Regularizing GRACE mascon solutions using river run-off data

O. (Obbe) Lucassen
(4377729)

Submitted in partial fulfillment of the requirements for
the degree Master of Science in Geoscience & Remote Sensing

Department of Geoscience and Remote Sensing
Faculty of Civil Engineering & Geosciences
Delft University of Technology

April 8, 2021

SUPERVISORS:
Dr. P. (Pavel) Ditmar
Prof. R. (Roland) Klees
Prof. Dr. ir. S. (Susan) Steele-Dunne
Dr. M. (Markus) Hrachowitz

LOCATION:
Faculty of Civil Engineering & Geosciences
Stevinweg 1, 2628 CN Delft
The Netherlands

TIME FRAME:
April 8, 2021

1 FOREWORD

Before you lies the report "Regularizing GRACE mascon solutions using river run-off data", which tries to make a contribution to the field of GRACE data processing. The research focuses on a new type of data combination to improve the quality of terrestrial water storage variations estimations from GRACE data. This thesis is written as completion of the Geoscience and Remote Sensing Master's program at the Delft University of Technology. Research on this topic was conducted from April 2020 to April 2021.

In this period I learned a great deal, not only on the topic of GRACE data processing but also about areas I was less familiar with, like hydrology. The research was really exciting, dealing with a large variety of data sets and the complicated underlying physical processes made it a real challenge. So, being able to answer all research questions, which were formulated at the start of the research, gives me a sense of fulfillment.

The research questions were formulated together with my supervisors, Dr. P. (Pavel) Ditmar, Prof. R. (Roland) Klees, Prof. Dr. Ir. S. (Susan) Steele-Dunne, and Dr. M. (Markus) Hrachowitz, and were based on an idea of Dr. P. (Pavel) Ditmar. I would like to thank all the supervisors for their support and guidance throughout the entire process. A special thanks to Dr. P. (Pavel) Ditmar, who spend much time and effort by attending weekly meetings and sharing his profound knowledge on the topic. I would also like to thank my family and friends, who kept me motivated this entire period.

I hope you will enjoy reading my thesis.

Obbe Lucassen

Rotterdam, April 8, 2021

This thesis is submitted in partial fulfillment of the requirements for the degree Master of Science in Geoscience and Remote Sensing at the Delft University of Technology.

ABSTRACT

Accurate estimates of terrestrial water storage variations (TWSV) are critical for a variety of applications, e.g., model calibration and climate studies. This study aims to find the added value of river run-off data for regularizing GRACE mascon solution, from which TWSV can be estimated. Most subbasins of the Mississippi Basin show an exponential relationship between run-off and TWSV with moderate to good predictive value. The mean value of the explained variance (R^2) for the models is 0.4, when excluding the Lower Mississippi and Middle Missouri subbasins the mean value increases to 0.5. The Lower Mississippi and Middle Missouri subbasins are the only subbasins with no clear correlation between run-off and storage, most likely due to aquifer depletion.

Since ready-to-used GRACE mascon solutions cannot be modified, the mascon method developed at the Geoscience and Remote Sensing Department of the Delft University of Technology is used in this study. This study proposes modifications to tailor this method to the Mississippi Basin and mass anomalies of hydrological origin. After applying standard Tikhonov regularization, the exponential relationship is used to determine the most probable model outcome for each subbasin, adding a new term to the estimator. The resulting GRACE mascon solutions (Tikhonov and run-off regularized solutions) have a root-mean-square deviation (RMSD) of around 3.5 cm when compared to ready-to-use mascon solutions from the Jet Propulsion Laboratory and Goddard Space Flight Center. The run-off regularized solutions show smaller deviation (on average 7 %) percent) than the Tikhonov regularized solutions with respect to an independent validation data set. The largest difference between the Tikhonov and run-off regularized solutions occur in the dryer subbasins (Platte, Middle, and Upper Missouri, and Upper Mississippi subbasin). These areas show a stronger inter-annual trend in TWSV, these trends are captured better by the run-off regularized solutions. Additional research, co-estimating the regularization bias, shows that the run-off regularized solutions induce less bias into the solutions. This study shows that using run-off data when processing GRACE data could be of added value, especially in semi-humid to arid areas, since the TWSV in these areas is more susceptible to inter-annual variations.

CONTENTS

2 INTRODUCTION

Having accurate estimations of the terrestrial water storage variations (TWSV) is very important for a wide range of applications. Terrestrial water storage (TWS) is all water on the surface and in the subsurface of the Earth. These estimations may be used to predict floods and droughts, to study the effects of climate change on the water cycle, or may act as boundary conditions in climatological and hydrological models. So, creating a method to observe TWSV with a high spatial and temporal resolution is of high relevance.

Satellite Gravimetry is a remote sensing technique that measures the state of the Earth's gravitational field. After repeated measurements, changes in the gravity field can be inferred. Changes in the water storage lead to changes in the gravitational field and this way the Gravity Recovery and Climate Experiment (GRACE) mission can be used to detect TWSV. The GRACE mission operated from 2002 to 2017, after this mission ended the GRACE Follow-On was launched, which is still operational at this moment. The spatial resolution of GRACE products is limited to around 300 kilometers, this is mainly due to the altitude at which the satellites orbit (Wouters et al., 2014). Only for the few months that are characterized by a short repeat period, ground track coverage has a minor influence on the spatial resolution. Currently, a temporal resolution of one month is the norm, however, there is a wide effort to increase the temporal resolution (Wouters et al., 2014). There is a trade-off between spatial and temporal resolution, an increase in temporal resolution results in a decrease of spatial resolution and vice versa. Another problem of the GRACE mission is the difference in sensitivity in the East-West direction compared to the North-South direction. The sensitivity in the East-West direction is worse since the satellites are in a near-polar orbit. As a result, the noise in GRACE data has a strong non-isotropic North-South character, this causes the well-known 'stripes' in GRACE solutions (Wouters et al., 2014).

To minimize the noise a filter or regularization technique is used. The idea of a filter is that it drops the variance in the solutions, but as a consequence, it adds bias. So, when calculating mass or gravity variations from GRACE data filtering is needed to smoothen the solution, which leads to biased amplitude estimates. Filtering can be applied after an initial solution by inversion is obtained. Regularization adds additional information within the least-squares inversion procedure to prevent over-fitting. By replacing traditional constraints with data-based constraints within the regularization procedure, the bias is expected to drop with respect to traditional Tikhonov regularization. This is expected because the data-based constraints are based on the psychical processes themselves, whereas the traditional constraints are purely mathematical.

From early on GRACE data has been used to improve global hydrological models (Güntner, 2008). It is used to estimate TWSV for basins as small as 50,000 km^2 (Biancamaria et al., 2019) up to the Amazon and Mississippi basin (Klosko et al., 2009; Xavier et al.,

2010). The link between GRACE-derived TWSV and run-off is studied for the Lena River (Suzuki et al., 2016) and for many other large rivers (Riegger and Tourian, 2014). In the early stage of the GRACE missions, the Earth's gravity field was parameterized using spherical harmonics (SH) basis functions. A later developed approach parameterizes the mass anomalies using regional mass concentration (mascon) functions (Bridget R. Scanlon1, Zizhan Zhang1, Himanshu Save2, David N. Wiese3, Felix W. Landerer3, Di Long4, Laurent Longuevergne5 and Chen2, 2016). A method at TU Delft is developed that transforms monthly spherical harmonic solutions into mass anomaly estimates. This method was successfully applied to Greenland and Antarctica to study regional mass loss trends (Engels et al., 2018; Ran et al., 2018).

Previous work from Riegger and Tourian (2014) and MacEdo et al. (2019) showed that run-off and GRACE-derived storage show a relationship. Riegger and Tourian (2014) already showed that river run-off and variations in the terrestrial water storage estimated by off-the-shelf GRACE solutions show a high correlation after processing the run-off data. Their study was focused on a number of the world's largest river basins, including the Amazon, Niger, and Mackenzie Basin with drainage areas from more than 1.5 to 4 million km^2. Here mass changes are easier to detect given the spatial resolution of GRACE solutions ($\pm$300km). MacEdo et al. (2019) demonstrated that there is a relationship between run-off and storage for different subbasins of the Mississippi Basin. The hypothesis is that run-off data could be used to constrain the GRACE solutions and this way create high-resolution data-only estimates of the TWSV, if the correlation is also existing for smaller basins.

Jet Propulsion Laboratory (JPL) and Goddard Space Flight Center (GSFC) produce level-3 mass concentration block (mascon) solutions (Luthcke et al., 2013; Watkins et al., 2016). A mascon is a predefined area for which gravity solutions in terms of mass anomalies are calculated. As soon as off-the-shelf solutions are concerned these can not be modified, so the variant of the mascon method created at Delft University of Technology will be used to study the effect of run-off data when applied to regularize the estimates (Ran et al., 2018).

The main goal of this project is to analyze the added value of run-off data for high-resolution estimation of TWSV in the study area from GRACE data. The sub-goals of the project are:

- selection of a study area;
- adoption of the GRACE data processing technique developed earlier at the Geoscience and Remote Sensing department at TU Delft;
- fine-tuning and validation of that technique in the context of the selected study area;
- collection of independent data to estimate TWSV in the study area;
- quantification of the relationships between TWSV and run-off variations in the study area using GRACE-based TWSV estimates;
- implementation of regularization into the GRACE data processing technique to combine GRACE data with prior information; and

- validation of the obtained results using independent data.

The structure of the thesis is as follows: first, related previous work and background information is discussed in Chapter 3. This is followed by a deeper look into the requirements for a suitable study area, the properties it should possess, and the final decision made regarding the study area in Chapter 4. Which data is used will be laid out as well as how these data are processed, in Chapter 5. This chapter also includes the methodology of this project, how the run-off storage relationships are obtained, and how these are implemented into the mascon approach. Numerical experiments were performed to find the optimal parameter settings, these are discussed in Chapter 6. Next, the results will be shown in Chapter 7. In the discussion (Chapter 8) an interpretation of the obtained results is provided and a discussion on how they compare to JPL and GSFC mascon solutions. Also, recommendation for further work will be given. Finally, a conclusion will be drawn on the added value of the proposed technique in Chapter 9.

3 PREVIOUS WORK

In this chapter previous work is discussed that is of interest for this research. First, satellite gravimetry will be introduced, followed by an introduction to the mass concentration block approach. Next, the theoretical and experimental relationship between river run-off and water storage will be touched upon. Finally, independent data sets that could be used for validation will be presented.

3.1 SATELLITE GRAVIMETRY

Gravity is one of the fundamental forces which results from the presence of mass. One mass attracts another mass and vice versa, according to Newton's law of attraction:

$$|\mathbf{F}_1| = |\mathbf{F}_2| = G\frac{m_1 m_2}{d^2}, \tag{3.1}$$

where F_1 and F_2 are the forces exerted on the two point masses, m_1 and m_2, d is the distance between those masses and G is the gravitational constant. From this and the assumption that the Earth is a spherical symmetric object, the acceleration a mass experiences due to mass M to can be written as:

$$g = \frac{GM}{r^2}, \tag{3.2}$$

where g is the gravitational acceleration, M the second mass (Earth) and r the distance between $m1$ and the center of the Earth. At the Earth's surface the average gravity (g) is 9.81 m/s^2. However, this attraction is not constant over the Earth's surface. A reason for these regional differences in the rotation of the Earth, this is a process that is well understood and can be accurately predicted. Water distribution and movement in the atmosphere also cause differences in gravitational attraction. Just like the Earth's rotation these can be accurately measured and modeled. The non-homogeneous distribution of mass is another reason for these spatial differences. This is much more complicated since it requires knowledge of the Earth's interior to model and predict this. These processes that occur in the subsurface and interior of the Earth, like plate tectonics and changes in the density of the interior, are very slow. This means that on a relatively short time scale they can be assumed to be constant. All this information makes time-variable gravity measurements very valuable since know process can be subtracted and information on less known processes can be obtained (Wouters et al., 2014). This research focuses on mass changes due to water in terrestrial areas, i.e. water stocks that are held in a specific river basin. Another example of physical processes that trigger changes in gravitational attraction is the mass loss of glaciers and ice sheets. NASA's and DLR's joint Gravity And Climate Experiment (GRACE) satellite mission was launched in 2002 to observe these changes.

The processes due to the redistribution of water can cause variation in the gravitational attraction of the order 10^{-8} m/s^2. To be able to measure these changes very accurate instruments are needed. On a global scale, this was done by the Gravity Recovery And

Climate Experiment (GRACE) satellites, which were launched in 2002 and operated up to 2017. GRACE mission consisted of two satellites (separated by about 220 kilometers) that were in a near-polar (inclination of 89 degrees) orbit. It constantly tracked the distance between the two satellites and from that information on the Earth's gravity field is extracted. Mass anomalies cause the satellite to change the gravitational acceleration, and since the two satellites are in different positions this changes the inter-satellite distance. This distance is measured using a very accurate K-band microwave ranging system (KBR). The satellites were also equipped with GPS receivers to track their position at all times. From this range of data, variation in the Earth's gravity field can be recovered. Different background models are used to remove the known signal from; ocean tides, polar tides, Earth tides, atmospheric effects, and the Earth's static gravity field. The processing of this Level-1 data (range-data) is done at different institutes. The resulting GRACE data is generally distributed as Stokes coefficients of spherical harmonic function which can be related to variations in equivalent water height at the Earth's surface. The next section discusses the main equations exploited during the conversion of Stokes coefficients into equivalent water heights (Wouters et al., 2014).

The Earth's gravity field is described by the gravitational potential (V). Equation 3.3 shows gravitational potential in terms of Stokes coefficients:

$$V(r,\theta,\lambda) = \frac{GM}{R}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\left(\frac{R}{r}\right)^{l+1}\bar{P}_{lm}(\theta,\lambda)\left(\bar{C}_{lm}\cos m\lambda + \bar{S}_{lm}\sin m\lambda\right), \tag{3.3}$$

where M is the mass of the Earth, r us the radial coordinate of the observation point, θ and λ are the longitude and colatitude of the observation point, R is the semi-major axis of the reference ellipsoid, P are the Normalized associated Legendre functions, l and m the degree and order of the Stokes coefficients and C and S are the coefficients. To convert this to units of equivalent water height, Fourier coefficients of equivalent water height are computed:

$$\begin{aligned}\left(\bar{C}_{lm}\cos m\lambda + \bar{S}_{lm}\sin m\lambda\right)^{(\delta h_w)} &= \frac{\left(\bar{C}_{lm}\cos m\lambda + \bar{S}_{lm}\sin m\lambda\right)^{(\delta s)}}{\rho_w} = \\ &\frac{R(2l+1)}{3\left(1+k'_l\right)}\frac{\rho_{av}}{\rho_w}\left(\bar{C}_{lm}\cos m\lambda + \bar{S}_{lm}\sin m\lambda\right),\end{aligned} \tag{3.4}$$

where k'_l are the load Love numbers, ρ_w the density of water, ρ_{av} the average density of the Earth and $\bar{C}_{lm}$ and $\bar{S}_{lm}$ are the temporal variations of the Stokes coefficients. Then,

$$\delta h_w(\theta,\lambda) = \frac{R\rho_{av}}{3\rho_w}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\bar{P}_{lm}(\theta,\lambda)\frac{2l+1}{1+k'_l}\left(\bar{C}_{lm}\cos m\lambda + \bar{S}_{lm}\sin m\lambda\right). \tag{3.5}$$

Equation 3.5 shows how Stokes coefficients are linked to mass changes in equivalent water height. The spherical harmonic approach is the most common method of obtaining information on the gravity field using GRACE data. This method is the standard method at the beginning of the GRACE era and represents the gravity field as spherical harmonics. Later, another approach was developed where range data is directly related to homogeneous density changes in predefined mass concentration blocks. The mascon approach is used to determine mass anomalies and is not used to estimate the gravity field. This approach, which has some major advantages over the traditional spherical harmonics approach, is discussed below.

3.2 GRACE MASCON METHODS

As briefly explained in the introduction, a mascon (mass concentration block) is a form of mass anomaly basis function which can be derived from GRACE inter-satellite range observations, in combination with other data. So, each block that is defined gets a certain mass concentration and this mass concentration is homogeneous throughout the whole block. This means that a discrete solution is obtained. In the last ten years, the mascon approach caught more momentum, and nowadays it is used by Jet Propulsion Laboratory (NASA), Goddard Space Flight Center (NASA), and CSR (Wouters et al., 2014). There are some advantages when using the mascons approach instead of the spherical harmonic approach. One of them is that the implementation of geophysical constraints is much easier. Furthermore, leakage is reduced from land to ocean by the configuration of mascons (Chen et al., 2016).

3.2.1 *GSFC Mascon solutions*

Luthcke et al. (2013) describes how Level-1 data is processed to get GSFC mascon solutions (Loomis et al., 2019). These mascons are created through SH expansion (Chen et al., 2016). Inter-satellite K-band range-rate (KBRR), orbit location, accelerometer, and attitude data are used to get global estimates of mass anomalies. Different background models are used to model and subtract different processes from the data, these are shown in Table 3.1. The GGM02C model is used to model the static gravity field. The 'GOT4.7' tide model is used to model the ocean tides. Variations in atmospheric masses are modeled using data from the European Centre for Medium-Range Weather Forecasts (ECMWF). The baroclinic Ocean Model for Circulation and Tides (OMCT) is used to model and subtract ocean mass variations. ICE-5G (VM2) is used the model the glacial isostatic adjustment (Chen et al., 2016).

Table 3.1: Background models used in GSFC mascon approach from Chen et al. (2016)

Effect	**Model**
Solid Earth Gravity	GGM02C
Solid tides	IERS2003
Ocean tides	GOT4.7
Solid pole tide	IERS2003
Ocean pole tide	IERS2010
Nontidal atmosphere and ocean dealiasing	ECMWF
GIA	ICE–5G

3.2.2 *JPL Mascon solutions*

JPL mascons are created through analytical partial derivatives (Chen et al., 2016). The detailed data processing steps to obtain mascon solutions are provided by Watkins et al. (2016). Similar to the GSFC mascons, background models are used to compensate for known gravity variations. Table 3.2 provides a list of subtracted effects and which models are used. The mean (static) gravity field is modeled using the GIF48 model. Solid Earth and pole tides are modeled by the IERS2010 nonelastic Earth model. Ocean tides are modeled by the GOT4.7 model. Variations of atmospheric mass are subtracted using

ECMWF data. Effects from other planets, the Sun and the Moon are modeled using the DE-421 model. Finally, the IERS2010 background model is used to mitigate the effects of the ocean pole tides.

Table 3.2: Background models used in JPL mascon approach from Watkins et al. (2016).

Effect	Model
Solid Earth Gravity	GIF48
Solid tides	IERS2010 nonelastic Earth
Ocean tides	GOT4.7
Solid pole tide	IERS2010 nonelastic Earth
Ocean pole tide	IERS2010
Nontidal atmosphere and ocean dealiasing	ECMWF atmosphere and baroclinic OMCT ocean model
Third body perturbations	DE-421

3.2.3 *TU Delft Mascon approach*

Since off-the-shelf GRACE mascon solutions can not be modified in the desired way the method developed at Delft University of Technology is used in the thesis. The mascon approach that was developed at TU Delft and applied to Greenland forms the skeleton of this research (Ran et al., 2018). This method was later successfully applied to Antarctica (Engels et al., 2018). This mascon approach transforms monthly spherical harmonic coefficients (SHC) into mass anomaly estimates. The SHCs are used to get gravity anomalies at satellite altitude, which are then transformed into mass anomalies per mascon at the Earth's surface. Below the exploited functional model, where this method is based, is laid out.

The following steps are derived from Engels et al. (2018). The computations can be split into three steps. The first step deals with the computation of gravity disturbances at a satellite altitude from monthly spherical harmonic coefficients. These disturbances are then linked to mascons at the Earth's surface and finally, mascon parameters are estimated using least-squares techniques.
The gravity disturbance at a specific point (δg_p) and SHCs (ΔC_{lm} and ΔS_{lm}) are related as

$$\delta g_p = \frac{GM}{r_p^2} \sum_{l=1}^{L} \frac{l+1}{1+k_l'} \left(\frac{a}{r_p} \right)^l \sum_{m=0}^{l} \bar{P}_{lm} \times \left(\sin \phi_p \right) \left(\Delta C_{lm} \cos m\lambda_p + \Delta S_{lm} \sin m\lambda_p \right), \quad (3.6)$$

where r (radial distance),ϕ (latitude) and λ (longitude) define the location of the point, L is the maximum spherical harmonic degree, l the spherical harmonic degree, m the corresponding order, GM is the geocentric constant, a is the semi-major axis of the reference ellipsoid, k_l' is the load Love of degree l and $\bar{P}_{lm}$ is the normalized associated Legendre function of a specific degree and order.

For the second part, we want to link those gravity disturbances (at satellite altitude) to mascons at the Earth's surface. Suppose there are N mascons M_i, each with a surface den-

sity of ρ_i. Then, using almost identical quantities as Equation 3.6, the relationship can be expressed as

$$\delta g_p = -\frac{\partial}{\partial r}\left(G\sum_{i=1}^{N}\rho_i\int_{M_i}\frac{\mathrm{d}s}{l_p}\right) = -\frac{\partial}{\partial r}\left(G\sum_{i=1}^{N}\rho_i I_{i,p}\right), \tag{3.7}$$

where G is the universal gravitational constant and

$$I_{i,p} = \int_{M_i}\frac{\mathrm{d}s}{l_p} \tag{3.8}$$

with l_p as the distance between an integration point (at the Earth's surface) and a data point (at a satellite altitude). $I_{i,p}$ needs to be computed using numerical integration, for which purpose a composed Newton-Cotes formula is used. The nodes within a mascon are located on a Fibonacci gird. When K_i is the number of nodes, $I_{i,p}$ can be written as

$$I_{i,p} \approx w_i\sum_{j=1}^{K_i}\frac{1}{l_{ij,p}}, \tag{3.9}$$

where $w_i = S_i/K_i$ with S_i as the surface area of the mascon i. $l_{ij,p}$ is the distance between the node and the data point. This can be calculated as

$$l_{ij,p} = \left(r_{ij}^2 + r_p^2 - 2r_{ij}r_p\cos\psi_{ij,p}\right)^{\frac{1}{2}} \tag{3.10}$$

where

$$\cos\psi_{ij,p} = \sin\phi_p\sin\phi_{ij} + \cos\phi_p\cos\phi_{ij}\cos(\lambda_p - \lambda_{ij}),$$

with $\cos\psi_{ij,p}$ the angle between the node and the data point.
Combining expressions 2.1-2.5 results in the functional model that relates gravity disturbances at satellite altitude to the surface densities of the mascons. This functional model is then written as

$$\delta g_p \approx G\sum_{i=1}^{N}\rho_i\sum_{j=1}^{K_i}w_{ij}\left(r_{ij}^2 + r_p^2 - 2r_{ij}r_p\cos\Psi_{ij,p}\right)^{-\frac{3}{2}}\times(r_{ij} - r_p\cos\psi_{ij.p}). \tag{3.11}$$

This functional model is used together with the 'pseudo observations' at satellite altitude to estimate the unknown densities. These 'pseudo observations' are the gravity disturbances synthesized with Equation 3.6. In matrix-form, this can be written as

$$y = A'x \tag{3.12}$$

where A' is the design matrix, y the vector with gravity disturbances, and x the vector with surface densities. The design matrix is low-pass filtered to get a spectral consistent functional model. Then the best-linear unbiased estimator can be extracted by

$$\hat{x} = (A^T C_d^{-1} A)^{-1} A^T C_d^{-1} y, \tag{3.13}$$

where A' is the low-pass-filtered design matrix and C_d is the data noise covariance matrix, defined as a unit matrix in this study. So this gives the estimator that is used in this research.

