



HAL
open science

Total water storage variability from GRACE mission and hydrological models for a 50,000 km² temperate watershed: the Garonne River basin (France)

Sylvain Biancamaria, Moussa Mballo, Patrick Le Moigne, J.M. Sánchez-Pérez, Grégory Espitalier-Noël, Youen Grusson, Roxelane Cakir, Vincent Häfliger, Florian Barathieu, Marhiu Trasmonte, et al.

► To cite this version:

Sylvain Biancamaria, Moussa Mballo, Patrick Le Moigne, J.M. Sánchez-Pérez, Grégory Espitalier-Noël, et al.. Total water storage variability from GRACE mission and hydrological models for a 50,000 km² temperate watershed: the Garonne River basin (France). *Journal of Hydrology: Regional Studies*, 2019, 24, pp.100609. 10.1016/j.ejrh.2019.100609 . hal-02272137

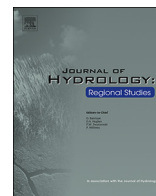
HAL Id: hal-02272137

<https://hal.science/hal-02272137v1>

Submitted on 16 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Total water storage variability from GRACE mission and hydrological models for a 50,000 km² temperate watershed: the Garonne River basin (France)

Sylvain Biancamaria^{a,*}, Moussa Mballo^a, Patrick Le Moigne^b, José Miguel Sánchez Pérez^c, Grégory Espitalier-Noël^c, Youen Grusson^c, Roxelane Cakir^c, Vincent Häfliger^d, Florian Barathieu^a, Marhiu Trasmonte^a, Aaron Boone^b, Eric Martin^{b,d}, Sabine Sauvage^{c,*}

^a LEGOS, Université de Toulouse, CNES, CNRS, IRD, UPS - 14 avenue Edouard Belin, 31400, Toulouse, France

^b CNRM, Université de Toulouse, Météo-France, CNRS - 42 avenue Gaspard Coriolis, 31057, Toulouse, France

^c ECOLAB, Université de Toulouse, CNRS, INPT, UPS - Avenue de l'Agrobiopole, 31326, Castanet Tolosan, France

^d IRSTEA, UR RECOVER, Aix-en-Provence, France

ARTICLE INFO

Keywords:

Garonne
GRACE
Hydrological model
Basin water balance
Cross-Validation
GRACE spatial resolution

ABSTRACT

Study Region: Garonne Basin, France.

Study Focus: This study analyses water mass variations for the whole Garonne basin (50,000 km² drainage area). To do so, Total Water Storage Anomalies (TWSA) from seven global solutions based on the Gravity Recovery And Climate Experiment (GRACE) satellite mission measurements (~300 km spatial resolution) are inter-compared with TWSA from two hydrological models, SAFRAN-ISBA-MODCOU (SIM) and Soil and Water Assessment Tool (SWAT), between January 2003 and December 2010.

New Hydrological Insights for the Region: Despite the small size of the Garonne basin compared to GRACE spatial resolution, good agreement between GRACE solutions and hydrological model TWSA has been found (maximum correlation coefficient ~0.9 and Nash-Sutcliffe Efficiency, NSE, ~0.7). These datasets showed that TWSA in the Garonne basin is mainly due to water stored in the first dozen meters of soil and in the shallow aquifer. To a smaller extent, snow also influences Garonne TWSA. Open surface water TWSA is quite small and TWSA from deep aquifer is negligible. The most important drought period occurred in 2011/2012, due to low precipitation during the two hydrological years and ETR close to previous years. Important precipitation in 2013/2014 helps to refill the water stocks. This study also showed that GRACE and models mismatches should be due to GRACE poor spatial resolution, but also to its monthly time resolution (rarely shown in previous studies).

1. Introduction

Studying the water cycle at global and regional scale requires a set of observation tools and models to estimate water and energy stored in different compartments of the water cycle (i.e. atmosphere, continental surfaces and oceans) and their fluxes. Over

* Corresponding authors.

E-mail addresses: sylvain.biancamaria@legos.obs-mip.fr (S. Biancamaria), sabine.sauvage@univ-tlse3.fr (S. Sauvage).

<https://doi.org/10.1016/j.ejrh.2019.100609>

Received 7 December 2018; Received in revised form 26 April 2019; Accepted 25 May 2019

2214-5818/ © 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

continents, it is also important to know water stocks and their variations at the watershed scale and seasonal to multi-annual time scales, for planning required actions from water managers to satisfy water and energy needs by human population within the basin. Traditionally, it is done with in situ measurements of river discharge, water table for ground water, snow depth, etc. These data remain the most precise and accurate measurements available. However, they are only partial measurement of all water fluxes within a watershed, as gauge network are at best heterogeneous in space and time (Pavelsky et al., 2014). Since few decades, some remote sensing sensors have proven their capabilities to complement in situ measurements and/or provide estimates of variables not measured in situ (Lettenmaier et al., 2015). Especially, the Gravity Recovery And Climate Experiment (GRACE) satellite mission, launched in 2002, can be used to estimate total water storage change variations at coarse spatial resolution (~300 km) and monthly time scale.

To get homogenous estimates of these fluxes in space and time, hydrological models have been developed during the last decades to compute water storage, dynamic and fluxes within river watersheds (Döll et al., 2016). However; these numerical models are inherently limited by the implemented physical processes, the approximation of the chosen equations to model these processes, the numerical approximation used to compute solutions to these equations, uncertainties in the parameters of the models and in the boundary conditions used as model inputs (Döll et al., 2016). Nowadays, in situ and satellite observations are commonly used to calibrate, validate and even correct models outputs or/and parameters (Lettenmaier et al., 2015).

Estimating the total water storage variations within watershed remains a difficult task, as hydrological models are often validated or calibrated only for some variables (more frequently against in situ river discharge, sometimes using satellite estimates of snow extent or, less frequently, against water table in situ measurements) but not for all variables and storage compartments.

The use of both GRACE data and hydrological modeling at watershed scale can help 1) to cross validate a combination of hydrology model outputs and 2) to identify modeled compartment(s) where the most important variations are occurring. However, due to GRACE coarse spatial resolution, many previous studies using GRACE data considered watershed with drainage area much higher than 100,000 km² (e.g. Wouters et al., 2014). It is not sure that GRACE data could be used for medium-size river basins, like the Garonne basin located in South West of France (50,000 km² drainage area). The present study investigates the potential of GRACE data for this kind of basin. Especially, it focuses on the following two questions: How to estimate and validate total water storage variations over the medium-size Garonne basin using GRACE data and hydrological model outputs at monthly scale? In which (hydrology) compartments those monthly variations are the most important? They are addressed in the following sections, which present the study domain (section 2.1), GRACE products (section 2.2) and hydrological models inputs and outputs (sections 2.3) used, the methodology applied to tackle these two questions (section 2.4), and finally the results obtained (section 3).

2. Study basin, data used and methodology

2.1. Garonne basin

The study focuses on the Garonne River basin (Fig. 1), which is one of the principal fluvial systems in France mainland and located in the southwest of the country. The basin (black polygon on Fig. 1) drains 50,000 km² at Tonneins, the last gauging station uninfluenced by the tidal action (120 km upstream the estuary). Its watershed consists of three main geographic entities: the Pyrenees Mountains to the south with some peaks exceeding 3000 m, the plateau of the Massif Central to the northeast between 1400 and 1700 m, and the plain between them, whose elevation is less than a few hundred metres.

The basin is also at the boundary of different climate zones. The Mediterranean climate in the east of the basin is characterized by hot and dry southeastern winds. The Garonne basin near its estuary is impacted by an Oceanic climate, with western winds causing high precipitation and cool temperatures. Watershed flows are impacted by these weather conditions, especially flows in the central part of the basin are generated by precipitation over the Massif Central and the Pyrenees mountains. The average annual precipitation is around 900 mm.y⁻¹ over the basin (according to <https://donneespubliques.meteofrance.fr>). Upstream Garonne and Ariège rivers discharges during spring are highly influenced by snowmelt from the Pyrenees Mountains (Caballero et al., 2007).

At the outlet of the basin (Tonneins), according to the national gauge measurements database (Banque Hydro, <http://www.hydro.eaufrance.fr/>), the average discharge between 1913 and 2018 is equal to 598 m³.s⁻¹. The highest discharge on record reached 5700 m³.s⁻¹ (March 5th, 1930) and the lowest, 37.5 m³.s⁻¹ (August 4th, 1989). The hydrological year goes from October to September of the subsequent year. Monthly discharge is maximum in February (1913/2018 mean: 970 m³.s⁻¹) and minimum in August (1913/2018 mean: 177 m³.s⁻¹).

According to the CORINE Land Cover 2012 (Büttner et al., 2014), the mountainous regions are dominated by forest and alpine grassland (37% of the watershed), while the plain is dominated by agricultural activities (60%). The remaining land is either artificialized land (2.5%) or water bodies (0.5%). Cambisols are the most frequent soil classes, covering 55% of the total area and spreading across the watershed. Different classes of shallow soils are also present in the upper part of the watershed (19%), such as lithosols, regosols, andosols, rendzinas and rankers. Luvisols cover 18% of the watershed, mostly on hillsides and in the plain. The Garonne River mostly flows on fluvisols (7%). Influence of human activities on monthly discharge values are not transmitted to the downstream hydrologic regime significantly and remain contained in the upper part of the basin (Grusson et al., 2017a).

2.2. GRACE satellites measurements

2.2.1. GRACE mission

The Gravity Recovery And Climate Experiment (GRACE) is a National Aeronautics and Space Administration (NASA) and

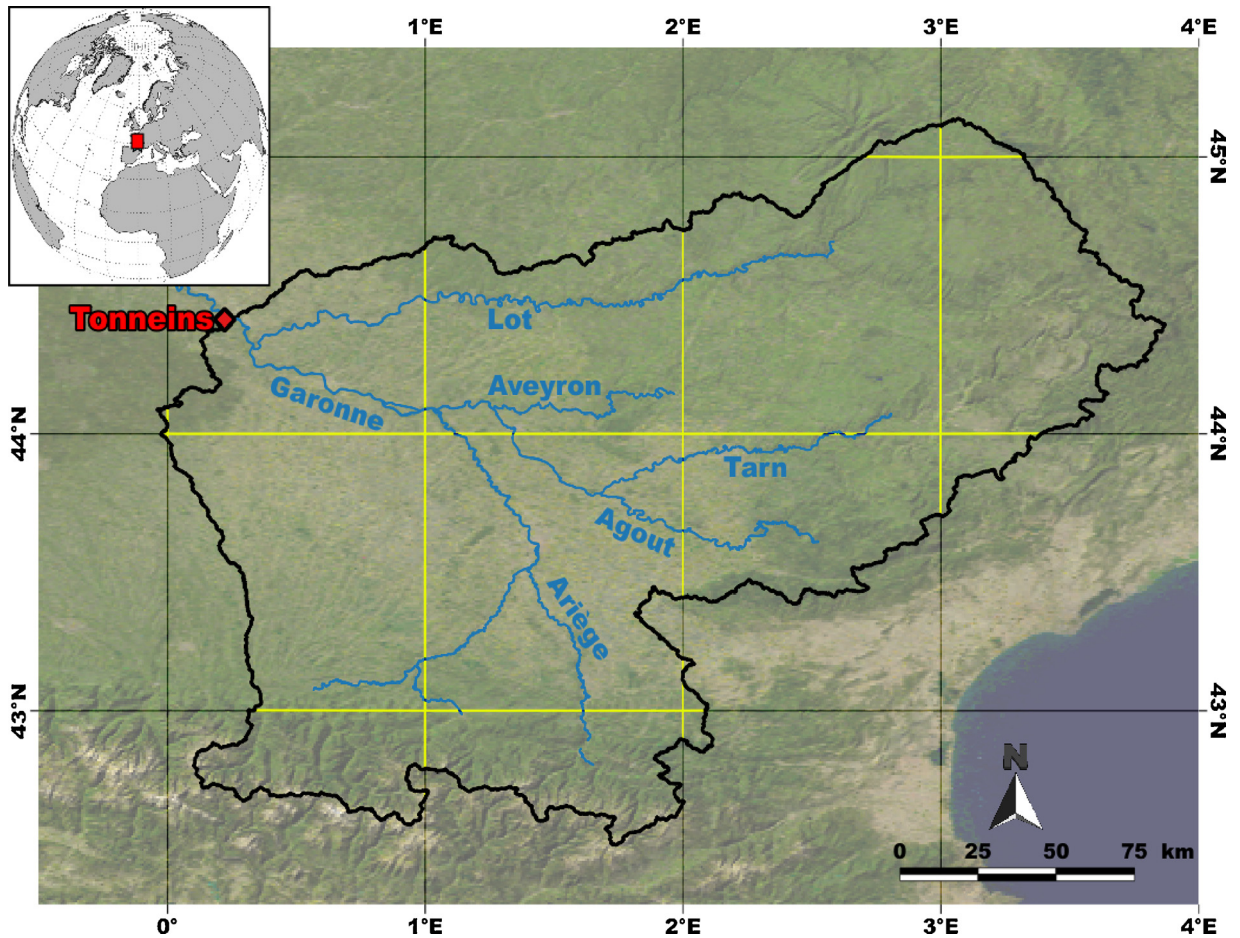


Fig. 1. Map of the Garonne basin up to Tonneins (black polygon) and river network (blue lines, from IGN BD Carthage database, available at www.sandre.eaufrance.fr). Vertical and horizontal lines correspond to a $1^\circ \times 1^\circ$ longitude/latitude grid (used by most GRACE products) and yellow lines correspond to pixels intersecting the Garonne. Background image is the NASA MODIS “Blue Marble Next Generation” image (Stöckli et al., 2005) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

Deutsches Zentrum für Luft- und Raumfahrt (DLR) satellite mission to map the global Earth gravity field every 30 days (Tapley et al., 2004). It is composed of two twin satellites following each other, 220 km apart. As the orbit of each satellite is determined by the Earth gravity field, local variations of this field will affect the satellites motion slightly differently and therefore the distance between them. That is why the distance between the two satellites is constantly monitored on-board and is then converted by different processing centers to Earth gravity field variations, called “GRACE solutions” (for more information and reviews on GRACE, see, for example, Wouters et al., 2014). If the principle of GRACE measurements is quite simple, in practice the inversion process is complex and source of multiple errors. Some GRACE solutions are computed as Stokes coefficients of spherical harmonic functions. It is widely assumed that GRACE solutions should observe physical phenomena with a spatial resolution of about 300 km, due to the satellites orbit altitude, instrument accuracy, the maximum degree of spherical harmonics used and the spatial scale of the filters applied to GRACE measurements at different processing steps (e.g. Landerer and Swenson, 2012; Wouters et al., 2014; Scanlon et al., 2016). Most (if not all) processing centers now provide their solutions rather in spherical harmonic coefficients or recomputed on regular geographic grid. Because of the different hypotheses and approaches used to inverse gravity fields from GRACE satellites distance measurements, currently available solutions are not identical. To get sufficient spatial coverage by the satellite, measurements are often integrated over a month to do the inversion (Wouters et al., 2014), even if sub-monthly solutions have recently been made available (e.g. Kurtenbach et al., 2012; Ramillien et al., 2015).

