by the spatial/temporal coverage and the accuracy of the forcing data (e.g., see E-RUN dataset, Gudmundsson and Seneviratne, 2016; GRUN dataset, Ghiggi et al., 2019; FLO1K dataset, Barbarossa et al., 2018). Additional limitations stem from the employed method to estimate runoff. Indeed, random forests such as employed in Gudmundsson and Seneviratne, 2016, like other machine learning techniques, are powerful tools for data driven modeling, but they are prone to overfitting, implying that noise in the data can obscure possible signals (Hastie et al., 2009). Moreover, the influence of land parameters on continental-scale runoff dynamics is not taken into account as the underlying hypothesis is that the hydrological response of a basin exclusively depend on present and past atmospheric forcing. It is easy to understand that this assumption will only be valid in certain circumstances and might lead to problems, e.g., over complex terrain (Orth and Seneviratne, 2015) or in cases of human river flow regulation (Ghiggi et al., 2019).

Remote sensing can provide estimates of nearly all the climate variables of the global hydrological cycle including soil moisture (e.g., Wagner et al., 2007; Seneviratne et al., 2010), precipitation (Huffman et al., 2014) and total terrestrial water storage (e.g., Houborg et al., 2012; Landerer and Swenson, 2012; Famiglietti and Rodell, 2013). It has undeniably changed and improved dramatically the ability to monitor the global water cycle and, hence, runoff. By taking advantage of satellite information, some studies tried to develop methodologies able to optimally produce multivariable datasets from the fusion of in situ and satellite-based observations (e.g., Rodell et al., 2015; Zhang et al., 2018; Pellet et al., 2019). Other studies exploited satellite observations of hydrological variables, e.g., precipitation (Hong et al, 2007), soil moisture (Massari et al., 2014), and geodetic variables (e.g., Sneeuw at. al., 2014; Tourian et al., 2018) to monitor single components of the water cycle in an independent way.

Although the majority of these studies provide runoff and river discharge data at basin scale and monthly time step, they deserve to be recalled here as important for the purpose of the present study. In particular, Hong et al. (2007) presented a first attempt to obtain an approximate but quasi-global annual streamflow dataset, by incorporating satellite precipitation data in a relatively simple rainfall-runoff simulation approach. Driven by the multiyear (1998-2006) Tropical Rainfall Measuring Mission Multi-satellite Precipitation Analysis, runoff was independently computed for each global land surface grid cell through the Natural Resources Conservation Service (NRCS) runoff curve number (CN) method (NRCS, 1986) and subsequently routed to the watershed outlet to simulate streamflow. The results, compared to the in situ observed discharge data, demonstrated the potential of using satellite precipitation data for diagnosing river discharge values both at global scale and for medium to large river basins. If, on the one hand, the work of Hong et al. (2007) can be considered as a pioneer study, on the other hand it presents a serious drawback within the NRCS-CN method that lacks a realistic definition of the soil moisture conditions of the catchment before flood events. This aspect is not negligible, as it is well established that soil moisture is paramount in the partitioning of precipitation into surface runoff and infiltration inside a catchment (Brocca et al., 2008). In particular, for the same rainfall amount but different values of initial soil moisture conditions, different flooding effects can occur (see e.g. Crow et al., 2005; Brocca et al., 2008; Berthet et al., 2009; Merz and Bloschl, 2009; Tramblay et al., 2010). On this line following Brocca et al. (2009), Massari et al. (2016) presented a very first attempt to estimate global streamflow data by using

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development
Discussions
Open Access

satellite Soil Moisture Active and Passive (SMAP) and Global Precipitation Measurement (GPM) products. Although the validation was carried out by routing the monthly surface runoff only in a single basin in Central Italy, the obtained results suggested to dedicate additional efforts in this direction.

Among the studies that use satellite observations of hydrological variables for runoff estimation, the hydro-geodetic approaches are undoubtedly worth mentioning, see e.g., (Sneeuw et al.,2014) for a comprehensive overview or Lorenz et al. (2014) for an analysis of satellite-based water balance misclosures with discharge as closure term. In particular, the satellite mission Gravity Recovery And Climate Experiment (GRACE), which observed the temporal changes in the gravity field, has given a strong impetus to satellite-driven hydrology research (Tapley et al., 2019). Since temporal gravity field variations over the continents imply water storage change, GRACE was the first remote sensing system to provide observational access to deeper groundwater storage. The relation between GRACE groundwater storage change and runoff was characterized by Riegger and Tourian (2014), which even allowed the quantification of absolute drainable water storage over the Amazon (Tourian et al., 2018). In essence the storage-runoff relation describes the gravity-driven drainage of a basin and, hence, the slow-flow processes. Due to GRACE's spatial-temporal resolution, runoff and river discharge are generally available for large basins (>160'000 km^2) and at monthly time step.

Based on the above discussion, it is clear that each approach presents strengths and limitations that enable or hamper the runoff and river discharge monitoring at finer spatial and temporal resolutions. In this context, this study presents an attempt to find an alternative method to derive daily river discharge and runoff estimates at ¼ degree spatial resolution exploiting satellite observations and the knowledge of the key mechanisms and processes that act in the formation of runoff, i.e., the role of soil moisture in determining the response of a catchment to precipitation. For that, soil moisture, precipitation and terrestrial water storage anomalies (TWSA) observations are used as input into a simple modelling framework named STREAM (SaTellite based Runoff Evaluation And Mapping). Unlike classical land surface models, STREAM exploits the knowledge of the system states (i.e., soil

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

moisture and TWSA) to derive river discharge and runoff, and thus it 1) skips the modelling of the evapotranspiration fluxes which are known to be a non-negligible source of uncertainty (Long et al. 2014), 2) limits the uncertainty associated with the over-parameterization of soil and land parameters and 3) implicitly takes into account processes, mainly human-driven (e.g., irrigation, change in the land use), that might have a large impact on the hydrological cycle and hence on runoff.

The detailed description of the STREAM model is given in section 4. The collected datasets and the experimental design for the Mississippi River Basin (section 2) are described in sections 3 and 5, respectively. Results, discussion and conclusions are drawn in section 6, 7 and 8, respectively.

2. STUDY AREA

The STREAM model presented here has been tested and validated over the Mississippi River basin. With a drainage area of about 3.3 million km2, the Mississippi River basin is the fourth largest watershed in the world, bordered to the West by the crest of the Rocky Mountains and to the East by the crest of the Appalachian Mountains. According to the Köppen climate classification, the climate is subtropical humid over the southern part of the basin, continental humid with hot summer over the central part, continental humid with warm summer over the eastern and norther parts, whereas a semiarid cold climate affects the western part. The average annual air temperature across the watershed ranges from 4°C in the West to 6°C in the East. On average, the watershed receives about 900 mm/year of precipitation (77% as rainfall and 23% as snowfall), more concentrated in the eastern and southern portions of the basin with respect to its northern and western part (Vose et al., 2014).

The river flow has a clear natural seasonality mainly controlled by spring snowmelt in the mountainous areas of the basins and by heavy rainfall exceeding the soil moisture storage capacity in the central and southern part of the basin (Berghuijs et al., 2016), but it is also heavily regulated by the presence of about 1000 large dams (Global Reservoir and Dam Database GRanD, Lehner et al., 2011) spread-out across the basin. The annual average of Mississippi river discharge at the Vicksburg outlet section is equal to 17’500 m3/s (see Table 1). Given the variety of climate and topography

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

across the Mississippi River basin, it is a good candidate to test the suitability of the STREAM model for river discharge and runoff simulation.

3. DATASETS

The datasets used in this study include in situ observations, satellite products and model outputs. The first two datasets have been used as input data to the STREAM model. Conversely, the model outputs are used as a benchmark to validate the performance of the STREAM model.

3.1 In situ Observations

In situ observations comprise air temperature (T_{air}) and river discharge data (Q).

For T_{air} data the Climate Prediction Center (CPC) Global Temperature data developed by the American National Oceanic and Atmospheric Administration (NOAA) using the optimal interpolation of quality-controlled gauge records of the Global Telecommunication System (GTS) network (Fan et al., 2008) have been used. The dataset, downloadable at (https://psl.noaa.gov/data/gridded/data.cpc.globaltemp.html) is available on a global regular 0.5°×0.5° grid, and provides daily maximum (T_{max}) and minimum (T_{min}) air temperature data from 1979 to present. The daily average air temperature data have been generated as the mean of T_{max} and T_{min} of each day.

Daily Q data over the study basins have been taken from the Global Runoff Data Center (GRDC, https://www.bafg.de/GRDC/EN/Home/homepage_node.html). In particular, 11 gauging stations located along the main river network of the Mississippi River basin have been selected to represent the spatial distribution of runoff over the basin. The location of these gauging stations along with relevant characteristics (e.g., the upstream basin area, the mean annual river discharge and the presence of upstream dams) are summarized in Table 1. As it can be noted, mean annual river discharge ranges from 141 to 17’500 m^3/s, and 3 out 11 sections are located downstream big dams (Lehner et al., 2011).

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

3.2 Satellite Products

Satellite products include observations of precipitation (*P*), soil moisture and TWSA.