3.3 TERRESTRIAL WATER STORAGE VARIATIONS AS ESTIMATED BY GRACE

Water can be stored in many different components of the Earth's system; the atmosphere, biosphere, cryosphere, and lithosphere. Water is in continuous movement through the atmosphere, over the surface, and in the subsurface of the Earth. The total amount of water remains fairly constant, however, the location where this water is stored is highly variable. Hydrological analyses are done to predict where water is at a certain moment and in what form. Almost all hydrological analyses involve the application of a mass balance. Reynolds transport theorem can be used to derive a mass balance for any sort of system (Margulis, 2017). After defining a certain control volume the mass balance can be written as

$$\frac{\delta S}{\delta t} = \Sigma_i I_i - \Sigma_i O_i, \tag{3.14}$$

where $\frac{\delta S}{\delta t}$ is the derivative of the total mass with respect to time, $\Sigma_i I_i$ is the sum of all incoming mass and $\Sigma_i O_i$ is the sum of all outgoing mass. This balance simply states that the difference in storage comes from the difference in in- and outflow of the system.

A mass balance is only relevant when a certain control volume is defined since the total water storage on Earth is constant. Often basin, subbasin, or other watersheds are used as control volumes in hydrological studies since the groundwater flow between them is assumed to be negligible. The main reason that a watershed is a convenient control volume is that it has only one lateral flux, the run-off at the outlet. The other, vertical fluxes, that influence the total water storage are precipitation and evapotranspiration. Considering these fluxes, Equation 3.14 can be rewritten into Equation 3.15, where S is the total amount of water stored in the system, P indicates the total amount of precipitation into the system, E is the amount evapotranspiration out of the system and R is the total run-off of water out of the system.

$$\frac{\delta S}{\delta t} = P - E - R \tag{3.15}$$

Water can be stored in many different ways when considering a river basin. Mostly water is stored on the surface (rivers, lakes, snow accumulation), in the soil (soil moisture), or the deeper ground (groundwater).

One way of detecting variations in terrestrial water storage is by detecting changes in the Earth's gravity field. This is because water is the main cause of temporal changes in the Earth's gravity field on a short timescale, leaving out the effect of earthquakes and assuming the possible GIA effects are accounted for. Most of the other processes, like tectonic plate movement or erosion of mountainous areas, occur over a way longer period (Rodell et al., 2007). NASA's and DLR's joint Gravity And Climate Experiment (GRACE)

satellite mission was launched in 2002 to observe these changes. The spatial resolution of GRACE products is limited by the altitude at which the two satellites orbit.

The changes that the satellites detect are the result of vertical integration of all mass changes, including the Earth's interior, crust, surface, and atmosphere. After applying atmospheric, tide, and glacial isostatic adjustment corrections, the assumption can be made that short-term changes are due to the redistribution of water. Gravimetry can only measure mass changes, from the GRACE estimates it is not clear which component of the hydrological cycle caused the mass change. Only information is retrieved regarding the total mass change within an area. The mass anomalies are found by subtracting the mean of a given period and are often represented in equivalent water height (EWH).

Two different objectives exist when using GRACE for hydrological purposes. One uses Equation 3.15, where the obtained storage variations are compared to the sum of the different processes (precipitation, evaporation, and run-off) influencing the basin. The other technique focuses on the different components that together make up the terrestrial water storage term. The different components are groundwater, soil moisture, snow water equivalent, surface water, and water in the biosphere, as shown in Equation 3.16.

$$\Delta TWS = \Delta GW + \Delta SM + \Delta SWE + \Delta SW + \Delta BIO \tag{3.16}$$

When the second technique is used an assumption is often made that the temporal changes in the surface water and biosphere component are negligibly small. For some warmer areas, the snow component (ΔSWE) is also absent or negligibly small. Independently of the technique used, the GRACE TWSV estimates form a mathematical constraint for both these balances.

As mentioned in the introduction, GRACE has a limited spatial resolution. This results in an amplification of noise as the spatial scale reduces, this requires filtering, which results in signal leakage, when a signal at a specific location is spread over neighboring geographical locations. For large basins, GRACE has proven to give accurate estimates regarding the total storage in the basin (Xavier et al., 2010). Seasonal variations in the stored water mass, as well as inter-annual trends, are visible from GRACE data. For smaller basins, GRACE estimates become less reliable due to the limited spatial resolution of the GRACE products. Over the last decade, the spatial resolution has improved due to advanced processing techniques (Biancamaria et al., 2019; Zaitchik et al., 2008). Biancamaria et al. (2019) even showed that GRACE data can be used in combination with hydrological models to obtain good estimates of TWSV for a $50.000 km^2$ basin. For this small basin, there was a very high correlation between the GRACE estimates and independent model outcomes (Biancamaria et al., 2019).

GRACE data are noisy, this noise increases with higher spherical harmonic degrees or smaller mascon blocks. So, GRACE TWSV estimates for smaller basins are less reliable. Different methods are used during processing to filter out this noise, this always leads to the loss of valuable signal as a side effect.

3.4 THEORETICAL RELATIONSHIP BETWEEN RUN-OFF AND STORAGE

The simple mass balance equation (Equation 3.15) shows that lateral run-off influences the storage. The way the run-off and storage are linked depends partly on the characteristics of precipitation (P) and evapotranspiration (E) for specific basins. Different climates have a different amount of precipitation and evaporation, and different ratios between them, and thus will have a different R-S relationship.

The superposition of coupled and uncoupled storage components plays a huge role in identifying the relationship. For the coupled storage there is a rapid conversion through different storage components, from canopy to soil or from soil to groundwater. Snow accumulation is an uncoupled storage component and an example of a long-term conversion; snow into meltwater. For a simple system, only considering coupled storage components, storage will increase until a certain threshold is reached when no more water can be stored. When there is still an inflow into the system this will then result in an increased run-off, assuming that evaporation stays constant. Once the inflow starts to decrease, this will first be visible as a drop in the run-off. Later the storage will start to decrease too. So, there will be a delay between a change in run-off and the response in storage. For a tropical climate, with rain all year round, there is no storage of water in the form of snow or ice. So finding the relationship for those areas will be easier than finding the relationship for (partly) snow-covered areas.

3.5 EXPERIMENTAL RELATIONSHIP BETWEEN RUN-OFF AND STORAGE AS ESTIMATED BY GRACE

Riegger and Tourian (2014) were one of the first to characterize the relationship between run-off and storage as estimated by GRACE. They compared monthly mass changes and run-off values for different catchments in different climatic zones. Figure 3.1 shows the R-S relationship for tropical, boreal, and seasonally dry basins. For fully humid catchments (Amazon basin), hysteresis reveals a time-invariant temporal delay from storage to runoff, as shown in Figure 3.1. The R-S relationship, for these specific catchments, can be characterized as a Linear Time-Invariant (LTI) system (Riegger and Tourian, 2014). For boreal catchments (Mackenzie basin), MODIS snow data is used to separate coupled and uncoupled components. The coupled parts of these catchments show the same relationship as the tropical catchments. The non-linear part of the R-S relationship in boreal catchments can thus be fully linked to the uncoupled storage components (snow and ice) (Riegger and Tourian, 2014).

Many researchers used their work to further study the relationship between run-off and storage estimates from GRACE data (Sproles et al., 2015; Tourian et al., 2018). Research of Sproles et al. (2015) focused on a watershed of the Columbia River Basin to find the effect of climates, but also topography and geology, on the run-off storage relationship. The results show a hysteresis relationship for all three subbasins of the Columbia River Basin. The size and form of the 'loop' are determined by the climate, topography, and geology. Climate has the main control over the hysteresis loops, determining the hydrologic inputs into the system: more water results in a bigger loop. However, topography influences processes like precipitation and melting. Sproles et al. (2015) conclude: "The climatic, topographic, and geological characteristics of each watershed explain the S–R

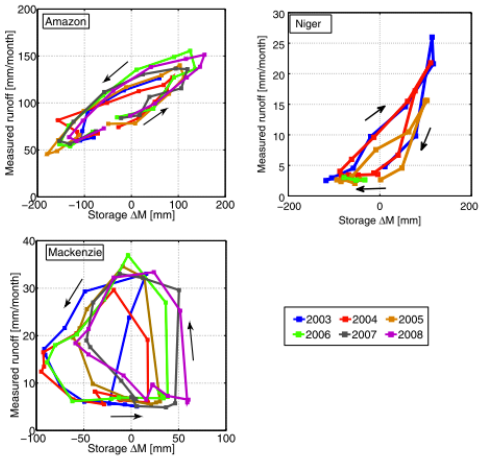


Figure 3.1: Characteristics of R-S relationships displayed for several years for fully humid tropical (Amazon), seasonally dry tropical (Niger), and boreal (Mackenzie) catchments from Riegger and Tourian (2014).

relationship that governs the shape and size of its respective hysteresis curve."

MacEdo et al. (2019) characterized the run-off storage relationship for different subbasins of the Mississippi River. The stream gauges that are used in that research are shown in Figure 3.2.

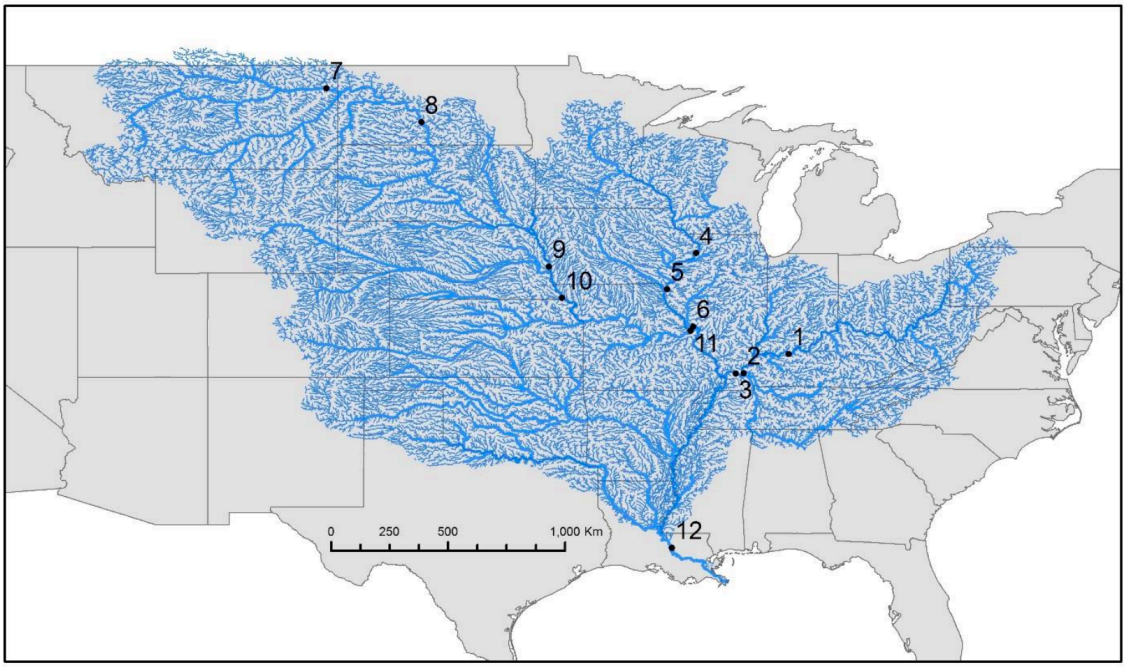


Figure 3.2: Study region with the location of selected USGS streamflow gauges from MacEdo et al. (2019)

GRACE data is applied in a streamflow recession analysis with river discharge measurements across several subbasins of the Mississippi River basin. This is a technique that

uses discharge measurements to estimate other watershed properties. This is especially interesting since most of these sub-basins do not have a tropical or boreal climate. Exponential relationships are developed between monthly, non-winter discharge, and GRACE TWSV estimates for 12 subbasins of the Mississippi for 12.5 years. For a specific outlet, the whole drainage area is used in the analysis. The assumption is made that the baseflow portion of the run-off reflects the differences in storage. To separate base and event flow, a forward-looking low-flow filter is used. Baseflow is the portion of the total flow that is mostly driven by groundwater and event flow is the portion of the total flow that is mostly driven by rainfall events (Margulis, 2017). Finally, exponential relationships are developed between TWSV and monthly mean baseflow. The obtained relationships show a good fit, see Figure 3.3. Blue curves indicate the relationship between total flow and storage and red indicated the relationship between baseflow and storage. For all subbasins, the relationship showed a better fit when using baseflow instead of total flow. R^2 values ranged from 0.46 to 0.92, with a mean of 0.83. The larger subbasins had better fitting relationships since small-scale influences on the water cycle are less dominant. The three subbasins that performed the worst (lowest R^2 values) are all in the Missouri area, which is a cold area. That probably means that even with winter months (December-March) extracted, snow and ice stills play a significant role. That paper indicates that subbasins of the Mississippi basin show a good relationship with run-off. This research does not consider differential run-off for subbasins that are not upstream, so this should be further investigated. Differential run-off the difference between run-off in and out of a specific subbasin, so the effect of the specific subbasins is visible.

Climate, topography, and geology all play a role when it comes to defining the R-S relationship. However, for areas with sufficient data the relation can be derived empirically. Empirical exponential relations will form the basis of this study.

3.6 INDEPENDENT DATA SOURCES FOR VALIDATION

To draw valid conclusions, an independent data set is needed for the results to be validated. Previous studies used different approaches for validating the TWSV estimates from either a model or GRACE data. Below some of these approaches are further described for the Mississippi and European basins.

The North American Land Data Assimilation System (NLDAS) and Global Land Data Assimilation System (GLDAS) collects and reanalyses the best possible observations to support modeling activities. It is used as forcing data for many different land surface models. Spatial and temporal resolution is lower for the global data set, but still sufficient for the model to produce estimates of soil moisture and snow water equivalent that can be used for validation of GRACE estimates. These forcing data form the basis of most land surface models that are used.

Tangdamrongsub et al. (2014) and Li et al. (2012) both assimilated GRACE data into a model for the Rhine basin and Western and Central Europe, respectively. The model was validated by comparing groundwater estimates from the model to in-situ groundwater measurements. Seitz et al. (2008), which considered Central Europe, compared GRACE

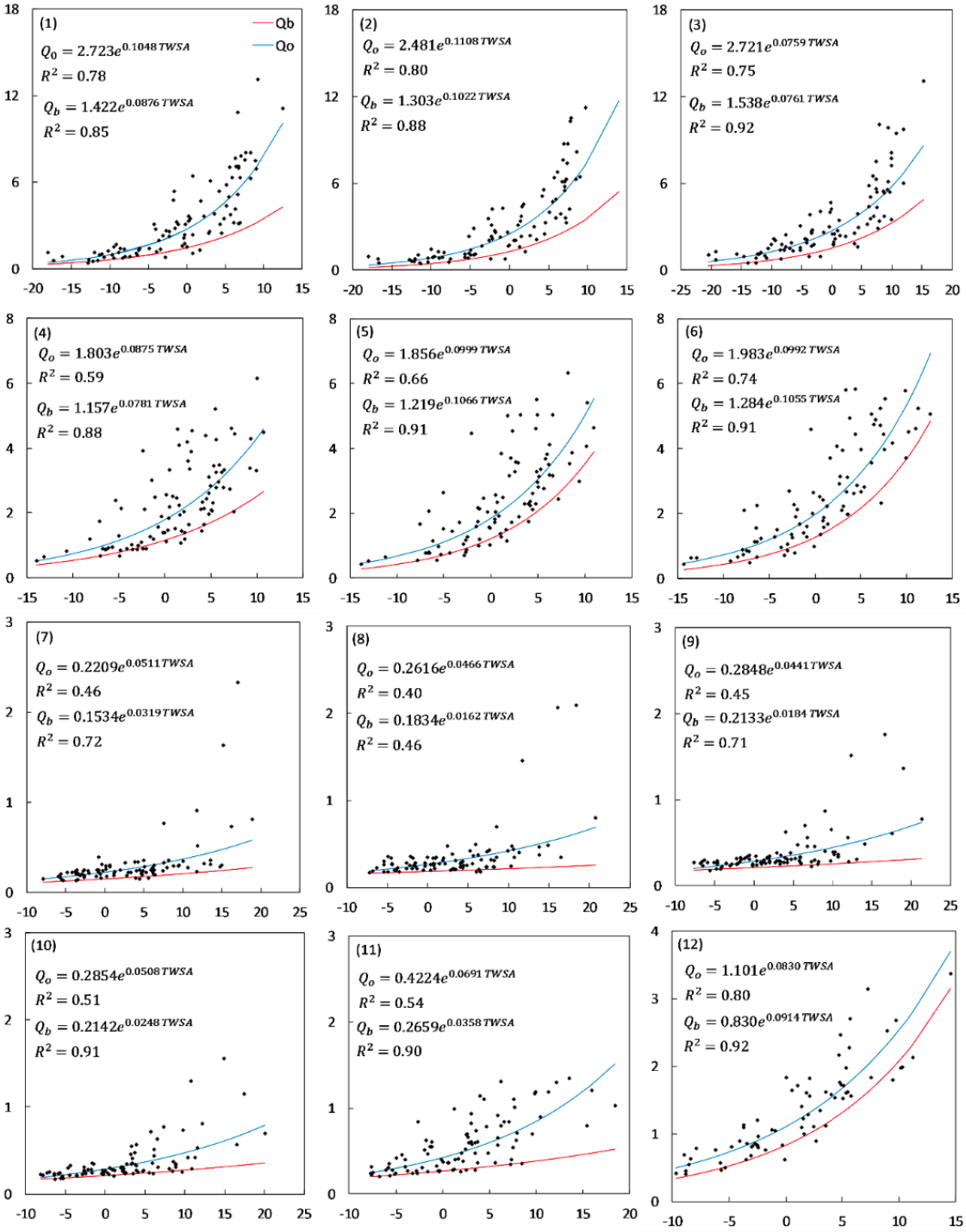


Figure 3.3: Non-winter (April–October) monthly observed discharge (Qo, y axis, in cm) and storage (S, x axis, in cm represented by TWSVs); the lines represent the relationship between observed discharge (blue) or baseflow (red) and storage from MacEdo et al. (2019).

EWHs with water storage variations computed from independent atmospheric and hydrological models. Results can be validated using Equation 3.15, when information on run-off, precipitation, and evapotranspiration is obtained. The same method was used by

Springer et al. (2014) and Springer et al. (2017) where the approach was to look at the closure of the water budget using Equation 3.15.

Zaitchik et al. (2008) validates GRACE TWSV estimates for the Mississippi basin using a network of in-situ groundwater measurements in combination with the Catchment Land Surface Model from NASA. This approach gives a very high similarity between the two independent estimates of TWSV. A more or less similar approach is used by Klosko et al. (2009), where GRACE mascon solutions were compared to a Mississippi basin land surface model (NOAH) and groundwater data from wells. The model soil moisture and snow as parts of TWSV.

Using groundwater measurements and a soil moisture model seems like the best way to validate the GRACE estimates. Soil moisture values from the model can be converted to equivalent water level and combined with groundwater anomalies as done in Cai et al. (2014). For groundwater, in-situ measurements can be used since the water table is less variable in space and time, compared to soil moisture (Gruber et al., 2013; Zaitchik et al., 2008). This in combination with the fact that more groundwater in-situ measurements are present than in-situ soil moisture measurements explains why this part of the validation in-situ measurements can be used. In-situ groundwater measurements are often used for validation.

4 STUDY AREA

This chapter discusses the similarities and differences between the Mississippi Basin and European basins as possible study areas. First, the requirements for a suitable study area for this project are introduced and the reasoning behind these will be explained. This is followed by a section that goes into the character of the Mississippi and European basins and discusses which one is a more suitable study area.

4.1 REQUIREMENTS

The requirements that are discussed can be spilt in the following categories:

- drainage area;
- climate;
- boundary configuration; and
- data availability.

Drainage area is a very important factor in this study. For very large areas GRACE has proven to give very accurate estimates (Zaitchik et al., 2008). So, for very large areas using additional run-off data to regularize the GRACE solution will not have a significant effect, due to the already high accuracy of the estimates. A good consideration needs to be made between having a small enough area to be able to see the effect of regularization and having a large enough area to ensure that GRACE measurements are sensitive to mass variations there.

As described in Chapter 3, the climate of a basin or subbasin (watershed) has a very large influence on the overall water cycle in the area. This means that a climate will also have a huge influence on the terrestrial water storage variations. To investigate what this newly proposed method will result in, a significant TWSV signal needs to be present. A very arid climate may not be suitable, since there will not be large seasonal variations in the terrestrial water storage. Furthermore, the presence of frozen water as snow, ice, or permafrost complicates the relationship between run-off and storage variations (TWSV) (Riegger and Tourian, 2014). Since the main focus of this research is to investigate run-off regularized GRACE solutions, looking at boreal and polar catchments is not advisable.

The GRACE satellites were in a nearly circular polar orbit. This means that they orbited the globe in a North-South direction. The result of this orbit is that the spatial sensitivity in the East-West direction is worse than the sensitivity in the North-South direction (Wouters et al., 2014). So additional data (in this case run-off data) is most useful when this data contains information about mass anomaly variations in the East-West direction. This means that basins at similar longitude are less useful than basins at similar latitude.

A very obvious, but not always fulfilled requirement, is the presence of sufficient run-off data. This run-off data should be available for the same period when the GRACE mission operated (2002-2017). In-situ groundwater measurements in combination with soil moisture and snow water equivalent estimates will be used for validation, due to the high spatial variability. For groundwater, the spatial variability is smaller than for soil moisture, so a less dense network will be sufficient to give an accurate spatial representation. So, the availability and quality of these measurements and models are important.

4.2 MISSISSIPPI BASIN

Figure 4.1 shows the Mississippi Basin. The main rivers that form this basin are the Missouri, Ohio, Arkansas, and Mississippi Rivers.

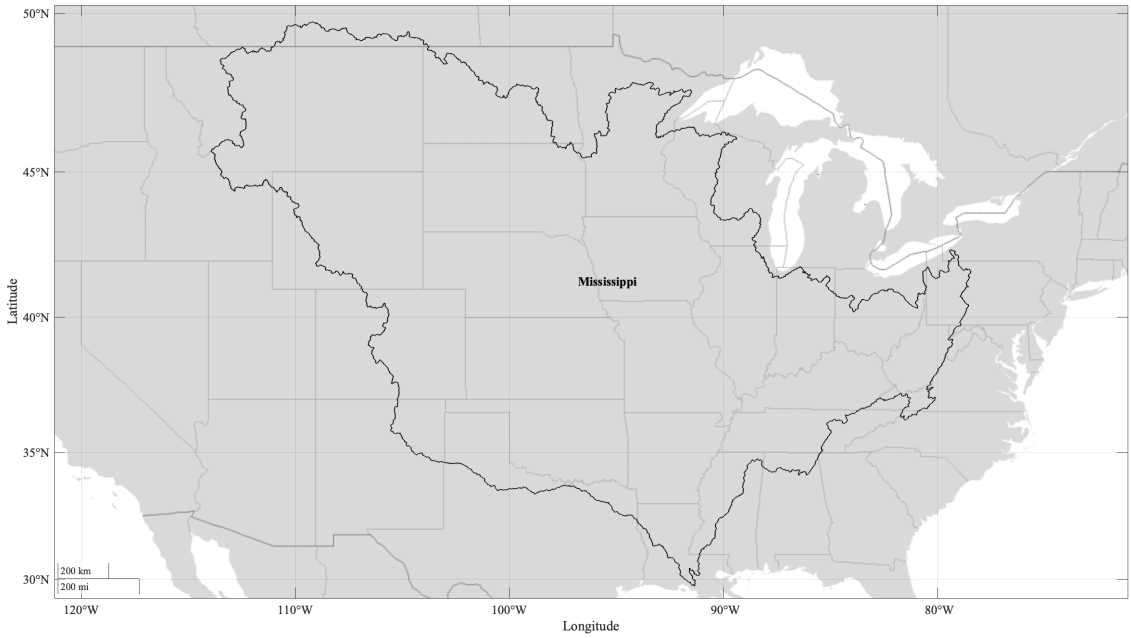


Figure 4.1: Mississippi Basin in the United States of America.

