universitätbonn igg

Institut für Geodäsie und Geoinformation

Schriftenreihe

52

Maike Schumacher

Methods for assimilating remotely-sensed water storage changes into hydrological models

Maike Schumacher • **Methods for assimilating remotely-sensed water storage changes into hydrological models**

universitätbonn igg

Institut für Geodäsie und Geoinformation

Schriftenreihe

52

Maike Schumacher

Methods for assimilating remotely-sensed water storage changes into hydrological models

Diese Arbeit wurde am 14. Juli 2016 als Dissertation zur Erlangung des Grades Doktor der Ingenieurwissenschaften (Dr.-Ing.) der Landwirtschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn vorgelegt.

Referent:	Prof. Dr.-Ing. Jürgen Kusche
Korreferent:	Prof. Dr. techn. Wolf-Dieter Schuh
Korreferent:	Prof. Dr. rer. nat. Petra Döll

Tag der mündlichen Prüfung: 12. September 2016

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
http://hss.ulb.uni-bonn.de
elektronisch publiziert.

Schriftenreihe des Instituts für Geodäsie und Geoinformation
der Rheinischen Friedrich-Wilhelms-Universität Bonn

Herausgeber:
Prof. Dr.-Ing. Theo Kötter
Prof. Dr.-Ing. Heiner Kuhlmann
Prof. Dr.-Ing. Jürgen Kusche
Prof. Dr. Lutz Plümer
Prof. Dr. techn. Wolf-Dieter Schuh
Prof. Dr. Cyrill Stachniss

Die Aufnahme dieser Arbeit in die Schriftenreihe wurde von den Herausgebern der Reihe einstimmig beschlossen.

Dieses Werk ist einschließlich aller seiner Teile urheberrechtlich geschützt.
Abdruck auch auszugsweise nur mit Quellenangabe gestattet.
Alle Rechte vorbehalten.

Bonn 2016 ISSN 1864-1113

This thesis is dedicated to Matthias Schürheck
whose words always enlightened my life
and consoled me in difficult times

and to Elisabeth Schürheck
who is like a shelter for me
to find calm and peace.

Ein Kinderherz soll sein --
wie die Lilie, so rein,
wie der Tau so klar,
wie der Spiegel so wahr,
wie der Quell so frisch,
wie das Vöglein im Gebüsch
so froh! ---

“Die Philosophie steht in diesem großen Buch geschrieben, dem Universum, das unserem Blick ständig offen liegt. Aber das Buch ist nicht zu verstehen, wenn man nicht zuvor die Sprache erlernt und sich mit den Buchstaben vertraut gemacht hat, in denen es geschrieben ist. Es ist in der Sprache der Mathematik geschrieben, und deren Buchstaben sind Kreise, Dreiecke und andere geometrische Figuren, ohne die es dem Menschen unmöglich ist, ein einziges Wort davon zu verstehen; ohne diese irrt man in einem dunklen Labyrinth herum.”

- Galileo Galilei -

Acknowledgments

First of all, I would like to thank my supervisor Prof. Dr. Jürgen Kusche who offered me the great chance to undertake my PhD studies in the Astronomical, Physical and Mathematical Geodesy (APMG) group at the Institute of Geodesy and Geoinformation, University of Bonn. I am very grateful for all the fruitful dialogs, the guidance and freedom that gave me the opportunity to build my own research profile, and the valuable insights and suggestions that without doubts helped me to improve the quality of my work. Our dialogs were always very open, and they were forthright with both criticism and praise.

My gratitude also goes to Prof. Dr. Wolf-Dieter Schuh (University of Bonn, Germany) and Prof. Dr. Petra Döll (Frankfurt University, Germany) for being a referee of this thesis, and for providing valuable discussions and kind supports. Prof. Schuh has been my great mathematics teacher, whose understanding of mathematical formulations and precise comments on the parameter estimations considerably improved the quality of my work. I cannot thank enough Prof. Döll for giving me the unique chance to work with the WaterGAP Global Hydrology Model (WGHM) source code, and constantly supporting me to cope with technical-hydrological issues. Without her generous collaboration, the proposed calibration and data assimilation framework of this thesis could not be completed. Furthermore, my thanks go to PD Dr. Axel Nothnagel and Prof. Dr. Sven Lautenbach (University of Bonn, Germany) who agreed to be the chairperson and member of my examination board.