The satellite *P* dataset used in this study is the Multi-satellite Precipitation Analysis 3B42 Version 7 (TMPA 3B42 V7) estimate produced by the National Aeronautics and Space Administration (NASA) as the 0.25°×0.25° quasi-global (50°N-S) gridded dataset. The TMPA 3B42 V7 is a gauged-corrected satellite product, with a latency period of two months after the end of the month of record, available at 3h sampling interval from 1998 to present (2020). Major details about the *P* dataset, downloadable from http://pmm.nasa.gov/data-access/downloads/trmm, can be found in Huffman et al. (2007).

Soil moisture data have been taken from the European Space Agency Climate Change Initiative (ESA CCI) Soil Moisture project (https://esa-soilmoisture-cci.org/) that provides a product continuously updated in term of spatial-temporal coverage, sensors and retrieval algorithms (Dorigo et al., 2017). In this study, the daily combined ESA CCI SOIL MOISTURE product v4.2 is used, that is available at global scale with a grid spacing of 0.25°, for the period 1978-2016.

TWSA have been obtained from the Gravity Recovery And Climate Experiment (GRACE) satellite mission. Here we employ the NASA Goddard Space Flight Center (GSFC) global mascon model, i.e., Release v02.4, (Luthcke et al. 2013). It has been produced based on the mass concentration (mascon) approach. The model provides surface mass densities on a monthly basis. Each monthly solution represents the average of surface mass densities within the month, referenced at the middle of the corresponding month. The model has been developed directly from GRACE level-1b K-Band Ranging (KBR) data. It is computed and delivered as surface mass densities per patch over blocks of approximately 1°×1° or about 12’000 km^2. Although the mascon size is smaller than the inherent spatial resolution of GRACE, the model exhibits a relatively high spatial resolution. This is attributed to a statistically optimal Wiener filtering, which uses signal and noise covariance matrices. The coloured (frequency-dependent) noise characteristic of KBR data was taken in to account when compiling the model, which has allowed for a reliable computation of these noise and signal

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

covariance matrices. They play a crucial role when filtering and allow to achieve a higher spatial resolution compared to commonly applied GRACE filtering methods such as Gaussian smoothing and/or destriping filters. GRACE data are available for the period 01 January 2003 to 15 July 2016.

3.3 Model Outputs

To establish the quality of the STREAM model in runoff simulation, monthly runoff (R) data obtained from the Global Runoff Reconstruction (GRUN_v1, https://doi.org/10.3929/ethz-b-000324386) have been used for comparison. The GRUN dataset (Ghiggi et al., 2019) is a global monthly R dataset derived through the use of a machine learning algorithm trained with in situ Q observations of relatively small catchments (<2500 km^2) and gridded precipitation and temperature derived from the Global Soil Wetness Project Phase 3 (GSWP3) dataset, (Kim et al., 2017). The dataset covers the period from 1902 to 2014 and it is provided on a 0.5° ×0.5° regular grid.

4. METHOD

4.1 STREAM Model: the Concept

The concept behind the STREAM model is that river discharge is a combination of hydrological responses operating at diverse time scales (Blöschl et al., 2013; Rakovec et al., 2016). In particular, river discharge can be considered made up of a *slow-flow component*, produced as outflow of the groundwater storage and of a *quick-flow component*, i.e. mainly related to the surface and subsurface runoff components (Hu and Li, 2018).

While the high spatial and temporal (i.e., intermittence) variability of rainfall and the highly changing land cover spatial distribution significantly impact the variability of the *quick-flow component* (with scales ranging from hours to days and meters to kilometres depending on the basin size), *slow-flow river discharge* reacts to precipitation inputs more slowly (i.e., months) as water infiltrates, is stored, mixed and is eventually released in times spanning from weeks to months. Therefore, the two components can be estimated by relying upon two different approaches that involve different types

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development Discussions

of observations. Based on that, within the STREAM model, satellite soil moisture, precipitation and TWSA will be used for deriving river discharge and runoff estimates. The first two variables are used as proxy of the *quick-flow* river discharge component while TWSA is exploited for obtaining its complementary part, i.e., the *slow-flow river discharge* component. Firstly, we exploit the role of the soil moisture in determining the response of the catchment to the precipitation inputs, which have been soundly demonstrated in more than ten years of literature studies (see e.g., Brocca et al., 2017 for a comprehensive discussion on the topic). Secondly, we consider the important role of terrestrial water storage in determining the slow-flow river discharge component as modelled in several hydrological models (e.g., Sneeuw et al., 2014).

It is worth noting that this *modus operandi*, i.e. to model the *quick-flow* and *slow-flow* discharge component separately exploring their process controls independently, has been largely applied and tested in recent and past studies, e.g., for the estimation of the flow duration curve (see e.g, Botter et al., 2007a, b; Yokoo and Sivapalan 2011; Muneepeerakul et al., 2010; Ghotbi et al., 2020).

4.2 STREAM Model: the Laws

The STREAM model is a conceptual hydrological model that, by using as input observation of P, soil moisture, TWSA and T_{air} data, simulates continuous R and Q time series.

The model entails three main components (Figure 1): 1) a snow module to separate precipitation into snowfall and rainfall, 2) a soil module to simulate the evolution in time t of the quick and slow runoff responses, Qfu [mm] and Qsl [mm], and 3) a routing module that transfers these components through the basins and the rivers for the simulation of the *quick-flow* river discharge, QF [m^3/s], and the *slow-flow* river discharge, SF [m^3/s] components.

The soil module is composed of two storages, Su and Sl as illustrated in Figure 1. The upper storage receives inputs from P, released through a snow module (Cislaghi et al., 2020) as rainfall (r) or stored as snow water equivalent (SWE) within the snowpack and on the glaciers. In particular, according to

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development Discussions

Cislaghi et al. (2020), *SWE* is modelled by using as input T_{air} and a degree-day coefficient, C_{m}, to be estimated by calibration.

Once separated, *r* input contributes to the quick runoff response while the *SWE* (like other fluxes contributing to modify the soil water content into *Su*) is neglected as already considered in the satellite TWSA. Therefore, the first key point of the STREAM model is that the water content in the upper storage is directly provided by the satellite soil moisture observations and the loss processes like infiltration or evaporation do not need to be explicitly modelled to simulate the evolution in time *t* of soil moisture. Consequently, the quick runoff response, Qfu from the first storage can be computed through equation (1) as follows:

$$Qfu(t) = r(t)\, SWI(t,T)^{\alpha} \qquad (1)$$

where:

- SWI is the Soil Water Index (Wagner et al., 1999), i.e., the root-zone soil moisture product referred to the first layer of the model, derived by the surface satellite soil moisture product, θ, by applying the exponential filtering approach in its recursive formulation (Albergel et al., 2009):

  $$SWI_n = SWI_{n-1} + K_n(\theta(t_n) - SWI_{n-1}) \qquad (2)$$

  with the gain K_n at the time t_n given by:

  $$K_n = \frac{K_{n-1}}{K_{n-1} + e^{\left(\frac{t_n - t_{n-1}}{T}\right)}} \qquad (3)$$

- T [days] is a parameter, named characteristic time length, that characterizes the temporal variation of soil moisture within the root-zone profile and the gain K_n ranges between 0 and 1;
- α[-] is a coefficient linked to the non-linearity of the infiltration process and it takes into account the characteristics of the soil;
- for the initialization of the filter $K_1 = 1$ and $SWI_1 = \theta(t_1)$.

The second key point of STREAM approach concerns the estimation of the slow runoff response, Qsl, from the second storage. The hypothesis here, shared also with other studies (e.g., Rakovec et al., 2016), is that the dynamic of the slow runoff component can be represented by the monthly TWSA data. Indeed, the time scale of slow runoff response is typically in the range of seasons to years and it is almost independent upon the water that is contained in that upper storage. For that, the slow runoff response Qsl, from the second storage, can be computed through equation (4) as follows:

$$Qsl(t) = \beta \, (TWSA^{*}(t))^{m} \tag{4}$$

where:

- $TWSA^{*}$ [-] is the TWSA estimated by GRACE normalized by its minimum and maximum values. The assumption behind this equation is that TWSA can be assumed as a proxy of the evolution in time, t, of the Sl, i.e., the storage of the lower storage.
- β [mm h^{-1}] and m [-] are two parameters describing the nonlinearity between slow runoff component and $TWSA^{*}$.

Note that, being based on a conceptual framework, we assume that soil moisture acts both on the generation of the quick flow part (mainly) and is partly responsible of the slow flow contribution indirectly via TWSA observations (indeed TWSA already contains the soil moisture signal in themselves).

The STREAM model runs in a semi-distributed version in which the catchment is divided into s elements, each one representing either a subcatchment with outlet along the main channel or an area draining directly into the main channel. Each element is assumed homogeneous and hence constitutes a lumped system.

The routing module (controlled by a γ parameter) conveys the Qfu and Qsl response components at each element outlet (subcatchments and directly draining areas, Brocca et al., 2011) and successively at the catchment outlet of the basin. Specifically, the quick component Qfu is routed to the element

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

outlet by the Geomorphological Instantaneous Unit Hydro-graph (GIUH, Gupta et al., 1980) for subcatchments or through a linear reservoir approach (Nash, 1957) for directly draining areas; the Qsl slow component is transferred to the outlet section by a linear reservoir approach. Finally, a diffusive linear approach (controlled by the parameters *C* and *D*, i.e., Celerity and Diffusivity, Troutman and Karlinger, 1985) is applied to route the quick and slow runoff components at the outlet section of the catchment (Brocca et al., 2011). In the first case we obtain the *quick-flow* river discharge component, *QF* [m^3/s], and in the second case the *slow-flow* river discharge component, *SF* [m^3/s] (see Figure 1).