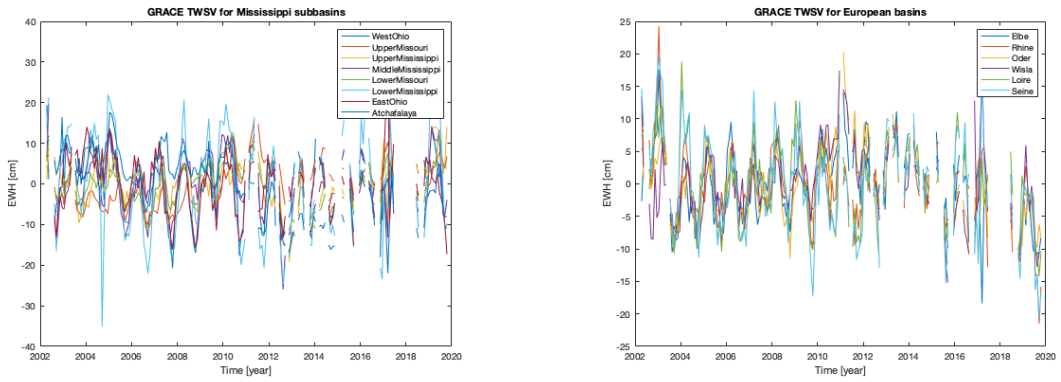
The Mississippi Basin is a huge basin with a size of around 3,000,000 km^2. So, for the total basin GRACE estimates will already be very accurate. A basin can be split into subbasins of arbitrary size, from very small to very large. So taking the Mississippi Basin as the starting point gives the ability to extract subbasins of any desired size, where the effect of regularization is visible.

Table 4.1: Root mean square differences (TWS) for neighboring sub-basin of the Mississippi river.

Boundary	RMSD (2002-2019) [cm]
Upper Missouri-Upper Mississippi	6.42
Lower Missouri-Middle Mississippi	5.46
Middle Mississippi-West Ohio	6.68
West Ohio-East Ohio	6.29
Atchafalaya-Lower Mississippi	10.17

The Mississippi basin encounters multiple climates since the area it covers is very large. This will result in different behavior of TWSV within the different subbasins. This is a favorable situation as described in the previous section. When neighboring (sub)basins experience similar TWSV behavior leakage in and out of specific subbasins cancel each other. So, the results seem very good even though a lot of signal leakage might have occurred. The Mississippi Basin encounters many different climates, the only class of climates that does not occur is polar climates. In the East and South part of the basin, temperate climates are dominant, in the North part continental climates are the most common, and in the West part, dry climates are dominant. In the Northern part snow and ice have a large impact on the overall water cycle. Using the GRACE plotter tool the difference in signal is investigated (The GRACE Plotter, by CNES/GRGS, February 9th, 2021). When signals in neighboring basins are very similar the differences between high and low-resolution estimates may not be very different. Figure 4.2a shows TWSV for the subbasins of the Mississippi river when split up into the 8 subbasins. Table 4.1 indicates the root mean square differences of TWSV estimates (using GRACE and GRACE-FO data) for neighboring subbasins for the period 2002-2019. The neighboring subbasins show a mean RMSD of 7 cm, which indicates a significant difference in terrestrial water storage variations.

The orientation of the boundaries separating subbasins is of course dependent on the configuration of the subbasin. As previously explained (sub)basins that are separated by North-South boundaries are favorable. The choice of the subbasins is quite flexible, however, it is limited by the orientation of the main tributaries of the Mississippi River, the Ohio River, and the Missouri River. Roughly speaking the basin can be split up into three sections; Missouri basin (left), Ohio basin (right), and the Mississippi basin (middle). This way it will be possible the generate subbasins separated by North-South boundaries. These areas will be split further to get areas of the desired size and climates. This is to illustrate that this requirement can be met.



(a) Mississippi basin

(b) European basins

Figure 4.2: TSWV signal of the Mississippi and European (sub)basins.

Run-off data from different outlets of the sub-basins of the Mississippi are available from the United States Geological Survey (USGS). A total of 3264 stations are present in the Mississippi Basin, which has available data for the period 2002-2019. The time interval for this data is 15 minutes. So this will be sufficient for this project.

The United States, together with Europe, has some of the most advanced hydrological models, including soil moisture models, these are further discussed in Chapter 3. The same holds for the groundwater monitoring network, which is also operated by USGS. So, regarding the validation of the outcome of this project, the Mississippi Basin is one of the best-monitored river basins in the world.

4.3 EUROPEAN BASINS

For the European basins, the basins that are considered are the Rhine, Elbe, Oder, Vistula, Loire, and Seine basin. The river basins are shown in Figure 4.3. The size of these basins ranges from 79,000 km^2 (Seine) to 185,000 km^2 (Rhine). This is significantly smaller than the main subbasins of the Mississippi River. As stated above basins can be subdivided into infinitely smaller subbasins, however, the opposite is not the case. These European basins can not be increased in the drainage area.

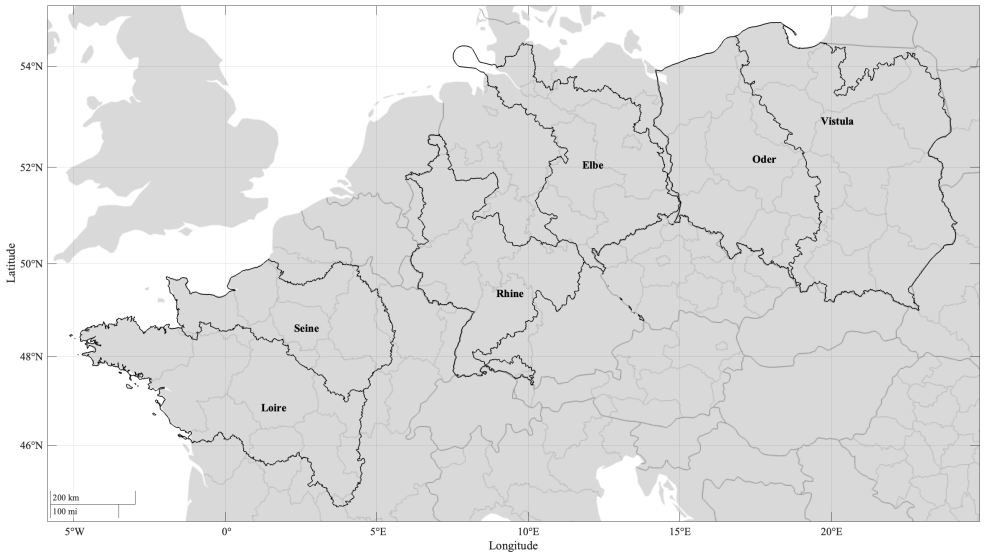


Figure 4.3: River basins in Europe.

The Rhine, Oder, Elbe, Loire, and Seine basin have the same moderate climate without dry seasons and moderate summer. The Vistula basin has a continental climate without dry seasons and warm summer, this difference from the neighboring Elbe basin. So, the behavior of the TWSV in the European basins is more or less similar. Figure 4.2 shows the signals for the European basins. Table 4.2 gives the root mean square differences for neighboring basins. The mean RMSD for neighboring basins is 4 centimeters. The signals show not much variability, which will mean that leakage in and out of the basin will cancel each other.

The basins in Europe are separated by North-South boundaries, this means that the run-off data will give additional information mostly in the East-West direction. This is favorable since the GRACE sensitivity in the East-West direction is worse than the sensitivity in the North-South direction, as mentioned before.

Run-off data from the European rivers can be obtained from Global Runoff Data Base

Table 4.2: Root mean square differences (TWSV) for neighboring European basins.

Boundary	RMSD (2002-2019) [cm]
Loire-Seine	3.37
Seine-Rhine	4.33
Rhine-Elbe	4.87
Elbe-Oder	3.88
Oder-Vistula	3.60

(GRDB). This data has to be requested. All European basins have data available from the period 2002-2019.

Europe has some of the most advanced hydrological models, including soil moisture models, these are further discussed in Chapter 3. However, Europa has no central point for collecting groundwater data like the United States of America. In-situ groundwater data should be collected from different countries independently.

4.4 COMPARISON

Different climates within the Mississippi basin cause different TWSV behavior as indicated by Figure 4.2. The European basins show more similarity in TWSV behavior. So, adding run-off data to regularize the GRACE TWSV estimates may have a bigger effect when applied to the subbasins of the Mississippi River. Since we want to study the effect of using run-off data to regularize GRACE solutions, the Mississippi seems more suitable to do so looking at climates.

Considering the boundaries of the (sub)basins both areas can be used. For the European basins, this is even easier, since the boundaries are mostly North-South already. For the Mississippi basin, a certain configuration of subbasins needs to be made. Creating different pairs of subbasins with North-South boundaries will not be a problem.

As far as the run-off data availability is concerned both the Mississippi basin and European basins show very good coverage. Data from USGS can be directly downloaded, data from GRDC needs to be requested through a few steps. Since USGS data is close to real-time almost all data can be used directly. In the GRDC Station Catalogue, the basic metadata of all stations is listed, including the available data.

Using a combination of in-situ measurements and model data to validate the GRACE estimates is the most suitable option. When snow months for specific areas are not considered, the TWSV signal is the result of two main components, variations in groundwater storage and variations in soil moisture. To pick the most suitable area, it is important to consider the availability of groundwater measurements and soil moisture data for the Mississippi and European basins. Lots of different measurement techniques are used to measure soil moisture, even within the International Soil Moisture Network (ISMN). The coverage of ISMN is different for Europe and the US. There are more in-situ soil moisture measurements present in the United States than in Europe. It will be difficult to extract soil moisture characteristics for European basins, from this sparse network. Since the amount in-situ measurements in the US is larger, models representing soil moisture will be more

reliable, since these measurements are used for calibration. Groundwater measurements will be sufficiently available for both areas. Since there are separate networks in different European countries, it will require additional effort to obtain those data.

Finally, the subbasins of the Mississippi River are more suitable to reach the project's goal, mainly because of the difference in the behavior of TWSV in neighboring (sub)basins. Since the TWSV behavior in Europe is so similar, it will be more difficult to see the effect of regularization using run-off data.

5 DATA AND METHODOLOGY

In this chapter, all data that is needed to regularize GRACE solutions will be discussed. Next, the methodology of this research is described in detail. This part consists of three main parts; finding the relationship between run-off and GRACE estimated terrestrial water storage, the use of the obtained relationship to regularize GRACE estimates, and the validation method.

5.1 DATA

This section deals with all data sets that are used. It indicates what kind of data is used, where it is obtained, and what properties it has. Figure 5.1 shows the configuration of subbasins within the Mississippi basin as used in this research. These subbasins were identified based on size, the orientation of the boundaries and climate, and bounded by drainage divides, this is explained in more detail in Chapter 4. The area size is between 180.000 (Atchafalaya) and 480.000 (Upper Missouri) km^2. Table 5.1 gives an overview of all the data used in this research and its purpose.

Table 5.1: Overview of data used in the study.

Data type	Source	Purpose
Run-off	USGS	Input: R-S relationship
GRACE TWSV estimates (storage)	GSFC	Input: R-S relationship
GRACE TWSV estimates (storage)	JPL	Input: R-S relationship
Soil Moisture	NOAH-LSM (NASA)	Validation
Soil Moisture	VIC-LSM (NASA)	Validation
Soil Moisture	MOSAIC-LSM (NASA)	Validation
Snow Water Equivalent	NOAH-LSM (NASA)	Validation
Snow Water Equivalent	VIC-LSM (NASA)	Validation
Snow Water Equivalent	MOSAIC-LSM (NASA)	Validation
Groundwater	USGS	Validation
GRACE SHCs	ITSG	Input: in-house approach

For these subbasins, two main types of data need to be collected: run-off data and GRACE mascon solutions. The final results will give terrestrial water storage estimates for the Mississippi subbasins. To validate these results independent data is used. Soil moisture and snow water equivalent data are collected from different land surface models and groundwater data are collected from in-situ well measurements performed by United States Geological Survey (USGS). These three variables will be combined into one single time series for each subbasin, which is used to validate GRACE TWSV estimates.

5.1.1 *Run-off*

River run-off data is collected from the USGS, this is a scientific bureau that is part of the United States government. Stream gauges were selected by corresponding drainage area, data availability, and location. Figure 5.1 indicates the selected measurement sites

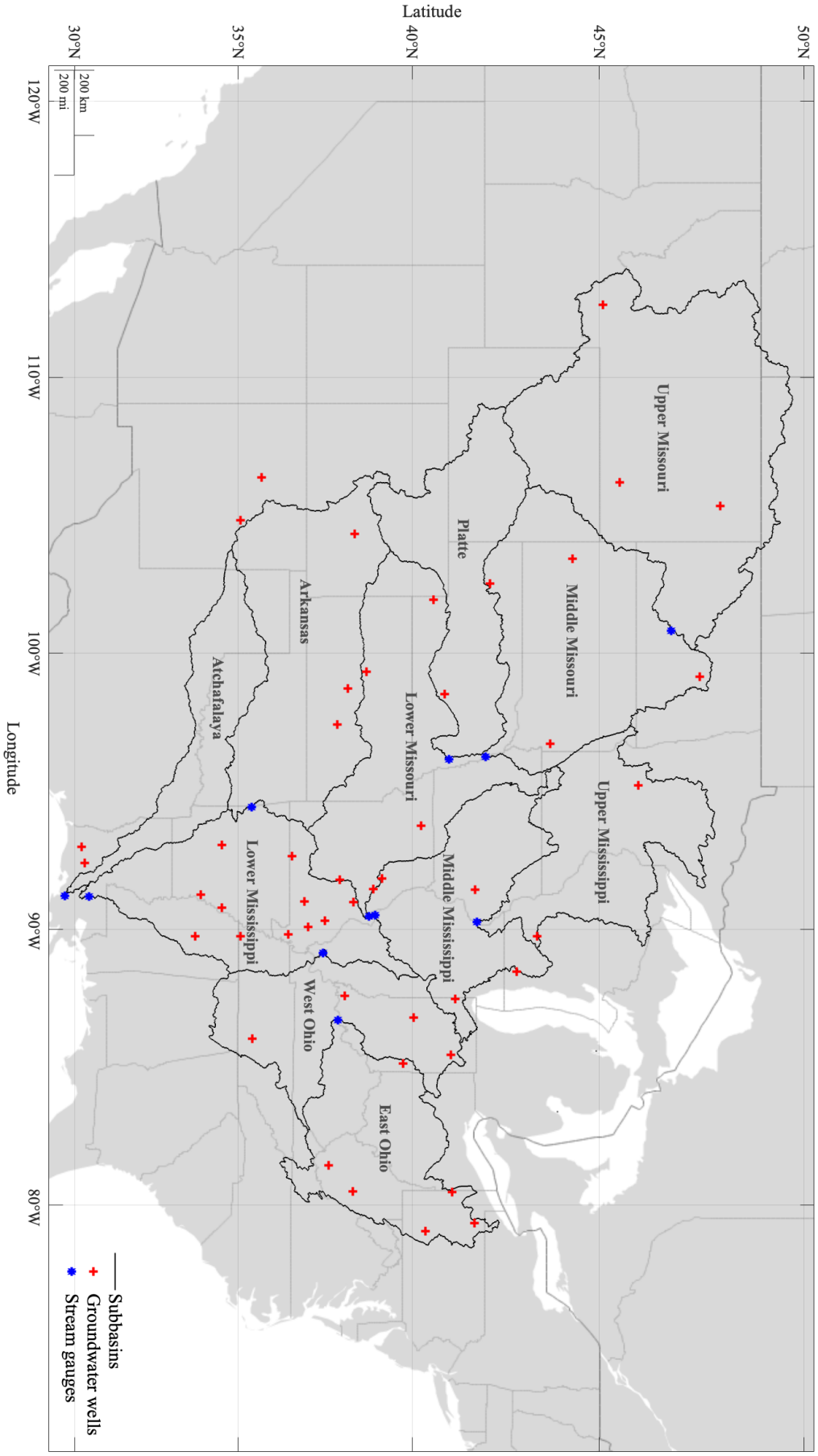


Figure 5.1: The study area (Mississippi Basin) including the configuration of subbasins and run-off data points.

Table 5.2: Metadata stream gauge stations.

Station Number	River	Period of record		Coordinates [deg]		Drainage area [km^2]
		form	until	lat	lon	
7374000	**Mississippi**	Mar/2004	Apr/2020	30.4457	-91.1916	328000
3381700	**Ohio**	Jan/2002	Jun/2020	37.6842	-88.1333	224000
3303280	**Ohio**	Oct/1975	Jan/2020	37.8994	-86.7056	230000
5420500	**Upper Mississippi**	Jun/1973	Mar/2020	41.7806	-90.2519	203000
6935965	**Missouri**	Apr/2000	Jun/2020	38.7889	-90.4707	314000
6805500	**Platte**	Jun/1963	Jun/2020	41.0153	-96.1578	230000
6601200	**Missouri**	Oct/1987	Jun/2020	42.0072	-96.2412	322000
6342500	**Missouri**	Oct/1927	Jun/2020	46.8142	-100.8232	483000
7249455	**Arkansas**	Oct/1997	Jun/2020	35.3917	-94.4322	387000
5587450	**Mississippi**	Apr/1933	Jun/2020	38.9680	-90.5123	220000
7381600	**Atchafalaya**	Sep/1995	Sep/2015	29.6926	-91.2118	182000

of the time series collected. Each stream gauge is the outlet of a specific subbasin. Table 5.2 shows the metadata of the run-off time series that are used; their location, the period, and drainage area. The time series consists of daily averaged discharge in cubic feet per second (ft^3/s) and are verified by USGS.

5.1.2 *GRACE*

To find the preliminary relationship between run-off and GRACE estimated storage, two different ready-to-use GRACE mascon solutions were used. A 'mascon' is defined as an area of constant mass concentration on the Earth's surface, as explained in Chapter 3. In contrast to the traditional spherical harmonics, mascons have a specific geophysical location. This makes it possible to apply prior information (constraints) during data inversion to remove the correlated errors (North-South stripes) from the solution. For spherical harmonics, these constraints can only be applied after inversion. For the spherical harmonics, approach filters need to be applied after a solution of the gravity field is obtained. When these filters are applied after computation a lot of 'real' geophysical signal is lost as well (Girotto and Rodell, 2019). This holds for all types of constraints, however when constraints are applied during data inversion, as is the case for mascons, this is less. Furthermore, mascon solutions have an improved separation of ocean and land signal, which is especially important when looking at terrestrial water storage variations (Girotto and Rodell, 2019). Both the Jet Propulsion Laboratory (JPL) and Goddard Space Flight Center (GSFC) at NASA offer ready-to-use GRACE mascon solutions. Both consist of grids with mass anomalies in equivalent water heights (m) with respect to a long time mean.

The mascon solutions from JPL are derived from Level-1B data (2003/01-2020/04). In the beginning, these solutions were computed from the GRACE data and later on from the GRACE Follow-On data. Represented by 4551 equal-area, 3 arc-degree mascon cells located across the surface of the earth, these solutions capture variations in the gravity field at each localized cell (Colorado Center for Astrodynamics Research, 2020). All reported data are anomalies relative to the 2004.0-2009.999 time-mean baseline.

The mascon solutions from GSFC are derived from GRACE Level-1B data (2003/01-2016/07). This mascon solution is comprised of 41,168 equal-area, 1 arc-degree mascon cells. The

effective resolution is similar to the JPL GRACE mascon solutions ($\pm$300km), the smaller cells with respect to the JPL cells are used to accurately define the constraint regions (Colorado Center for Astrodynamics Research, 2020). All reported data are anomalies relative to the 2004.0-2016.0 time-mean baseline. From these two mascon solutions time-series, the regions of interest (subbasins) can be extracted and coupled to the river run-off of these subbasins.

For the in-house mascon approach at TU Delft sets of spherical coefficients (SHCs) (2003/01-2016/08) from GRAZ University of Technology are used as input. SHCs are level-2 data, which are used to compute the in-house GRACE mascon solution (level-3 data). Chapter 3 discusses in more detail how exactly the in-house mascon solutions are created.

5.1.3 *Soil moisture and snow water equivalent*

To get reliable soil moisture and snow water equivalent estimates different model outcomes are obtained. Three different land surface models are averaged to get a valid estimate of both quantities throughout the Mississippi basin. The models that are used are NOAH, MOSAIC, and VIC, which all use NLDAS as input to model different land and subsurface properties. In Chapter 3 the decision to use these three models for validation was further elaborated upon.
The Mosaic, NOAH, and VIC monthly data sets contain several land surface parameters that are the outcome of each individual land-surface model, which uses NLDAS-2 as input. Each monthly value represents a monthly average of the parameter of interest within a calendar month. All the models are executed over the whole contiguous United States on a 0.125-degree equiangular grid, corresponding to the grid of the input data (NLDAS-2). For the Mosaic data set each grid point can contain up to 10 tiles which consist of three soil layers. The total column depth is 200 centimeters. The NOAH model has four soil layers, with a total column thickness of 200 centimeters. Vic uses three soil layers and the thickness (up to 360 centimeters) of these layers varies spatially. The models differ in the equations and parameterizations that are applied, however, the models compute all most identical parameters (GSFC, 2020).

5.1.4 *Groundwater*

Groundwater measurements are needed for validation as an addition to the soil moisture and snow water equivalent estimates. Groundwater measurements are obtained from USGS National Water Information System. Variation in groundwater level and reservoir characteristics were checked to decide whether a specific well is representative of the groundwater in the surroundings. When a well showed too much variability or when it was located in a small regional aquifer the well was not selected. This is based on earlier work by Li and Rodell (2015), Rodell et al. (2007), and Zaitchik et al. (2008). Each of the wells used was determined to be open to an unconfined aquifer, to be representative of the surrounding groundwater level. Most of the wells in the database did not meet these criteria so that a total of only 42 wells were selected. The locations of the selected wells are shown in Figure 5.1. The obtained time-series consist of daily averaged depths to water level (ft).

5.2 METHODOLOGY

This section outlines the methodology that underlies this research. First, the delineation of the subbasins is discussed. Followed by the methods used to find the relationship between run-off and terrestrial water storage variations. Next, the process of regularizing the GRACE mascon solution is laid out. Finally, the approach for making an independent validation data set is explained.

5.2.1 *Delineation of the subbasins*

The focus of this research is on the subbasin-level since GRACE estimates for larger areas have proven to be accurate enough (Wouters et al., 2014). Masks are needed to extract the correct areas from the JPL and GSFC mascon solutions. To get the correct outlines and masks (polygons and TIFFS) of these subbasins, these subbasins need to be delineated from the Mississippi Basin. The GRASS tool in QGIS is used to get the corresponding subbasin for each stream gauge site. The process consists of 5 main steps. First, a Digital Elevation Model (DEM) needs to be obtained for the study area. Using this DEM the flow direction for each grid cell can be determined. Followed by the calculation of the flow accumulations. Finally, when defining specific outlet points, in this case, the stream gauge sites, the corresponding watershed (subbasin) can be delineated.
These watersheds are converted into TIFF, matrices, and polygons to be used throughout the whole project. The eleven resulting subbasins are shown in Figure 5.1.

5.2.2 *Relationship between run-off and GRACE TWSV estimates for the Mississippi subbasins*

To investigate the specific relationship between run-off variations and mass anomalies for the Mississippi subbasins different data sets are needed. Eleven measurement sites were selected when defining the subbasins, these are the outlets of the watersheds (subbasins). More on the configuration of subbasins and the study area as a whole can be found in Section 5.2.1 and in Chapter 4. Daily averaged discharge (m^3/s) data is obtained for each of the sites from the year 2000 up to 2020. To get a valuable data set, that could show a relation to the TWSV, these data need to be processed. The method used is largely based on earlier research from MacEdo et al. (2019). Monthly means need to be computed, area-weighted averages need to be computed and months influenced by snow need to be mitigated. These time series will be compared to the time series of the GRACE-derived TWSV estimates for the individual subbasins.