GRACE satellites are on a quasi-polar orbit (with an 89° inclination) at an altitude around 500 km, allowing a global coverage of continental surfaces. The satellites were launched simultaneously March 17, 2002. The total mission lifetime largely exceeded the initially planned 5 years lifetime, as the mission ended science operations in October 2017 (<https://www.jpl.nasa.gov/news/news.php?feature=6984>), due to a battery failure on one satellite. Yet, since 2011, issue with satellites batteries has led the mission team to collect data only when the sun is positioned favorably to the solar arrays (Wouters et al., 2014). Therefore, gaps in GRACE time series are quite important since then. There had also been some few missing measurements before 2011, especially in the early times of the mission.

2.2.2. GRACE data used in hydrology and uncertainties

At the monthly to interannual time scales observed by the mission, variations of earth gravity field can be attributed mainly to redistribution of water in its fluid envelope (e.g. Tapley et al., 2004; Wouters et al., 2014). Therefore, GRACE solutions are also provided as Total Water Storage (TWS) anomalies (TWSA) at monthly time scale. It corresponds to the sum of all water mass variations at the continents surface and in the soil (i.e. the sum of snow water equivalent, surface water, soil water and groundwater; Chen et al., 2016).

GRACE solutions have been widely used to compute at basin and subbasin scale TWSA for big river watersheds like the Amazon, Congo or Mekong basins to cite a few (see, for example, Wouters et al., 2014, for a review). These observations have had a wide range of applications in continental hydrology since the last decade: to validate/improve hydrological models (e.g. Niu and Yang, 2006), correct model's outputs with data assimilation techniques (e.g. Zaitchik et al., 2008), compute basin-scale water budget in combination with other products (e.g. Gao et al., 2010), estimate groundwater storage change using independent computation of soil and snow water storage variations (see Chen et al., 2016, for a review), or investigate trends in TWS and their relationships with climate variability and human activities (Rodell et al., 2018).

GRACE data near its spatial resolution limit (~300 km) has important signal to noise ratio, as the noise increase with degrees of the spherical harmonic decomposition (toward higher spatial resolution). Processing centers reduce noise level by applying smoothing filters, which tend to reduce and smooth the amplitude of the water mass variations. Besides, because of the processing treatments and the truncation of the spherical harmonic degree, GRACE TWS suffers of "leakage effects", i.e. water mass tends to be spatially spread and can "leak" toward nearby other water mass (Longuevergne et al., 2010; Wouters et al., 2014). Because of the orbit orientation, GRACE measurements are more sensitive to variation of mass in the satellite along track ("North/South at low/mid-latitudes) direction than the ones in the across track (i.e. "East/West) direction, leading to "stripes" in the initial release of GRACE solutions. Some "destriping" algorithms have then been developed to significantly reduce this source of error, which in turn tend to affect the amplitude of the signal and spatial resolution (Wouters et al., 2014). Ancillary data are needed by processing centers to invert TWSA over continents, like atmosphere and ocean models outputs, Glacial Isostatic Adjustment (GIA) model(s)... which have their own source of uncertainties. Finally, Long et al., 2017 analyzed sixty river basins in the world and found more disparity in TWSA trend for medium basins (area < 200,000 km²) than for bigger ones, indicating that differences in GRACE solutions are more likely to be observed for medium and small watersheds.

That's why, it remains unsure that GRACE measurements can provide useful data for a 50,000 km² basin like the Garonne River basin.

2.2.3. GRACE solutions

Seven publicly available monthly global GRACE solutions have been considered in this study. Downloaded products correspond to liquid water equivalent thickness in centimeters on regular grid. The three first solutions are the GRACE Tellus – Land release RL05 monthly mass 1°x1° grids (available at <https://grace.jpl.nasa.gov/data/get-data/monthly-mass-grids-land/>) computed by the University of Texas – Center for Space Research (CSR, version DSTvSCS1409), GeoForschungsZentrum (GFZ, version DSTvSCS1409) and Jet Propulsion Laboratory (JPL, version DSTvSCS1411) processing centers (Swenson and Wahr, 2006; Landerer and Swenson, 2012; Swenson, 2012). The filtering used during the processing steps of these GRACE Tellus – land solutions tends to reduce retrieved TWSA amplitude. To overcome this issue, Landerer and Swenson, 2012 computed a global grid of time invariant scaling factors, using CLM4 global hydrological model outputs. These scaling factors have been applied to the three solutions in our study. Hereinafter, these solutions are labeled 'CSR', 'GFZ' and 'JPL', respectively.

The fourth solution corresponds to GRACE/LAGEOS Equivalent Water Heights 1°x1° grids from the Centre National d'Etudes Spatiales/Groupe de Recherche en Géodésie Spatiale (CNES/GRGS, release 03 – version 3 available at <http://grgs.obs-mip.fr/grace/variable-models-grace-lageos/grace-solutions-release-03>; Lemoine et al., 2016). This solution is labeled 'GRGS' in the following sections.

Finally, the three last solutions correspond to so-called global 'mascons' solutions. Contrarily to the previous solutions, mascons solutions parameterize the Earth gravity field using regional mass concentration functions, which has some advantages, like the decrease of ocean/land leakage (Scanlon et al., 2016). The mascons solutions have been computed by CSR (over a 0.5°x0.5° grid, version 1; Save et al., 2016), NASA Goddard Space Flight Center (GSFC, 1°x1° grid, version 2.2; Luthcke et al., 2013; Loomis and Luthcke, 2014) and JPL (0.5°x0.5° grid, RL06 M.MSCNv01 dataset with Coastal Resolution Improvement Filtering; Watkins et al., 2015) processing centers, called hereafter 'CSR mascon', 'GSFC mascon' and 'JPL mascon', respectively. Time invariant scaling factors are recommended and provided only for the JPL mascon solution. Documentation for this solution also clearly mentioned that "the native resolution is the size of a single mascon which are 3 degrees in size" (https://grace.jpl.nasa.gov/data/get-data/jpl_global_mascons/), despite the solution is provided on a 0.5°x0.5° grid.

The time mean value removed to the provided GRACE Tellus – Land solutions, CSR/JPL mascon solutions TWSA has been computed over January 2004 to December 2009 time span. For GRGS solution and GSFC mascon solution, the mean removed has been computed over January 2004 to December 2015, and over January 2003 to December 2012 time periods, respectively. Furthermore, GIA corrections have not been considered, as GIA is negligible on the Garonne basin (Guo et al., 2012).

Fig. 1 shows the 1°x1° grid used by most solutions. Yellow lines correspond to portions of the grid that lie within the studied Garonne basin. As the intrinsic GRACE spatial resolution is actually coarser than 1°x1°, this figure shows how small the basin is compared to GRACE resolution. This issue is highlighted in many GRACE solutions documentations. For example, the CSR mascon documentation clearly states that "the users must exercise caution when using these solutions in basins smaller than approximately 200,000 km². Moreover, these solutions should be used to perform basin level time-series analysis and never be used for analysis at a

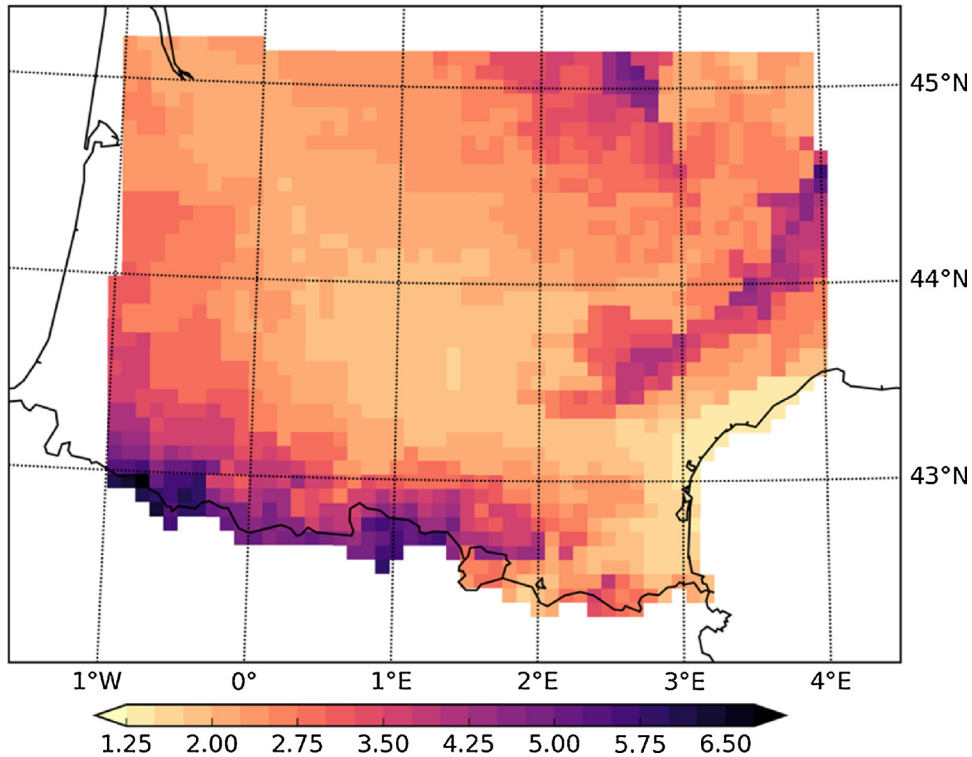


Fig. 2. 2000–2014 mean annual total precipitation rate ($\text{mm}\cdot\text{day}^{-1}$) over the Garonne basin from SAFRAN analysis.

single grid point”.

2.3. Hydrological models

This study inter-compared TWSA from seven global GRACE solutions with TWSA from two hydrological models, SAFRAN-ISBA-MODCOU (SIM) and Soil and Water Assessment Tool (SWAT), between 2003 and 2010. SIM and SWAT are the only two operational distributed and physically based hydrological models currently deployed over the Garonne watershed that can simulate the different compartments of the water cycle.

2.3.1. Climate forcing dataset

In this study, climate dataset used to force the hydrological models comes from the Système d’Analyse Fournissant des Renseignements A la Neige (SAFRAN; Durand et al., 1993; Quintana-Seguí et al., 2008) analysis system. Raimonet et al. (2017) have shown that SAFRAN currently provides the best estimate of near-surface variables over France, compared to other existing datasets. Especially, reconstruction of the liquid and solid precipitation fields benefits from the high-density network of Meteo-France local measurements. Moreover, Grusson et al. (2017a) showed that using SAFRAN dataset in a hydrological modeling of the Garonne basin led to better performances. SAFRAN performs an analysis of near-surface variables by optimal interpolation (OI) on homogeneous climatic areas, i.e. without strong horizontal climatic gradient. The OI method combines upper-air atmospheric analysis and in situ observations to calculate the best estimate of total daily precipitation, and 6-hly 2-meter air temperature and humidity, 10-meter wind speed, downward solar and infrared radiation. These analyzed variables are then interpolated in space, on an 8km-mesh regular grid to better take into account topography, and in time, at an hourly time step. More detailed description of SAFRAN analyses can be found in Durand et al. (1993) and Quintana-Seguí et al. (2008). SAFRAN analyses are computed from 1958 onwards (Vidal et al., 2010). Fig. 2 shows SAFRAN 2000–2014 mean annual total precipitation in South West of France. The two hydrological models used in this study, which were forced with SAFRAN datasets, are described in the following sections.