4.3 STREAM Parameters

The STREAM model uses 8 parameters of which 5 are used in the soil module (α, T [days], β [mm h^{-1}], m, Cm) and 3 in the routing module (γ, C [km h^{-1}] and D [km^2 h^{-1}]). These parameters are calibrated by maximizing the Kling-Gupta Efficiency index (KGE, Gupta et al., 2009; Kling et al., 2012, see paragraph 5.1 for more details) between observed and simulated river discharge.

5. EXPERIMENTAL DESIGN

5.1 Modelling Setup for Mississippi River Basin

The modelling setup is carried out in four steps (Figure 2):

1. *Input data collection.* Two different groups of data have to be collected to setup the model, i.e., topographic information and hydrological variables. Concerning the topographic information, the SHuttle Elevation Derivatives at multiple Scales (HydroSHED, https://www.hydrosheds.org/) DEM of the basin at the 3″ resolution (nearly 90 m at the equator) as well as the location of the gauging stations where the model should be calibrated/validated, are collected. Concerning the hydrological variables, gridded precipitation, T_{air}, soil moisture and TWSA are collected. In addition, in situ Q time series for the sections where the model should be calibrated/validated as well as modelled runoff datasets are required.

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development Discussions

2. *Sub-basin delineation.* STREAM model is run in the semi-distributed version over the Mississippi River basin. The TopoToolbox (https://topotoolbox.wordpress.com/), a tool developed in Matlab by Schwanghart et al. (2010), and the DEM of the basin have been used to derive flow directions, to extract the stream network and to delineate the drainage basins over the Mississippi River basin. In particular, by considering only rivers with Horton-Strahler order greater than 3, the Mississippi watershed has been divided into 53 sub-basins as illustrated in Figure 3. Red dots in the figure indicate the location of the 11 discharge gauging stations selected for the study area.

3. *Extraction of input data.* Precipitation, T_{air}, soil moisture and TWSA datasets data have to be extracted for teach sub-basin of the study area. If characterized by different spatial resolution, these datasets need to be resampled over a common spatial grid prior to be used as input into the model. To run the STREAM model over the Mississippi river basin, input data have been resampled over the precipitation spatial grid at 0.25° resolution through a bilinear interpolation. For each of the 53 Mississippi subbasins, the resampled precipitation, soil moisture, T_{air} and TWSA data have been extracted.

4. *STREAM model calibration.* In situ river discharge data are used as reference data for the calibration of STREAM model. For Mississippi, the STREAM model has been calibrated over five sections as illustrated in Figure 3: the inner sections 4, 6, 9, 11 and the outlet section 10, are used to calibrate the model and all sub-basins contributing to the respective sections are highlighted with the same colour. This means that, for example, the sub-basins labelled as 1, 2, 5 to 15, 17, 22, 23, and 30 contribute to section 4, sub-basins 31, 37, 38 and 41 contribute to section 6 and so on. Consequently, the sub-basins highlighted with the same colour are assigned the same model parameters, i.e. the parameters that allow to reproduce the river discharge data observed at the related outlet section. Once calibrated, the STREAM model has been run to provide continuous daily Q and R time series, at the outlet section of each subbasin and over each grid pixel, respectively. By considering the spatial/temporal availability of both in situ and satellite observations, the entire analysis period covers the maximum common observation period, i.e., from 01 January 2003 to 15 July 2016 at daily time

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

scale. To establish the goodness-of-fit of the model, the simulated river discharge and runoff timeseries are compared against in situ river discharge and modelled runoff data.

5.2 Model Evaluation Criteria and Performance Metrics

The model has been run over a 13.5-year period split into two sub periods: the first 8 years, from January 2003 to December 2010, have been used to calibrate the model successively validated over the remaining 5.5 years (January 2011 - July 2016).

In particular, three different validation schemes have been adopted to assess the robustness of the STREAM model:

1. Internal validation aimed to test the plausibility of both the model structure and the parameter set in providing reliable estimates of the hydrological variables against which the model is calibrated. For this purpose, a comparison between observed and simulated river discharge time series on the sections used for model calibration has been carried out for both the calibration and validation sub periods.
2. Cross-validation testing the goodness of the model structure and the calibrated model parameters to predict hydrological variables at locations not considered in the calibration phase. In this respect, the cross-validation has been carried out by comparing observed and simulated river discharge time series in gauged basins not considered during the calibration phase;
3. External validation aimed to test the capability of the model "*to get the right answers for the right reasons*" (Kirchner 2006). In this respect, the capability of the model to reproduce variables (e.g., fluxes or state variables) other than discharge and not considered in the calibration phase, should be tested. As runoff is a secondary product of the STREAM model, obtained indirectly from the calibration of the river discharge (basin-integrated runoff), the comparison in terms of runoff can be considered as a further external validation of the model. Runoff, differently from discharge, cannot be directly measured. It is generally modelled through land surface or hydrological models. Its validation requires a comparison against

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development Discussions
Open Access
EGU

modelled data that, however, suffer from uncertainties (Beck et al., 2017). Based on that, in this study the GRUN runoff dataset described in the section 3.3 has been used for a qualitative comparison.

5.3 Performance Metrics

To measure the goodness-of-fit between simulated and observed river discharge data three performance scores have been used:

- the relative root mean square error, RRMSE:

$$RRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Qsim_i - Q_{obs_i}\right)^2}}{\frac{1}{n}\sum_{i=1}^{n}\left(Q_{obs_i}\right)} \qquad (5)$$

where Q_{obs} and Q_{sim} are the observed and simulated discharge time series of length n. RRMSE values range from 0 to +∞, the lower the RRMSE, the better the agreement between observed and simulated data.

- the Pearson correlation coefficient, R, measures the linear relationship between two variables:

$$R = \frac{\sum_{i=1}^{n}(Qsim_i - \overline{Qsim_i})(Qobs_i - \overline{Qobs_i})}{\sqrt{\sum_{i=1}^{n}(Qsim_i - \overline{Qsim_i})^2 (Qobs_i - \overline{Qobs_i})^2}} \qquad (6)$$

where $\overline{Q_{obs}}$ and $\overline{Q_{sim}}$ represent the mean values of Q_{obs} and Q_{sim}, respectively. The values of R range between -1 and 1; higher values of R indicate a better agreement between observed and simulated data.

- the Kling-Gupta efficiency index (KGE, Gupta et al., 2009), which provides direct assessment of four aspects of discharge time series, namely shape, timing, water balance and variability. It is defined as follows:

$$KGE = 1 - \sqrt{(R-1)^2 + (\delta - 1)^2 + (\varepsilon - 1)^2} \qquad (7)$$

where R is the correlation coefficient, δ the relative variability and ε the bias normalized by the standard deviation between observed and simulated discharge. The KGE values range between -∞ and 1; the higher the KGE, the better the agreement between observed and simulated data. Simulations characterized by values of KGE in the range -0.41 and 1 can be assumed as reliable;

values of KGE greater than 0.5 have been assumed good with respect to their ability to reproduce observed time series (Thiemig et al., 2013).

6. RESULTS

The testing and validation of the STREAM model is presented and discussed in this section according to the scheme illustrated in section 5.2.

6.1 Internal Validation

The performance of the STREAM model over the calibrated river sections is illustrated in Figure 4 and summarized in Table 2. Figure 4 shows observed and simulated river discharge time series over the whole study period (2003-2016); in Table 2 the performance scores are evaluated separately for the calibration and validation sub periods. It is worth noting that the model accurately simulates the observed river discharge data and is able to give the "right answer" with good modelling performances. Score values of KGE and R over the calibration (validation) period are higher than 0.62 (0.67) and 0.75 (0.75) (resp.) for all the sections; RRMSE is lower than 46% (51%) for all the sections except for section 9, where it rises up to 71% (77%). The performances remain good even if they are evaluated over the entire study period as indicated by the scores on the top of each plot of Figure 4.

6.2 Cross-validation

The cross-validation has been carried out over the six river sections illustrated in Figure 5 not used in the calibration step. The performance scores on the top of each plot refer to the entire study periods; the scores split for calibration and validation periods are reported in Table 2. For some river sections the performance is quite low (see, e.g., river section 1, 2 and 5) whereas for others the model is able to simulate the observed discharge data quite accurately (e.g., 7 and 8). In particular, for river sections 1, 2 even if KGE reaches values equal to 0.35 and 0.40 (for the whole period), respectively, there is not a good agreement between observed and simulated river discharge and the R score is lower than 0.55 for both river sections. The worst performance is obtained over section 5, with negative KGE and low R (high RRSME). These results are certainly influenced by the presence of dams located upstream to these river sections (see Table 1): the model, not having a specific module for modelling reservoirs, is not able to accurately reproduce the dynamics of river discharge over regulated river sections. Better performances are obtained over river sections 3 (slightly influenced by the presence of dams in section 1 and 2), 7 and 8. In particular, over river section 7, the STREAM model overestimates the observed river discharge highlighting that the model parameters estimated for river section 6 are not suitable to accurately reproduce river discharge for river section 7 (see Figure 3 and Figure 5). Conversely, the performances over river section 8, whose parameters have been set equal to the ones of river section 10, are quite high (KGE equal to 0.71, 0.80 and 0.77 for the entire, the calibration and the validation period, respectively; R equal to 0.83, 0.84 and 0.84 for the entire, calibration and validation periods, respectively).