Using the masks created in QGIS the different subbasins can be extracted from the 4551 or 41,168 grid cells (depending on the spatial sampling of mascon solution). The grid cells are averaged with respect to the cell size, which is included in the solution, to get one specific value (EWH) for each subbasin for a specific month. Doing this for all the monthly solutions results in a time series of mass anomalies (TWSV) for each subbasin with respect to a predefined mean. The subtracted mean for the JPL solution is converted to the same mean subtracted from the GSFC solution (2004.0-2016.0). This mean is used throughout the research. Apart from this, the method is the same for both mascon solutions. This is the only processing needed for the JPL and GSFC mascon solutions.

For run-off data, daily values of mean discharge per second are converted to mean discharge values in centimeters per month. This is done for each of the 11 subbasins, as:

$$\overline{R}_{month} = \frac{\sum_{d=1}^{d_{max}} \overline{R}_d}{d_{max}} * \frac{d_{max}}{area} * 24 * 60 * 60, \tag{5.1}$$

where $\overline{R}_{month}$ is the mean monthly run-off in $cm/month$, $\overline{R}_d$ the mean daily run-off in m^3/s, d_{max} the total number of days in a specific month and $area$ is the area of the specific subbasins in m^2.

The Middle Mississippi, Middle Missouri, Lower Mississippi, Lower Missouri, and West Ohio subbasin do not only have a river outflow, but also a river inflow. To get valuable information on one specific subbasin only the difference between in- and outflow needs to be determined. This is done by subtracting the total inflow for a certain month from the outflow, which is later divided by the area of the subbasins. As shown by Equation 5.2, where $\overline{R}^{diff}_{month}$ is the mean monthly differential run-off in $m/month$ and R^{tot}_{out} and R^{tot}_{in} the total out- and inflow in m^3 in a calendar month for a specific basin.

$$\overline{R}^{diff}_{month} = \frac{R^{tot}_{out} - R^{tot}_{in}}{area} \tag{5.2}$$

To make a valuable comparison between run-off and storage, a delay needs to be taken into account. An increase in storage will not immediately result in an increased run-off, as explained in Chapter 3. The correlation between the two time series (run-off and TWSV estimates) is used to indicate the delays for the different basins. Ninety-one different run-off time series are computed per stream gauge, from an offset of 0 days (no delay) to an offset of 90 days (three-month delay). The offset with the highest correlation corresponds to the specific delay for that subbasin and that specific time series is used for further processing.

The further processing of the run-off is based on previous work from MacEdo et al. (2019). Their research showed exponential relationships between run-off and total flow (Q_t) as well as baseflow (Q_b). Throughout the Mississippi basin, the relations between run-off and baseflow have higher R^2 values than the relationships between run-off and total flow. However, by MacEdo et al. (2019) GRACE TWSV estimates are used to separate total flow and baseflow. In this research that is not feasible since it will make the data added to the mascon solutions dependent on GRACE data. This is why total flow is used to determine the relation between run-off and storage in this research.

To extract the months influenced by snow and ice a data set is used that indicates the average first day of thaw and frost for a period of 30 years (1975-2005) from the United States Department of Agriculture (USDA). Using this data all subbasins were subdivided into three categories; cold, moderate, and warm. Atchafalaya and Lower Mississippi are considered warm, Upper Mississippi and Upper Missouri are considered cold and the other 7 subbasins are considered to have a moderate climate. For all the warm subbasins no months were excluded from further processing, for the moderate subbasins December, January and February were excluded and for the cold subbasins, October up to March

were excluded.

Previous studies showed the run-off and storage generally show an exponential relationship when looking at non-winter discharge (MacEdo et al., 2019), (Kirchner, 2009). The assumed relationship is as:

$$Q = \alpha e^{\beta S} + c, \tag{5.3}$$

where Q is discharge (run-off) in $cm/month$, α, β and c are constants that need to be determined for each basin and $S = TWSV$ in cm/month. The time series for all subbasins is fitted to this model and constants and the goodness of the fits are extracted, these are further discussed in Chapter 7.

5.2.3 *TU Delft mascon approach*

The mascon approach that is developed at TU Delft and applied to Greenland forms the skeleton of this approach (Ran et al., 2018). This method is later successfully applied to Antarctica, where the so-called *dynamic patch approach* was introduced (Engels et al., 2018). This approach calculates mass anomalies for a large number of patch (mascon) configurations. The final result is an average of the different scenarios. In this research, this mascon approach is adapted for the Mississippi Basin. The exploited functional model used in this approach is described in Chapter 3. Here the adaptations made to the approach are explained. Next, the way the subbasins are divided in patches of smaller sizes is laid out. This is followed by a description of the steps taken to distribute the pseudo-observation points at satellited altitude and the inversion points on the Earth's surface. Next, the method of applying two different types of regularization to these solutions is described. Finally, the way the validation time series is computed is explained.

5.2.3.1 *Adaption of the TU Delft mascon approach*

The TU Delft mascon approach needs to be adapted for this specific research. The approach needs to be tailored for the Mississippi basins. Furthermore, instead of trends, monthly solutions need to be calculated, this requires significant changes in the existing scripts. Finally, the delineation of the basins is not arbitrary in this research, they need to coincide with the subbasins that were defined. Adaption to the scripts was made to make this possible. More in-depth details of the changes made to the existing approach can be found in Appendix A.1.

First a short overview will be given of the steps taken in this approach. First gravity disturbances at satellite altitude are computed from SHCs. These are called pseudo-observations, the number of pseudo-observations depends on the resolution and the buffer zone. In Figure 5.2, a resolution of 1 degree and a buffer zone of 100 km are used. When the buffer zone increases the number of pseudo-observations increases as well.

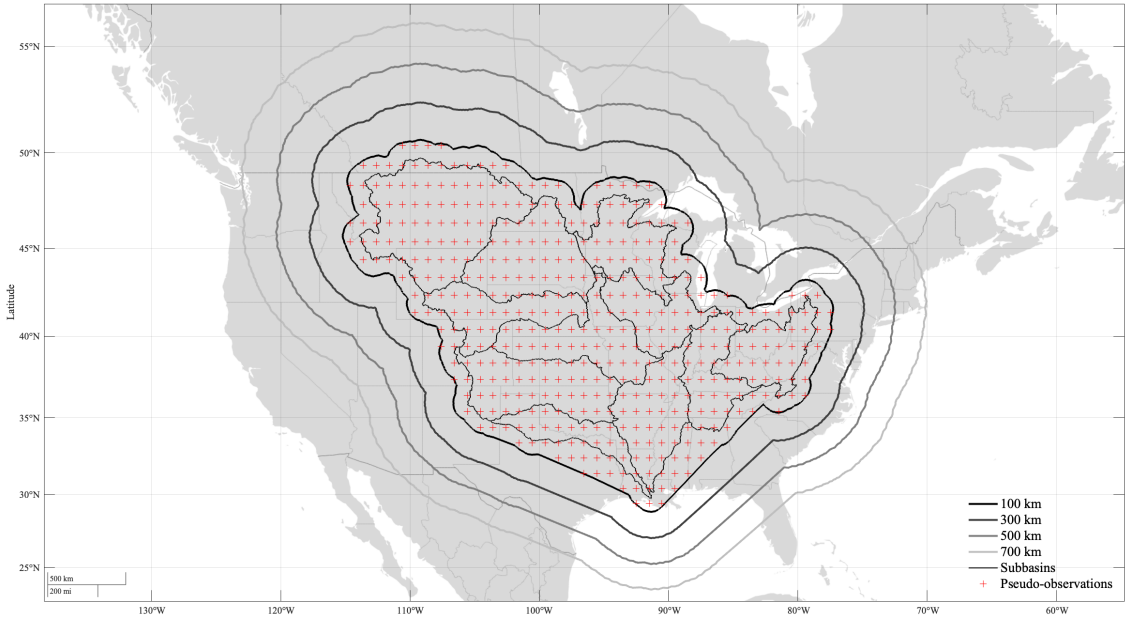


Figure 5.2: Configuration of extra basins around the area of interest in order to mitigate the signal leakage into the Mississippi basin.

Then, there is a grid on the Earth's surface for which solution are computed by inversion. Figure 5.3 shows this grid, zoomed in on all small part of the area to see the distinctive grid points.

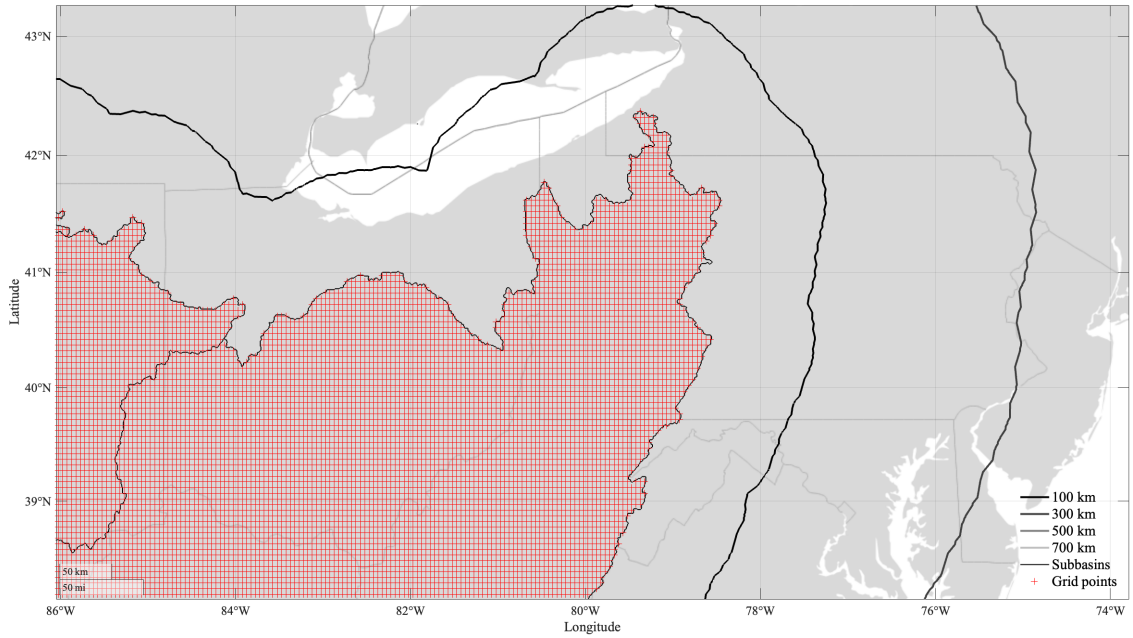


Figure 5.3: Configuration of extra basins around the area of interest in order to mitigate the signal leakage into the Mississippi basin.

Using a Fibonacci grid with a specific spacing (50 km in this case) different patches are created, shown in Figure 5.4. The patches form areas of constant mass change. From the grid points the mass anomalies for all patches can be determined by interpolation.

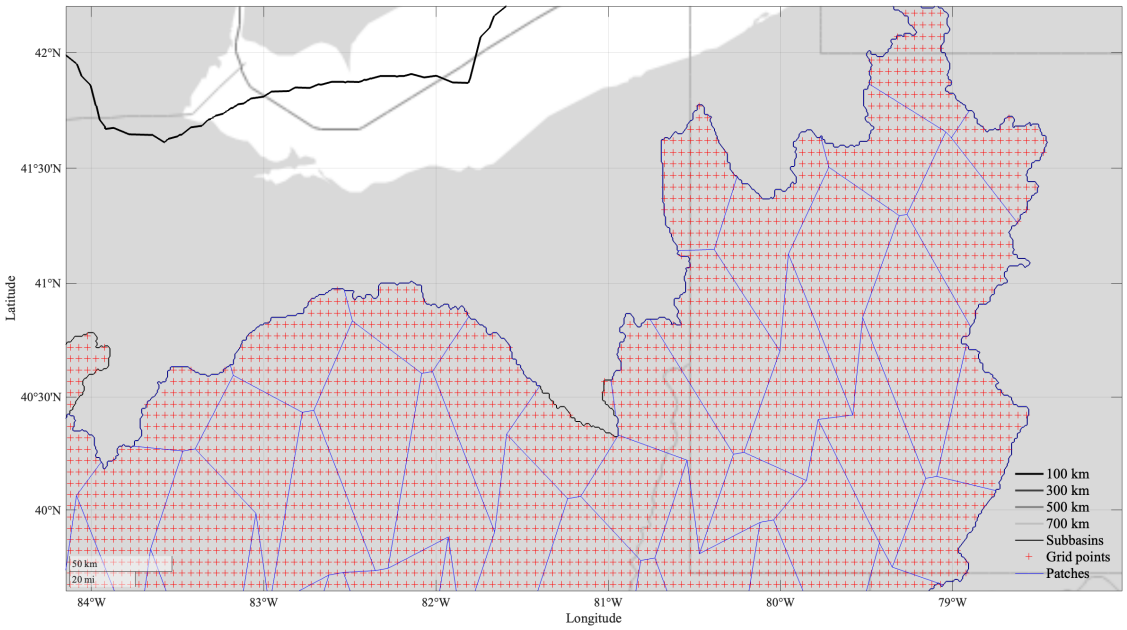


Figure 5.4: Configuration of extra basins around the area of interest in order to mitigate the signal leakage into the Mississippi basin.

5.2.3.2 *Parameterization*

Eleven subbasins are defined and used as input, see Figure 5.1, each of these subbasins is subdivided into multiple smaller patches. These smaller patches are the mascons, they represent an area of uniform mass change. This is done by creating a Fibonacci grid with a specific spacing in and around the subbasins. From these points, the Voronoi diagram is extracted. The optimal setting of this grid spacing is determined using numerical experiments. The results of these numerical experiments are discussed in Chapter 6. The number of mascons in each subbasin is dependent on the Voronoi diagram and thus on the Fibonacci grid spacing.

To limit signal leakage into the subbasins extra mascons are selected to model and mitigate their potential effect. All surrounding areas are covered as shown in Figure 5.5. Five of these mascons cover the Great Lakes since fluctuation in their water level can cause a signal significant enough to cause leakage in the subbasins. The other 25 areas cover all sides of the Mississippi basin. Most of the signal leakage into the Mississippi basin will be accounted for in this way.

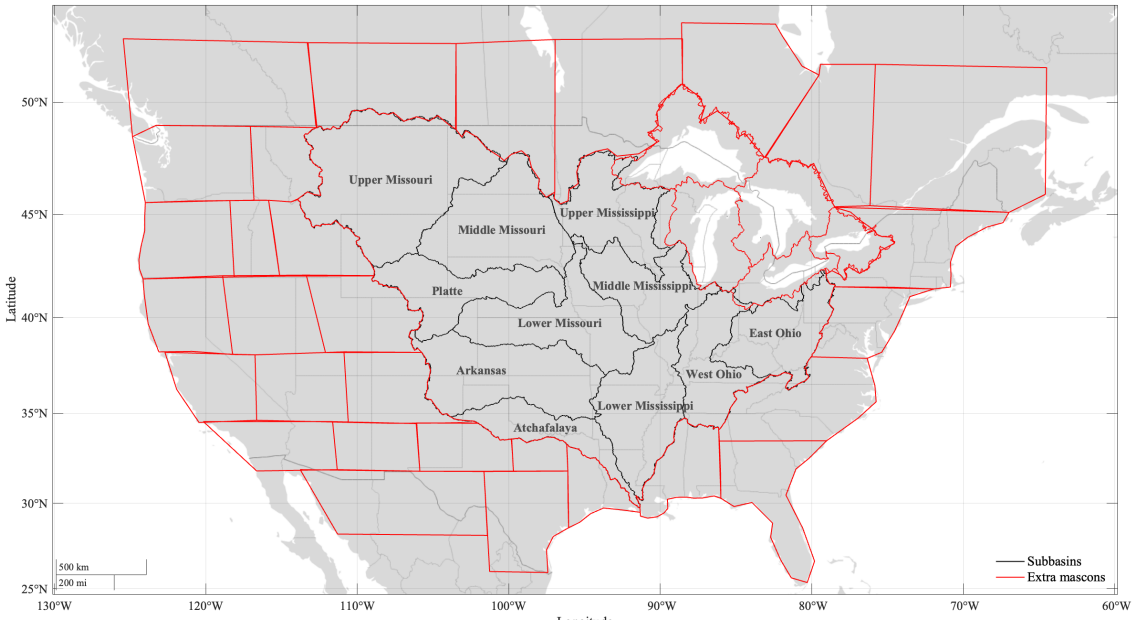


Figure 5.5: Configuration of extra basins around the area of interest in order to mitigate the signal leakage into the Mississippi basin.

5.2.3.3 *Distribution of pseudo-observation points*

This approach computes gravity disturbances at satellite altitude from SHCs. These are the 'pseudo-observation' that are used to compute the mascons at the Earth's surface. The height of the data grid needs to be determined, in a perfect situation, this would correspond to the actual orbital height of the GRACE mission. This height decreased over time from around 500 km to 480 km. Another option is to set a constant satellite altitude. The research found that both methods produce similar results. In this research the suggestion from Baur and Sneeuw (2011) is used: 500 kilometers. A buffer zone of 300 kilometers (see Figure 6.2) is used, a buffer zone is needed since data points at satellite altitude are influenced by a wider area (Baur and Sneeuw, 2011). The size of the buffer zone was determined using numerical experiments, which are discussed in Chapter 6. The pseudo-observation points are computed on a 1-degree 'Gauss-Neumann' (GN) grid over the Mississippi basin including the corresponding buffer zone area (Engels et al., 2018). A GN-grid is a grid for which the latitude circles coincide with zeros of the Legendre polynomial of degree $L + 1$, where L is the maximum degree of SHCs (Sneeuw, 1994).

5.2.4 *Regularization of the TU Delft mascon approach*

Two different types of regularization are applied to the in-house mascon solution. Tikhonov regularization is applied and regularization using run-off data is applied. The two methods of regularization will be described after the general approach is discussed.

As discussed in Chapter 3 the best linear unbiased estimated (BLUE) is used in order to obtain the solution. The BLUE of the mass anomalies is

$$\hat{x} = (\mathbf{A}^T {\mathbf{C_d}}^{-1} \mathbf{A})^{-1} \mathbf{A}^T {\mathbf{C_d}}^{-1} \mathbf{d}, \tag{5.4}$$

where $\mathbf{A}$ is the design matrix, $\mathbf{C_d}$ is the data noise covariance matrix and d is the data vector. The rows of $\mathbf{A}$ are the pseudo-observation points and the columns are the mascons (patches) on the Earth's surface for which the mass anomaly is determined. The data vector ($\mathbf{d}$) holds the gravity disturbances ate the pseudo-observation points. Equation 5.4 gives the best linear unbiased estimate of the inverse problem, which is unregularized. The concept of *statistical regularization* can be used to incorporate knowledge of data noise and the model into the estimation (Turchin et al., 1970). The full derivations can be found in Appendix A.2 these are based on **PDitmar2011**. The estimator that is obtained after applying regularization is:

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{P_d}\mathbf{A} + \alpha\mathbf{P_x}\right)^{-1}\left(\mathbf{A}^T\mathbf{P_d}\mathbf{d} + \alpha\mathbf{P_x}\mathbf{x_0}\right), \tag{5.5}$$

where, P_d and P_x are predefined matrices and α is the regularization parameter which is computed by:

$$\alpha = \frac{\sigma_d^2}{\sigma_x^2}. \tag{5.6}$$

5.2.4.1 *Tikhonov regularization*

Tikhonov regularization is a form of regularization where a different type of prior information on the model outcome is used. Constraints are applied based on the location in the solution space. The most probable model correction ($\mathbf{x_0}$) in Equation 5.13 is equal to:

$$\mathbf{x_0} = \mathbf{m_0} - \mathbf{m_{in}}, \tag{5.7}$$

where $\mathbf{m_0}$ is the priori model and $\mathbf{x_{in}}$ the initial model. In the case of first-order Tikhonov regularization it is assumed that the priori model is equal to the initial model $\mathbf{x_{in}}$. The result of this assumption is that $\mathbf{x_0}$ equals zero. So, Equation A.11 is rewritten into:

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{P_d}\mathbf{A} + \alpha\mathbf{P_x}\right)^{-1}\left(\mathbf{A}^T\mathbf{P_d}\mathbf{d}\right). \tag{5.8}$$

Finally, by setting $\mathbf{P_x}$ equal to $\mathbf{R}$, the so-called regularization matrix, the final formula is obtained.

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{P_d}\mathbf{A} + \alpha R\right)^{-1}\left(\mathbf{A}^T\mathbf{P_d}\mathbf{d}\right). \tag{5.9}$$

The spatial constraints that are applied when using first-order Tikhonov regularization force unknowns that are close to each other in the solution space to exhibit similar behavior. Which holds for this application, it is a reasonable assumption that the TWSV of two areas next to each other is fairly similar. To determine the regularization matrix ($\mathbf{R}$) the procedure from Loomis et al. (2019) is used. Which only implies spatial constraints, no temporal constraints. First a difference operator ($\mathbf{D}$) is created with size $N(N-1)/2\text{x}N$, where N is equal to the number of unknowns, which also includes the additional mascons. This difference operator has the following properties:

$$\mathbf{D}_{ki} = 1, \quad \mathbf{D}_{kj} = -1, \quad \mathbf{D}_{kq} = 0, \tag{5.10}$$

where k is the row that constrains mascons i and j. Then the constrain equations are written as:

$$0 = \mathbf{D}x + e, \tag{5.11}$$

where e is a Gaussian distributed random error with zero mean and covariance matrix Γ. The entries of the covariance matrix are defined by

$$\Gamma_{ij} = \begin{cases} exp(1 - \frac{d_{ij}}{z}) & \text{if } \phi_i = \phi_j \\ 0 & \text{otherwise} \end{cases}$$

where z is the correlation distance, d_{ij} the distance between the points in the centers of the patches and ϕ_i and ϕ_j designate the constraint region for the i^{th} and j^{th} mascons. The constrain regions coincide with the subbasins in this approach to further limit leakage from one subbasin into another. Conservation of mass is ensured by adding an additional constrain to Equation 5.11, this results in:

$$\begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{D} \\ \mathbf{1^T} \end{pmatrix} x + \begin{pmatrix} \mathbf{e} \\ e \end{pmatrix}$$

Then the final $\mathbf{R}$ is defined as:

$$\mathbf{R} = \overline{\mathbf{D}}^T \overline{\Gamma \mathbf{D}}, \tag{5.12}$$

where $\overline{\mathbf{D}}$ is $\mathbf{D}$ with a vector of ones added at the bottom and $\overline{\Gamma}$ is equal to Γ with an addition 10 in the lower right corner to insure the inversion.

5.2.4.2 *Run-off regularization*

Run-off is used as additional information in order to regularize the solutions. A specific exponential relationship between run-off and storage for each subbasin is used. Using these relationships and run-off data, expectations of TWSV van be estimated and added as additional information. In this case the priori model ($\mathbf{m_0}$) is not equal to the initial model ($\mathbf{m_{in}}$), this means that the most probable model correction ($\mathbf{x_0}$) in not zero. The regularized least-squares adjustment formula, stated as:

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T \mathbf{P_d} \mathbf{A} + \alpha \mathbf{P_x}\right)^{-1} \left(\mathbf{A}^T \mathbf{P_d} \mathbf{d} + \alpha \mathbf{P_x} \mathbf{x_0}\right), \tag{5.13}$$

where $\mathbf{P_x}$ is a priori specified weight matrix and x_0 is the expected model outcome, which is in this case derived from the run-off data and the exponential relationship between run-off and storage. $\mathbf{P_x}$ is chosen to be equal to $\mathbf{R}$, which is developed in the previous section. The most probable model correction is given as:

$$\mathbf{x_0^i} = \frac{ln(\frac{Q^i_{tot} - c}{\alpha})}{\beta}, \tag{5.14}$$

where i indicates monthly time intervals and α, β and c are subbasin-specific parameters found in the previous part of the research.