2.3.2. SIM model

SAFRAN-ISBA-MODCOU (SIM) is a combination of a meteorological analysis system (SAFRAN), a land surface model (Interaction-Soil-Biosphere-Atmosphere, ISBA; Noilhan and Planton, 1989; Noilhan and Mahfouf, 1996) and a hydrogeological model (MODèle COUplé, MODCOU; Ledoux et al., 1989; David et al., 2011). It is used operationally at Meteo-France to monitor water resources over the French territory and can therefore be used to forecast flood risk and to monitor drought risks. SIM simulates energy and water budgets, river discharge and height of the water table in presence of aquifers. Over natural areas, ISBA land surface model computes water storage, based on input SAFRAN total precipitation and modeled total evaporation, surface runoff and soil infiltration. Total

evaporation is composed of plant evapotranspiration, evaporation of bare soil, evaporation of water intercepted by the vegetation, and snow and ice sublimation. The partitioning of local precipitation into runoff and infiltration is parameterized following Dümenil and Todini (1992). Runoff is routed by the MODCOU surface routing network towards the river network (Häfliger et al., 2015). Soil infiltration is diffused through the soil layers up to the vegetation root depth and contributes to the soil water storage and can contribute to aquifers water supply when they are simulated, which is not the case in the Garonne basin. In the current version of the ISBA model, the soil is discretized into 14 layers, resulting in a total depth of 12 m. The vertical discretization (bottom depth of each layer in meters) is the following: 0.01, 0.04, 0.1, 0.2, 0.4, 0.6, 0.8, 1, 1.5, 2, 3, 5, 8, 12, as described in Decharme et al. (2013). The heat transfer is solved over the total depth, whereas the moisture transfer is solved over the root depth only, which depends on the vegetation types.

Concerning the input data, apart from the SAFRAN climate dataset, SIM needs parameters that define the Garonne basin in terms of natural ecosystems, orography and soil texture. For that purpose, the ECOCLIMAP database (Faroux et al., 2013) is used to represent land cover at a 1-km horizontal resolution. Among the hundreds of ecosystems available in ECOCLIMAP, vegetation in the model is represented by only 12 plant functional types (PFTs) combining several ecosystems. The orography is derived from the Shuttle Radar Topography Mission (SRTM) at 90-meters (Farr et al., 2007), upscaled at 1-km resolution to be consistent with the ecosystems representation. Soil texture at 1-km resolution comes from the Harmonized World Soil Database (Nachtergaele et al., 2012). It is used to define model parameters that depend on clay or sand concentration, like the field capacity or wilting point soil moisture contents, which are commonly used to define thresholds between which plants evapotranspire at potential rate or are stressed due to a lack of water.

The evapotranspiration is calculated based on the ISBA-A-gs module (Calvet et al., 1998) where a simplified photosynthesis module is applied to represent the exchanges of water and carbon dioxide at the leaf scale.

A realistic representation of the snowpack behavior is required to accurately represent snow accumulation and melting, which are key components of the water cycle, especially in mountainous areas. For that purpose, snow is parameterized using Boone and Etchevers (2001) explicit multi-layer snow model. This model has demonstrated its ability to simulate accurately the main snow processes, such as the freezing of water that would have percolated within the snow, due to surface melting or liquid precipitation.

Surface runoff parameterization requires the calibration of one single parameter, b , representing the slope of the retention curve: the higher b , the faster water runs off. This parameter was setup uniformly for the Garonne basin in the SIM operational set up.

Although no specific calibration was made in the current study, whether for the land surface model (ISBA) or the hydrogeological model (MODCOU), the parameters of this latter that were calibrated over the 1974–1982 period (Boukerma, 1987) were not modified.

Table 1 sums up SIM input parameters and forcing sources, and validation datasets used for this study.

2.3.3. SWAT model

Soil and Water Assessment Tool (SWAT) is a hydro-agro-environmental model for river and watershed studies jointly developed by USDA Agricultural Research Service (USDA-ARS) and Texas A&M AgriLife Research from Texas A&M University. SWAT is a public domain model, which simulate quantity and quality of surface and ground water and predict environmental impact of human practices, like agriculture, point source, dams... It has been widely used to simulate distributed hydrology variables at a watershed scale (Gassman et al., 2007). The model is based on Hydrologic Response Units (HRU aggregated at subbasin level). A HRU is characterized by a topography slope, and soil and land use combination within a subbasin. The SWAT model integrates soil types and properties, land cover and agriculture management including automatic irrigation and fertilization, where water and fertilizer are applied to avoid plant stress. HRUs are the base unit to compute the water balance, which is composed of four compartments: snow, soil, shallow aquifer, and deep aquifer. Fluxes between these compartments correspond to the following hydrological processes: infiltration, runoff, evapotranspiration, lateral flow and percolation. Computation is performed at the HRU level and routed within the river network to downstream subbasins down to the outlet. The project has been set up with ArcSWAT version 10.4.19, a GIS-based graphical interface based on ArcMap 10.4, helping users to define HRUs and generate the associated input files for SWAT2012 version (Olivera et al., 2006). SWAT documentation available online explained the theory and details of hydrologic processes integrated in SWAT model (Neitsch et al., 2009).

Concerning the input data, all SWAT input files are listed in Table 1. The Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model (ASTER GDEM; Tachikawa et al., 2011) from NASA and the Ministry of Economy, Trade and Industry (METI) of Japan, available on 90 m global grid, was chosen to delineate the watershed and compute the river system.

Table 1
SWAT and SIM input parameters and forcing sources and validation datasets.

Data Type	SWAT		SIM	
	Data Source	Scale	Data Source	Scale
Atmospheric forcing	SAFRAN	8 km x 8 km	SAFRAN	8 km x 8 km
DEM	ASTER GDEM version 2 from NASA/METI	Grid cell 90 m x 90 m	SRTM GDEM from NASA and NGA	Grid cell 90 m x 90 m
Land Use	Corine Land Cover	Grid cell 100 m x 100 m	ECOCLIMAP (2013)	Grid cell 1 km x 1 km
Soil	European Soil Database	Grid cell 1 km x 1 km	Harmonized World Soil Database	Grid cell 1 km x 1 km
River discharge	Banque Hydro	20 gauges	Banque Hydro	34 gauges

Land uses are from the CORINE Land Cover 2012 (Büttner et al., 2014) map with a 100 m resolution and the soil data originate from the European Soil Database (EEA, 2007) with a 1 km resolution map. Soil properties have been adjusted by expertise. The SWAT model setup includes 52 different types of soil over the watershed. Each soil consists in 1 to 4 layers according each type of soil, each layer having a depth of 10 to 1400 cm and its own layer characteristics.

From these land use, soil databases and watershed topography, 22 land use, 13 soil and 5 slope classes were defined in SWAT. With these classes, SWAT model identified 1320 subbasins with 12 834 HRUs. For the HRUs definition, we selected only classes which occupy at least 10% on the subbasin to avoid anecdotal classes.

The SAFRAN atmospheric fields (see section 2.3.1) have been used to force SWAT.

River discharge observations used for calibration and validation come from Banque Hydro database. Monthly stream flow data from 20 selected gauging stations, including the outlet Tonneins, along the river network were used to calibrate and validate the model. The selection was made to sample all the hydrology and catchment diversity in soil, land use and climate conditions as proposed by Grusson et al. (2017a). For each step (validation and calibration), three years of initialization are performed. The SWAT model was calibrated over ten years from 2000 to 2010 and validated on the available previous period from 1980 to 1997. Calibration procedure are based on previous set-up of the model over the same watershed (Grusson et al., 2017a, b; Grusson et al., 2015). Same sensitive parameters have been considered and the calibrated values from those previous studies have been used as first guess in our calibration procedure. As the spatial definition (subwatershed and HRU) is slightly different in the present version of the model, calibration has been manually adjusted. This manual adjustment was based on watershed characteristics, bibliography and expertise.

2.3.4. Models evaluation

The evaluation of SIM and SWAT models performances has been conducted using the Nash-Sutcliffe Efficiency (NSE; Nash and Sutcliffe, 1970), by comparing simulation and observation at monthly and daily time step for SWAT and SIM, respectively. NSE range between minus infinite and 1. A NSE of 1 indicates a perfect fit between observation and simulation. A value of 0 indicates that the simulation produced by the model is equivalent to the average of the observed data. A negative value indicates very poor agreement between simulation and observation. For each model an estimation of the bias is also conducted to estimate if discharges are overestimated or underestimated by models.

Concerning SIM, it was run for the period 1958–2016 over France and was spun up by repeating twice the first year. SIM evaluation against daily measured discharges was performed over the Garonne basin for 34 gauging stations available in the Banque Hydro database and located along the Garonne River and its main tributaries. In addition to the NSE, and in order to investigate over/underestimation, percent bias (Pbias) has been computed (one minus the ration between modeled discharge and in situ discharge). It measures the average tendency of the simulated data to be larger or smaller than their observed counterparts. The optimal value of Pbias is 0, with low-magnitude values indicating accurate model simulation. Positive values indicate model underestimation bias, and negative values indicate model overestimation bias (Gupta et al., 1999). On average over the entire basin, SIM gives satisfactory results with an averaged NSE of 0.55 and a Pbias of -0.04. Fig. 3.a shows the percentage of SIM pixels containing in situ stations as a function of NSE. 75% of pixels with stations have an NSE greater than 0.5, which can be considered as a rather good result. Moreover, more than 50% of pixels have a NSE greater than 0.6, which shows the good agreement between SIM and in situ measurements. Pbias (Fig. 3.b) indicates that, for one third of the stations, SIM tend to overestimate discharge, whereas for one fourth of the stations SIM tends to underestimate discharge. On average, SIM is unbiased.

For the SWAT model, the Nash and Pbias are calculated for monthly discharge, at 20 different gauging stations spread over the watershed (see Fig. 4). The stations have been selected to encompass the diversity of hydrological regime existing over the watershed.

A soft calibration was led to calibrate the SWAT model based on previous studies and the expertise of specialists. Grusson et al. (2017a) and Grusson et al. (2017b) performed parameters sensitivity analysis over 20 gauging stations within the Garonne watershed. The uncertainty analysis was done using the SWAT Calibration and Uncertainty Procedure (SWAT-CUP) tool, which is based

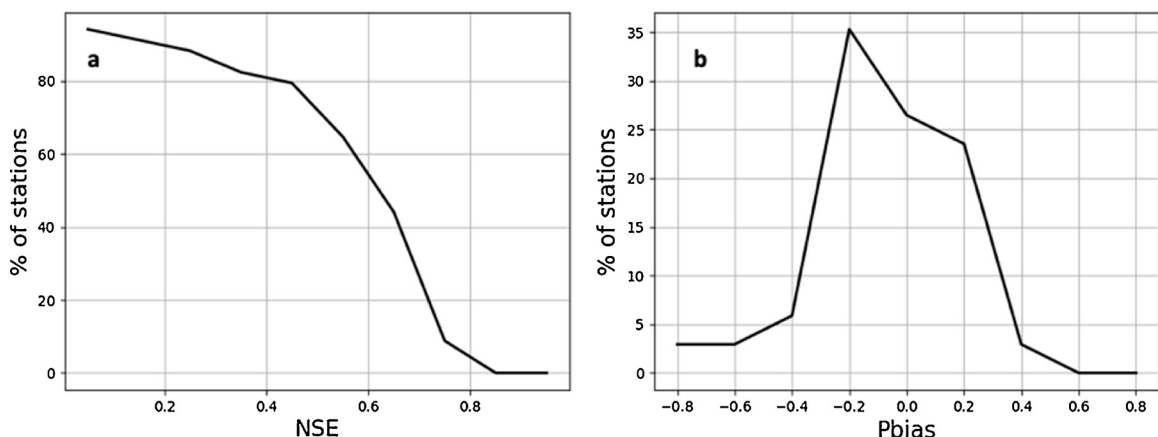


Fig. 3. Percentage of SIM pixels co-located with validation stations versus NSE (a) and Pbias (b), for the 34 gauging stations over the 1958–2016 period.