This finding, which could be due to different/similar interbasin characteristics, raises doubts about the robustness of model parameters and whether it is actually possible to transfer model parameters from one river section to another. A more in-depth investigation about the model calibration procedure will be carried out in future studies.

6.3 External Validation

For the external validation, the monthly runoff time series provided by the GRUN datasets have been compared against the ones computed by the STREAM model. For that, STREAM daily runoff time series have been aggregated at monthly scale and re-gridded at the same spatial resolution of the GRUN dataset (0.5°). The comparison is illustrated in Figure 6 for the common period 2003–2014. Although the two datasets consider different rainfall inputs, the two models agree in identifying two distinct zones, i.e., the western and the eastern area. Likely due to the calibration procedure, the STREAM runoff map appears patchier with respect to GRUN and discontinuities along the sub-basin boundaries (identified in Figure 3) can be noted. This should be ascribed to the automatic calibration procedure of the model that, differently from other calibration techniques (e. g., regionalization procedures), does not consider the basin physical attributes like soil, vegetation, and geological

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development Discussions

properties that govern spatial dynamics of hydrological processes. This calibration procedure can generate sharp discontinuities even for neighbouring subcatchments individually calibrated. It leads to discontinuities in model parameter values and consequently in the simulated hydrological variable (runoff).

7. DISCUSSION

In the previous sections, the ability of the STREAM model to accurately simulate river discharge and runoff time series has been presented. In particular, Figures 4, 5 and 6 demonstrate that satellite observations of precipitation, soil moisture and terrestrial water storage anomalies can provide accurate daily river discharge estimates for near-natural large basins (absence of upstream dams), and for basins with draining area lower than 160’000 km^2 (see section 7), i.e., at spatial/temporal resolution lower than the ones of the TWSA input data (monthly, 160’000 km^2). This is an important result of the study as it demonstrates, on one hand, that the model structure is appropriate with respect to the data used as input and, on the other hand, the great value of information contained into TWSA data that, even if characterized by limited spatial/temporal resolution, can be used to simulate runoff and river discharge at basin scale. Hereinafter, the strengths and the main limitations of the STREAM approach are discussed.

Among the strengths of the STREAM model it is worth highlighting:

1. **Remote sensing-based data-driven model**. Discharge and runoff estimates are obtained through a remote sensing-based data-driven model, simpler than classical hydrological models or LSMs. In particular, discharge and runoff estimates are obtained by exploiting as much as possible satellite observations and by keeping the modelling component at a minimum. The knowledge of the key mechanisms and processes that act in the formation of runoff, like the role of the soil moisture in determining the response of the catchment to precipitation, played a major role in the definition of the model structure. Being an observational-based approach, the STREAM model presents two main advantages: 1) possibility to directly ingest observations (soil moisture and terrestrial water storage

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

data) into the model structure, allowing to take implicitly into account some processes, mainly human-driven (e.g., irrigation, change in the land use), which might have a large impact on the hydrological cycle and hence on total runoff; 2) the independence with respect to existing large scale hydrological models such as, e.g., the evapotranspiration is not explicitly modelled.

2. **Simplicity**. The STREAM data-driven structure: 1) limits the input data required (only precipitation, T_{air}, soil moisture and TWSA data are needed as input; LSM/GHMs require many additional inputs such as wind speed, shortwave and longwave radiation, pressure and relative humidity); 2) limits and simplifies the processes to be modelled for runoff/discharge simulation. Processes like evapotranspiration, infiltration or percolation, are not modelled therefore avoiding the need of using sophisticated and highly parameterized equations (e.g., Penman-Monteith for evapotranspiration, Allen et al.,1998, Richard equation for infiltration, Richard, 1931); 3) limits the number of parameters (only 8 parameters have to be calibrated) thus simplifying the calibration procedure and potentially reduce the model uncertainties related to the estimation of parameter values.

3. **Versatility**. The STREAM model is a versatile model suitable for daily runoff and discharge estimation over sub-basins with different physiographic characteristics. The results obtained in this study clearly indicate the potential of this approach to be extended at the global scale. Moreover, the model can be easily adapted to ingest input data with spatial/temporal resolution different from the one tested in this study (0.25°/daily). For instance, satellite missions with higher space/time resolution, or near real time satellite products could be considered. As an example, the Next Generation Gravity Mission design studies all encompass double-pair scenarios, which would greatly improve upon the current spatial resolution of single-pair missions like GRACE and GRACE-FO (> 100’000 km²).

4. **Computationally inexpensive**. Due to its simplicity and the limited number of parameters to be calibrated, the computational effort for the STREAM model is very limited.

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

However, some limitations have to be acknowledged for the current version of the STREAM model:

1. **Presence of reservoir, diversion, dams or flood plain**. As the STREAM model does not explicitly consider the presence of discontinuity elements along the river network (e. g, reservoir, dam or floodplain), discharge estimates obtained for sections located downstream of such elements might be inaccurate (see, e.g., river sections 1 and 2 in Figure 5).

2. **Need of in situ data for model calibration and robustness of model parameters.** As discussed in the results section, parameter values of the STREAM model are set through an automatic calibration procedure aimed at minimizing the differences between simulated and observed river discharge. The main drawback of this parameterization technique is that the models parameterized with this technique may exhibit (1) poor predictability of state variables and fluxes at locations and periods not considered in the calibration, and (2) sharp discontinuities along sub-basin boundaries in state flux, and parameter fields (e.g., Merz and Blöschl, 2004).

To overcome these issues, several regionalization procedures, as for instance summarized in Cislaghi et al. (2020), could be conveniently applied to transfer model parameters from hydrologically similar catchments to a catchment of interest. In particular, the regionalization of model parameters could allow to: i) estimate discharge and runoff time series over ungauged basins overcoming the need of discharge data recorded from in–situ networks; ii) estimate the model parameter values through a physically consistent approach, linking them to the characteristics of the basins; iii) solve the problem of discontinuities in the model parameters, avoiding to obtain patchy unrealistic runoff maps.

8. CONCLUSIONS

This study presents a new data-driven model, STREAM, for runoff and river discharge estimation. By using as input satellite data of precipitation, soil moisture and terrestrial water storage anomalies, the model has been able to provide accurate daily river discharge and runoff estimates at the outlet

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

river section and the inner river sections and over a 0.25°×0.25° spatial grid of the Mississippi river basin. In particular, the model is suitable to reproduce:

1. river discharge time series over the calibrated river section with good performances both in calibration and validation periods;
2. river discharge time series over river sections not used for calibration and not located downstream dams or reservoirs;
3. runoff time series with a quite good agreement with respect to the well-established GRUN observational-based dataset used for comparison.

The integration of observations of soil moisture, precipitation and terrestrial water storage anomalies is a first alternative method for river discharge and runoff estimation with respect to classical methods based on the use of TWSA-only (suitable for river basins larger than 160’000 km², monthly time scale) or on classical LSMs (Cai et al., 2014).

Moreover, although simple, the model has demonstrated a great potential to be easily applied over subbasins with different climatic and topographic characteristics, suggesting also the possibility to extend its application to other basins. In particular, the analysis over basins with high human impact, where the knowledge of the hydrological cycle and the river discharge monitoring is very important, deserves special attention. Indeed, as the STREAM model is directly ingesting observations of soil moisture and terrestrial water storage data, it allows the modeller to neglect processes that are implicitly accounted for in the input data. Therefore, human-driven processes (e.g., irrigation, land use change), that are typically very difficult to simulate due to missing information and might have a large impact on the hydrological cycle, hence on total runoff, could be implicitly modelled. The application of the STREAM model on a larger number of basins is also required to investigate the possibility to regionalize the model parameters and overcome the limitations of the automatic calibration procedure highlighted in the discussion section.

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development
Discussions

AUTHOR CONTRIBUTION

S.C. performed the analysis and wrote the manuscript. G.G. collected the data and helped in performing the analysis; C.M, L.B., A.T., N.S., H.H.F., C.M., M.R. and J.B. contributed to the supervision of the work. All authors discussed the results and contributed to the final manuscript.

CODE AVAILABILITY

The STREAM model code will be made available once the manuscript will be published.

DATA AVAILABILITY

All data and codes used in the study are freely available online. Air temperature data are available at https://psl.noaa.gov/data/gridded/data.cpc.globaltemp.html (last access 25/11/202). In situ river discharge data have been taken from the Global Runoff Data Center (GRDC, https://www.bafg.de/GRDC/EN/Home/homepage_node.html (last access 25/11/202). Precipitation and soil moisture data are available from http://pmm.nasa.gov/data-access/downloads/trmm and https://esa-soilmoisture-cci.org/, respectively.

COMPETING INTERESTS

The authors declare that they have no conflict of interest.

ACKNOWLEDGMENTS

The authors wish to thank the Global Runoff Data Centre (GRDC) for providing most of the streamflow data throughout Europe. The authors gratefully acknowledge support from ESA through the STREAM Project (EO Science for Society element Permanent Open Call contract n° 4000126745/19/I-NB).