5.2.4.3 *Regularization parameter*

To determine the regularization parameter the variance component estimation (VCE) method is used. A regularization parameter for each month individually is computed first. Later the regularization parameters are averaged and one value is used for all months to prevent fake temporal variations. As described by Odolinski and Teunissen (2019) this method estimates the variance of the data and model components to determine the statistically optimal regularization parameter as:

$$\alpha = \frac{\sigma_d^2}{\sigma_x^2}. \tag{5.15}$$

The approach from Ran et al. (2018) is used to obtain the regularization parameters. The full explanation of the approach can be found in Appendix A.3. The regularization parameter is updated until a certain threshold is reached. When the new regularization parameter differs no more than 1 percent from the previous regularization parameter the process is stopped and this is chosen as the final regularization parameter. When choosing appropriate initial values the regularization parameter converged to non-negative values.

To further improve TWSV estimates, the regularization parameter is also determined differently. Solutions are computed for a wide variety of regularization parameters. From these the RMSE (later called $RMSE_{co}$) remaining after co-estimating the specific yield and regularization bias are computed. Since the specific yield and regularization bias are not well known, co-estimating these makes sure that those factors are eliminated from the remaining errors. The resulting error is a better representation of performance. In this research, TWSV is composed of soil moisture, snow water equivalent, and groundwater variation. Specific yield (see Table 5.3) is used to convert the water table variation in a well to variations in groundwater level, it is a constant factor that links the two. Estimating the two coefficients as:

$$C_1 * GR(t) - C_2 * GW(t) = SM(t) + SWE(t), \tag{5.16}$$

where GR is the GRACE estimated TWSV, GW the groundwater variations, SM the soil moisture variations, SWE the snow water equivalent, and C_1 and C_2 the constants that capture the regularization bias and specific yield error, respectively. Using the least-square adjustment the constants are estimated and can be used to compute the resulting error ($RMSE_{co}$). These constants and errors give a better understanding of the different error sources and the performance of the different solutions. Then, the regularization parameter is chosen to minimize the $RMSE_{co}$.

5.2.5 *Validation*

To validate the result of this new approach an independent data set is needed. Soil moisture and snow water equivalent model data are used in combination with in-situ groundwater measurements, this is based on previous work discussed in Chapter 3. The data that is used for this validation data set is written about at the beginning of this chapter.

Soil moisture and snow water equivalent are extracted from the grid using the masks

of the subbasins. Both variables are given in centimeters water equivalent, so no conversion is needed. Area weighted monthly averages are computed for both variables for all subbasins individually, resulting in two separated monthly time series for each subbasin. Now, the only term that needs to be added is groundwater. Only groundwater fluctuations are of interest, so the long-term mean (2004.0-2016.0) is subtracted for all the wells separately. This coincides with the mean subtracted from the GRACE solutions. In order to convert water level in well to groundwater level in the surround specific yield is used, as is discussed in Chapter 3. Regional averages of specific yield are used, based on previous work from Li et al. (2015). Since there are large scale spatial differences in groundwater levels that are of interest, this assumption yields. The specific yields used are 0.09, 0.14, 0.14, 0.16 and 0.17 for the Ohio region, Missouri region, Lower Mississippi region, Arkansas/Atchafalaya region and Upper Mississippi region, respectively. The number of well and specific yield for each basin is shown in Table

Table 5.3: Specific yield and number of wells for each subbasin from Li et al., 2015.

Subbasin	Number of wells	Specific yield
Arkansas	4	0.16
Atchafalaya	3	0.16
East Ohio	5	0.09
Lower Mississippi	6	0.14
Lower Missouri	4	0.14
Middle Mississippi	4	0.17
Middle Missouri	3	0.14
Platte	3	0.14
Upper Mississippi	3	0.17
Upper Missouri	3	0.14
West Ohio	3	0.09

Using the 42 in-situ groundwater measurement sites groundwater levels are interpolated over the whole basin (same grid as model data) using Triangulation-based linear interpolation. From this point, the same method is applied for extracting soil moisture and snow water equivalent. Masks are used to extract the correct area and then area-weighted averages are applied. Equation 5.17 is used to compute the final validation set.

$$TWSV = SM + SWE + GW \tag{5.17}$$

6 NUMERICAL EXPERIMENTS

Several experiments are carried out to investigate the performance of the approach and to calibrate different parameters, such as the buffer zone, correlation distance, resolution of observations, resolution of grid points, and to motivate certain decisions. Optimal settings need to be determined to fine-tune the in-house approach in the context of the study area. First, the set-up of the experiments is described in Section 6.1. In Section 6.2 the results are discussed and the optimal choices are determined.

6.1 SET-UP OF THE EXPERIMENTS

The validation data set is used as a 'true' signal. The process of creating this validation data set is explained in Chapter 5. In these experiments, Tikhonov regularization is used in combination with an in-house mascon approach. Mascon solutions are compared to this data set and correlation and RMSD are computed and used as indicators of performance. Two types of RMSDs are calculated, the first is the standard RMSD and the second RMSD (later called $RMSD_{co}$) is the error remaining after co-estimating the specific yield and regularization bias. Since the specific yield and regularization bias are not well-known co-estimating these makes sure that those factors are eliminated from as much as possible from the remaining errors. The resulting error is a better representation of performance. The errors reported in this chapter are the mean errors overall 11 subbasins. The parameters and decisions that need to be optimized are the buffer zone, grid spacing, resolution of the pseudo observations, resolution of the grid points, and correlation distance. The 'true' signal is shown in Figure 6.1, which is composed of soil moisture, groundwater, and snow water equivalent variations. Finally, the different parameter selection methods are tested and the optimal configuration is based on the outcome of those experiments.

Table 6.1: Optimal parameter settings.

Parameter	Value
Buffer zone	300 km
Grid spacing	50 km
Resolution data points	1 °
Resolution grid points	0.05 °
Correlation distance	100 km

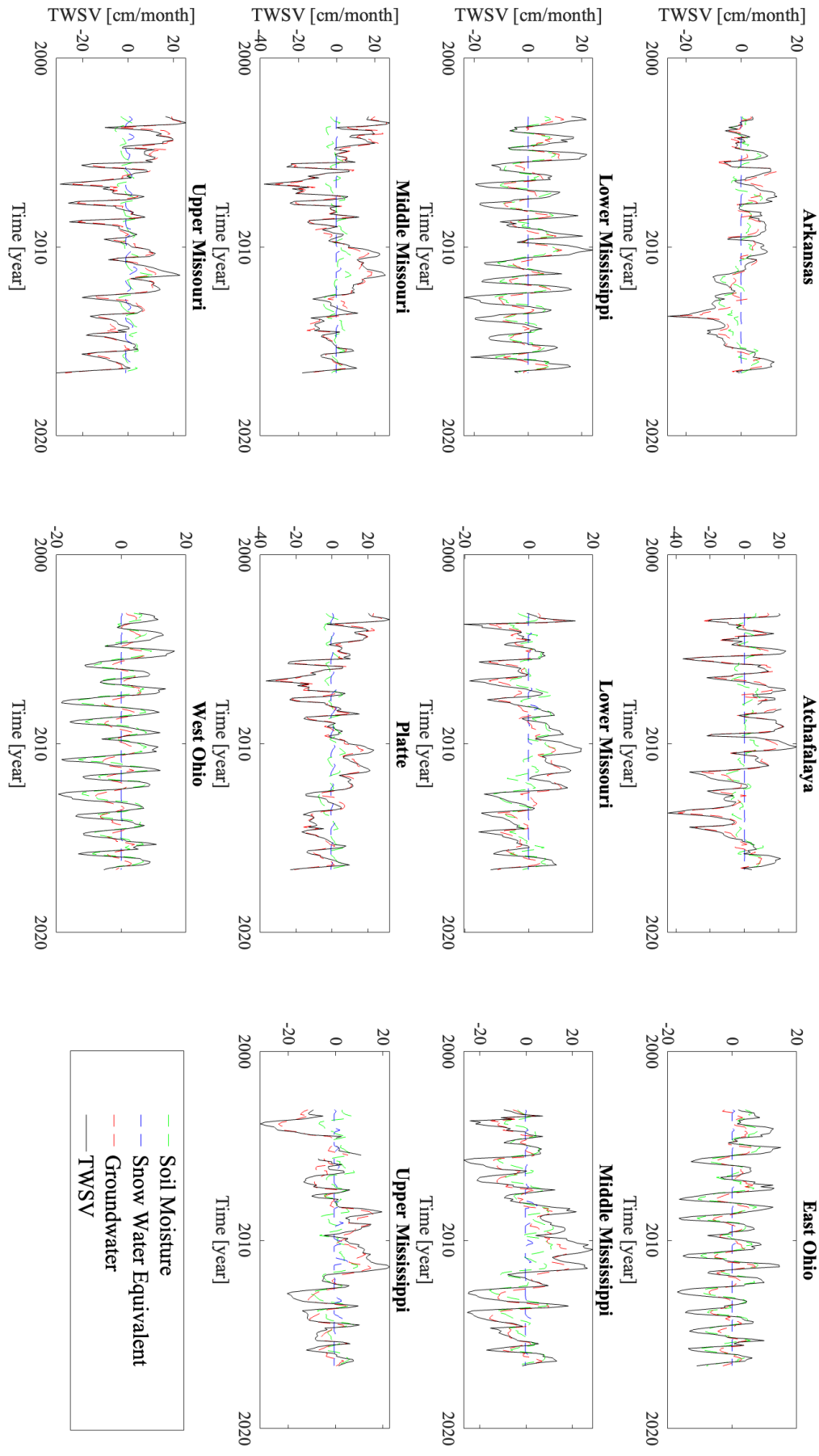


Figure 6.1: The TWSV (black) validation data for all subbasins consisting of snow water equivalent (blue), soil moisture (green) and groundwater (red) variations.

6.2 RESULTS

A buffer zone is needed to capture all signals from a specific basin (Baur and Sneeuw, 2011). The impact of the width of the buffer zone is investigated by considering five different sizes: 0, 100, 300, 500, and 700 km. The buffer zones are shown in Figure 6.2, only four are shown since a 0 km buffer zone is equal to the outline of the subbasins. For all five buffer zones, Tikhonov regularized least-squares solutions are computed. All other parameters are set to the values listed in Table 6.1. The results are shown in Table 6.2. From there it is clear that a 300 kilometers buffer zone gives the smallest errors, for both *RMSD* and $RMSD_{co}$, and the highest correlation to the validation set.

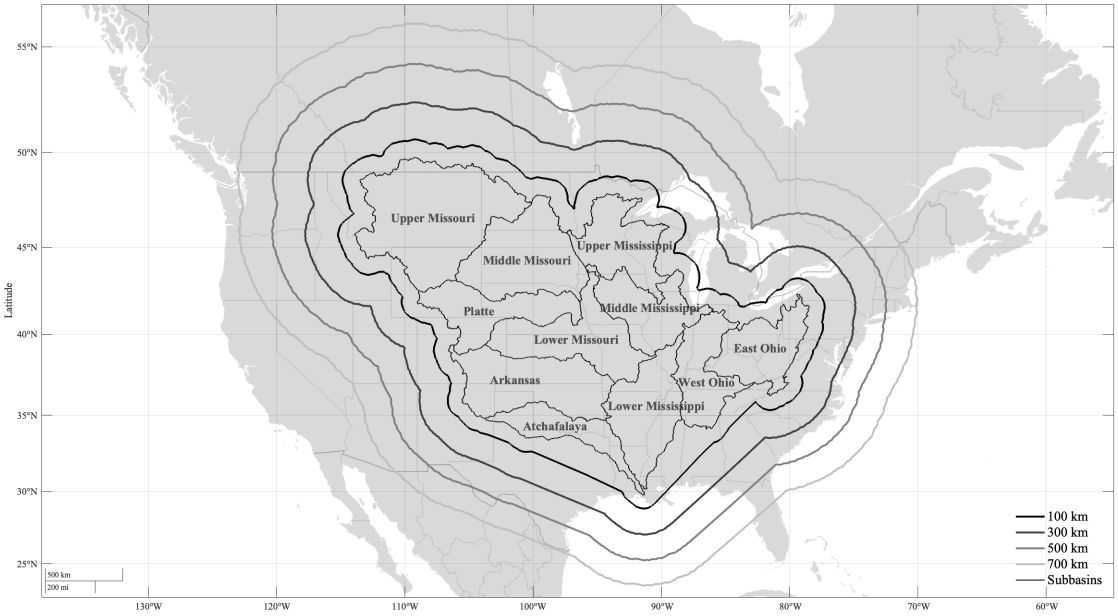


Figure 6.2: The different buffer zones considered in the numerical experiments.

Table 6.2: Metrics of performance for different buffer zones.

Buffer zone [km]	RMSD [cm]	$RMSD_{co}$ [cm]	Correlation [-]
0	7.88	3.95	0.69
100	7.85	3.97	0.70
300	7.70	3.97	0.71
500	9.27	4.04	0.56
700	9.36	4.02	0.54

From a Fibonacci grid the Vonoroi diagram gets extracted, these are the patches on the Earth's surface for which the mass anomalies are computed by inversion of gravity disturbances at the data points at 500-km altitude. Figure 6.3 shows how these points and corresponding patches are defined for different Fibonacci grids for the Arkansas subbasins, as an example. The points are shown in blue and the corresponding Vonoroi decomposition is shown in red. The grid points that are inside a certain patch determine the mass anomaly for that patch. Table 6.3 shows the results, these indicated that a grid spacing of 50 kilometers gives the best results.

Table 6.3: Metrics of performance for different grid spacing.

Resolutions of grid points [degree]	RMSD [cm]	$RMSD_{co}$ [cm]	Correlation [-]
40	11.12	4.03	0.54
50	7.70	3.97	0.71
75	11.12	4.13	0.52

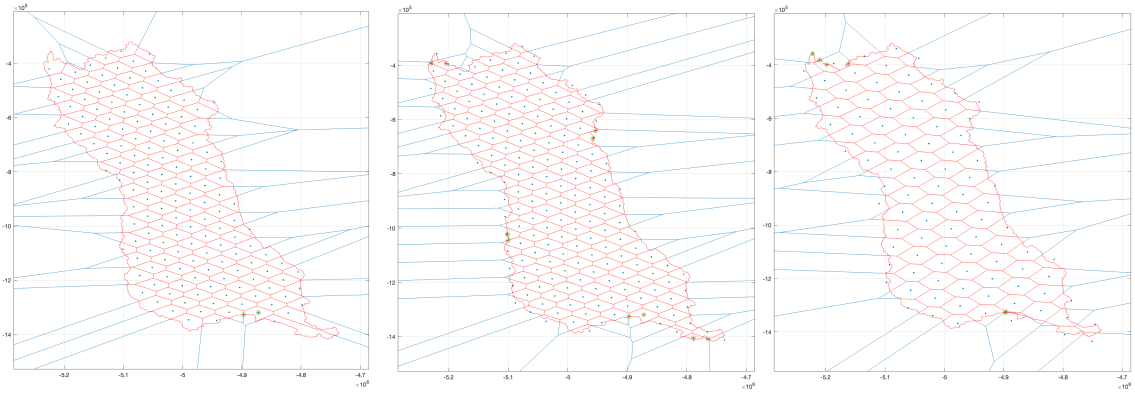


Figure 6.3: Configuration of the Fibonacci grid with 40 (left), (50) middle and 70 (right) kilometer grid spacing (blue) for the Arkansas subbasin, with corresponding Vonoroi decomposition (red).

The resolution of the grid points on the Earth's surface is set to values of 0.1 0.05 and 0.01 degrees. Table 6.4 shows that the different resolutions produce similar errors. A resolution of 0.05 degrees is chosen to limit computational time, this is also the resolution that was used when this method was applied to Greenland and Antarctica (Engels et al., 2018; Ran et al., 2018).

Table 6.4: Metrics of performance for different resolutions of grid points.

Resolutions of grid points [degree]	RMSD [cm]	$RMSD_{co}$ [cm]	Correlation [-]
0.1	7.70	3.97	0.71
0.05	7.70	3.97	0.71
0.01	7.71	3.97	0.71

For the pseudo observations that are generated at satellite altitude, the optimal resolution has to be determined. Resolutions of 1, 0.8, and 0.5 degrees were investigated. Smaller resolutions were not considered due to the huge increase in computation time. Table 6.5 shows the results of these numerical experiments, from this table it can be concluded that there is no significant difference between different resolutions of the pseudo observations. A resolution of 1 degree is chosen, to limit computational time.

Table 6.5: Metrics of performance for different resolutions of pseudo observations.

Resolution of pseudo observations [degree]	RMSD [cm]	$RMSD_{co}$ [cm]	Correlation [-]
1	7.70	3.97	0.71
0.8	7.69	3.98	0.71
0.5	7.69	3.98	0.71

The correlation distance in Equation **??**, is set to 25, 50, 100 and 150 km. The results of these experiments are shown in Table 6.6. From these, 100 kilometers showed the lowest

RMSD and $RMSD_{co}$ values. The correlation is the same for a 100 and 150-kilometer correlation distance and the difference in *RMSD* is not significant. However, 100 kilometers is used to prevent over-regularization. This value corresponds to the value used in research by Loomis et al. (2019).

Table 6.6: Metrics of performance for different correlation distances.

Correlation distance [km]	**RMSD [cm]**	$RMSD_{co}$ **[cm]**	**Correlation [-]**
25	7.73	4.01	0.71
50	7.70	3.98	0.71
100	7.70	3.97	0.71
150	7.72	3.98	0.71

As described in Chapter 5, two different approaches are used to find the optimal regularization parameter. Table 6.7 shows the metrics of performance for the different methods of selecting the regularization parameter. The co-estimation method produced better result, looking at the differences and correlations. So, the co-estimation approach is used to produce further results.

Table 6.7: Metrics of performance for different methods of determining optimal regularization parameter.

Method	**RMSD [cm]**	$RMSD_{co}$ **[cm]**	**Correlation [-]**	λ [-]
Variance Component	8.50	4.27	0.66	$2.87 * 10^{-}21$
Co-estimation	7.70	3.97	0.71	$5.34 * 10^{-}23$

Table 6.1 summarizes the optimal settings resulting from these numerical experiments. These settings result in 459 pseudo-observation over the basin and buffer zone and 1287 unknowns that need to be determined.

7 RESULTS

In this chapter, the results of this study will be presented. First, the run-off-storage relationships will be laid out. This is followed by the results of the Tikhonov and run-off regularized mascon methods and how these are linked to hydrological processes in the subbasins. All outcomes will be presented in this chapter, these will later be used to draw conclusions in Chapter 9.

7.1 RELATIONSHIP BETWEEN RUN-OFF AND GRACE MASS VARIATION ESTIMATES

Monthly differential run-off time series were computed up to a delay of 90 days (3 months) since there is no instant response in run-off when storage changes. Table 7.1 shows the metrics of the relationship between run-off and storage. The optimal delay, in days, between differential run-off and storage changes is shown, these were found by looking at the maximum correlation. Most subbasins show no or a very small delay, which is the result of using monthly mean values. The data points are fitted to Equation 5.3. R^2 gives the ratio between the variance explained by the model and the total variance. This indicates the predictive value of the model, close to 1 means that it has a very good predictive value, and close to 0 means that it has no predictive value. For the subbasins that show no good fit (low R^2 value), the delay is not seen as valuable information, since there is no significant relationship between differential run-off and storage for those basins. The corresponding correlation between the resulting differential run-off time series and storage time series is also shown in Table 7.1, as well as the calendar months that are considered for each subbasin. For the colder regions, only the months from April to October are considered. For the moderate areas, March to November are considered and for the warmest areas, all months are used.

Table 7.1: Metrics of the relationship between differential run-off and storage for all subbasins.

Subbasin	Delay [days]	R^2	Correlation	RMSD [cm/month]	Months	Mean monthly run-off [cm/month]
Arkansas	2	0.33	0.57	0.45	Jan-Dec	0.76
Atchafalaya	2	0.35	0.60	1.80	Mar-Nov	5.07
East Ohio	14	0.75	0.86	1.48	Mar-Nov	4.87
Lower Mississippi	0	0.00	0.17	1.88	Mar-Nov	1.51
Lower Missouri	0	0.69	0.82	0.61	Apr-Oct	1.34
Middle Mississippi	0	0.63	0.78	1.27	Mar-Nov	2.78
Middle Missouri	0	0.04	0.04	0.11	Apr-Oct	0.28
Platte	10	0.45	0.68	0.09	Apr-Oct	0.28
Upper Mississippi	0	0.42	0.65	0.97	Apr-Oct	2.32
Upper Missouri	0	0.27	0.51	0.07	Mar-Nov	0.34
West Ohio	0	0.55	0.74	0.81	Mar-Nov	1.87

The R^2 value ranges from 0.00 for the Lower Mississippi subbasin to 0.75 for the East Ohio subbasin. In general, a higher R^2 means a better fit for the data. It shows that for the subbasins with larger mean monthly run-off the models have a better predictive value. The correlation between R^2 and the mean monthly run-off is 0.5, which indicates a mod-

erate dependence. So, the goodness of the model fit is not only dependent on the mean monthly run-off, other processes like evapotranspiration and precipitation play a role as well. In Chapter 8, precipitation and evapotranspiration will be investigated as possible causes for these moderate R^2 values.

There are two subbasins where differential run-off and storage show no relationship at all, these are the Lower Mississippi and Middle Missouri subbasins. The processes that are mentioned above can explain part of this bad fit, however, the anthropogenic influences in those subbasins are most likely to blame for the majority of this, this is further elaborated upon in Chapter 8. The other subbasins, excluding the two mentioned above, have an average R^2 of 0.5. These subbasins show a reasonable to good relationship between differential run-off and storage. This means that, for those subbasins, a good estimate of TWSV can be made using only run-off data, especially in the context of this research. This finding is further exploited by regularizing the GRACE mascon solutions using the run-off data.

Figure 7.1 shows the differential run-off data points and the exponential fit to this data. The large scatter of data points that are visible around the fitted model is the reason for the moderate R^2 values. The hydrological processes that play a role and can explain this scatter are evapotranspiration and precipitation, as indicated above. The values of α range from 0.08 to 9.00 cm/month, these values are highly correlated to the mean monthly run-off. The Lower Mississippi and Middle Missouri subbasins are left out of this analysis, due to the lack of a clear relationship for those subbasins. The α values for the subbasins in the Missouri region significantly lower the α values in the Ohio region. Most of the models are almost linear, this is determined by the value of β. The values of β determine the exponential behavior, these range from 0.00 for the Platte subbasin to 0.052 for the Upper Missouri subbasin. This coefficient indicates how the run-off reacts to a change in storage. So, for subbasins with a lower β value (more linear), subbasins storage can increase with only a small increase in run-off. This is the case for subbasins where there are a lot of lakes and reservoirs where this water can be stored before run-off changes (MacEdo et al., 2019). Subbasins with a higher α value will show a more significant change in run-off when storage changes since fewer reservoirs and lakes can dampen this effect.