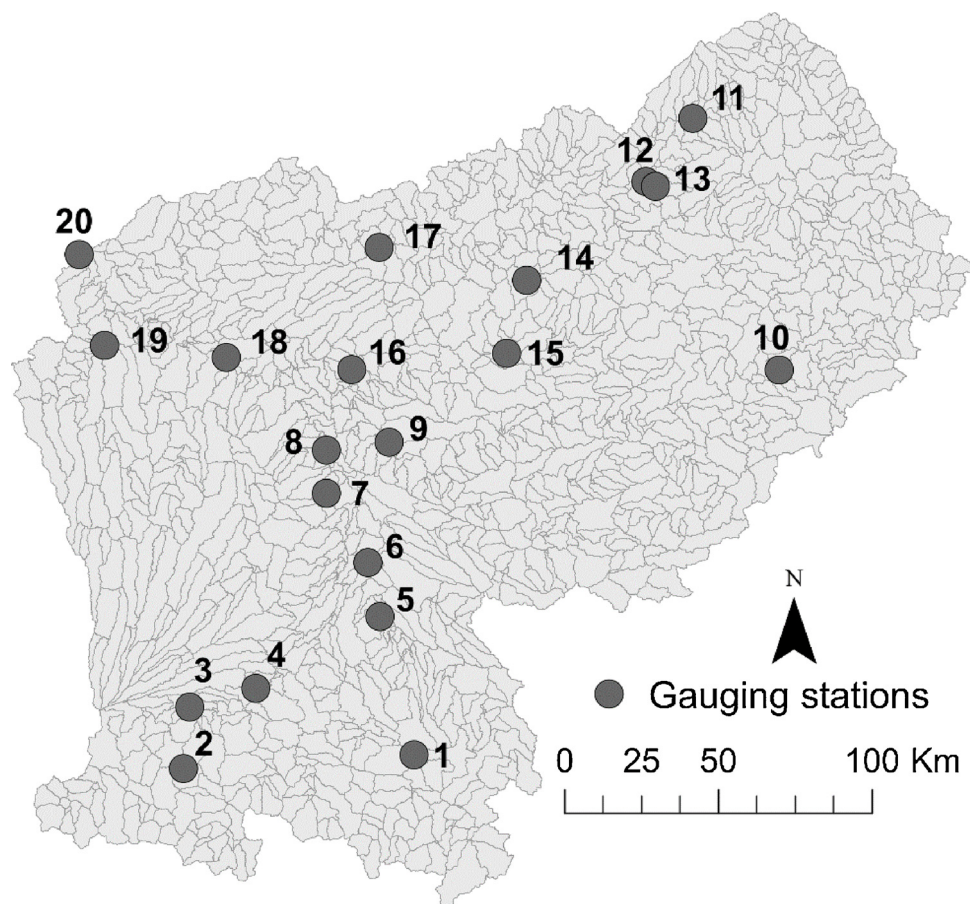


Fig. 4. Localization of the 20 gauging stations used to validate and calibrate the SWAT model.

on the Sequential Uncertainty Fitting version 2 (SUFI-2) algorithm. All sources of uncertainties (driving variables, conceptual model, parameters, measured data) were considered. Uncertainties in the model output variables are driven by parameters uncertainties, which can be expressed as the 95% probability distribution, also named 95PPU (95% prediction uncertainty). To get the best performance, the 95PPU envelop has to capture most of observation values and have a small spread. In order to evaluate 95PPU performance, two statistics are used: P-factor and R-factor (Abbaspour et al., 2004; Abbaspour, 2012). P-factor is the percentage of observed data enveloped by our modeling and need to be over 70% to consider a good discharge modeling. R-factor represents the thickness of the 95PPU envelop and need to be around 1. Fig. 5 illustrates the 95PPU intervals of simulation and observation during calibration and validation period at the outlet of the basin. In this case, 73% of observed monthly values were within 95PPU area and R-factor is equal 0.66.

An average NSE value of 0.76 was obtained for the calibration period (2000–2010). Table 2 presents NSE and Pbias for the 20 gauges. The rank of Nash and Pbias evaluation is given according the ranking proposed by Moriasi et al. (2007) and based on monthly discharge evaluation. For the calibration period, 14 out of the 20 stations show good to very good performances (Table 2). Only four stations show unsatisfactory performances with an overall average NSE of 0.70. The positive Pbias indicates that, like the SIM model, SWAT model tends to globally overestimate the volume of water. Performances remain steady over the validation period with a good to very good model performance at 14 out of 20 stations (Table 2). Still 4 stations are presenting NSE below 0.5. These four stations are not necessarily the same than the four gauges with the lowest NSE during the calibration period. Two stations presenting unsatisfactory results during the calibration are performing well during the validation period. The overall tendency of the model is also to overestimate the volume of water compared to observations.

2.4. GRACE and hydrological models cross-validation methodology

As presented in previous sections, on one side, GRACE data are quite coarse and its capability to observe TWSA over the Garonne basin is questionable. On the other side, SIM and SWAT hydrological modeling of the basin are the two only operational distributed and physically based hydrological models currently deployed over the Garonne watershed, and some compartments of their mass budget cannot be validated based on in situ data alone. Therefore, and as stated in the introduction, the purpose of this study is to compare GRACE solutions and SIM/SWAT models outputs, in order to cross-validate them and to compute a hydrology water balance

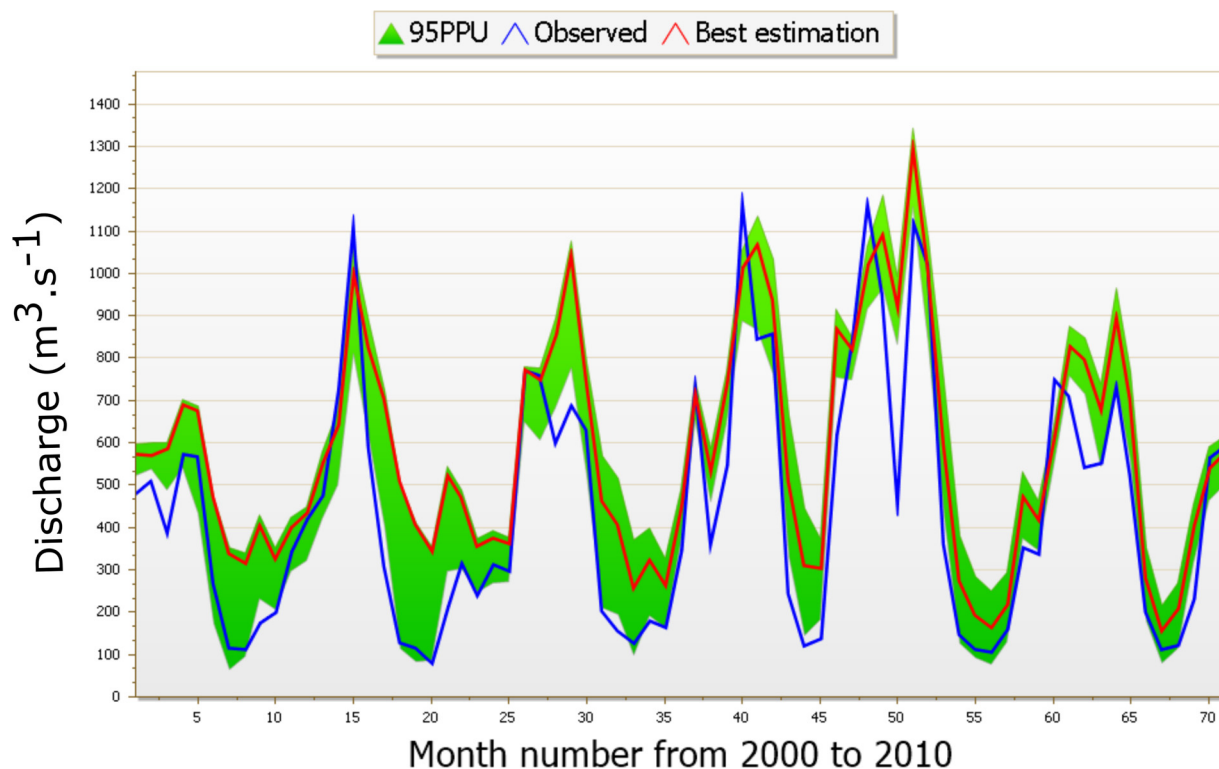


Fig. 5. Monthly calibration and validation simulation at Garonne outlet showing the 95% prediction uncertainty intervals determined with recording discharge.

Table 2

SWAT model performance metrics for both validation and calibration periods. Bold numbers correspond to good and very good performances (i.e. NSE > 0.60, as recommended by Moriasi et al., 2007). See Fig. 4 for the station location.

Name	Calibration		Validation		
	NSE	Pbias (%)	NSE	Pbias (%)	
1	Foix	0.84	12.1	0.65	-3.8
2	Saint-Béat	0.75	17.2	0.25	26.5
3	Valentine	0.45	33.1	0.22	29.6
4	Roquefort	0.73	0.3	0.76	-9.4
5	Auterive	0.60	15.2	0.83	0.7
6	Portet	0.84	5.4	0.79	17.5
7	Larra	0.25	45.4	0.02	97.0
8	Verdun	0.82	16.9	0.77	22.7
9	Villmure	0.88	0.1	0.46	26.9
10	Millau	0.83	-2.5	0.76	14.0
11	Sarrans	0.57	24.8	0.77	13.6
12	Truyère-Aval	0.83	15.1	0.70	5.2
13	Truyère-Amont	0.39	38.9	0.40	48.6
14	Villefranche	0.71	16.5	0.81	28.4
15	Viaur	0.78	0.2	0.57	58.9
16	Loubéjac	0.45	21.5	0.72	43.6
17	Cahors	0.83	0.9	0.73	25.9
18	Lamigistère	0.87	10.5	0.75	27.0
19	Nérac	0.78	0.2	0.82	27.2
20	Tonneins	0.87	13.6	0.66	33.5

at the Garonne basin scale.

The following sections present the methodology used to compare and cross-validate GRACE TWSA with SIM and SWAT TWSA (section 2.4.1), and to compute Garonne basin water balance (section 2.4.2).

2.4.1. GRACE and hydrological models outputs comparisons

TWSA from GRACE are computed by selecting all pixels that lie, even partially, in the Garonne basin. Then, for each of these pixels, GRACE solutions liquid water equivalent thicknesses are multiplied by the area of the pixel portion which intersects the basin polygon to compute pixel TWSA (in m^3). Then, all pixels TWSA are summed, to get GRACE TWSA over the whole basin. For each GRACE solution time series, its time mean over the time period January 2003 to December 2010 is removed. This removal is needed, as the reference of TWSA is not common for all solutions (see section 2.2.3) and this time period has very few missing GRACE measurements (as explained in section 2.2.1).

Modeled TWSA are computed in three steps. First, for each pixel, liquid water-equivalent volumes stored in all modeled hydrology compartments are summed. For SIM, these compartments correspond to snow, water intercepted by the vegetation, river network and all soil layers. For SWAT, they correspond to snow, river network, all soil layers, shallow aquifer connected to the river and the deep confined aquifer. Then TWS of all pixels within the basin are summed (weighted by the proportion of pixels area within the basin polygon). Finally, the mean of this sum between January 2003 and December 2010 (for consistency with Garonne GRACE TWSA time series) is removed to obtain TWSA for the whole basin.

Quantitative comparison between monthly GRACE solutions and monthly hydrological models outputs has been performed only from January 2003 to December 2010 time span, as before and after GRACE time series have important gaps. For all datasets, the mean annual cycle has also been computed: for each month (January to December), the mean TWSA is computed for all corresponding month between 2003 and 2010. Time series without the mean annual cycle have also been calculated: for each month, the corresponding mean month value has been removed. The comparison is done by computing the correlation coefficient and the NSE between GRACE solutions and models full time series, their mean annual cycle and time series without their mean annual cycle.

TWSA time series and hydrology budget have been interpreted qualitatively for all datasets over their whole common time period, i.e. August 2002 to July 2014 (corresponding to twelve hydrological years). Hydrological model outputs have also been analyzed to identify the main hydrology compartment(s) contributing to TWSA observed by GRACE.

2.4.2. Garonne basin water balance

After the intercomparison of all datasets, the basin hydrology water balance has been estimated. It corresponds to a water mass balance or continuity equation for a closed system, translating the fact that variations of TWS during a certain period of time is equal to the quantity of water received and lost during that period of time. For a watershed, water is gained via precipitation (P) and lost via evapotranspiration (sum of the soil evaporation, evaporation of the precipitation intercepted by the canopy and the vegetation transpiration; this sum is noted ETR) and water discharge at the outlet of the basin (Q, often approximated to river outflows, as aquifer outflows is neglected and not known). Therefore, the generic water balance equation for a watershed is:

$$\frac{dTWS}{dt}(t) = P(t) - ETR(t) - Q(t) \quad (1)$$

For a hydrological model, which provides only discrete estimations of variables in time and space, Eq. (1) can be integrated during the month m (whose first and last time steps are noted t_1 and t_2 , respectively and the model numerical time step is noted Δt) over the N pixels or HRU within the watershed (Eq. 2).

$$\Delta TWS(m) = TWS(t_2) - TWS(t_1) = \sum_{s=1}^N \sum_{t=t_1}^{t_2} [P(s, t) - ETR(s, t) - Q(s, t)] \cdot \Delta t \quad (2)$$

Some studies (see Wouters et al., 2014 for a review) have used this equation to compare TWS change (ΔTWS), estimated from GRACE and P-ETR-Q from models or from other satellite or in situ observations, to rather assess the quality of GRACE measurements or to get bounds from GRACE in estimated P-ETR-Q. However, as GRACE provides mean monthly TWSA and not TWSA between the first and last day of the month, it is not possible to compute directly ΔTWS from GRACE, even if its measurements were perfects. GRACE ΔTWS is often approximated using a TWSA difference between two consecutive months. In this study, we used a second order central difference (Eq. 3) to compute it, as it is more numerically stable than forward or backward difference. This approximated ΔTWS is noted $\Delta T\bar{W}S$ hereafter, to clearly show that it is based on monthly average of TWS.

$$\Delta T\bar{W}S(m) = \frac{TWSA(m+1) - TWSA(m-1)}{2} \quad (3)$$

The purpose of this work is to compare $\Delta T\bar{W}S$ from GRACE solutions to the right-hand side of Eq. 2 computed from SIM/SWAT models inputs/outputs. It is done to evaluate the usefulness of GRACE solutions to provide information for computing the water balance of a basin like the Garonne and to assess how P, ETR and Q affect TWS variations for the Garonne.