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

REFERENCE

Albergel, C., Rüdiger, C., Carrer, D., Calvet, J. C., Fritz, N., Naeimi, V., Bartalis, Z., & Hasenauer, S. (2009). An evaluation of ASCAT surface soil moisture products with in-situ observations in southwestern France. Hydrology and Earth System Sciences,13, 115–124, doi:10.5194/hess-13-115-2009.

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration — guidelines for computing crop water requirements. FAO Irrigation & Drainage Paper 56. FAO, Rome.

Balsamo, G., A. Beljaars, K. Scipal, P. Viterbo, B. vanden Hurk, M. Hirschi, and A. K. Betts (2009). A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the integrated forecast system, J. Hydrometeorol., 10(3), 623–643, doi:10.1175/2008JHM1068.1.

Barbarossa, V., Huijbregts, M. A., Beusen, A. H., Beck, H. E., King, H., & Schipper, A. M. (2018). FLO1K, global maps of mean, maximum and minimum annual streamflow at 1 km resolution from 1960 through 2015. Scientific data, 5, 180052.

Beck, H. E., van Dijk, A. I., de Roo, A., Dutra, E., Fink, G., Orth, R., & Schellekens, J. (2017). Global evaluation of runoff from ten state-of-the-art hydrological models. Hydrology and Earth System Sciences, 21(6), 2881-2903. doi; doi.org/10.5194/hess-21-2881-2017.

Berghuijs, W. R., Woods, R. A., Hutton, C. J., and Sivapalan, M. (2016). Dominant flood generating mechanisms across the United States, Geophys. Res. Lett., 43, 4382–4390, https://doi.org/10.1002/2016GL068070.

Berthet, L., Andréassian, V., Perrin, C., & Javelle, P. (2009). How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. Hydrology and Earth System Sciences, 13(6), 819-831.

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., & Savenije, H. H. G. (Eds.) (2013). Runoff predictions in ungauged basins: A synthesis across processes, places and scales. Cambridge: Cambridge University Press.

Botter, G., Peratoner, F., Porporato, A., Rodriguez-Iturbe, I., and Rinaldo, A. (2007b). Signatures of large-scale soil moisture dynamics on streamflow statistics across U.S. Climate regimes, Water Resour. Res., 43, W11413, doi:10.1029/2007WR006162.

Botter, G., Porporato, A., Daly, E., Rodriguez-Iturbe, I., and Rinaldo, A. (2007a). Probabilistic characterization of base flows in river basins: Roles of soil, vegetation, and geomorphology, Water Resour. Res., 43, W06404,doi:10.1029/2006WR005397.

Brocca, L., Melone, F., Moramarco, T. (2008). On the estimation of antecedent wetness conditions in rainfall-runoff modelling. Hydrological Processes, 22 (5), 629-642, doi:10.1002/hyp.6629. http://dx.doi.org/10.1002/hyp.6629.

Brocca, L., Melone, F., Moramarco, T., & Morbidelli, R. (2009). Antecedent wetness conditions based on ERS scatterometer data. *Journal of Hydrology*, *364*(1-2), 73-87

Brocca, L., Melone, F., & Moramarco, T. (2011). Distributed rainfall-runoff modelling for flood frequency estimation and flood forecasting. Hydrological processes, 25(18), 2801-2813.

Brocca, L., Ciabatta, L., Massari, C., Camici, S., & Tarpanelli, A. (2017). Soil moisture for hydrological applications: open questions and new opportunities. Water, 9(2), 140.

Cai, X., Yang, Z. L., David, C. H., Niu, G. Y., & Rodell, M. (2014). Hydrological evaluation of the Noah-MP land surface model for the Mississippi River Basin. Journal of Geophysical Research: Atmospheres, 119(1), 23-38.

Cislaghi, A., Masseroni, D., Massari, C., Camici, S., & Brocca, L. (2020). Combining a rainfall–runoff model and a regionalization approach for flood and water resource assessment in the western Po Valley, Italy. Hydrological Sciences Journal, 65(3), 348-370.

Crochemore, L., Isberg, K., Pimentel, R., Pineda, L., Hasan, A., & Arheimer, B. (2020). Lessons learnt from checking the quality of openly accessible river flow data worldwide. Hydrological Sciences Journal, 65(5), 699-711

Crow, W. T., Bindlish, R., & Jackson, T. J. (2005). The added value of spaceborne passive microwave soil moisture retrievals for forecasting rainfall-runoff partitioning. Geophysical Research Letters, 32(18).

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

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

Famiglietti, J. S., & Rodell, M. (2013). Water in the balance. Science, 340(6138), 1300-1301.

Fan, Y. & Van den Dool, H. A (2008). Global monthly land surface air temperature analysis for 1948–present. Journal of Geophysical Research: Atmospheres 113, D01103.

Fekete, B. M., Looser, U., Pietroniro, A., and Robarts, R. D. (2012). Rationale for monitoring discharge on the ground, J. Hydrometeorol.,13, 1977–1986.

Ghiggi, G., Humphrey, V., Seneviratne, S. I., & Gudmundsson, L. (2019). GRUN: an observation-based global gridded runoff dataset from 1902 to 2014. Earth System Science Data, 11(4), 1655-1674.

Ghotbi, S., Wang, D., Singh, A., Blöschl, G., & Sivapalan, M. (2020). A New Framework for Exploring Process Controls of Flow Duration Curves. Water Resources Research, 56(1), e2019WR026083.

Gudmundsson, L., & Seneviratne, S. I. (2016). Observation-based gridded runoff estimates for Europe (E-RUN version 1.1). Earth System Science Data, 8(2), 279-295.

Gudmundsson, L., Wagener, T., Tallaksen, L. M., & Engeland, K. (2012a). Evaluation of nine large-scale hydrological models with respect to the seasonal runoff climatology in Europe. Water Resources Research, 48(11).

Gudmundsson, L., Tallaksen, L. M., Stahl, K., Clark, D. B., Du-mont, E., Hagemann, S., Bertrand, N., Gerten, D., Heinke, J.,Hanasaki, N., Voss, F., and Koirala, S. (2012b). Comparing Large-Scale Hydrological Model Simulations to Observed Runoff Percentiles in Europe, J. Hydrometeorol., 13, 604–62.

Gupta VK, Waymire E, Wang CT. (1980). A representation of an instantaneous unit hydrograph from geomorphology. Water Resources Research 16: 855–862, doi: 10.1029/WR016i005p00855.

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology, 377(1-2), 80-91.

Haddeland, I., Heinke, J., Voß, F., Eisner, S., Chen, C., Hagemann, S., & Ludwig, F. (2012). Effects of climate model radiation, humidity and wind estimates on hydrological simulations. Hydrology and Earth System Sciences, 16(2), 305-318.

Hastie, T., Tibshirani, R., and Friedman, J. H. (2009). The Elements ofStatistical Learning – Data Mining, Inference, and Prediction,Second Edition, Springer Series in Statistics, Springer, NewYork, 2nd Edn., available at: http://www-stat.stanford.edu/~tibs/ElemStatLearn/ (last access: 5 July 2016).

Hong, Y., Adler, R. F., Hossain, F., Curtis, S., & Huffman, G. J. (2007). A first approach to global runoff simulation using satellite rainfall estimation. Water Resources Research, 43(8).

Houborg, R., Rodell, M., Li, B., Reichle, R., & Zaitchik, B. F. (2012). Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. Water Resources Research, 48(7).

Hu GR., Li XY. (2018). Subsurface Flow. In: Li X., Vereecken H. (eds) Observation and Measurement. Ecohydrology. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-47871-4_9-1

Huffman, G. J., R. F. Adler, D. T. Bolvin, G. J. Gu, E. J. Nelkin, K. P. Bowman, Y. Hong, E. F. Stocker, and D. B. Wolff. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. Journal of Hydrometeorology 8 (1): 38–55. doi:10.1175/jhm560.1.

Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., & Adler, R. F. (2014). TRMM Version 7 3B42 and 3B43 Data Sets. NASA/GSFC, Greenbelt, MD.

Kim, H., Watanabe, S., Chang, E. C., Yoshimura, K., Hirabayashi, J., Famiglietti, J., and Oki, T. (2017). Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1) [Data set], Data Integration and Analysis System (DIAS), https://doi.org/10.20783/DIAS.501.

Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. Water Resources Research, 42(3).

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

Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. Water resources research, 48(4).

Lehner, B., C. Reidy Liermann, C. Revenga, C. Vörösmarty, B. Fekete, P. Crouzet, P. Döll, M. Endejan, K. Frenken, J. Magome, C. Nilsson, J.C. Robertson, R. Rodel, N. Sindorf, and D. Wisser. 2011. High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. Frontiers in Ecology and the Environment 9 (9): 494-502.

Long, D., Longuevergne, L., & Scanlon, B. R. (2014). Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites. Water Resources Research, 50(2), 1131-1151.

Lorenz, C., H. Kunstmann, B. Devaraju, M. J. Tourian, N. Sneeuw, and J. Riegger (2014). Large-Scale Runoff from Landmasses: A Global Assessment of the Closure of the Hydrological and Atmospheric Water Balances. J. Hydrometeor., 15, 2111–2139, doi:10.1175/JHM-D-13-0157.1.