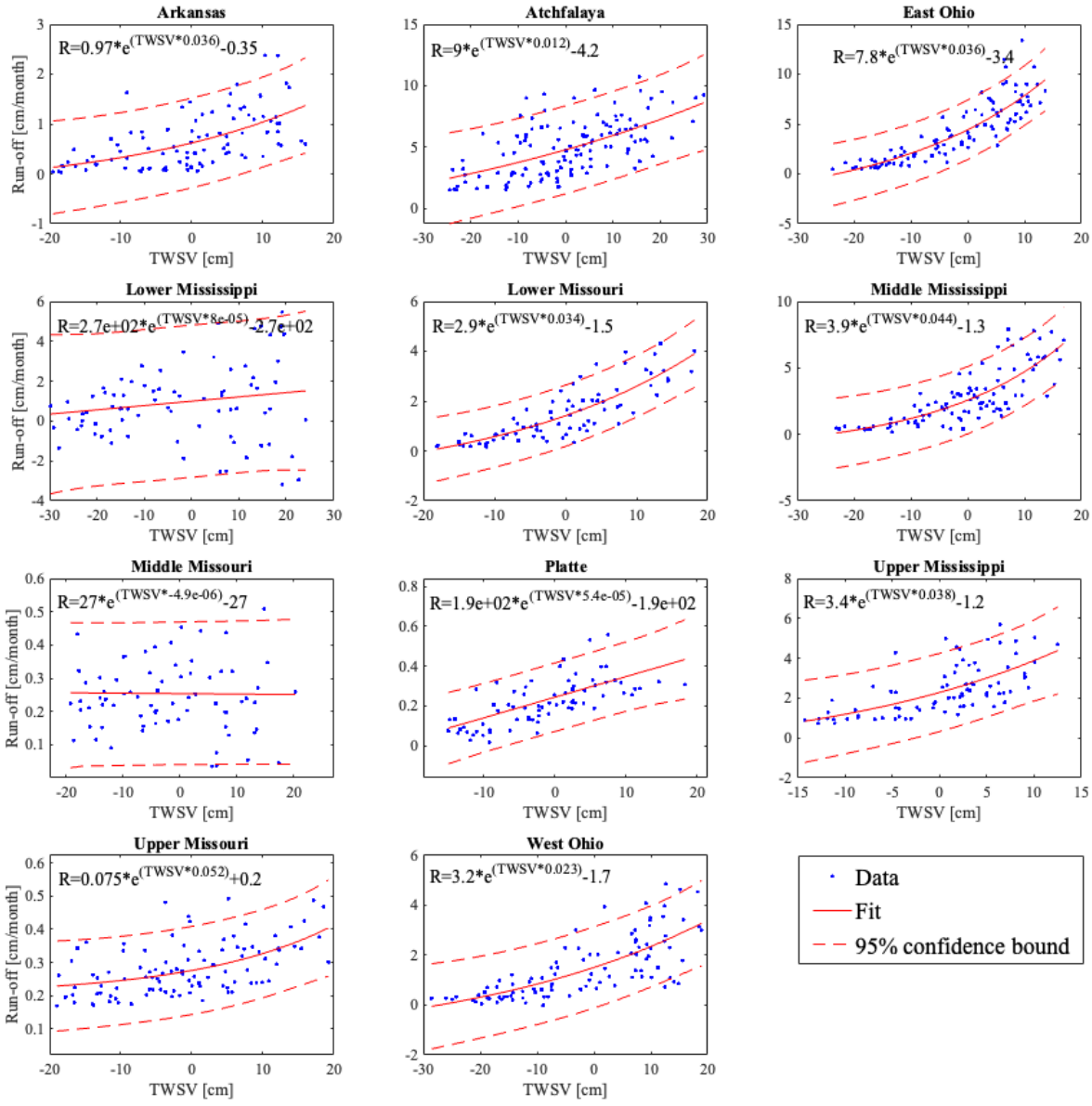


Figure 7.1: Experimental run-off-storage relationships for the Mississippi subbasins.

Figure 7.2 shows the goodness of the fits in a spatial context. No clear geographical dependence is visible from this map. Relating this figure to the climates occurring in the Mississippi Basin gives the insight that the subbasins with colder climates (Upper Missouri, Middle Missouri, and Upper Mississippi subbasin) do not necessarily give relationships with a lower predictive value. Another factor that can have a large influence on the relationship between run-off and storage is human activity. In the Lower Mississippi subbasin, 12% of the land is used for agricultural purposes, this is 8.5% for the Lower Missouri/Platte subbasins. Irrigation uses both surface water and groundwater, which affects the hydraulic cycle. Possible causes for these variations in model fit are further discussed in Chapter 8.

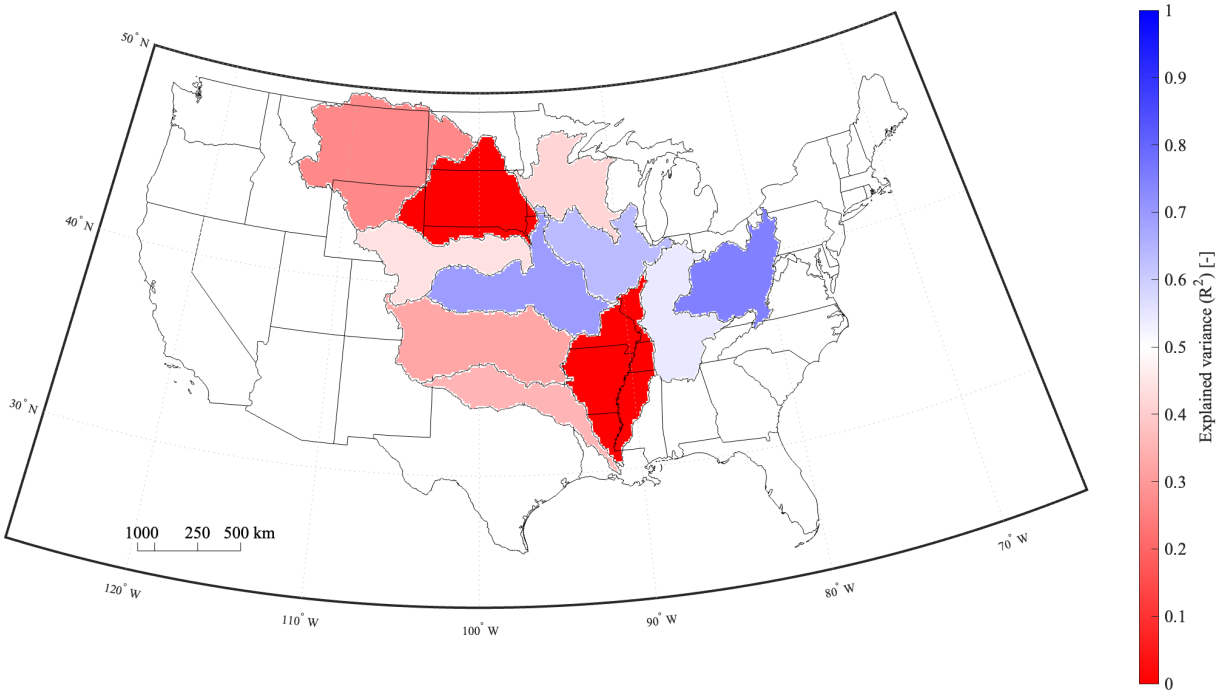


Figure 7.2: Portion of the explained variance by the of the run-off storage fitted relationship.

7.2 IN-HOUSE GRACE MASCON SOLUTIONS

As an initial check in-house mascon solutions are compared to JPL and GSFC mascon solutions, to make sure that the in-house approach does not result in clearly inappropriate estimates of TWSV. Figure 7.3 shows the in-house TWSV estimates, for both Tikhonov and run-off regularized mascons, and the mean of the JPL and GSFC TWSV estimates. Generally speaking, the in-house solutions show good similarity to the JPL/GSFC mascon solution. The *RMSD* between the JPL and the GSFC solution is 2.01 *cm* (correlation 0.96), the difference between the in-house Tikhonov solution and the JPL mascon solution is 3.44 (correlation 0.89) *cm* and the difference between the in-house Tikhonov solution and the GSFC mascon solution is 3.61 *cm* (correlation 0.86). The difference between the in-house run-off solution and the JPL mascon solution is 3.55 (correlation 0.90) *cm* and the difference between the in-house Tikhonov solution and the GSFC mascon solution is 3.74 *cm* (correlation 0.88). On average, the difference between the in-house solution and the ready-to-use solutions is larger than between the JPL and GSFC mascon solutions. The inter-annual variations that are visible in the JPL/GSFC mascon solutions are also visible in the in-house solutions. Looking at annual amplitudes, the peak values are higher for the in-house solutions, in most cases. This can be a sign of an over-regularization ('dampening of physical signal') of the JPL and GSFC mascon solutions. This is not the case for the Lower Mississippi and Middle Missouri subbasins, here the JPL/GSFC mascon solution shows higher peak amplitudes, these are the basins for which no clear relationship between storage and run-off data was found.

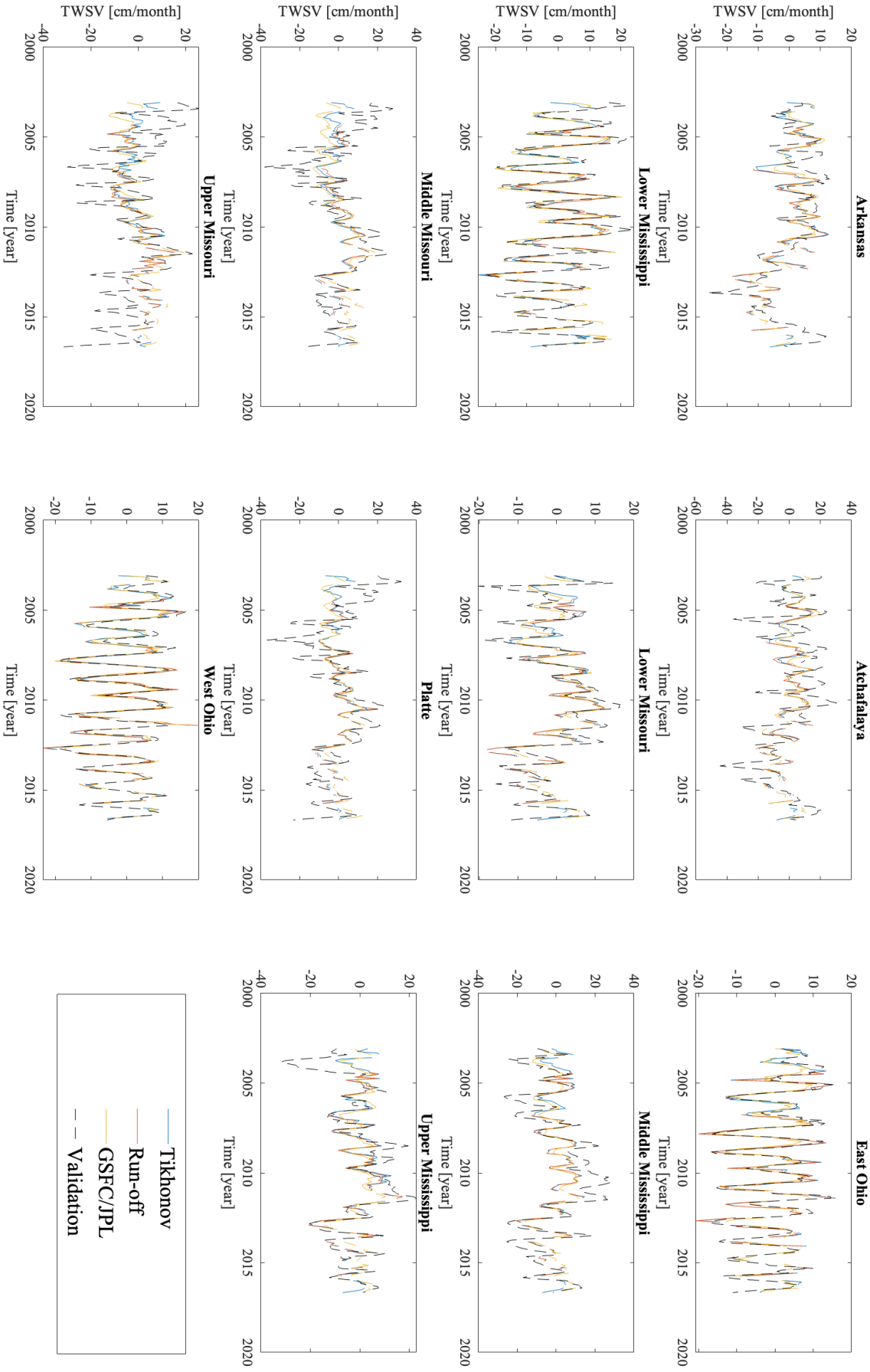


Figure 7.3: The Tikhonov (blue) and run-off (red) regularized mascon solutions and GSFC/JPL mascon solutions compared to the validation data (black).

As described before, this comparison can not be used to identify the solution performance. To do so the TWSV estimates have to be compared to the independent TWSV data shown in Figure 6.1. To get a more valuable comparison the amplitude and phase of the annual variations and the long-term trend are estimated, in addition to a comparison of the full time-series.

Table 7.2: Differences and correlation values for the Tikhonov, run-off and GSFC/JPL mascon solutions.

	Tikhonov			Run-off			GSFC/JPL		
Subbasin	RMSD	Correlation	RMSD_co	RMSD	Correlation	RMSD_co	RMSD	Correlation	RMSD_co
Arkansas	6.40	0.58	2.70	6.33	0.64	2.76	5.33	0.70	2.93
Atchafalaya	13.61	0.48	4.95	14.45	0.46	4.92	12.31	0.60	5.33
East Ohio	3.34	0.92	3.06	3.12	0.94	2.87	3.21	0.94	2.54
Lower Mississippi	4.87	0.92	4.19	4.74	0.94	4.14	4.35	0.94	3.75
Lower Missouri	4.27	0.84	2.84	4.12	0.86	2.69	3.68	0.90	2.55
Middle Mississippi	9.16	0.81	5.60	8.83	0.85	5.46	8.52	0.90	4.82
Middle Missouri	10.26	0.65	4.85	8.93	0.73	4.81	11.38	0.46	4.05
Platte	10.75	0.61	4.99	8.80	0.69	4.81	11.12	0.50	4.12
Upper Mississippi	7.97	0.70	4.47	5.93	0.82	3.95	8.45	0.70	4.07
Upper Missouri	10.97	0.35	3.17	10.25	0.36	2.69	11.22	0.29	1.80
West Ohio	3.13	0.94	2.87	3.01	0.95	2.76	2.98	0.95	2.50

First, the complete time series are compared to the validation data. Table 7.2 shows the $RMSD$, $RMSD_{co}$ and correlation for the in-house mascon solutions and the GSFC/JPL mascon solution wit respect to the validation data. The $RMSD$ and $RMSD_{co}$ values of the in-house solutions are also shown in Figure 7.4. GSFC/JPL mascon solutions have smaller deviations ($RMSD$) and higher correlations for most subbasins. However, for the Upper Missouri, Middle Missouri, Platte, and Upper Mississippi subbasins the in-house solutions have smaller errors and higher correlations. These are the subbasins with the strongest inter-annual variation in TWSV and the lowest mean monthly run-off. Possible causes for this difference between the in-house and ready-to-use solutions are discussed in Chapter 8. When comparing the run-off regularized solutions to the Tikhonov regularized solutions a few interesting results are visible. The run-off methods show significantly smaller RMSDs and higher correlation for the Upper Missouri, Middle Missouri, Platte, and Upper Mississippi subbasins. For the other subbasins, the two methods perform more or less the same, with often slightly smaller deviations for the run-off regularized solution. Only for the Atchafalaya subbasins, the Tikhonov regularized solutions have smaller errors. The same patterns are visible for the correlation between the in-house solutions and the validation data. The mean RMSD for the run-off regularized solutions is 7.13 cm and for the Tikhonov regularized solutions 7.70 cm, so looking on basin-scale the run-off regularized solutions perform better. The mean correlation for the Tikhonov regularized solution is 0.71 and for the run-off regularized solution this is 0.75.

The $RMSD_{co}$ are significantly smaller than the original deviations for most subbasins, which implies that a large portion of the difference could be dedicated to the uncertainty in specific yields and the bias that is induced by regularizing. For the GSFC/JPL solutions, the absolute difference is smaller since the deviations in the first place were smaller. Looking at the Tikhonov and run-off solutions, the difference that was visible for $RMSD$ is mostly gone when $RMSD_{co}$ is considered. Figure 7.4 shows the $RMSD$ and $RMSD_{co}$ values for the two methods for all subbasins. It also shows the values of C_1 and C_2 and

their corresponding 95 percent confidence bounds. As indicated in Chapters 5 and 6, specific yield and regularization bias are two large uncertainties in the validation data. To limit the influence of these two factors in the performance analysis of the different solutions, they are co-estimated using the different components of the TWSV. Figure 7.4 shows that for all subbasins, except the Arkansas subbasin, the 95 percent confidence interval intersects the value of 1. 1 is an important number, since 1 means that there is no bias induced and all values larger than 1 indicated that the solutions are dampened. So, the values of C_1 in Figure 7.4 indicate that the signal in all subbasin could be dampened, except the Arkansas subbasin, where the 95 % interval is below 1. For most subbasins, the run-off regularized solutions are closer to the value of 1, which means less regularization bias in the run-off regularized solutions. This is in line with the earlier finding. The mean value of C_1 is 1.11 for the Tikhonov regularized solutions and 1.05 for the run-off regularized solutions. So, overall the run-off regularized solutions contain less regularization bias. Looking at the confidence intervals of the C_2 estimates shows that most of these estimates are not significant. However, this does not influence the earlier findings regarding C_1. Also, the negative C_2 estimates for the Arkansas and Atchafalaya subbasins are not physically possible. For all subbasins, the run-off regularized solutions show values closer to 1. This, in this case, means that the 'fitted' specific yield is closer to the assumed specific yield from literature. The mean values of C_2, are 0.32 and 0.36 for the Tikhonov and run-off estimates, respectively.

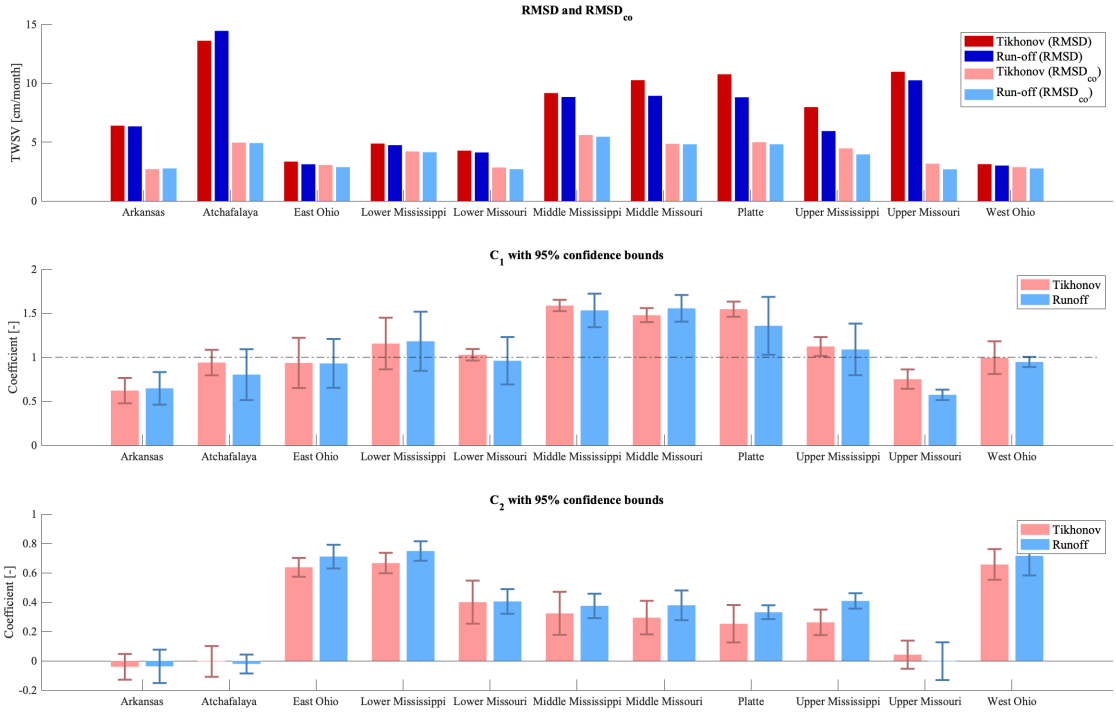


Figure 7.4: The *RMSD* and $RMSD_{co}$ for the Tikhonov (red) and run-off (blue) regularized solutions with respect to the validation data. Also, the values of C_1 and C_2 are shown per subbasin with corresponding confidence intervals.

Looking at the $RMSD_{co}$, the performance of the two different regularization techniques does not differ significantly. This means that most of the difference between the two methods could be explained by the regularization bias and uncertainty in a specific field.

To find out which of those two factors is dominant the regularization bias and specific yield were estimated separately. Figure 7.5 shows the results when only estimating the regularization bias. It is clear that the differences between the *RMSD* and $RMSD_{co}$ are small, this indicates that the regularization bias is small.

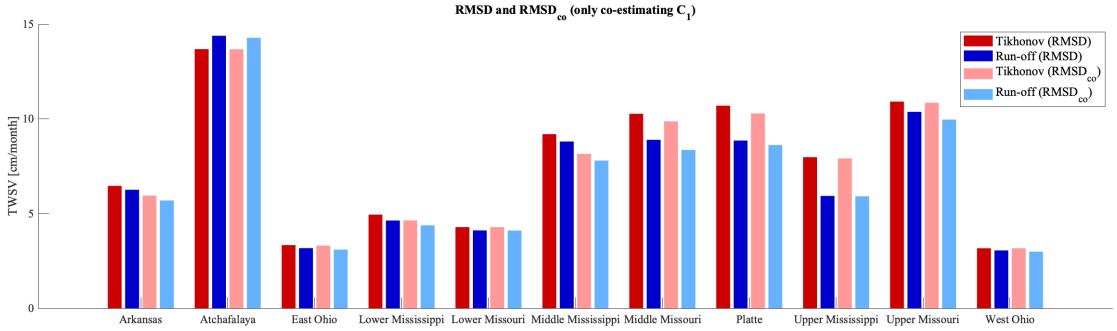


Figure 7.5: The *RMSD* and $RMSD_{co}$ for the Tikhonov (red) and run-off (blue) regularized solutions with respect to the validation data. $RMSD_{co}$ is determined after co-estimating the regularization bias (C_1) separately.

Figure 7.6 shows the results when only estimating the specific yield. Here a clear drop in deviation is visible between *RMSD* and $RMSD_{co}$. This indicates that most of the error can be accounted to the specific yield used in this research, or the the uncertainty of the groundwater time-series in general.

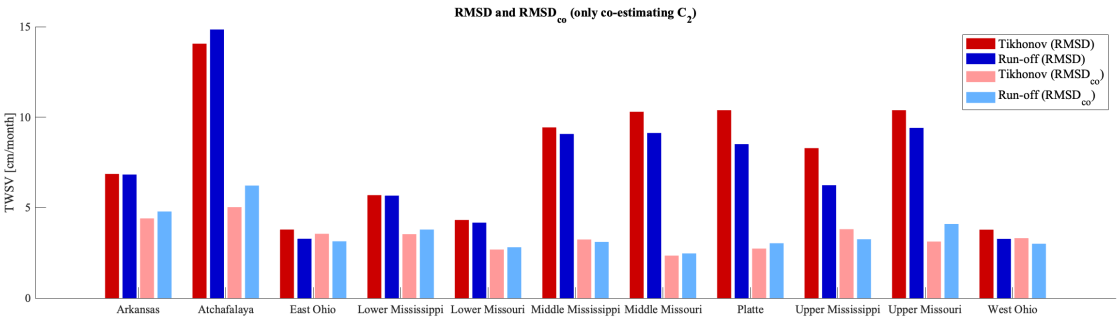


Figure 7.6: The *RMSD* and $RMSD_{co}$ for the Tikhonov (red) and run-off (blue) regularized solutions with respect to the validation data. $RMSD_{co}$ is determined after co-estimating the specific yield (C_2) separately.

Run-off regularized solutions show less regularization bias than the Tikhonov regularized solutions, although this is not the main element in the error budget of the solutions. A large portion of the error is accounted for by the co-estimation of the specific yield, however, the relatively low values (around 0.3) and uncertainty make it hard to draw any conclusions from this. This is discussed further in Chapter 8. A large difference could come from the better estimation of annual peak amplitude, which is looked into next.

The annual amplitude and long term trend are estimated by fitting the mascon TWSV estimates to:

$$TWSV = A + Bt + C * sin(\omega t + \phi), \tag{7.1}$$

where A is the offset, B is the long-term trend, C the amplitude and ϕ the phase and ω the frequency of the annual variations which is determined by:

$$\omega = \frac{2\pi}{T},$$

where T is 1 year. Table 7.3 gives the results for the GSFC/JPL, in-house mascon solutions and the validation data set. The seasonal amplitudes are generally speaking higher for the in-house TWSV estimates than for the GSFC/JPL TWSV estimates. The only subbasins for which this is not the case are the Lower Mississippi and West Ohio subbasins. Validation data capture stronger inter-annual variations (trends), as shown in Figure 7.3, this results in underestimated annual amplitudes, since this model can only estimate a single value over a time period. So, large negative and positive peaks will cancel each other. The fit can only explain linear trends, when there is a trend on a smaller time scale these are not captured. Visible in Figure 6.1 is that these short-term trends are stronger in the validation data.