3. Results and discussions

3.1. GRACE solutions and models outputs cross-validation

GRACE solutions and hydrological models TWSA time series (in m^3) over the Garonne basin, their mean annual cycle (2003–2010 average) and TWSA without this mean annual cycle time series are presented on Fig. 6. They show very good agreements between GRACE solutions and hydrological models (Fig. 6a). Especially, their mean annual cycles (Fig. 6b) have similar amplitude and phasing. However, it should be noted that GRACE solutions mean annual cycles are 1-month delayed compared to hydrological

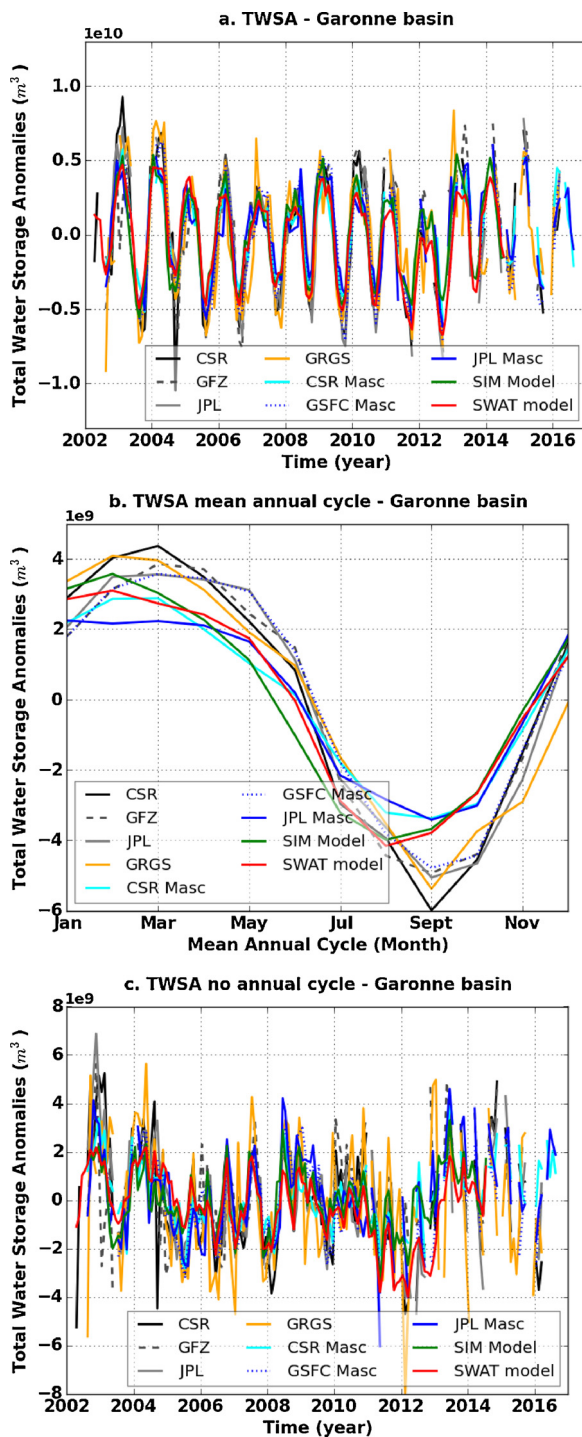


Fig. 6. TWSA (in m^3) over the Garonne basin (a.), 2003–2010 TWSA mean annual cycle (b.) and TWSA without their 2003–2010 mean annual cycle (c.) for the seven GRACE solutions (Tellus Land CSR, Tellus Land GFZ, Tellus Land JPL, CNES/GRGS, CSR mascon, GSFC mascon, JPL mascon) between April 2002 and August 2016 and the two hydrological models (SIM and SWAT) between April 2002 and August 2014.

models. GRACE time series have their maximum values in March (except for GRGS solution, which is in February) and minimum values in September, whereas for both SIM and SWAT the minimum and maximum occur in February and August, respectively. Time series without their mean annual cycle (Fig. 6c) are also quite coherent between satellite and models time series, even if GRACE solutions (especially GRGS solution) have more variations at highest frequencies than hydrological models. GRGS solution does not apply some post-processing smoothing than other solutions use, which could explain its more important variability.

Table 3

Correlation (Corr.) coefficient and Nash-Sutcliff Efficiency (NSE) between GRACE solutions and hydrological models TWSA over the Garonne basin. Bold figures correspond to highest correlation coefficient and NSE between GRACE solutions and SIM or SWAT models.

		Models			
		SIM		SWAT	
		Corr.	NSE	Corr.	NSE
GRACE Tellus Land	CSR	0.91	0.60	0.92	0.62
	GFZ	0.86	0.52	0.84	0.46
	JPL	0.86	0.57	0.91	0.68
	Mean	0.90	0.67	0.92	0.71
GRACE GRGS		0.77	0.27	0.79	0.32
GRACE Global Mascons	CSR	0.93	0.86	0.92	0.83
	GSFC	0.90	0.71	0.93	0.77
	JPL	0.89	0.78	0.88	0.74
Models	SIM	1.00	1.00	0.96	0.93
	SWAT	0.96	0.92	1.00	1.00

Tables 3 and 4 present correlation coefficient and NSE computed over 2003/2010 time period between GRACE solutions and models time series, with and without their mean annual cycle, respectively. As recommended in the GRACE Tellus – Land documentation, statistics for the mean of the three solutions are also provided. Table 3 shows the overall good agreement between GRACE solutions and model outputs. GRGS solution has the lowest correlation and NSE (0.77 and 0.27 with SIM, respectively, and 0.79 and 0.32 with SWAT, respectively). Other solutions have correlation and NSE above 0.80 and 0.50, respectively, with both SIM and SWAT. Especially, the best comparison is obtained for mascon solutions. Correlation and NSE between CSR mascon and SIM/SWAT outputs are above 0.90 and 0.80, respectively. GSFC mascon is the second solution that best match model outputs, with slightly lower coefficients. Concerning time series without mean annual cycle, CSR and GSFC mascon also provides the best correlation and NSE, which are between 0.78 and 0.80 for correlation, and 0.33 and 0.54 for NSE. The JPL mascon solution also compares well to model outputs, but with lower coefficients than the two other mascon solutions, especially for time series without mean annual cycle. Overall, GSFC mascon seems to provide generally slightly better results than CSR mascon, but the best correlation coefficient is obtained between CSR mascon and SIM time series. It was expected to get better results with mascon solutions, as they decrease leakage errors and are more adapted locally than solution based on spherical harmonics (Scanlon et al., 2016). It is consistent with the results obtained by Long et al. (2017) over sixty basins. Worse results were also expected with time series without mean annual cycle (Table 4), as noise is proportionally more important in these time series than with the ones with the annual cycle.

These comparisons cross-validate GRACE solutions and model outputs. All datasets are quite coherent between themselves, even without their mean annual cycle, indicating similar interannual variability. The mean annual cycles are very close, with similar amplitudes, but with 1-month delay between GRACE and hydrological models. It shows that GRACE is actually capable to provide TWSA for a 50,000 km² temperate basin, like the Garonne basin. Especially, CSR mascon and GSFC mascon solutions are the most suitable for this basin. Similarly, these good results seem to indicate that hydrological models have enough physics and good enough forcing data to match independent satellite measurements of TWSA.

Quite logically, for the Garonne basin, the mean annual cycle (Fig. 6.b) is positive during winter and spring (maxima in February/March), when the precipitation is the highest, and negative in summer and autumn (minima in August/September), when the

Table 4

Correlation (Corr.) coefficient and Nash-Sutcliff Efficiency (NSE) between GRACE solutions and hydrological models TWSA without their respective 2003–2010 mean annual cycle over the Garonne basin. Bold figures correspond to highest correlation coefficient and NSE between GRACE solutions and SIM or SWAT models.

		Models			
		SIM		SWAT	
		Corr.	NSE	Corr.	NSE
GRACE Tellus Land	CSR	0.73	−0.40	0.73	−0.15
	GFZ	0.56	−0.57	0.41	−0.85
	JPL	0.63	−0.51	0.74	0.07
	Mean	0.72	0.06	0.71	0.15
GRACE GRGS		0.49	−1.91	0.48	−1.27
GRACE Global Mascons	CSR	0.78	0.54	0.73	0.44
	GSFC	0.80	0.33	0.79	0.44
	JPL	0.77	0.14	0.69	0.04
Models	SIM	1.00	1.00	0.82	0.65
	SWAT	0.82	0.60	1.00	1.00

precipitation is at its lowest rate (Martin et al., 2016). Fig. 6.c shows that TWSA on the Garonne basin has important interannual variability, with high positive TWSA hydrological years (2003, 2004 and 2013) and low negative TWSA years (2005, 2006, 2008, 2011 and 2012). It should also be noted that the most important depletion of TWS during the time period occurred in 2012, followed by an important refill of TWS during hydrological year 2013.

These results complement previously published studies. If GRACE spatial resolution is widely recognized to be around 300 km, the minimum size of the river basin to which GRACE data can be applied has been investigated by few studies. Rodell et al. (2007) investigated the potential to estimate Mississippi groundwater mass variations with GRACE data, and soil moisture and snow water equivalent outputs from Global Land Data Assimilation System (Rodell et al., 2004) models. They found that they could not retrieve the groundwater seasonal cycle correctly for basins smaller than 900,000 km². However, as highlighted by Wouters et al. (2014), this study was done with the first release of GRACE products and Rodell et al. (2007) considered that the main source of error was due to the GRACE product itself. Since then, GRACE data processing has been quite improved and new techniques have been developed to study smaller basins. For example, Longuevergne et al. (2010) have developed a methodology to derive GRACE TWS for some basins with drainage area close to mission resolution limit. This method has shown improvements on TWS estimation for a 200,000 km² aquifer in the US High Plains Aquifer, compared to GRACE global solutions. Lorenz et al. (2014) estimated that smaller catchments could be observed, as long as the TWS annual cycle has important amplitude variations. Vishwakarma et al. (2018) investigated the issue of GRACE spatial resolution for hydrology studies. They demonstrated that if, theoretically, the spatial resolution is related to the band-limit of the spherical harmonic spectrum of the GRACE datasets, in practice, because of the noise and corrections applied to GRACE data, the complete band-limited signal cannot be retrieved and the spatial resolution of GRACE cannot be clearly set. They estimated that for a catchment size around 63,000 km², observation error should be at a level of 2 cm in equivalent water height (so accuracy around 1.26 km³ of equivalent water volume). Our results are coherent with Lorenz et al. (2014) and Vishwakarma et al. (2018) findings: TWSA important seasonal and interannual variability is the reason why GRACE solutions compare well to SIM and SWAT model outputs over the 50,000 km² Garonne basin.

3.2. Analysis of modeled hydrology compartments

Contrarily to GRACE, hydrological models provide estimates of water stored in all modeled hydrology compartments. Fig. 7 shows contribution of water stored in all modeled compartments to the 2003/2010 TWSA mean annual cycle for SIM (Fig. 7.a) and SWAT (Fig. 7.b). Most TWSA is due to soil water (among which, the shallow aquifer connected to the river stores an important quantity of water according to SWAT). SWAT Deep aquifer contribution to TWSA is one order of magnitude lower than other compartments, despite the fact that its mean absolute water storage is the most important one. Rivers contribution to TWSA is quite small compared to other compartments, as floodplain storage is quite limited and does not last more than a few days. SIM and SWAT do not model lakes and reservoirs in this study. However, lakes and reservoirs contribution to TWSA might not be that important, as there are no big reservoirs or lakes within the basin. Snow has more impact on TWSA during winter. Logically, water intercepted by vegetation is negligible. So, in summary, most of the Garonne TWSA is due to the first dozen meters of soil and to the shallow aquifer connected to the river.

The small contribution of surface water (mainly from snow) to TWSA for the Garonne basin must be due to the size of the basin, which cannot store important mass of water in floodplains or in the river network at monthly time scale, unlike big river basins. Some previous studies using GRACE data on big river basins indicated that surface water has more impact on TWSA. For example, Becker et al. (2018) estimated that water stored in surface water bodies (floodplains, lakes, rivers and wetlands) contributes to $19 \pm 5\%$ of the annual variations of GRACE TWSA in the Congo River basin during the time period 2003/2007. As summarize by Frappart and Ramilien (2018), surface water bodies could even be responsible of 40% to 50% of the TWSA annual variations for basin with important floodplains, like the Amazon (Papa et al., 2013), Orinoco (Frappart et al., 2015) or Ganges-Brahmaputra (Papa et al., 2015) basins. These results are actually coherent with the results obtained for the Garonne basin. Indeed, the main compartments contributing to Garonne TWSA are the first dozen meters of soil and the shallow aquifer connected to the river, so compartments closely

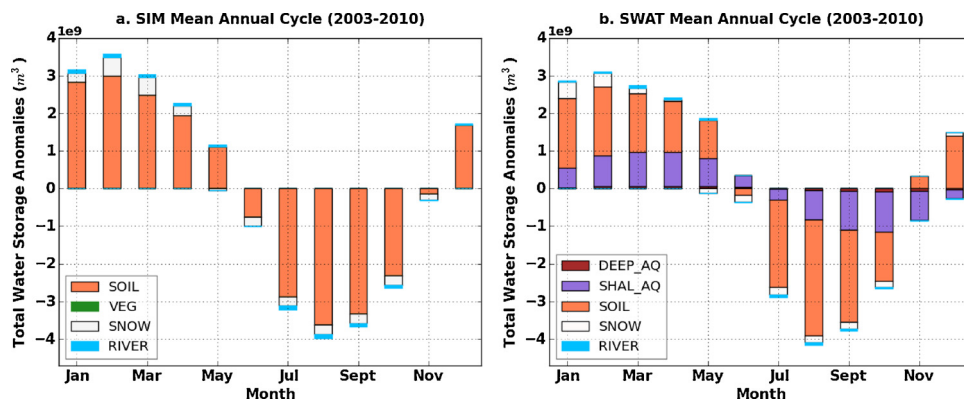


Fig. 7. Contribution of each modeled compartments to the 2003–2010 mean annual cycle TWSA for SIM (a.) and SWAT (b.) models.