Luthcke, S.B., Sabaka, T.J., Loomis, B.D., Arendt, A.A., McCarthy, J.J., Camp, J. (2013) Antarctica, Greenland and Gulf of Alaska land-ice evolution from an iterated GRACE global mascon solution, Journal of Glaciology, Vol. 59, No. 216, 2013 doi:10.3189/2013JoG12J147.

Massari, C., Brocca, L., Tarpanelli, A., Hong, Y., Crow, W., Ciabatta, L, Camici, S., Barbetta, S., Moramarco, T. (2016). Global surface runoff estimation in near real time by using SMAP and GPM, poster at SMAP conference.

Massari, C., Brocca, L., Barbetta, S., Papathanasiou, C., Mimikou, M., & Moramarco, T. (2014). Using globally available soil moisture indicators for flood modelling in Mediterranean catchments. Hydrology and Earth System Sciences, 18(2), 839.

Merz, R., & Blöschl, G. (2009). A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria. Water Resources Research, 45(1).

Mueller Schmied, H., Adam, L., Eisner, S., Fink, G., Flörke, M., Kim, H., ... & Song, Q. (2016). Variations of global and continental water balance components as impacted by climate forcing uncertainty and human water use. Hydrology and Earth System Sciences, 20(7), 2877-2898.

Muneepeerakul, R., Azaele, S., Botter, G., Rinaldo, A., & Rodriguez-Iturbe, I. (2010). Daily streamflow analysis based on a two-scaled gamma pulse model. Water Resources Research, 46(11).

Nash, J. E. (1957). The form of the instantaneous unit hydrograph, IASH publication no. 45, 3–4, 114–121.

Natural Resources Conservation Service (NRCS) (1986), Urban hydrology for small watersheds, Tech. Release 55, 2nd ed., U.S. Dep. of Agric., Washington, D. C. (available at ftp://ftp.wcc.nrcs.usda.gov/downloads/hydrology_hydraulics/tr55/tr55.pdf)

Orth, R., & Seneviratne, S. I. (2015). Introduction of a simple-model-based land surface dataset for Europe. Environmental Research Letters, 10(4), 044012.

Pellet, V., Aires, F., Munier, S., Fernández Prieto, D., Jordá, G., Dorigo, W. A., ... & Brocca, L. (2019). Integrating multiple satellite observations into a coherent dataset to monitor the full water cycle–application to the Mediterranean region. Hydrology and Earth System Sciences, 23(1), 465-491.

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., ... & Hagemann, S. (2014). Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. Proceedings of the National Academy of Sciences, 111(9), 3262-3267.

Rakovec, O., Kumar, R., Attinger, S., & Samaniego, L. (2016). Improving the realism of hydrologic model functioning through multivariate parameter estimation. Water Resources Research, 52(10), 7779-7792.

Richards, L.A. (1931). Capillary conduction of liquids through porous mediums. Physics. 1 (5): 318–333. Bibcode:1931Physi.1.318R. doi:10.1063/1.1745010.

Riegger, J., and M. J. Tourian (2014), Characterization of runoff-storage relationships by satellite gravimetry and remote sensing, Water Resour. Res., 50, 3444–3466, doi:10.1002/2013WR013847.

Rodell, M., Beaudoing, H. K., L'Ecuyer, T. S., Olson, W. S., Famiglietti, J. S., Houser, P. R., Adler, R., Bosilovich, M. G.,Clayson, C. A., Chambers, D., Clark, E., Fetzer, E. J., Gao, X., Gu, G., Hilburn, K., Huffman, G. J., Lettenmaier, D. P., Liu, W. T., Robertson, F. R., Schlosser, C. A., Sheffield, J. and Wood, E. F. (2015). The observed state of the water cycle in the early 15twenty-first century, J. Clim., 28(21), 8289–8318, doi:10.1175/JCLI-D-14-00555.1.

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. (2017). A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset, Earth Syst. Sci. Data, 9, 389–413,https://doi.org/10.5194/essd-9-389-2017.

Schwanghart, W., & Kuhn, N. J. (2010). TopoToolbox: A set of Matlab functions for topographic analysis. Environmental Modelling & Software, 25(6), 770-781.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., ... & Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. Earth-Science Reviews, 99(3-4), 125-161.

Sneeuw, N., Lorenz, C., Devaraju, B., Tourian, M. J., Riegger, J., Kunstmann, H., & Bárdossy, A. (2014). Estimating runoff using hydro-geodetic approaches. Surveys in Geophysics, 35(6), 1333-1359.

Solomatine, D. P., & Ostfeld, A. (2008). Data-driven modelling: some past experiences and new approaches. Journal of hydroinformatics, 10(1), 3-22.

Tapley, B.D., Watkins, M.M., Flechtner, F. et al. (2019). Contributions of GRACE to understanding climate change. Nat. Clim. Chang. 9, 358–369, doi:10.1038/s41558-019-0456-2.

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., & De Roo, A. (2013). Hydrological evaluation of satellite rainfall estimates over the Volta and Baro-Akobo Basin. Journal of Hydrology, 499, 324-338.

Tourian, M. J., Reager, J. T., & Sneeuw, N. (2018). The total drainable water storage of the Amazon river basin: A first estimate using GRACE. Water Resources Research, 54. https://doi.org/10.1029/2017WR021674.

Tramblay, Y., Bouvier, C., Martin, C., Didon-Lescot, J. F., Todorovik, D., & Domergue, J. M. (2010). Assessment of initial soil moisture conditions for event-based rainfall–runoff modelling. Journal of Hydrology, 387(3-4), 176-187.

Troutman, B. M., Karlinger, M.B. (1985). Unit hydrograph approximation assuming linear flow through topologically random channel networks. Water Resources Research, 21: 743 – 754, doi: 10.1029/WR021i005p00743.

Vose, R.S., Applequist, S., Durre, I., Menne, M.J., Williams, C.N., Fenimore, C., Gleason, K., & Arndt, D. (2014). Improved Historical Temperature and Precipita on Time Series For U.S. Climate Divisions. Journal of Applied Meteorology and Climatology, 53(May), 1232–1251. DOI: 10.1175/JAMC-D-13-0248.1

Vörösmarty C. J., and Coauthors (2002). Global water data: A newly endangered species. Eos, Trans. Amer. Geophys. Union, 82, 54.

Wagner, W., Blöschl, G., Pampaloni, P., Calvet, J. C., Bizzarri, B., Wigneron, J. P., & Kerr, Y. (2007). Operational readiness of microwave remote sensing of soil moisture for hydrologic applications. Hydrology Research, 38(1), 1-20.

Wagner, W., Lemoine, G., & Rott, H. (1999). A method for estimating soil moisture from ERS scatterometer and soil data. Remote Sensing of Environment,70, 191–207, doi:10.1016/S0034-4257(99)00036-X.

Wisser, D., Fekete, B. M., Vörösmarty, C. J., and Schumann, A. H. (2010). Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network- Hydrology (GTN-H), Hydrol.Earth Syst. Sci., 14, 1–24, doi:10.5194/hess-14-1-2010.

Yokoo, Y., & Sivapalan, M. (2011). Towards reconstruction of the flow duration curve: Development of a conceptual framework with a physical basis. Hydrology and Earth System Sciences, 15(9), 2805–2819. https://doi.org/10.5194/hess-15-2805-2011.

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Zhang, Y., Pan, M., Sheffield, J., Siemann, A. L., Fisher, C. K., Liang, M., ... & Zhou, T. (2018). A Climate Data Record (CDR) for the global terrestrial water budget: 1984–2010. Hydrology and Earth System Sciences (Online), 22(PNNL-SA-129750).

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development
Discussions
Open Access
EGU

773 Table 1. Location of gauging stations over the Mississippi basins and upstream contributing area. Red
774 colored text indicates stations where the STREAM model has been calibrated.

#	**River**	**Station name**	**Latitude (°)**	**Longitude (°)**	**Upstream area (km^2)**	**Mean annual river discharge (m^3/s)**	**Presence of dam**
1	Missouri	Bismarck, ND	-100.82	46.81	481’232	633	Garrison dam
2	Missouri	Omaha, NE	-95.92	41.26	814’371	914	Gavins Point Dam
3	Missouri	Kansas City, MO	-94.59	39.11	1’229’427	1499	---
4	Missouri	Hermann, MO	-91.44	38.71	1’330’000	2326	---
5	Kansas	Wamego, KS	-96.30	39.20	143’054	141	Kanopolis
6	Mississippi	Keokuk, IA	-91.37	40.39	282’559	1948	---
7	Rock	Near Joslin, IL	-90.18	41.56	23’835	199	---
8	Mississippi	Chester, IL	-89.84	37.90	1’776’221	6018	---
9	Arkansas	Murray Dam Near Little Rock, AR	-92.36	34.79	408’068	1249	---
10	**Mississippi**	**Vicksburg, MS**	**-90.91**	**32.32**	**2’866’590**	**17487**	**---**
11	Ohio	Metropolis, ILL.	-88.74	37.15	496’134	7931	---

775
776

Table 2. Performance scores obtained over the Mississippi river sections during the calibration and validation periods.