During my bachelor and master studies, I had great supports from Prof. Kusche, Prof. Schuh, PD Dr. Klaus Börger, Dr. Lutz Roese-Koerner, Dr. Enrico Kurtenbach, Prof. Dr. Torsten Mayer-Gürr (Graz University, Austria), and Dr. Annette Eicker. During this time, I wrote my BSc and MSc theses, and worked as a student assistant in the Theoretical Geodesy (TG) and APMG groups. Their kindness and scientific inputs were essentially the main motivation behind my decision to continue my research towards a PhD.

I would like to acknowledge the financial supports awarded by the German Research Foundation (DFG) under the DFG SPP1257 "Mass transport and mass distribution in the Earth system" project "REGHYDRO" and under the "BAYES-G" (Bayesian Methods in Geodetic Earth System Research) project. Further, I am thankful for the exchange grant (2015/16 57044996) awarded by the German Academic Exchange Service (DAAD) to visit the Australian National University (ANU).

A great deal of the data that I used in my research was provided mostly freely by international and national organizations such as NASA, ESA, University of Bonn and Graz University. I was also fortunate to receive data sets and great advices on my research, for which particularly, I would like to thank Dr. Matthew Rodell (NASA, USA) for the groundwater well time series in the Mississippi River Basin (USA), Dr. Russell Crosbie (CSIRO, Australia) for the groundwater well observations in the Murray-Darling River Basin (Australia), and Dr. Fabrice Papa (LEGOS-IRD, France) for the Global Inundation Extent from Multi-Satellite (GIEMS) data set. In addition, the river discharge data from the Global Runoff Data Centre (GRDC, Bundesanstalt für Gewässerkunde (BfG), Koblenz, Germany) are acknowledged. I cannot thank Hannes Müller Schmied (Frankfurt University, Germany) enough for taking his time and providing technical and scientific

supports regarding the WGHM's input data, source code modifications, and analysis of the model outputs. His research contribution has had a tremendous positive impact on my experience.

The calibration and data assimilation software of this thesis is written using MATLAB and C++ with the Open Source HPS computing libraries ATLAS, BLAS, LAPACK and OpenMPI. I would like to thank Dr. Roelof Rietbrock and Dr. Jan Martin Brockmann (University of Bonn, Germany) for their technical supports, and helping me with various computer issues.

My special thanks go to Dr. Ehsan Forootan (Cardiff University, UK), Hannes Müller Schmied, PD Dr. Klaus Börger, Matthieu Talpe (University of Colorado at Boulder, USA), and Sebastian Halsig (University of Bonn, Germany) for proof-reading my thesis.

During the last year of my PhD studies, I had the unique chance to spend two months as a guest researcher at ANU, Australia. I am very grateful for the kind supports of Prof. Dr. Albert van Dijk during my research stay at his Water and Landscape Dynamics group at ANU, for sharing his hydrological expertise with me, and the valuable guidance throughout my graduate studies. Thanks to Dr. Paul Tregoning (ANU, Australia), Dr. Luigi Renzullo (CSIRO, Australia), and Dr. Russell Crosbie for the fruitful discussions during my stay in Canberra. I also had a chance to meet Prof. Dr. Jeffrey Walker and Prof. Dr. Valentijn Pauwels at Monash University (Melbourne), Prof. Dr. Shin-Chan Han, Prof. Dr. George Kuczera, and Prof. Dr. Garry Willgoose at the University of Newcastle, Australia, as well as Prof. Dr. Joseph Awange and Prof. Dr. Ir. Peter Teunissen (Curtin University, Australia), whose helpful discussions inspired my work. Especially, I would like to thank Prof. Dr. Mark Thyer (University of Adelaide, Australia) for his invitation to visit the University of Adelaide. This chance gave me the great opportunity to discuss my research with Prof. Dr. Dmitri Kavetski, Prof. Dr. Seth Westra, and Dr. Michael Leonard. I would also like to thank them for their time and scientific supports. My thanks go to the PhD students Siyuan Tian (ANU, Australia), Khandu (Curtin University, Australia), David Wright and Bree Bennett (University of Adelaide, Australia) for their warm welcomes during my visits.

In April 2015, I attended a short course on "Model building, inference, and hypothesis testing in hydrology" given by Prof. Dr. Mark Thyer, Prof. Dr. Dmitri Kavetski, Dr. Fabrizio Fenicia (Eawag, Switzerland), and Dr. Benjamin Renard (Irstea, France), which considerably broadened my knowledge about hydrological modeling and inference. This course inspired me to use the simple one-bucket model, discussed in the course, as an example to demonstrate the functionality of the methods in this PhD thesis.