Table 5

Correlation (Corr.) and Nash-Sutcliffe (NS) coefficients between $\Delta T\bar{W}S$ computed from CSR or GSFC GRACE mascon solution and precipitation (P) minus evapotranspiration (ETR) minus outlet discharge at Tonneins (Q) from SIM or SWAT outputs only, a combination of modeled evapotranspiration (SIM, SWAT and the mean of the two models) outputs and in situ discharge at Tonneins from Banque Hydro, over 2003–2010. Bold figures correspond to highest correlation and NSE values between P-ETR-Q and GRACE $\Delta T\bar{W}S$ solutions. Correlation and Nash-Sutcliffe coefficient between $\Delta T\bar{W}S$ from CSR or GSFC GRACE mascon solutions and $\Delta T\bar{W}S$ from models and CSR mascon solution are also provided on the last table lines. Italic bold figures correspond to highest correlation and NSE values between CSR/SIM/SWAT $\Delta T\bar{W}S$ and CSR/GSFC $\Delta T\bar{W}S$.

	CSR mascon $\Delta T\bar{W}S$		GSFC mascon $\Delta T\bar{W}S$	
	Corr.	NSE	Corr.	NSE
SAFRAN P- SIM ETR- SIM Q	0.71	-0.03	0.76	0.44
SAFRAN P- SWAT ETR- SWAT Q	0.72	-0.6	0.82	0.31
SAFRAN P - SIM ETR - in situ Q	0.72	-0.15	0.78	0.44
SAFRAN P - SWAT ETR - in situ Q	0.73	-0.22	0.80	0.43
SAFRAN P-mean(SIM,SWAT) ETR-in situ Q	0.73	-0.15	0.79	0.46
$\Delta T\bar{W}S$ SIM	0.86	0.67	0.89	0.79
$\Delta T\bar{W}S$ SWAT	0.86	0.70	0.92	0.84
$\Delta T\bar{W}S$ CSR	1.00	1.00	0.88	0.59

connected to surface waters and the atmosphere interface.

3.3. Basin scale water balance

As described in section 2.4.2, the possibility to compute Garonne basin water balance using both GRACE data and SIM/SWAT inputs and outputs has been investigated. Correlation coefficient and NSE between GRACE solutions $\Delta T\bar{W}S$ and P-ETR-Q from models have been computed to assess their consistency (Table 5), as they should be similar (Eq. 2). Different P-ETR-Q have been tested: only from SIM inputs/outputs (first line of Table 5), only from SWAT inputs/outputs (second line of Table 5), P-ETR from SIM or SWAT and in situ Q at Tonneins from Banque Hydro database (third and fourth line of Table 5, respectively) and a composite estimate of P-ETR-Q with P from SAFRAN, ETR the mean between SIM and SWAT ETR and Q from in situ gauge (fifth line of Table 5). Only $\Delta T\bar{W}S$ from CSR and GSFC mascons solutions have been considered, as section 3.1 showed they best compare to SIM and SWAT TWSA. $\Delta T\bar{W}S$ from SIM and SWAT is also computed and compared to GRACE ones (last lines of Table 5). $\Delta T\bar{W}S$ from CSR mascon solutions has a correlation coefficient around 0.7 with the different P-ETR-Q estimates, but exhibits poor NSE (they are all negative). Much better results are obtained with $\Delta T\bar{W}S$ from GSFC mascon solution, which has correlation coefficient around 0.8 and NSE around 0.4 with the different P-ETR-Q estimates. Correlation and NSE are pretty similar for the four P-ETR-Q estimates. Correlation and NSE are better between GRACE mascon solutions $\Delta T\bar{W}S$ and models $\Delta T\bar{W}S$ (around 0.86 and 0.68 for CSR mascon solution, respectively and 0.9 and 0.8 for GSFC mascon solution, respectively). Surprisingly, correlation and NSE are slightly worse between CSR mascon and GSFC mascon (0.88 and 0.59, respectively) than between the mascon solutions and the models. This highlights the spread of the GRACE solutions because of differences in their processing, even if they are based on the same raw measurements. These differences are exacerbated by the small size of the basin compared to the GRACE products resolutions. It also shows that, at least for the Garonne basin, the GSFC mascon solution seems to be the best suitable solution for basin scale water balance study.

Fig. 8.a shows SIM and SWAT P-ETR-Q time series along with CSR and GSFC mascon solutions $\Delta T\bar{W}S$. It shows the overall good agreement between the datasets, and some high frequency variations in SWAT P-ETR-Q, which are not present in other datasets. This could explain why NSE between SWAT and GSFC mascon is lower than for SIM.

Fig. 8.b presents monthly accumulated SAFRAN precipitation time series over the basin, along with SIM and SWAT ETR and Q. If models' ETR and Q are similar, SIM Q and ETR are slightly more correlated to SAFRAN precipitation. That's why SWAT P-ETR-Q has more high frequency variations than SIM P-ETR-Q.

In order to analyze more easily interannual variations of precipitation, evapotranspiration and discharge, their hydrological yearly mean has been computed for both models (Fig. 8.c). SWAT evapotranspiration, in average, is always lower than SIM. Their discharges have similar values, even if SIM discharge is more correlated to precipitation than SWAT. From Fig. 8.c, it appears that important precipitation that occur during hydrological years 2004, 2009 and 2013 have relatively low impact on evapotranspiration, whereas river discharge increased during these years. On the contrary, important precipitation during year 2007 was concomitant with higher evapotranspiration and lower discharge. Concerning 2005, 2011 and 2012 dry years, evapotranspiration did not exhibit important increase or drop compared to other years, even if its lowest value occurred during hydrological year 2011. It should be highlighted that small variations of ETR at yearly time scale does not mean that ETR could not have important variations at seasonal time scale (Grusson et al., 2018). Besides, some important rain events, if they occur over a short time period might saturate the soil very quickly and/or might increase runoff with few impacts on vegetation and ETR, explaining the difference between ETR and precipitation on Fig. 8.c. Concerning discharge, it drops significantly in 2011 and 2012. So, the minimum TWSA observed in 2012 (see section 3.1 and Fig. 6) is mainly due to an important precipitation drop, an evapotranspiration rate that was close to other years and the fact that preceding year was also very dry (i.e. memory effect of TWS). During this drought period (the most important one within the studied period), the situation was so critical, that many localities in the basin have been officially recognized by the French government as having experienced "natural disaster" (République Française, 2012).

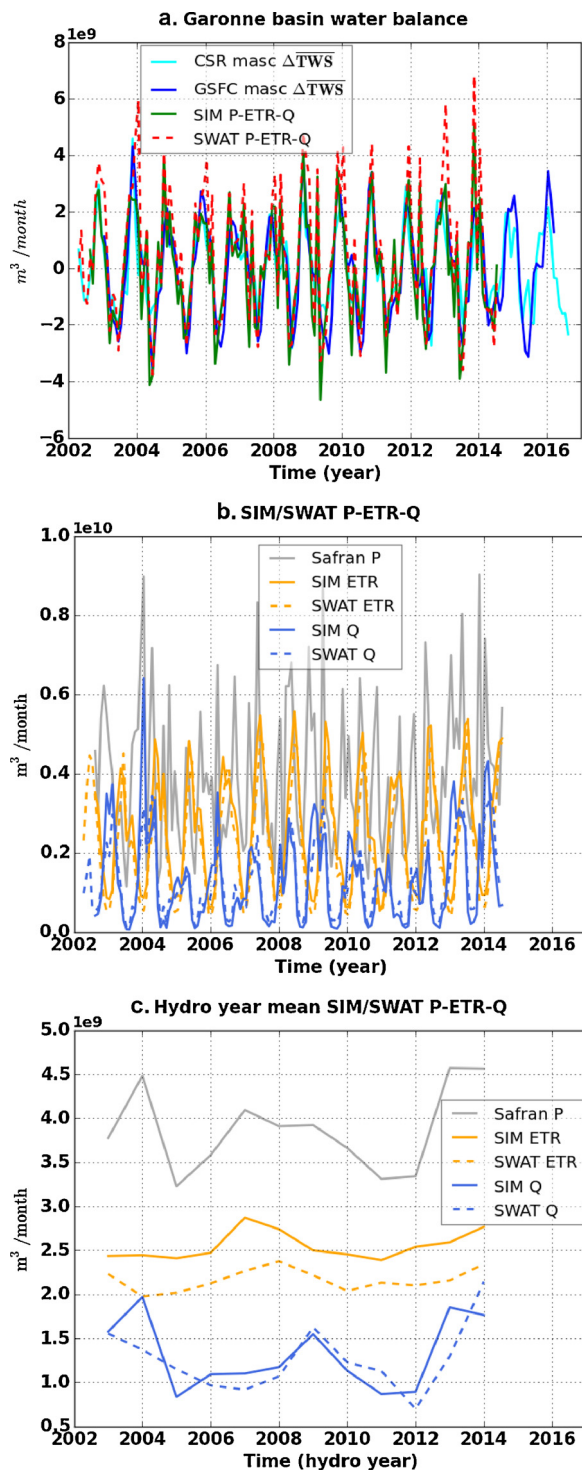


Fig. 8. (a.) Garonne basin water balance: GRACE ΔTWS from CSR mascon (cyan) and GSFC mascon (dark blue) solutions in $m^3/month$ compared to Precipitation (rain + snow from SAFRAN model), minus Evapotranspiration and discharge from SIM (green) and SWAT (dashed red line) models. (b.) SAFRAN precipitation (gray), SIM and SWAT evapotranspiration (solid and dashed orange lines, respectively), and SIM and SWAT river discharge at Tonneins (solid and dashed blue lines, respectively). (c.) Similar to panel (b.) for hydrological year mean (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

Table 6

Similar to Table 5 for time series without their respective 2003–2010 mean annual cycle over the Garonne basin.

	CSR mascon $\Delta T\bar{W}S$		GSFC mascon $\Delta T\bar{W}S$	
	Corr.	NSE	Corr.	NSE
SAFRAN P- SIM ETR- SIM Q	0.36	-2.18	0.52	-0.98
SAFRAN P- SWAT ETR- SWAT Q	0.36	-2.61	0.57	-1.16
SAFRAN P - SIM ETR - in situ Q	0.36	-2.36	0.54	-1.07
SAFRAN P - SWAT ETR - in situ Q	0.39	-2.29	0.56	-0.98
SAFRAN P-mean(SIM,SWAT) ETR-in situ Q	0.38	-2.28	0.55	-1.00
$\Delta T\bar{W}S$ SIM	0.22	-3.81	0.79	0.60
$\Delta T\bar{W}S$ SWAT	0.59	0.28	0.79	0.63
$\Delta T\bar{W}S$ CSR	1.00	1.00	0.64	0.04

Table 6 gives correlation and NSE for time series without their mean annual cycle between CSR and GSFC mascon solutions $\Delta T\bar{W}S$ and the different P-ETR-Q estimates. GSFC mascon solution still provides the best correlation coefficient (around 0.55), but the NSE remains quite low (~ -2.3 for CSR mascon and ~ -1 for GSFC mascon). GSFC mascon and models $\Delta T\bar{W}S$ remains well correlated even without their mean annual cycles, with a good NSE (~ 0.6). Again, GSFC mascon $\Delta T\bar{W}S$ better compares with models than with CSR mascon $\Delta T\bar{W}S$.

From these comparisons, it is clear that GSFC mascon solution is better suited than CSR mascon solution for computing $\Delta T\bar{W}S$ to study the Garonne basin water balance. They compare well to models P-ETR-Q. However, when the mean annual cycle is removed, the NSE is pretty low. This must be due to low spatial resolution of GRACE data (leading to noisy water mass estimates), which are quite coarse, compared to Garonne basin size. The fact that the temporal derivative from Eq. 1 is approximated by a second order central difference (Eq. 3) might also add errors.

But another source of error is the use of $\Delta T\bar{W}S$ instead of more accurate ΔTWS (see section 2.4.2). For example, correlation and NSE between SIM P-ETR-Q and SIM $\Delta T\bar{W}S$ are equal to 0.87 and 0.59, respectively. If these results are better than the ones obtained between SIM P-ETR-Q and GRACE mascon solutions, they are not that far from the comparison between GSFC mascon $\Delta T\bar{W}S$ and SIM P-ETR-Q (correlation coefficient and NSE equal to 0.76 and 0.44, respectively, see Table 5).