#	CALIBRATION PERIOD			VALIDATION PERIOD		
SCORE	KGE (-)	R (-)	RRMSE (%)	KGE (-)	R (-)	RRMSE (%)
CALIBRATED SECTIONS						
10	0.78	0.78	30	0.74	0.80	38
9	0.62	0.75	71	0.67	0.85	77
6	0.83	0.84	39	0.73	0.84	46
4	0.77	0.78	46	0.72	0.75	50
11	0.82	0.82	44	0.70	0.86	51
SECTIONS NOT USED FOR CALIBRATION						
1	-3.26	0.08	137	0.20	0.44	96
2	-0.57	0.48	118	0.40	0.53	89
3	0.16	0.71	83	0.39	0.70	72
5	-1.49	0.24	368	-1.26	0.31	358
7	0.53	0.68	71	0.20	0.70	81
8	0.80	0.84	36	0.77	0.84	39

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development
Discussions
Open Access
EGU

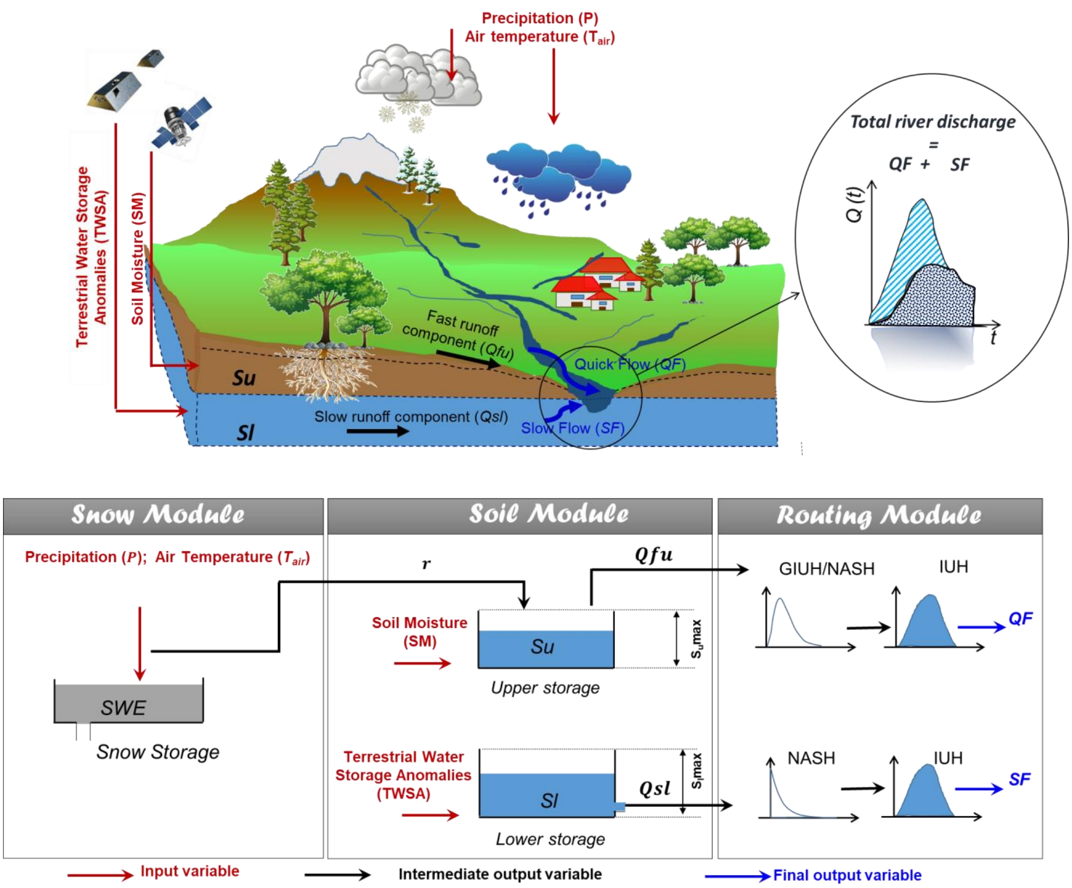


Figure 1. Configuration of the STREAM model adopted for total runoff estimation. The model includes three modules, the snow module allowing to separate snowfall from rainfall, the soil module that simulates the slow and quick runoff components (*Qsu* and Qfu, respectively) and the routing module for flood simulation. Red arrows indicate input variables; black arrows indicate intermediate output variables; blue arrows indicate final output variables. The components *Qfu* and *Qsu* are computed by using satellite *P*, soil moisture and TWSA data as input to the soil module. Please refer to text for symbols.

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

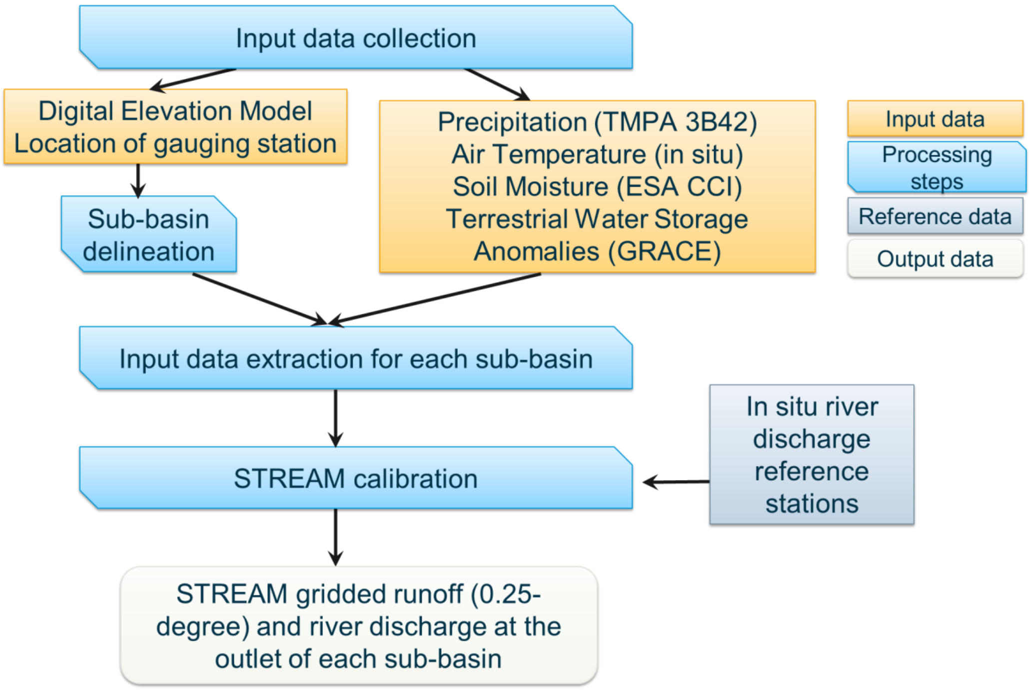


Figure 2. Processing steps of the STREAM approach.

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development
Discussions
Open Access
EGU

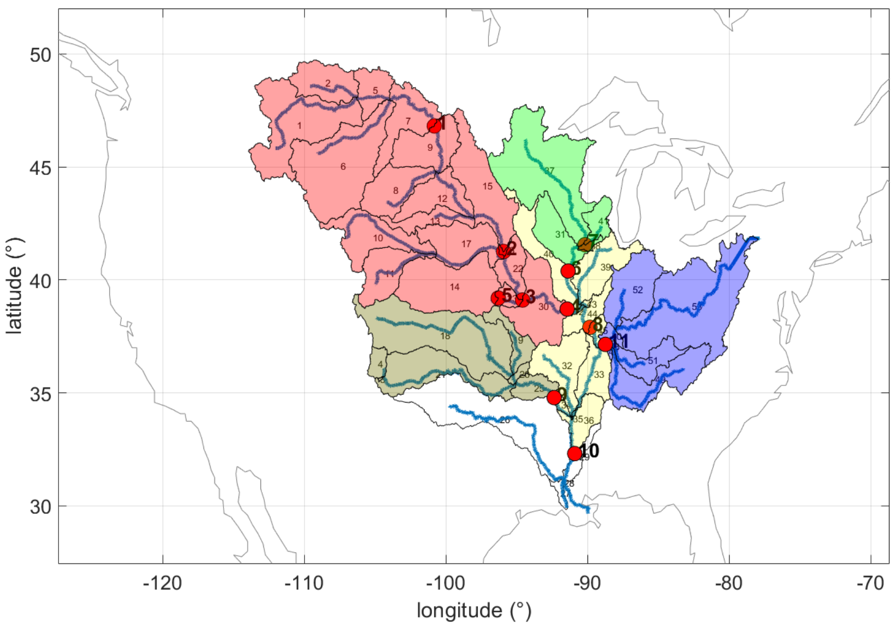


Figure 3. Mississippi sub-basin delineation. Red dots indicate the location of the discharge gauging stations; different colours identify different inner sections (and the related contributing sub-basins) used for the model calibration.