Table 7.3: Comparison of annual variations and linear trends for mass variation estimates for the Tikhonov, run-off, GSFC/JPL mascon solutions and the validation data.

	Annual		
Subbasin	Amplitude [cm]	Phase [months]	**Trend [cm/year]**
Arkansas: *GSFC/JPL*	2.11	-0.46	-1.04
Arkansas: *Tikhonov*	3.45	-0.47	-0.69
Arkansas: *Run-off*	4.85	-0.40	-1.35
Arkansas: *Validation*	1.53	0.58	-0.63
Atchafalaya: *GSFC/JPL*	3.50	-0.23	-1.09
Atchafalaya: *Tikhonov*	4.24	-0.41	-0.81
Atchafalaya: *Run-off*	6.90	-0.25	-1.45
Atchafalaya: *Validation*	7.21	0.59	-0.61
East Ohio: *GSFC/JPL*L	4.69	-0.08	-0.18
East Ohio: *Tikhonov*	5.35	-0.31	-0.63
East Ohio: *Run-off*	7.35	-0.15	-1.11
East Ohio: *Validation*	4.94	-0.04	-0.64
Lower Mississippi: *GSFC/JPL*	6.80	-0.23	0.04
Lower Mississippi: *Tikhonov*	5.18	-0.36	-0.67
Lower Mississippi: *Run-off*	6.52	-0.20	-1.27
Lower Mississippi: *Validation*	7.45	-0.13	-0.89
Lower Missouri: *GSFC/JPL*	2.62	-0.63	0.24
Lower Missouri: *Tikhonov*	3.65	-0.47	-0.51
Lower Missouri: *Run-off*	4.92	-0.37	-1.03
Lower Missouri: *Validation*	2.20	-0.36	-0.29
Middle Mississippi: *GSFC/JPL*	3.37	-0.65	0.50
Middle Mississippi: *Tikhonov*	4.25	-0.41	-0.45
Middle Mississippi: *Run-off*	5.93	-0.32	-1.00
Middle Mississippi: *Validation*	5.09	-0.71	0.16
Middle Missouri: *GSFC/JPL*	2.73		1.41
Middle Missouri: *Tikhonov*	2.88	-0.69	0.03
Middle Missouri: *Run-off*	3.29	-0.43	-0.26
Middle Missouri: *Validation*	3.60	-0.47	-0.59
Platte: *GSFC/JPL*	2.79	-0.77	0.69
Platte: *Tikhonov*	2.82	-0.66	-0.17
Platte: *Run-off*	3.90	-0.56	-0.47
Platte: *Validation*	3.80	-0.25	-1.23
Upper Mississippi: *GSFC/JPL*	2.72	-0.76	0.06
Upper Mississippi: *Tikhonov*	3.56	-0.51	-0.27
Upper Mississippi: *Run-off*	4.77	-0.40	-0.69
Upper Mississippi: *Validation*	3.08	-0.78	0.37
Upper Missouri: *GSFC/JPL*	2.89	-7.11	1.22
Upper Missouri: *Tikhonov*	2.96	-0.81	0.37
Upper Missouri: *Run-off*	3.13	-0.58	0.63
Upper Missouri: *Validation*	3.96	0.03	-1.13
West Ohio: *GSFC/JPL*	6.30	-0.20	-0.07
West Ohio: *Tikhonov*	5.20	-0.33	-0.56
West Ohio: *Run-off*	6.13	-0.16	-1.05
West Ohio: *Validation*	5.51	-0.20	-0.68

Run-off regularized solutions show higher annual amplitudes compared to Tikhonov regularized solutions, this is the case for all subbasins. Since the in-house solutions capture the same inter-annual variations this is an important finding. The phase and trend values do not indicate any other significant differences between the two solutions.

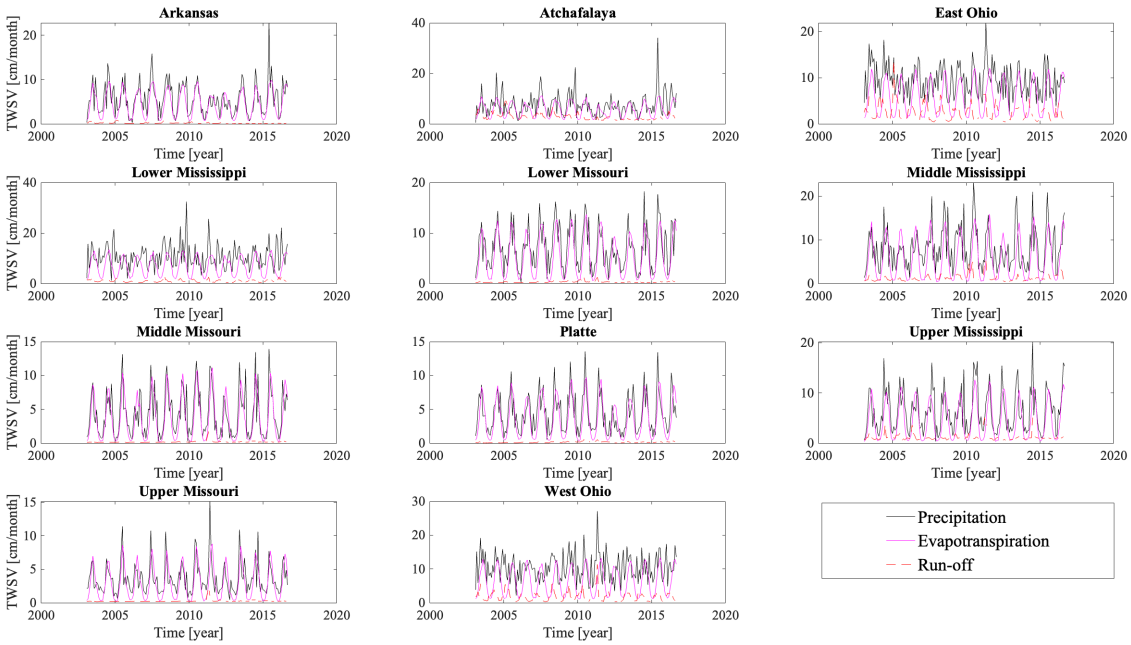


Figure 7.7: The time series of precipitation (black), evapotranspiration (pink), and run-off (red) for all the subbasins..

Figure 7.7 shows the time series of the hydrological processes for the different subbasins. To find out the favorable condition for applying run-off regularization, the results are compared to the different hydrological properties of the subbasins. When ignoring anthropogenic influences, the water storage in a subbasin is affected by precipitation, evapotranspiration, and run-off, as described in Chapter 3. The least amount of precipitation falls in the Middle and Upper Missouri and the Platte subbasins, around 4 cm/month. More precipitation falls in the East and West Ohio region, around 10 cm/month. The Lower Mississippi subbasin has the most precipitation, 10.9 cm/month. Lower Missouri, Middle Mississippi, Upper Mississippi, Arkansas, and Atchafalaya subbasins all have moderate precipitation. No inter-annual trends are visible for all subbasins, the amount of total precipitation remains fairly constant.

The smallest amount of evapotranspiration occurs at the Middle and Upper Missouri and the Platte subbasins, around 3 cm/month. More evapotranspiration occurs in the East and West Ohio subbasins, around 6 cm/month. The Lower Mississippi subbasins have the highest evapotranspiration, 6.8 cm/month. The Lower Missouri, Middle Mississippi, Upper Mississippi, Arkansas, and Atchafalaya subbasins all have moderate evapotranspiration. So, the same trends are visible as for precipitation. The areas with little precipitation have little evapotranspiration and vice versa. For evapotranspiration, there are no inter-annual variations visible, just like precipitation. It is visible that for the subbasins with less precipitation (Upper, Lower, Middle Missouri, Upper Mississippi, and Platte), the correlation between precipitation and evapotranspiration is higher. Meaning that for the more arid areas most of the water that enters the subbasin is immediately transpired or evaporated. Whereas for the more humid subbasins (East, West Ohio, Atchafalaya, and Lower Mississippi), this correlation is less, evapotranspiration is less dependent on precipitation due to the presence of water in the subbasin.

The lowest run-off occurs in the Upper, Middle, Lower Mississippi, and Platte subbasins, these make up the complete watershed of the Missouri River. The highest mean monthly run-off occurs in the East and West Ohio and Atchafalaya subbasins. Generally speaking, the amount of water circulating through a subbasin is the largest for the Ohio subbasins and the lowest for the Missouri subbasins. There is more precipitation in the Ohio subbasins which results in more evapotranspiration and run-off.

Looking at 7.2 shows that the run-off regularized solutions has a significantly lower *RMSD* for the Middle, Upper Missouri, Upper Mississippi, and Platte subbasins. For the same areas the correlation is significantly higher for the run-off regularized solutions, compared to the Tikhonov regularized solutions. In these subbasins the amount of water circulating through the subbasins is significantly lower than for the other subbasins within the Mississippi basin. The run-off values are lower, but also the mean monthly precipitation and evapotranspiration are lower. This is a sign that the run-off regularized approach might be useful in less wet areas since relatively small changes in annual precipitation can cause inter-annual variations in terrestrial water storage.

8 DISCUSSION

This chapter discusses the relevance of the obtained results. It looks into the similarities and differences between the results from this study and relevant previous studies, and the possible causes. It indicates the shortcomings of this research and what possible steps can be undertaken in order to proceed in this field of research. First, the results regarding the run-off storage relationships will be discussed. Followed by a discussion of the TWSV estimates from the adapted in-house approach. Finally, possible shortcomings and the relevance of the validation data are examined. Recommendations on further work will be given throughout each of the sections.

The values of mean monthly run-off found in this research are comparable to research from MacEdo et al. (2019). The highest discharge values are found in the Ohio region (East and West Ohio subbasins) and the lowest values are found in the Missouri region (Upper, Middle, and Lower Missouri and Platte subbasin). This research found no delay for most subbasins and a very short delay for the East Ohio and Platte subbasins. This seems reasonable since monthly averages are used and this way the effect of possible delays is mostly mitigated. The research of MacEdo et al. (2019) did not consider delay a factor when looking at the run-off storage relationship, which supports this reasoning. The R^2 values of the obtained relationships range from 0.04 to 0.75 in this research. In the research from MacEdo et al. (2019) the R^2 values ranged from 0.40 (in the Missouri region) to 0.80 (in the Ohio region). Important to note is that MacEdo et al. (2019) did not use differential run-off. The relationships were determined by relating the run-off time series of a stream gauge to storage variations in a complete watershed related to that outlet point. This means that not all relationships are comparable. The relationships that are comparable show high similarity. Both studies show the highest R^2 values in the Ohio region, which coincides with the highest mean monthly run-off values. The same holds for the opposite, the area with the lowest R^2 values is the Missouri region in both cases, where the mean monthly run-off is the lowest.

Winter months were excluded from the run-off time series to mitigate the effect of snow and ice. The success of this approach is measured by comparing the obtained results to results where snow and ice are accounted for following the approach of Riegger and Tourian (2014). In this approach, MODIS snow cover data is used to determine the fraction of a subbasin that is cover by snow or ice. This percentage is translated to a percentage of the TWSV that is frozen. The frozen storage part ratio is:

$$F(i,t) = 1 - S(i,t), \tag{8.1}$$

where S is the solid storage part ratio. This is determined as:

$$S(i,t) = \frac{A_S(i,t)}{A_i},$$

where A_S is the frozen area for a specific month and subbasin and A_i is the total area of the subbasin. The frozen area is determined as:

$$A_S(r,t) = \sum_{r=0}^{3600} \sum_{c=0}^{7200} i_{r,c} p_{r,c} a_{r,c}, \tag{8.2}$$

$$\text{with } i_{r,c} := \begin{cases} 1 & \text{if cell is within the subbasin } i \\ 0 & \text{if cell is outside the subbasin } i \end{cases}$$

This way the storage time series can be split into a solid and liquid part. The liquid part is used to fit the exponential relationships and determine the difference with the final results. Table 8.1 shows the results for both approaches, it indicates that there is no significant difference between the two. This means that excluding the winter months in the first place accounts for the effects of snow and ice. Note that MODIS snow cover data is solely used to validate this approach and not to be incorporated into the relationships.

Table 8.1: Goodness of fit from results compared to goodness of fit from the approach of Riegger and Tourian (2014)

Subbasin	R^2	R^2_{liquid}
Arkansas	0.33	0.33
Atchafalaya	0.35	0.33
East Ohio	0.75	0.71
Lower Mississippi	0.00	0.01
Lower Missouri	0.69	0.69
Middle Mississippi	0.63	0.62
Middle Missouri	0.04	0.03
Platte	0.45	0.46
Upper Mississippi	0.42	0.44
Upper Missouri	0.27	0.24
West Ohio	0.55	0.49

The Lower Mississippi and Middle Missouri are the two subbasins for which there is no clear dependence between run-off and storage. These are 2 of the 5 subbasins for which differential run-off is computed. The other 3 are the Lower Missouri, Middle Mississippi, and West Ohio subbasins. For the Lower Mississippi subbasin many months exist where the differential run-off is negative, this means that the outflow is smaller than the total inflow into the Lower Mississippi subbasin. Meaning that water is stored or leaves the subbasins through a different process. The latter can be either evapotranspiration or any human-induced process. Analysis of model data showed that evapotranspiration in the Lower Mississippi does not behave different to the other subbasins. The same holds for the precipitation in the Lower Mississippi subbasin. So, it is most likely that the absence of a clear correlation between run-off and storage is due to human activity, the same holds for the Middle Missouri subbasin, where there is also no abnormal behavior of precipitation and evapotranspiration.

Mississippi River Valley Alluvial Aquifer and the Ogallala Aquifer are two aquifers within the Mississippi basin. 12 percent of the land in the Lower Mississippi subbasin is used for agricultural purposes. This cropland needs to be irrigated and the main source of this water is the Mississippi River Valley Alluvial Aquifer (Massey et al., 2017). The Ogallala Aquifer is one of the world's largest aquifers, it underlies parts of the Arkansas, Platte,

Lower, and Middle Missouri subbasins. The Platte and Lower Missouri subbasin also have a high percentage of land used for agriculture, 10 and 7 percent, respectively. The aquifer has seen large depletion over the period 1990-2012 (Scanlon et al., 2012). Depletion of these natural aquifers can cause the run-off storage relationship to behave differently. For instance, when storage increases it may not results in an instant increase of run-off, since there is more water that can be stored due to the depletion of the aquifers. Many different processes play a role here and it is not a simple as the example states. However, it is most likely the main cause of why these areas show a weaker correlation between run-off and storage. In an attempt to increase the explained variance of the relationship for the Lower Mississippi and Middle Missouri subbasins data regarding this depletion can be gathered. Then, this effect can be accounted for in the storage time series of those subbasins. In this research an attempt was made in order to collect these data, however, this was not successful. The depletion of those aquifers occurs in different states and over many different sites, this highly complicates the collection of those data. There is no central institution that collects and distributes that data, as is the case for groundwater and run-off data which are distributed by USGS. Further work could be done on the collection of useful data regard aquifer depletion in the Mississippi Basin. Taking into account the purpose of these relationships it is very important to make sure this remains a data-only approach. Using model data or relationships that form the basis of models, will result in a biased TWSV estimate when using these to calibrate hydrological models or for other hydrological purposes.

More generally speaking, the data points do show a large scatter, for all subbasins. The two most probable causes of this are hysteresis and evapotranspiration. Riegger and Tourian (2014) discusses the effect of hysteresis on the run-off storage relationship. The hysteresis effect is caused by the superposition of different coupled and uncoupled storage components, which all have a unique temporal response. Some of the components involve fast transition, like surface water, and other components have a large time lag, like the melting process. For fully humid (Koppen class of Af) watersheds the hysteresis effect is very small. This means that the extremes of run-off storage occur at the same time. For boreal catchment the hysteresis effect is stronger, this is due to freezing, snow accumulation, and corresponding melting. Even though winter months are left out for most of the subbasins, melting is a process that can have an influence up to the summer months. Looking at the Mississippi basin, this hysteresis will have an influence on the run-off storage relationship as well. The West and East Ohio subbasins are the most humid subbasins, here the hysteresis effect is the smallest. For less humid and colder subbasins this effect becomes stronger. Since winter months are already extracted from the data points the main cause of this effect in this research will most likely be the melting process, which has an influence up to the start of the summer for some areas. Accounting for the melting process is something that could decrease the scatter in the data point used for extracting the run-off storage relationships, this could be something done when further researching this approach.

As discussed above, by eliminating the winter months most of the hysteresis effect should be accounted for, except the effect of melting water. So, the major process that causes this scatter is presumably evapotranspiration. Run-off and evapotranspiration are the two major processes causing water to leave the subbasin. As shown in Figure 7.7, evapotranspiration accounts for most of the water that leaves the subbasin. For the more arid areas

the correlation between precipitation and evapotranspiration is very high, in those areas the evapotranspiration flux is way higher than the run-off flux. There is no clear seasonality in run-off for those areas, this might be because actual evapotranspiration does not meet potential evapotranspiration. No seasonality in the run-off time-series is what partly causes the large scatter in the data points. Further research could look into the effect of evapotranspiration on the run-off storage relationship in more detail. This might result in better-fitting models.

Next, the GRACE TWSV estimates are discussed. As stated in Chapter 7, the in-house solution differ from the GSFC/JPL solutions. Looking at 7.3 shows that this is not the result of over regularization of the in-house solution. The annual peak values of the in-house solution often coincide or surpass, the peaks of the GSFC/JPL mascon solutions. The inter-annual variations that are captured by the GSFC/JPL mascon solutions are captured by the in-house solution as well. A clear distinction between the in-house and GSFC/JPL solution is visible for the Middle Missouri subbasin the in-house solutions estimates of TWSV are significantly higher in the period 2003-2006. Another difference in the two solutions is visible in the Arkansas subbasin, where the in-house solutions show much higher seasonal amplitude throughout the entire period. There are a couple of possible causes for this difference. Noise present in the SHCs obtained from ITSG could cause a difference in the solutions, additionally, these solutions could lack spatial resolution. Another reason for this difference can be the statistically sub-optimal data inversion. In the application of this method to the Greenland ice sheet, a data covariance matrix is computed on a month-by-month basis from the full noise covariance of the monthly SHCs (Ran et al., 2018). In this research, a unit matrix is used. Further work could also imply this statistically optimal data inversion in order to eliminate this as a possible cause. This difference between the in-house and GSFC/JPL mascon solution does not interfere with the main goal of this research, finding the added value of run-off in regularizing GRACE mascon solutions, which focuses on the difference between the two in-house solutions.

First, the results without co-estimating specific yield and regularization bias are discussed. It is clear from Table 7.2 that the GSFC/JPL mascons solution perform better in terms of *RMSE*. The only areas for which the in-house solution performs better are the Platte, Middle, and Upper Missouri, and Upper Mississippi subbasins, taking into account both correlation and *RMSE*. From these, the Middle Missouri subbasin showed different behavior in the period 2003-2006, as indicated above. This can be a sign that the in-house solution captures the 'true' signal better than the GSFC/JPL mascon solution for those areas during that period. The Platte, Middle, and Upper Missouri, and Upper Mississippi subbasins are the areas for which the run-off method preform significantly better than the Tikhonov method. As discussed in Chapter 7, those are the areas that have strong inter-annual trends of TWSV. This indicates that run-off is of added value in areas where there are strong inter-annual variations, which are more dampened when run-off data is not used.

Chapter 7 states that the higher correlation values for the run-off regularized solutions are a strong indicator of better performance of this method. By considering correlation, the effect of regularization bias and uncertainty in specific yield estimation is mostly mitigated. A study from Klosko et al. (2009) compared mascon recoveries to independent

flux estimates in the Mississippi basin. A similar validation method is used in this research. Their research divided the Mississippi basin into 4 large subbasins. The results showed correlations between the TWSV signal derived from GRACE and the validation data from 0.53 up to 0.92. Comparing the correlations from this research to the values found by Klosko et al. (2009) shows large similarities.

Next, the *RMSE* after co-estimating specific yield and regularization bias are discussed. The GSFC/JPL mascons perform better than the in-house solutions for all subbasins. However, an interesting difference is visible when comparing the two in-house solutions. A large portion of the *RMSD* was accounted for by the co-estimation of the specific yield. However the values of C_2 are relatively low and the uncertainty in those estimates very high. A C_2 value of 0.3 would mean that the used specific yield from Table 5.3 should be multiplied by a factor of 0.3. It is reasonable to assume that the actual specific yield will differ from the specific yield used in this case, however, these low values are not reasonable, this can be the result of over-fitting. This, in combination with the large confidence intervals for the estimates, make it impossible to draw a conclusion from these estimates. In further work more attention should go to the specific yield used for converting well measurements to groundwater changes. The specific yield can be determined individually for each well, this is time consuming, but could improve the certainty of the findings of further studies.

It is more reasonable to look at the peak amplitudes as main cause of the difference in *RMSDs*. Looking at 7.3 indicates that the run-off solution differs from the Tikhonov solution in the peak values. This might be an indicator that run-off data can be of added value for predicting seasonal peak values, which otherwise might be underestimated by GRACE mascon solutions. The fact that the validation data give the lowest annual amplitudes is counteractive, when looking at Figure 7.3. However, the error remaining after the fit is the largest for the validation data. This means that there is a lot of signals which is not explained by this simple fit in the validation data. This is due to the bigger inter-annual variation captured by the validation data, which can not be accounted for by this fitted model. This results in the fact that the annual amplitudes are underestimated. Looking at the three different mascon solutions shows that the run-off regularized solution shows the highest annual amplitudes for all subbasins. Klosko et al. (2009) showed that GRACE mascon estimates constantly underestimate the annual amplitudes. For the whole Mississippi Basin, the difference is 2.4 cm, for some subbasins this difference increases to 5.1 cm. The period used in that research was significantly shorter, which means that there is less effect of these inter-annual variations. Looking at the basin average of this research, the validation data gives a seasonal amplitude of 4.58 cm, the GSFC/JPL solution shows an amplitude of 3.58, the Tikhonov regularized solution shows an average amplitude of 4.04 cm and the run-off regularized solution shows an average amplitude of 4.63 cm. So, looking at the basin scale the run-off regularized solutions show the most similarities to the validation data when looking at annual amplitudes. The shortcoming of this approach is that it tries to capture amplitude and trends as single values. These can vary over time and that is why it could be better to use a stochastic model. This way the problem with strong inter-annual trends can also be mitigated. However, the simplicity of the used methods is lost in that case, subbasin characteristics can not be represented by a single number.

It is advisable to work with stochastic models for estimating the annual amplitudes and trends during further research.

9 CONCLUSIONS

This chapter summarizes the foremost findings and identifies the success of the different research objectives. First, the sub-goals of this research are examined. This is followed by a statement on the main goal: to identify the added value of run-off data for high-resolution estimation of TWSV in the study area from GRACE data.

First, a study area was selected. The Mississippi Basin in the United States of America was chosen over the different basins in Europe. This choice was mainly based on the differences in TWSV behavior. Since TWSV behavior in Europe is very similar, it will be more difficult to see the effect of run-off data. The comparable TWSV behavior is because most of the European basins have the same climate. In this case, the impact of run-off will be hard to detect, since signal leakage in and out of the basin will cancel one another out. The subbasins in the Mississippi Basin did show very distinctive TWSV behavior, which was the main driver for this decision.

Secondly, the in-house mascon approach was adapted to apply to be applicable to mass anomalies of hydrological origin in general and extended with different kinds of regularization. Originally this method was developed to estimate the mass loss of ice sheets. This required a lot of effort and manipulations, not explained in detail in this report. Two different types of regularization were implemented into the method, which is further discussed below.