For GRACE solutions evaluated in this study, only $\Delta T\bar{W}S$ are available. However, for models, the more accurate ΔTWS could be computed, to check the uncertainty due to this approximation. For example, Fig. 9.a shows SIM P-ETR-Q (orange line) and SIM $\Delta T\bar{W}S$ (dashed blue line). Clearly, P-ETR-Q has more variability than $\Delta T\bar{W}S$, which is smoother. However, P-ETR-Q and ΔTWS are almost identical (Fig. 9.b), their correlation coefficient and NSE are almost 1 (0.999 and 0.998, respectively). Therefore, results from Table 5 and 6 are mainly due to GRACE coarse resolution both in space and time. The fact that $\Delta T\bar{W}S$ tends to be smoother than P-ETR-Q (Fig. 9.a) might not be an issue for big river basins analyzed in previous studies. However, for 50,000 km² basins like the Garonne, using monthly mean TWSA will prevent to properly close the water balance equation.

Some mismatch between GRACE mascon $\Delta T\bar{W}S$ and models P-ETR-Q could also be due to model issue. Especially, SIM and SWAT models do not consider actual irrigation, reservoirs and lakes. Finally, if precipitation comes from the most accurate database available over France mainland (SAFRAN), it still has some errors that could impact the comparison.

4. Conclusions and perspectives

This study inter-compared TWSA from seven GRACE satellite mission solutions with TWSA computed from two hydrological

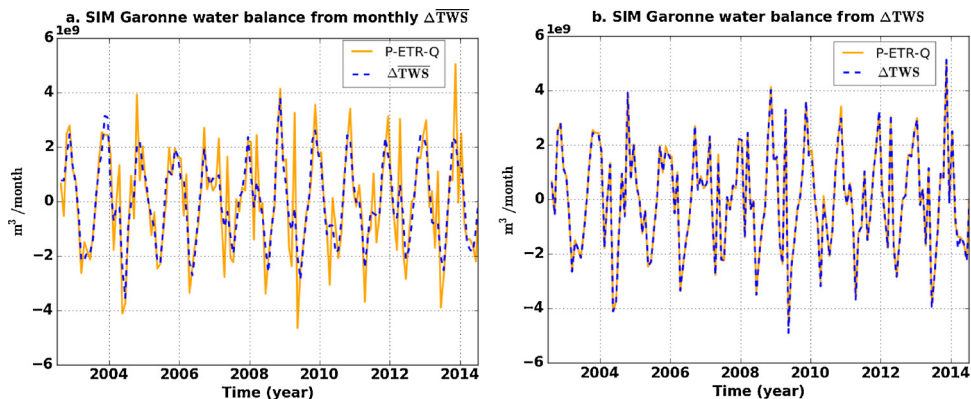


Fig. 9. SIM P-ETR-Q (orange lines) compared to $\Delta T\bar{W}S$ (blue dashed line) obtained from monthly mean SIM outputs (a.) and SIM ΔTWS (TWS difference between the last day of the month and the first day of the month; dashed blue line) (b.), over the Garonne basin at Tonneins (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

models (SIM and SWAT), over the 50,000 km² Garonne basin between August 2002 and July 2014. Despite that the Garonne basin is smaller than estimated GRACE spatial resolution, a good agreement between GRACE solutions and hydrological model TWSA has been found. Especially, GRACE mascon solutions from CSR and GSFC best match SIM and SWAT models TWSA, with correlation and NSE above 0.9 and 0.7, respectively. It also validates that SIM and SWAT models have enough physical processes to model correctly TWSA at basin scale. This good agreement is due to the important amplitude of the Garonne annual cycle, even if GRACE solutions and models annual cycle are 1 month shifted. Overall, it seems that some GRACE solutions provide meaningful observation for the Garonne basin, despite its small size compared to GRACE resolution.

The analysis of models outputs showed TWSA in the Garonne basin is mainly due to water stored in the first dozen meters of soil and in the shallow aquifer connected to the river. To a smaller extent, snow also influences Garonne TWSA. Open surface water (river network) TWSA is quite small and TWSA from deep aquifer and vegetation is negligible.

GRACE GSFC mascon solution ΔTWS best matches models P-ETR-Q. Yet, the basin-scale water balance is not close. It is due to GRACE poor spatial resolution compared to Garonne basin size, but also, as shown in this study, because of its monthly time resolution preventing the computation of accurate ΔTWS . This issue might not be important for big river basins, but for the Garonne basin ΔTWS is too smooth compared to monthly accumulated P-ETR-Q.

More work is needed to assess the reasons of the differences between GRACE and hydrological models. It is needed to better characterize sources of errors in GRACE solutions (like leakage error) and maybe to apply more local processing techniques, like the one proposed by Longuevergne et al. (2010). Using recently released sub-monthly GRACE solutions (e.g. Kurtenbach et al., 2012; Ramillien et al., 2015) might also help to compute a more accurate Garonne basin water balance. SIM and SWAT models used in this study currently represent state-of-the-art hydrological modeling of the Garonne basin. Yet, they do not take into account lakes, reservoirs and irrigation. Taking these effects into account would be important to improve our understanding of water fluxes within the basin. Doing a thorough sensitivity study on the models parameters and forcing should also help to better quantify models uncertainties. Garonne basin is quite impacted by human activities (reservoirs, irrigated agriculture), especially during low flow (Sauquet et al., 2009). Their impact on TWS at basin scale is still unclear and more ancillary data should be collected to investigate this link and to determine if GRACE data are accurate enough to observe these impacts. Finally, GRACE Follow-On (GRACE-FO), GRACE successor, was launched May 22nd, 2018. Validating GRACE-FO accuracy on the Garonne basin and its potential improvement compared to GRACE would be valuable to extend TWSA observation time series, as the basin is still experiencing important drought periods.

Conflict of interest

None.

Funding

This project was funded by the “Réseau Thématique de Recherche Avancée - Sciences et Technologies pour l’Aéronautique et l’Espace” (RTRA-STAE, Toulouse, France), through a grant attributed to the “Ressources en Eau sur le bassin de la GARonne: interaction entre les composantes naturelles et anthropiques et apport de la télédétection” (REGARD) project.

Acknowledgements

GRACE Land (i.e. global GRACE Tellus Land and JPL mascons) are available at <http://grace.jpl.nasa.gov>, supported by the NASA MEaSUREs Program. GSFC mascons solution is available at <https://neptune.gsfc.nasa.gov/gngphys/index.php?section=470> and also supported by the NASA MEaSUREs Program. GRACE and LAGEOS data from CNES/GRGS can be downloaded at <https://grace.obs-mip.fr>. CSR mascon solution can be found at http://www2.csr.utexas.edu/grace/RL05_mascons.html. All the institutions (i.e. CSR, CNES/GRGS, GFZ, GSFC and JPL) that provide freely these solutions are warmly thanked. The Banque Hydro database (<http://www.hydro.eaufrance.fr/>) is acknowledged for providing freely in situ discharge at Tonneins. Météo-France is thanked for providing freely SAFRAN atmospheric forcing data over the Garonne basin. Anny Cazenave, Alejandro Blazquez and Laurent Longuevergne are acknowledged for very fruitful discussions concerning GRACE data.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.ejrh.2019.100609>.

References

- Abbaspour, K.C., Johnson, C.A., van Genuchten, M.T., 2004. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone J.* 3, 1340–1352. <https://doi.org/10.2136/vzj2004.1340>.
- Abbaspour, K.C., 2012. SWAT-CUP 2012: SWAT Calibration and Uncertainty Programs - a User Manual. Swiss Federal Institute of Aquatic Science and Technology (EAWAG), Dübendorf, Switzerland, pp. 103.
- Becker, M., Papa, F., Frappart, F., Alsdorf, D., Calmant, S., Santos da Silva, J., Prigent, C., Seyler, F., 2018. Satellite-based estimates of surface water dynamics in the Congo Basin. *Int. J. Appl. Earth Obs. Geoinf.* 69, 196–209. <https://doi.org/10.1016/j.jag.2017.11.015>.

- Boone, A.A., Etchevers, P., 2001. An intercomparison of three snow schemes of varying complexity coupled to the same land surface model: local-scale evaluation at an alpine site. *J. Hydrometeorol.* 2, 374–394. [https://doi.org/10.1175/1525-7541\(2001\)002<0374:AOTSS>2.0.CO;2](https://doi.org/10.1175/1525-7541(2001)002<0374:AOTSS>2.0.CO;2).
- Boukerma, B., 1987. *Modélisation Des Écoulements Superficiels Et Souterrains Dans Le Sud-ouest De La France : Approche Du Bilan Hydrique*. PhD Thesis report. Ecole Nationale Supérieure des Mines de Paris.
- Büttner, G., Soukup, T., Kostra, B., 2014. CLC2012 Addendum to CLC2006 Technical Guidelines. ETC SIA. https://land.copernicus.eu/user-corner/technical.library/Addendum_finaldraft_v2_August82014.pdf.
- Caballero, Y., Voirin-Morel, S., Habets, F., Noilhan, J., Le Moigne, P., Lehenaff, A., Boone, A., 2007. Hydrological sensitivity of the Adour-Garonne river basin to climate change. *Water Resour. Res.* 43, W07448. <https://doi.org/10.1029/2005WR004192>.
- Calvet, J.-C., Noilhan, J., Roujean, J.-L., Bessemoulin, P., Cabelguenne, M., Olioso, A., Wigneron, J.-P., 1998. An interactive vegetation SVAT model tested against data from six contrasting sites. *Agric. For. Meteorol.* 92 (2), 73–95. [https://doi.org/10.1016/S0168-1923\(98\)00091-4](https://doi.org/10.1016/S0168-1923(98)00091-4).
- Chen, J., Famiglietti, J.S., Scanlon, B.R., Rodell, M., 2016. Groundwater storage changes : present status from GRACE observations. *Surv. Geophys.* 37 (2), 397–417. <https://doi.org/10.1007/s10712-016-9370-6>.
- David, C.H., Habets, F., Maidment, D.R., Yang, Z.-L., 2011. RAPID applied to the SIM France model. *Hydrol. Process.* 25 (22), 3412–3425. <https://doi.org/10.1002/hyp.8070>.
- Decharme, B., Martin, E., Faroux, S., 2013. Reconciling soil thermal and hydrological lower boundary conditions in land surface models. *J. Geophys. Res. Atmos.* 118, 1–16. <https://doi.org/10.1002/jgrd.50631>.
- Döll, P., Douville, H., Güntner, A., Müller Schmied, H., Wada, Y., 2016. Modelling freshwater resources at the global scale: challenges and prospects. *Surv. Geophys.* 37, 195–221. <https://doi.org/10.1007/s10712-015-9343-1>.
- Dümenil, L., Todini, E., 1992. A rainfall-runoff scheme for use in the Hamburg climate model. In: O’Kane, J.P. (Ed.), *Advances in Theoretical Hydrology: A Tribute to James Dooge*. Elsevier Science Publishers B.V., Amsterdam, pp. 129–157.
- Durand, Y., Brun, E., Merindol, L., Guyomarc’h, G., Lesaffre, B., Martin, E., 1993. A meteorological estimation of relevant parameters for snow models. *Ann. Glaciol.* 18, 65–71. <https://doi.org/10.3189/S0260305500011277>.
- European Environmental Agency, 2007. CLC2006 Technical Guidelines. EEA Technical Report N°17/2007, ISSN 1725-2237, 66p.
- Faroux, S., Kaptué Tchuenté, A.T., Roujean, J.-L., Masson, V., Martin, E., Le Moigne, P., 2013. ECOCLIMAP-II/Europe: a twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models. *Geosci. Model. Dev.* 6 (2), 563–582. <https://doi.org/10.5194/gmd-6-563-2013>.
- Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodríguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D., 2007. The shuttle radar topography mission. *Rev. Geophys.* 45 (2), R2004. <https://doi.org/10.1029/2005RG000183>.
- Frappart, F., Papa, F., Malbeteau, Y., León, J.G., Ramillien, G., Prigent, C., Seoane, L., Seyler, F., Calmant, S., 2015. Surface freshwater storage variations in the Orinoco floodplains using multi-satellite observations. *Remote Sens. (Basel)* 7, 89–110. <https://doi.org/10.3390/rs70100089>.
- Frappart, F., Ramillien, G., 2018. Monitoring groundwater storage changes using the gravity recovery and climate experiment (GRACE) satellite mission: a review. *Remote Sens. (Basel)* 10, 829. <https://doi.org/10.3390/rs10060829>.
- Gao, Y., Tang, Q., Ferguson, C.R., Wood, E.F., Lettenmaier, D.P., 2010. Estimating the water budget of major US river basins via remote sensing. *Int. J. Remote Sens.* 31 (14), 3955–3978. <https://doi.org/10.1080/01431161.2010.483488>.
- Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., 2007. The soil and water assessment tool: historical development, applications, and future research directions. *Trans. Asabe* 50 (4), 1211–1250. <https://doi.org/10.13031/2013.23637>.
- Guo, J.Y., Huang, Z.W., Shum, C.K., van der Wal, W., 2012. Comparisons among contemporary glacial isostatic adjustment models. *J. Geodyn.* 61, 129–137. <https://doi.org/10.1016/j.jog.2012.03.011>.
- Gupta, H.V., Sorooshian, S., Yapo, P.O., 1999. Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. *J. Hydrol. Eng.* 4 (2), 135–143. [https://doi.org/10.1061/\(ASCE\)1084-0699\(1999\)4:2\(135\)](https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135)).
- Grusson, Y., Sun, X., Gascoïn, S., Sauvage, S., Raghavan, S., Ancil, F., Sánchez-Pérez, J.M., 2015. Assessing the capability of the SWAT model to simulate snow, snow melt and streamflow dynamics over an alpine watershed. *J. Hydrol. (Amst)* 531 (3), 574–588. <https://doi.org/10.1016/j.jhydrol.2015.10.070>.
- Grusson, Y., Ancil, F., Sauvage, S., Sánchez Pérez, J.M., 2017a. Assessing the climatic and temporal transposability of the SWAT model across a large contrasted watershed. *J. Hydrol. Eng.* 22 (6). [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001491](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001491).
- Grusson, Y., Ancil, F., Sauvage, S., Sánchez Pérez, J.M., 2017b. Testing the SWAT model with gridded weather data of different spatial resolutions. *Water* 9 (1), 54. <https://doi.org/10.3390/w9010054>.
- Grusson, Y., Ancil, F., Sauvage, S., Sánchez Pérez, J.M., 2018. Coevolution of hydrological cycle components under climate change: the case of the Garonne River in France. *Water* 10 (12), 1870. <https://doi.org/10.3390/w10121870>.
- Häfliger, V., Martin, E., Boone, A., Habets, F., David, C.H., Garambois, P.-A., Roux, H., Ricci, S., Berthon, L., Thévenin, A., Biancamaria, S., 2015. Evaluation of regional-scale water level simulations using various river routing schemes within a hydrometeorological modelling framework for the preparation of the SWOT mission. *J. Hydrometeorol.* 16 (4), 1821–1842. <https://doi.org/10.1175/JHM-D-14-0107.1>.
- Kurtenbach, E., Eicker, A., Mayer-Gürr, T., Hohlschneider, M., Hayn, M., Fuhrmann, M., Kusche, J., 2012. Improved daily GRACE gravity field solutions using a Kalman smoother. *J. Geodyn.* 59–60, 39–48. <https://doi.org/10.1016/j.jog.2012.02.006>.
- Landerer, F.W., Swenson, S.C., 2012. Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resour. Res.* 48, W04531. <https://doi.org/10.1029/2011WR011453>.
- Ledoux, E., Girard, G., De Marsily, G., Deschenes, J., 1989. Spatially distributed modelling: conceptual approach, coupling surface water and ground-water. In: Morel-Seytoux, H.J. (Ed.), *Unsaturated Flow Hydrologic Modeling: Theory and Practice*. NATO Sciences Service, pp. 434–454.
- Lemoine, J.-M., Bourgogne, S., Biancale, R., Bruinsma, S., Gégout, P., 2016. CNES/GRGS solutions - focus on the inversion process. Oral Presentation at the GRACE Science Team Meeting.
- Lettenmaier, D.P., Alsdorf, D., Dozier, J., Huffman, G.J., Pan, M., Wood, E.F., 2015. Inroads of remote sensing into hydrologic science during the WRR era. *Water Resour. Res.* 51, 7309–7342. <https://doi.org/10.1002/2015WR017616>.
- Long, D., Pan, Y., Zhou, J., Chen, Y., Hou, X., Hong, Y., Scanlon, B.R., Longuevergne, L., 2017. Global analysis of spatiotemporal variability in merged total water storage changes using multiple GRACE products and global hydrological models. *Remote Sens. Environ.* 192, 198–216. <https://doi.org/10.1016/j.rse.2017.02.011>.
- Longuevergne, L., Scanlon, B.R., Wilson, C.R., 2010. GRACE Hydrological estimates for small basins: Evaluating processing approaches on the High Plains Aquifer, USA. *Water Resour. Res.* 46, W11517. <https://doi.org/10.1029/2009WR008564>.
- Loomis, B.D., Luthcke, S.B., 2014. Optimized signal denoising and adaptive estimation of seasonal timing and mass balance from simulated GRACE-like regional mass variations. *Adv. Adapt. Data Anal.* 06, 1450003. <https://doi.org/10.1142/S1793536914500034>.
- Lorenz, C., Kunstmann, H., Devaraju, B., Tourian, M.J., Sneeuw, N., Riegger, J., 2014. Large-scale runoff from landmasses: a global assessment of the closure of the hydrological and atmospheric water balances. *J. Hydrometeorol.* 15, 2111–2139. <https://doi.org/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. *J. Glaciol.* 59 (216), 613–631. <https://doi.org/10.3189/2013JoG12J147>.
- Martin, E., Gascoïn, S., Grusson, Y., Murgue, C., Bardeau, M., Ancil, F., Ferrant, S., Lardy, R., Le Moigne, P., Leenhardt, D., Rivalland, V., Sánchez Pérez, J.M., Sauvage, S., Therond, O., 2016. Hydrological modelling in highly anthropized river basins: examples from the Garonne basin. *Surv. Geophys.* 37 (2), 223–247. <https://doi.org/10.1007/s10712-016-9366-2>.
- Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. Asabe* 50 (3), 885–900.
- Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2009. *Soil and Water Assessment Tool Theoretical Documentation Version 2009*. Texas Water Resources