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development
Discussions

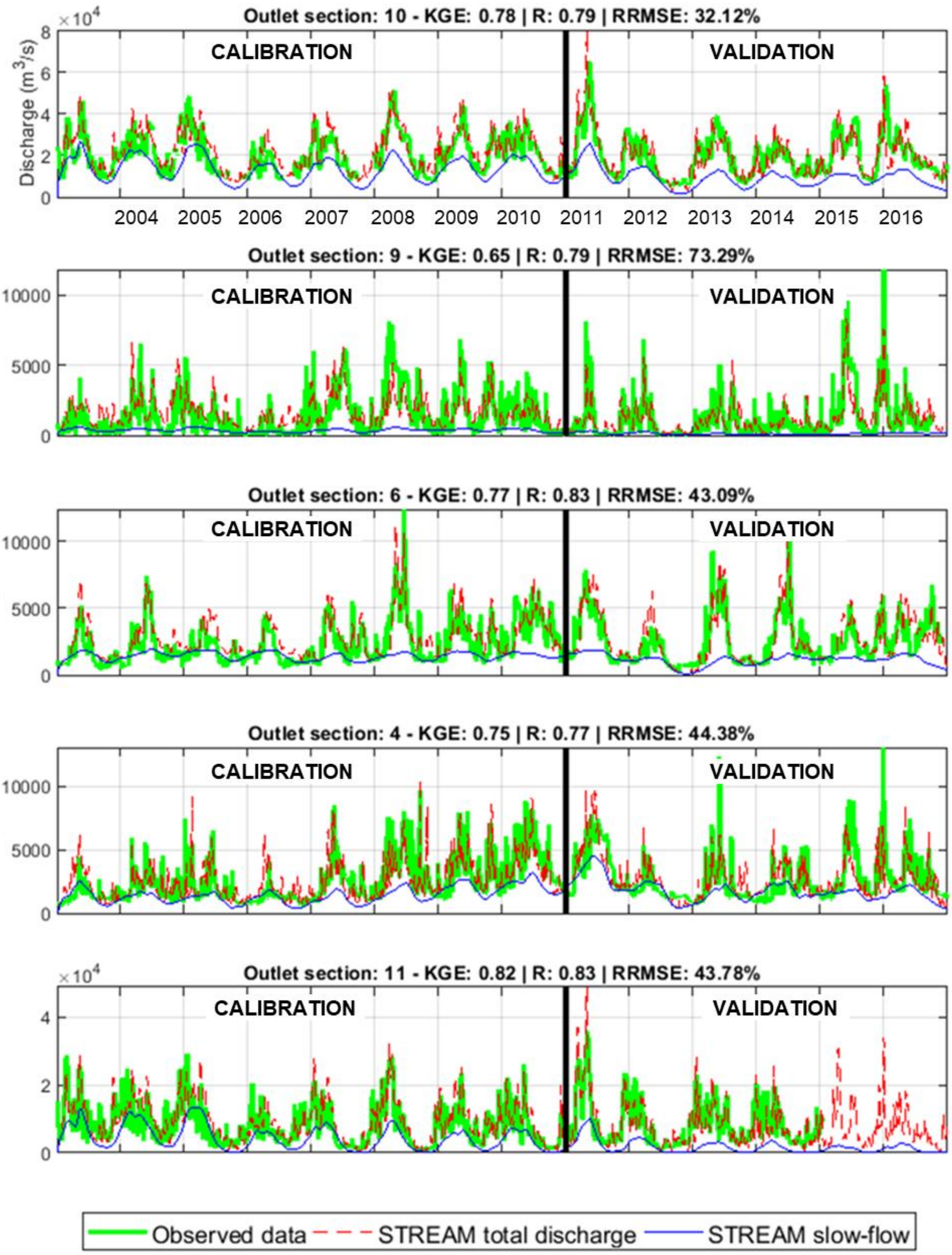
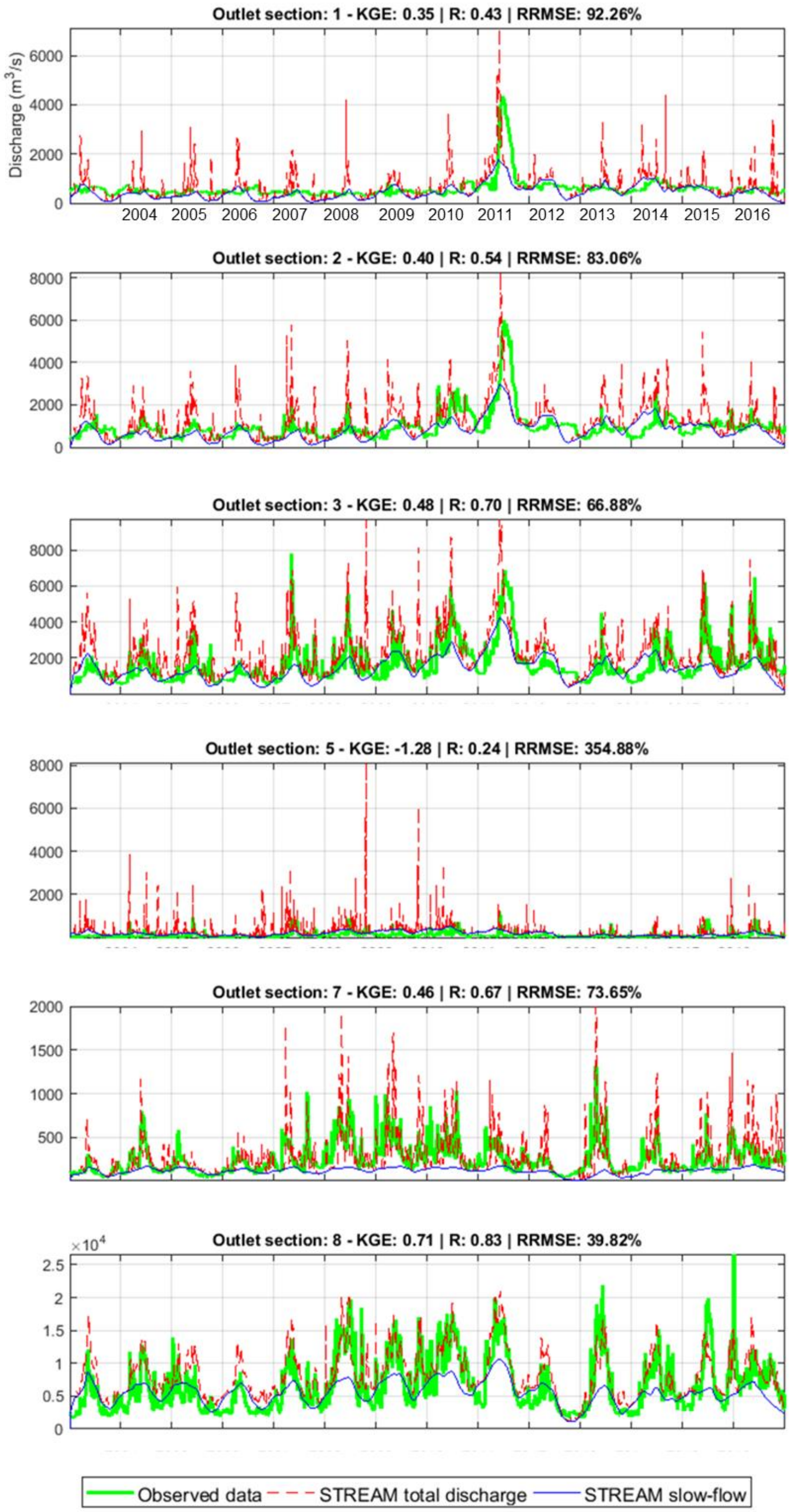


Figure 4. Comparison between observed and simulated river discharge time series over the five calibrated sections over Mississippi river basin. Performance scores at the top of each plot refer to the entire study period (2003–2016).


Outlet section: 1 - KGE: 0.35 | R: 0.43 | RRMSE: 92.26%
Discharge (m³/s)
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016
Outlet section: 2 - KGE: 0.40 | R: 0.54 | RRMSE: 83.06%
Outlet section: 3 - KGE: 0.48 | R: 0.70 | RRMSE: 66.88%
Outlet section: 5 - KGE: -1.28 | R: 0.24 | RRMSE: 354.88%
Outlet section: 7 - KGE: 0.46 | R: 0.67 | RRMSE: 73.65%
Outlet section: 8 - KGE: 0.71 | R: 0.83 | RRMSE: 39.82%
×10⁴
Observed data
STREAM total discharge
STREAM slow-flow

806

https://doi.org/10.5194/gmd-2020-399
Preprint. Discussion started: 25 January 2021

Geoscientific Model Development
Discussions
Open Access

Figure 5. Comparison between observed and simulated river discharge time series over the gauged sections not used in the calibration phase. Performance scores at the top of each plot refer to the entire study period (2003–2016).

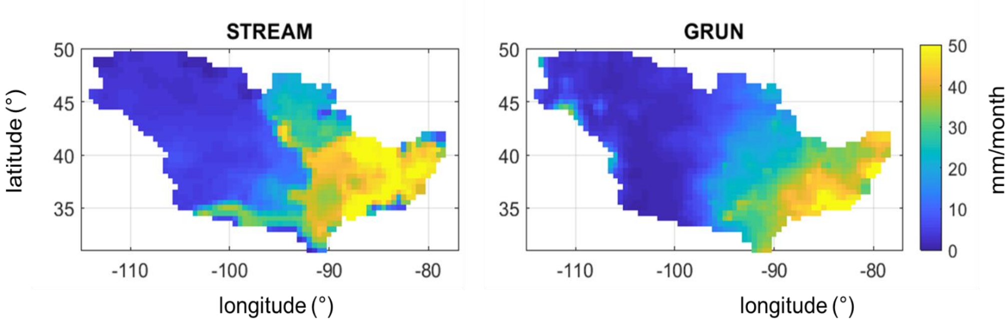


Figure 6. Mississippi river basin: mean monthly runoff for the period 2003–2014 obtained by STREAM and GRUN models.