Thirdly, this adapted approach was fine-tuned by several numerical experiments, which were executed to determine optimal parameter settings and to find the best method to determine the regularization parameter. A buffer zone of 300 km, a correlation distance of 100 km for regularization, a Fibonacci grid of 50 km, a pseudo-observation resolution of 1 degree, and a grid resolution of 0.05 degree are used. These are all optimized in the context of the Mississippi basin and show similarity to the previous applications of this approach.

Fourthly, a variety of independent data is collected to validate the results. Soil moisture and snow water equivalent data are gathered from different land surface models all driven by NLDAS. Groundwater is collected from USGS..... Together these three data sets make up the validation set, to validate the GRACE TWSV estimates. The largest uncertainty in the validation data set is the groundwater data. Taking regional averages for specific yield and the possibility of selecting a well that is not representable of the regional behavior induces some uncertainty.

Fifthly, the relationships between run-off and storage were quantified for all subbasins. Most subbasins of the Mississippi Basins displayed a clear exponential relationship between differential run-off and storage. Winter months are left out to minimize the effect

of snow and ice. The results from this show good similarity, looking at explained variance, to previous research in this area (MacEdo et al., 2019). The Lower Mississippi and Middle Missouri subbasins did not show a clear relationship, which is probably due to the influence of human activity as outlined in Chapter 8. The effect of evapotranspiration could explain the large relative moderated R^2 values for most subbasins. This process is not accounted for in this research, since it was not feasible for this data-only approach.

Two different types of regularization were incorporated into the in-house approach. Standard Tikhonov regularization is applied with a correlation distance of 100 km, which constrains the neighboring solution to not differ much. Thereafter run-off derived storage estimates are added in the regularization process to create a more physical-based regularization. The addition of run-off data into the regularization is a small adaption from standard Tikhonov regularization, resulting in one additional term in the estimator function.

The main goal of this research is to identify the added value of run-off data for high-resolution estimation of TWSV in the study area from GRACE data. Even though the GSFC/JPL mascon solution showed better results compared to the validation data, run-off data shows potential to be of added value when it comes to regularization of mascon solutions. Compared to standard Tikhonov regularization, run-off regularized solutions showed higher annual amplitudes. Taking into account that GRACE often underestimates the real hydrological signal, this might be an indication of a possible improvement. The run-off regularized solutions perform especially better in less humid areas. In those areas, in this research the subbasins of the Missouri basin, there are stronger inter-annual variations in TWS. Inter-annual variations are less dampened in the run-off regularized solutions. For humid areas, a larger absolute change in the hydrological process is needed to have inter-annual variations. Less dampening in the run-off regularized solutions is also supported by the larger annual amplitudes, compared to the standard Tikhonov regularized solutions. Furthermore, the run-off regularized solutions show less regularization bias. For most of the subbasins, the regularization bias is smaller for the run-off regularized solutions.

Regularization using run-off data shows potential particularly in areas with strong inter-annual variations in TWS, semi-humid and arid areas are more susceptible to such variations. For those areas, the run-off storage relationship is less explicit than for more wet areas. It shows that even though run-off storage relationships have moderate predictive value when it comes to regularizing GRACE mascon solutions these are still relevant. More research should be done to get a more definite conclusion regarding the added value of runoff data, however, the results of this research show great promise.

A APPENDIX

A.1 ADAPTION OF THE TU DELFT MASCON APPROACH

The approach consist of 6 main scripts: *part0.m*, *part1.m*, *part2.m*, *part3.m*, *computedgwn.m* and *estimatexflex.m*. Scripts *part0.m* and *part1.m* deal with the geometry of the different grids, *part1.m* and *part2.m* deal with the functional model and *computedgwn.m* and *estimatexflex.m* deal with the actual inversion.

For this project, the mascons must coincide with the subbasins that were defined. Whereas for the other regions this method was applied, e.g. Greenland, the configuration was not determined by (sub)basins. The script *part1.m* creates the outlines that form a specific mascon, these mascons consist of a set of smaller patches. This script was adapted and the result of this is that there are eleven mascons that all coincide with the subbasins that were defined. The output of this script is a cell structure that contains all the geometric information of the smaller patches within those mascons.

The script *part0.m* does the same thing, so creating small patches within a specific area, only for the extra mascons, this is further discussed in Chapter 6. These are the mascons created to absorb signal leaking into the basin, this script did not need adaption. So, *part0.m* and *part1.m* create the patches within the different mascons (subbasins) and store their locations.

The third script (*part2.m*) creates the functional model for all patches, using the geometric information of the patches created in the previous scripts. This model relates the gravity disturbance at satellite altitude to a mass anomaly for a specific patch. The grid that holds the observations at satellite altitude is created in *part0.m*, this is not changed from the previous version.

In *part3.m* the design matrices from both the external and main subbasins are combined as columns. So that there is one design matrix ($\mathbf{A}$) left.

In *computedgwn.m* the gravity disturbances at satellite altitude for all grid points are computed. This script is adapted for this project. In the original method mass trends needed to be estimated, however for this project monthly estimates need to be determined. This means the one-dimensional data vector ($\mathbf{d}$) is converted to a two-dimensional data matrix.

Finally, in *estimatexflex.m* inversion is applied to obtain a solution. To get a suitable solution, the inversion needs to be regularized. Different types of regularizations are applied as discussed in a separate section.

A.2 REGULARIZATION

The approach from **PDitmar2011**, which is based on Turchin et al., 1970, is explained below. Assuming that unknown model $\mathbf{x}$ satisfies the Gaussian distribution and the most probable model outcome is $\mathbf{x_0}$ with covariance matrix $\mathbf{C_x}$. This can be formulated as:

$$E(\mathbf{x}) = \mathbf{x_0}, \qquad D(\mathbf{x}) = \mathbf{C_x}, \tag{A.1}$$

with E and D as the expectation and dispersion operators. This lead to the following probability density function (pdf) of this model:

$$P(\mathbf{x}) \propto e^{-\frac{1}{2}(\mathbf{x}-\mathbf{x}_0)^T \mathbf{C}_\mathbf{x}^{-1}(\mathbf{x}-\mathbf{x}_0)}. \tag{A.2}$$

When assuming that noise in the data satisfies the Gaussian distribution. This assumption is justified by the fact that almost all data contain this type of noise. The second assumption made is that the noise had zero mean, when this is not the case the data is biased. If that is the case this bias should be estimates a model parameter as well. The data noise will have the following characteristics:

$$E(\mathbf{n}) = 0, \qquad D(n) = \mathbf{C_d}, \tag{A.3}$$

where $\mathbf{C_d}$ is the data noise covariance matrix. In the same way the data vector can also be represented as a random vector, with these propertied:

$$E(\mathbf{d}) = \mathbf{Ax}, \qquad D(\mathbf{d}) = \mathbf{C_d}. \tag{A.4}$$

When the model equals $\mathbf{x}$ the conditional pdf of the data vector becomes:

$$P(\mathbf{d} \mid \mathbf{x}) \propto e^{-\frac{1}{2}(\mathbf{d}-\mathbf{Ax})^T} \mathbf{C}_\mathbf{d}^{-1}(\mathbf{d}-\mathbf{Ax}). \tag{A.5}$$

Now using the Bayes theorem, which states that

$$P(\mathbf{d} \cap \mathbf{x}) = P(\mathbf{d} \mid \mathbf{x})P(\mathbf{x}) = P(\mathbf{x} \mid \mathbf{d})P(\mathbf{d}). \tag{A.6}$$

Then the pdf of the model, that is based on prior knowledge of the data noise and model, which is the conditional probability $P(\mathbf{x}|\mathbf{d})$, can be found as follows:

$$P(\mathbf{x} \mid \mathbf{d}) \propto P(\mathbf{d} \mid \mathbf{x})P(\mathbf{x}) \propto e^{-\frac{1}{2}(\mathbf{d}-\mathbf{Ax})^T} \mathbf{C}_\mathbf{d}^{-1}(\mathbf{d}-\mathbf{Ax}) - \frac{1}{2}\left(\mathbf{x}-\mathbf{x}_0\right)^T \mathbf{C}_\mathbf{x}^{-1}\left(\mathbf{x}-\mathbf{x}_0\right). \tag{A.7}$$

The model for which this pdf is maximum is the optimal model estimation $\hat{\mathbf{x}}$, this model estimation is found as the minimum of the quadratic objective function:

$$\hat{\mathbf{x}} = \arg\min_\mathbf{x} \left[(\mathbf{Ax}-\mathbf{d})^T \mathbf{C}_\mathbf{d}^{-1}(\mathbf{Ax}-\mathbf{d}) + \left(\mathbf{x}-\mathbf{x}_0\right)^T \mathbf{C}_\mathbf{x}^{-1}\left(\mathbf{x}-\mathbf{x}_0\right) \right]. \tag{A.8}$$

Differentiating with respect to $\mathbf{x}$ and assign the result to zero gives the following expression for the optimal model estimation:

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T \mathbf{C}_\mathbf{d}^{-1}\mathbf{A} + \mathbf{C}_\mathbf{x}^{-1}\right)^{-1} \left(\mathbf{A}^T \mathbf{C}_\mathbf{d}^{-1}\mathbf{d} + \mathbf{C}_\mathbf{x}^{-1}\mathbf{x_0}\right). \tag{A.9}$$

When assuming that the covariance matrices of the model and data noise can be written as a predefined matrix multiplied by a constant factor the following holds:

$$\mathbf{C}_\mathbf{d}^{-1} := \frac{1}{\sigma_d^2}\mathbf{P_d}, \quad \mathbf{C}_\mathbf{x}^{-1} := \frac{1}{\sigma_x^2}\mathbf{P_x}. \tag{A.10}$$

By inserting Equations A.10 into A.9 the following is obtained:

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{P_d}\mathbf{A} + \alpha\mathbf{P_x}\right)^{-1}\left(\mathbf{A}^T\mathbf{P_d}\mathbf{d} + \alpha\mathbf{P_x}\mathbf{x_0}\right), \tag{A.11}$$

where α is the regularization parameter which is computed by:

$$\alpha = \frac{\sigma_d^2}{\sigma_x^2}. \tag{A.12}$$

A.3 PARAMETER ESTIMATION

As described by Odolinski and Teunissen, 2019 this method estimates the variance of the data and model components in order to determine the statistically optimal regularization parameter as

$$\alpha = \frac{\sigma_d^2}{\sigma_x^2}. \tag{A.13}$$

The approach from Ran et al., 2018 is used to obtain the regularization parameters. After a first initial guesstimate of both variances, a first solution can be computed. Using this solution σ_d^2 and σ_x^2 can be computed again as described below. The data variance is computed as

$$\hat{\sigma}_d^2 = \frac{1}{\mathbf{N} - \mathrm{T}_d}(\mathbf{d} - \mathbf{A}\hat{\mathbf{x}})\mathbf{P_d}(\mathbf{d} - \mathbf{A}\hat{\mathbf{x}}) \tag{A.14}$$

with

$$\mathrm{T}_d = trace[\mathbf{N_d}\mathbf{N}^{-1}] \tag{A.15}$$

where $\mathbf{N}$ is the normal matrix, considered as the sum

$$\mathbf{N} = \mathbf{N_d} + \mathbf{N_x} \tag{A.16}$$

with $\mathbf{N_d}$ being the part of the normal matrix related to the data vector:

$$\mathbf{N_d} = \frac{1}{(\sigma_d^2)_0}\mathbf{A}^T\mathbf{P_d}\mathbf{A} \tag{A.17}$$

and $\mathbf{N_x}$ being the part of the normal matrix related to the initial model vector:

$$\mathbf{N_x} = \frac{1}{(\sigma_x^2)_0}\mathbf{R}. \tag{A.18}$$

The variance of the model vector can be estimated similarly:

$$\hat{\sigma}_x^2 = \frac{1}{M - \mathrm{T}_x}(\mathbf{x_0} - \hat{\mathbf{x}})\mathbf{R}(\mathbf{x_0} - \hat{\mathbf{x}}) \tag{A.19}$$

with

$$\mathrm{T}_x = trace[\mathbf{N_x}\mathbf{N^{-1}}]. \tag{A.20}$$

The regularization parameter is updated until a certain threshold is reached. When the new regularization parameter differs no more than 1 percent from the previous regularization parameter the process is stopped and this is chosen as the final regularization parameter.

BIBLIOGRAPHY

Baur, Oliver and Nico Sneeuw (2011). "Assessing Greenland ice mass loss by means of point-mass modeling: A viable methodology." In: *Journal of Geodesy* 85.9, pp. 607–615. ISSN: 09497714. DOI: 10.1007/s00190-011-0463-1.

Biancamaria, Sylvain et al. (2019). "Total water storage variability from GRACE mission and hydrological models for a 50,000 km2 temperate watershed: the Garonne River basin (France)." In: *Journal of Hydrology: Regional Studies* 24.June. ISSN: 22145818. DOI: 10.1016/j.ejrh.2019.100609.

Bridget R. Scanlon1, Zizhan Zhang1, Himanshu Save2, David N. Wiese3, Felix W. Landerer3, Di Long4, Laurent Longuevergne5 and Jianli Chen2 (2016). "Global evaluation of new GRACE mascon products for hydrologic applications." In: *Journal of the American Water Resources Association* 5.3, pp. 2–2. ISSN: 1093-474X. DOI: 10.1111/j.1752-1688.1969.tb04897.x.

Cai, Xitian, Zong Liang Yang, Cédric H. David, Guo Yue Niu, and Matthew Rodell (2014). "Hydrological evaluation of the noah-MP land surface model for the Mississippi River Basin." In: *Journal of Geophysical Research* 119.1, pp. 23–38. ISSN: 21562202. DOI: 10.1002/2013JD020792.

Chen, Jianli, James S. Famigliett, Bridget R. Scanlon, and Matthew Rodell (2016). "Groundwater Storage Changes: Present Status from GRACE Observations." In: *Surveys in Geophysics* 37.2, pp. 397–417. ISSN: 15730956. DOI: 10.1007/s10712-015-9332-4.

Engels, Olga, Brian Gunter, Riccardo Riva, and Roland Klees (2018). "Separating Geophysical Signals Using GRACE and High-Resolution Data: A Case Study in Antarctica." In: *Geophysical Research Letters* 45.22, pp. 12,340–12,349. ISSN: 19448007. DOI: 10.1029/2018GL079670.

Girotto, Manuela and Matthew Rodell (2019). *Terrestrial water storage*. Elsevier Inc., pp. 41–64. ISBN: 9780128148990. DOI: 10.1016/b978-0-12-814899-0.00002-x. URL: http://dx.doi.org/10.1016/B978-0-12-814899-0.00002-X.

Gruber, A., W.A. Dorigo, S. Zwieback, A. Xaver, and W. Wagner (2013). "Characterizing Coarse-Scale Representativeness of in situ Soil Moisture Measurements from the International Soil Moisture Network." In: *Vadose Zone Journal* 12.2, vzj2012.0170. ISSN: 1539-1663. DOI: 10.2136/vzj2012.0170.

Güntner, Andreas (2008). "Improvement of global hydrological models using GRACE data." In: *Surveys in Geophysics* 29.4-5, pp. 375–397. DOI: 10.1007/s10712-008-9038-y.Surveys.

Klosko, S., D. Rowlands, S. Luthcke, F. Lemoine, D. Chinn, and M. Rodell (2009). "Evaluation and validation of mascon recovery using GRACE KBRR data with independent mass flux estimates in the Mississippi Basin." In: *Journal of Geodesy* 83.9, pp. 817–827. ISSN: 09497714. DOI: 10.1007/s00190-009-0301-x.

Li, Bailing and Matthew Rodell (2015). "Evaluation of a model-based groundwater drought indicator in the conterminous U.S." In: *Journal of Hydrology* 526, pp. 78–88. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2014.09.027. URL: http://dx.doi.org/10.1016/j.jhydrol.2014.09.027.

Li, Bailing, Matthew Rodell, and James S. Famiglietti (2015). "Groundwater variability across temporal and spatial scales in the central and northeastern U.S." In: *Journal of Hydrology* 525, pp. 769–780. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2015.04.033. URL: http://dx.doi.org/10.1016/j.jhydrol.2015.04.033.

Li, Bailing, Matthew Rodell, Benjamin F. Zaitchik, Rolf H. Reichle, Randal D. Koster, and Tonie M. van Dam (2012). "Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe." In: *Journal of Hydrology* 446-447, pp. 103–115. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2012.04.035. URL: http://dx.doi.org/10.1016/j.jhydrol.2012.04.035.

Loomis, B. D., S. B. Luthcke, and T. J. Sabaka (2019). "Regularization and error characterization of GRACE mascons." In: *Journal of Geodesy* 93.9, pp. 1381–1398. ISSN: 14321394. DOI: 10.1007/s00190-019-01252-y. URL: https://doi.org/10.1007/s00190-019-01252-y.

Luthcke, Scott B., T. J. Sabaka, B. D. Loomis, A. A. Arendt, J. J. McCarthy, and J. Camp (2013). "Antarctica, Greenland and Gulf of Alaska land-ice evolution from an iterated GRACE global mascon solution." In: *Journal of Glaciology* 59.216, pp. 613–631. ISSN: 00221430. DOI: 10.3189/2013JoG12J147.

MacEdo, Heloisa Ehalt, Ralph Edward Beighley, Cédric H. David, and John T. Reager (2019). "Using GRACE in a streamflow recession to determine drainable water storage in the Mississippi River basin." In: *Hydrology and Earth System Sciences* 23.8, pp. 3269–3277. ISSN: 16077938. DOI: 10.5194/hess-23-3269-2019.

Margulis, Steven A. (2017). "Introduction To Hydrology." In: DOI: 10.2110/scn.94.32.0045.

Massey, Joseph H., C. Mark Stiles, Josh W. Epting, R. Shane Powers, David B. Kelly, Taylor H. Bowling, C. Leighton Janes, and Dean A. Pennington (2017). "Long-term measurements of agronomic crop irrigation made in the Mississippi delta portion of the lower Mississippi River Valley." In: *Irrigation Science* 35.4, pp. 297–313. ISSN: 14321319. DOI: 10.1007/s00271-017-0543-y.

Odolinski, Robert and Peter J.G. Teunissen (2019). "An assessment of smartphone and low-cost multi-GNSS single-frequency RTK positioning for low, medium and high ionospheric disturbance periods." In: *Journal of Geodesy* 93.5, pp. 701–722. ISSN: 14321394. DOI: 10.1007/s00190-018-1192-5. URL: https://doi.org/10.1007/s00190-018-1192-5.

Ran, J., P. Ditmar, R. Klees, and H. H. Farahani (2018). "Statistically optimal estimation of Greenland Ice Sheet mass variations from GRACE monthly solutions using an improved mascon approach." In: *Journal of Geodesy* 92.3, pp. 299–319. ISSN: 14321394. DOI: 10.1007/s00190-017-1063-5.

Riegger, J. and M. J. Tourian (2014). "Characterization of runoff-storage relationships by satellitegravimetry and remote sensing." In: *Journal of the American Water Resources Association*, pp. 3444–3466. ISSN: 1093-474X. DOI: 10.1111/j.1752-1688.1969.tb04897.x.

Rodell, Matthew, Jianli Chen, Hiroko Kato, James S. Famiglietti, Joe Nigro, and Clark R. Wilson (2007). "Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE." In: *Hydrogeology Journal* 15.1, pp. 159–166. ISSN: 14312174. DOI: 10.1007/s10040-006-0103-7.

Scanlon, Bridget R., Claudia C. Faunt, Laurent Longuevergne, Robert C. Reedy, William M. Alley, Virginia L. McGuire, and Peter B. McMahon (2012). "Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley." In: *Proceedings*

of the National Academy of Sciences of the United States of America 109.24, pp. 9320–9325. ISSN: 00278424. DOI: 10.1073/pnas.1200311109.

Seitz, Florian, Michael Schmidt, and C. K. Shum (2008). "Signals of extreme weather conditions in Central Europe in GRACE 4-D hydrological mass variations." In: *Earth and Planetary Science Letters* 268.1-2, pp. 165–170. ISSN: 0012821X. DOI: 10.1016/j.epsl.2008.01.001.

Sneeuw, Nico (1994). "Quadrature Methods in Historical Perspective." In: *Geophysical Journal International* 118, pp. 707–716.

Springer, Anne, Annette Eicker, Anika Bettge, Jürgen Kusche, and Andreas Hense (2017). "Evaluation of the water cycle in the European COSMO-REA6 reanalysis using GRACE." In: *Water (Switzerland)* 9.4, pp. 1–24. ISSN: 20734441. DOI: 10.3390/w9040289.

Springer, Anne, Jürgen Kusche, Kerstin Hartung, Christan Ohlwein, and Laurent Longuevergne (2014). "New estimates of variations in water flux and storage over Europe based on regional (Re)analyses and multisensor observations." In: *Journal of Hydrometeorology* 15.6, pp. 2397–2417. ISSN: 15257541. DOI: 10.1175/JHM-D-14-0050.1.

Sproles, E. A., S. G. Leibowitz, J. T. Reager, P. J. Wigington, J. S. Famiglietti, and S. D. Patil (2015). "GRACE storage-runoff hystereses reveal the dynamics of regional watersheds." In: *Hydrology and Earth System Sciences* 19.7, pp. 3253–3272. ISSN: 16077938. DOI: 10.5194/hess-19-3253-2015.

Suzuki, Kazuyoshi, Koji Matsuo, and Tetsuya Hiyama (2016). "Satellite gravimetry-based analysis of terrestrial water storage and its relationship with run-off from the Lena River in eastern Siberia." In: *International Journal of Remote Sensing* 37.10, pp. 2198–2210. ISSN: 13665901. DOI: 10.1080/01431161.2016.1165890. URL: http://dx.doi.org/10.1080/01431161.2016.1165890.

Tangdamrongsub, N., S. C. Steele-Dunne, B. C. Gunter, P. G. Ditmar, and A. H. Weerts (2014). *Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin*. Vol. 11. 10, pp. 11837–11882. ISBN: 1111837201. DOI: 10.5194/hessd-11-11837-2014.

Tourian, M. J., J. T. Reager, and N. Sneeuw (2018). "The Total Drainable Water Storage of the Amazon River Basin: A First Estimate Using GRACE." In: *Water Resources Research* 54.5, pp. 3290–3312. ISSN: 19447973. DOI: 10.1029/2017WR021674.

Turchin, V. F., Yu. A. Lazarev, M. Dakowski, and L. S. Turovtseva (1970). "RECONSTRUCTION OF PARTICLE MULTIPLICITY DISTRIBUTIONS USING THE METHOD OF STATISTICAL REGULARIZATION." In:

Watkins, Michael M., David N. Wiese, Dah-Ning Yuan, Carmen Boening, and Felix W. Landerer (2016). "Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons." In: *Journal of Geophysical Research: Solid Earth*, pp. 3782–3803. DOI: 10.1002/2015JB012608.Received.

Wouters, B., J. A. Bonin, D. P. Chambers, R. E.M. Riva, I. Sasgen, and J. Wahr (2014). "GRACE, time-varying gravity, Earth system dynamics and climate change." In: *Reports on Progress in Physics* 77.11. ISSN: 00344885. DOI: 10.1088/0034-4885/77/11/116801.

Xavier, Luciano, M. Becker, A. Cazenave, L. Longuevergne, W. Llovel, and O. C.Rotunno Filho (2010). "Interannual variability in water storage over 2003-2008 in the Amazon Basin from GRACE space gravimetry, in situ river level and precipitation data." In: *Remote Sensing of Environment* 114.8, pp. 1629–1637. ISSN: 00344257. DOI: 10.1016/j.rse.2010.02.005. URL: http://dx.doi.org/10.1016/j.rse.2010.02.005.

Zaitchik, Benjamin F., Matthew Rodell, and Rolf H. Reichle (2008). "Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River basin." In: *Journal of Hydrometeorology* 9.3, pp. 535–548. ISSN: 1525755X. DOI: 10.1175/2007JHM951.1.