- Institute. Technical Report N°406 <https://swat.tamu.edu/media/99192/swat2009-theory.pdf> (last access: 15 November 2018).
- Noilhan, J., Planton, S., 1989. A simple parameterization of land surfaces processes for meteorological models. *Mon. Weather. Rev.* 117, 536–549. [https://doi.org/10.1175/1520-0493\(1989\)117<0536:ASPOL5>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<0536:ASPOL5>2.0.CO;2).
- Noilhan, J., Mahfouf, J.-F., 1996. The ISBA land surface parameterization scheme. *Glob. Planet. Change* 13, 145–159. [https://doi.org/10.1016/0921-8181\(95\)00043-7](https://doi.org/10.1016/0921-8181(95)00043-7).
- Nachtergaele, F., van Velthuizen, H., Verelst, L., Wiberg, D., 2012. Harmonized World Soil Database, Version 1.2. FAO, IIASA, ISRIC, ISSCAS, JRC. available at: <http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/> (last access: 15 November 2018).
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models Part I – a discussion of principles. *J. Hydrol. (Amst)* 10 (3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Niu, G.-Y., Yang, Z.-L., 2006. Assessing a land surface model's improvements with GRACE estimates. *Geophys. Res. Lett.* 33, L07401. <https://doi.org/10.1029/2005GL025555>.
- Olivera, F., Valenzuela, M., Srinivasan, R., Choi, J., Cho, H., Koka, S., Agrawal, A., 2006. ARCGIS-SWAT: a geodata model and GIS interface for SWAT. *J. Am. Water Resour. Assoc.* 42 (2), 295–309. <https://doi.org/10.1111/j.1752-1688.2006.tb03839.x>.
- Papa, F., Frappart, F., Güntner, A., Prigent, C., Aires, F., Getirana, A.C.V., Maurer, R., 2013. Surface freshwater storage and variability in the Amazon basin from multi-satellite observations, 1993–2007. *J. Geophys. Res. Atmos.* 118 (21), 11951–11965. <https://doi.org/10.1002/2013JD020500>.
- Papa, F., Frappart, F., Malbeteau, Y., Shamsudduha, M., Venugopal, V., Sekhar, M., Ramillien, G., Prigent, C., Aires, F., Pandey, R.K., Bala, S., Calmant, S., 2015. Satellite-derived surface and sub-surface water storage in the Ganges-Brahmaputra river basin. *J. Hydro. : Reg. Stud.* 4 (A), 15–35. <https://doi.org/10.1016/j.ejrh.2015.03.004>.
- Pavelsky, T.M., Durand, M.T., Andreadis, K.M., Beighley, R.E., Paiva, R.C.D., Allen, G.H., Miller, Z.F., 2014. Assessing the potential global extent of SWOT river discharge observations. *J. Hydrol. (Amst)* 519, 1516–1525. <https://doi.org/10.1016/j.jhydrol.2014.08.044>.
- Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L., Morel, S., 2008. Analysis of Near-Surface Atmospheric Variables: Validation of the SAFRAN Analysis over France. *J. Appl. Meteorol. Climatol.* 47, 92–107. <https://doi.org/10.1175/2007JAMC1636.1>.
- Raimonet, M., Oudin, L., Thiéu, V., Silvestre, M., Vautard, R., Rabouille, C., Le Moigne, P., 2017. Evaluation of gridded meteorological datasets for hydrological modeling. *J. Hydrometeorol.* 18 (11), 3027–3041. <https://doi.org/10.1175/JHM-D-17-0018.1>.
- Ramillien, G.L., Frappart, F., Gratton, S., Vasseur, X., 2015. Sequential estimation of surface water mass changes from daily satellite gravimetry data. *J. Geod.* 89 (3), 259–282. <https://doi.org/10.1007/s00190-014-0772-2>.
- République Française, 2012. Arrêté du 11 juillet 2012 portant reconnaissance de l'état de catastrophe naturelle. *J. Off. de la République Française* 164, 11688. (last access 24 September 2018). <https://www.legifrance.gouv.fr/eli/arrete/2012/7/11/INTE1228647A/jo/texte>.
- Rodell, M., Houser, P.R., Jambor, U., Gottschalk, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J.K., Walker, J.P., Lohmann, D., Toll, D., 2004. The global land data assimilation system. *Bull. Am. Meteorol. Soc.* 85, 381–394. <https://doi.org/10.1175/BAMS-85-3-381>.
- Rodell, M., Chen, J., Kato, H., Famiglietti, J.S., Nigro, J., Wilson, C.R., 2007. Estimating groundwater storage changes in Mississippi River basin (USA) using GRACE. *Hydrogeol. J.* 15 (1), 159–166. <https://doi.org/10.1007/s10040-006-0103-7>.
- Rodell, M., Famiglietti, J.S., Wiese, D.N., Reager, J.T., Beaudoin, H.K., Landerer, F.W., Lo, M.-H., 2018. Emerging trends in global freshwater availability. *Nature* 557, 651–659. <https://doi.org/10.1038/s41586-018-0123-1>.
- Sauquet, E., Dupeyrat, A., Hendrickx, F., Perrin, C., Samie, R., Vidal, J.-P., 2009. IMAGINE 2030, Climate and Water Management: Uncertainties on Water Resources for the Garonne River Basin in 2030? (last access 24 November 2018). <http://cemadoc.irstea.fr/oa/PUB00028876-imagine-2030-climat-amenagements-garonne-quelles-i.html>.
- Save, H., Bettadpur, S., Tapley, B.D., 2016. High resolution CSR GRACE RL05 mascons. *J. Geophys. Res. Solid Earth* 121, 7547–7569. <https://doi.org/10.1002/2016JB013007>.
- Scanlon, B.R., Zhang, Z., Save, H., Wiese, D.N., Landerer, F.W., Long, D., Longuevergne, L., Chen, J., 2016. Global evaluation of new GRACE mascon products for hydrologic applications. *Water Resour. Res.* 52, 9412–9429. <https://doi.org/10.1002/2016WR019494>.
- Stöckli, R., Vermote, E., Saleous, N., Simmon, R., Herring, D., 2005. The Blue Marble Next Generation - a True Color Earth Dataset Including Seasonal Dynamics From MODIS. Published by the NASA Earth Observatory (accessed 25 April 2017). <https://earthobservatory.nasa.gov/Features/BlueMarble/>.
- Swenson, S.C., Wahr, J., 2006. Post-processing removal of correlated errors in GRACE data. *Geophys. Res. Lett.* 33, L08402. <https://doi.org/10.1029/2005GL025285>.
- Swenson, S.C., 2012. GRACE Monthly Land Water Mass Grids NETCDF RELEASE 5.0. Ver 5.0. PO.DAAC, CA, USA. Dataset accessed [2017-01-01] at: <https://doi.org/10.5067/TELND-NC005>.
- Tachikawa, T., Hato, M., Kaku, M., Iwasaki, A., 2011. Characteristics of ASTER GDEM version 2. *Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, 2011. pp. 3657–3660.
- Tapley, B.D., Bettadpur, S., Ries, J.C., Thompson, P.F., Watkins, M.M., 2004. GRACE measurements of mass variability in the Earth system. *Science* 305, 503–505. <https://doi.org/10.1126/science.1099192>.
- Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., Soubeyrou, J.-M., 2010. A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *Int. J. Climatol.* 30 (11), 1627–1644. <https://doi.org/10.1002/joc.2003>.
- Vishwakarma, B.D., Devaraju, B., Sneeuw, N., 2018. What is the spatial resolution of GRACE satellite products for hydrology. *Remote Sens. (Basel)* 10, 852. <https://doi.org/10.3390/rs10060852>.
- Watkins, M.M., Wiese, D.N., Yuan, D.-N., Boening, C., Landerer, F.W., 2015. Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *J. Geophys. Res. Solid Earth* 120, 264–2671. <https://doi.org/10.1002/2014JB011547>.
- Wouters, B., Bonin, J.A., Chambers, D.P., Riva, R.E.M., Sasgen, I., Wahr, J., 2014. GRACE, time-varying gravity, Earth system dynamics and climate change. *Rep. Prog. Phys.* 77 (11), 116801. <https://doi.org/10.1088/0034-4885/77/11/116801>.
- Zaitchik, B.F., Rodell, M., Reichle, R.H., 2008. Assimilation of GRACE terrestrial water storage data into a land surface model: results for the Mississippi River basin. *J. Hydrometeorol.* 9, 535–548. <https://doi.org/10.1175/2007JHM951.1>.