During my PhD studies, I attended several national and international conferences. This provided me the great chance to discuss my research with experts in hydrological sciences, geodesy, remote-sensing and statistical analysis. I am very thankful for all the impressions and ideas that I got and that often resulted in new collaborations, naming a few: Prof. Dr. Laurent Longuevergne (University Rennes, France), Dr. Lars Nerger (AWI, Germany), Dr. Rolf Reichle (NASA, USA), Prof. Dr. Gabriëlle De Lannoy (KU Leuven, Belgium), Prof. Dr. Bart Forman (NASA, USA), Andreas Kvas (Graz University, Austria), Dr. Ben Gouweleeuw (GFZ, Germany), Prof. Dr. Thorsten Wagener (Bristol University, UK), Prof. Dr. Andreas Güntner (GFZ, Germany), Prof. Dr. Michael Schmidt (DGFI, Germany), Dr.

Guillaume Ramillien (GRGS, France), Dr. Lucia Seoane (GRGS, France), Prof. Dr. Paul Bates (Bristol University, UK), and Dr. Lucia Plank (University of Tasmania, Australia).

I would like to thank my colleagues in the TG and APMG group for all the scientific and non-work related discussions, and the wonderful time in a great working environment: Ehsan Forootan, Ina Loth, Klaus Börger, Lutz Roese-Koerner, Cathrin Schnier, Silvia Müller, Annette Eicker, Roelof Rietbrock, Anne Springer, Judith Schall, Laura Jensen, Bernd Uebbing, Ribana Roscher, Jan Martin Brockmann, Boris Kargoll, Anno Löcher, Prof. Dr. Karl-Rudolf Koch, Christina Lück, Christina Esch, Joël Koehler, Michael Plümer, Sandra-Esther Brunnabend, Luciana Fenoglio-Marc, Fan Yang, Gerd May, and Basem Elsaka, as well as Prof. Dr. Hans-Peter Helfrich at the Institute for Numerical Simulation. My thanks also go to the students Tobias Jurek, Anne Braakmann-Folgmann, Katharina Franz, Anne Kockmeyer, Kerstin Schulze, and Laura Over, whose bachelor theses and student assistance jobs have been supervised by me. It is a pleasure for me to have a chance of working with these many nice people, learning from them, and passing on my own research experience.

I am very grateful to my friends Silvia, Klaus, Lara, Jenny, Ina, Katharina, Christine, Heribert, Verena, Lisa, Sebastian and Heera for distracting me from my thesis work. I would like to thank my sister Hanna, my mom Steffi and Gerold, my dad Kalle and Marion, and my grandmothers Elisabeth and Martha for supporting me and my decisions throughout my whole lifelong.

I would like to thank Ehsan for his love and immense support in all aspects of my life and for always keeping me grounded. Finalizing this thesis would not be possible without you. “Du bist zu Hause für immer und mich.”

Abstract

Understanding physical processes within the water cycle is a challenging issue that requires merging information from various disciplines. The Gravity Recovery And Climate Experiment (GRACE) mission provides a unique opportunity to measure time-variable gravity fields, which can be converted to global total water storage anomalies (TWSA). These observations represent a vertical integral of all individual water compartments, which is difficult to observe by in-situ or other remote-sensing techniques. Knowledge about interactions between hydrological fluxes and terrestrial water storage compartments is reflected in large-scale hydrological models that nowadays increase in complexity to simulate all relevant physical processes within the global water cycle. Hydrological models are driven by climate forcing fields and their parameters are usually calibrated against river discharge to ensure a realistic water balance on river basin scale. However, errors in climate forcing fields, model parameters, and model structure limit the reliability of hydrological models. Therefore, it is necessary to improve model simulations by introducing measurements, which is known as data assimilation or data-model fusion.

In this thesis, a novel calibration and data assimilation (C/DA) framework is developed to merge remotely-sensed large scale TWSA with hydrological models. To implement this framework, the WaterGAP Global Hydrology Model (WGHM) is chosen, which is a sophisticated $0.5^\circ \times 0.5^\circ$ conceptual model that simulates daily water changes in surface and sub-surface water compartments (including groundwater), and considers water consumption. In particular, a flexible approach is introduced to assimilate GRACE TWSA as (sub-)basin or gridded averages into WGHM, while for the first time, implementing the observation error correlations in the C/DA system. A sensitivity analysis is performed to identify significant parameters in the largest river basins world-wide. It is also investigated whether GRACE TWSA can be used to calibrate model parameters. To reduce sampling errors and to improve the computational efficiency, the classical ensemble Kalman filter (EnKF) technique is extended to a square root analysis (SQRA) scheme, and the singular evolutive interpolated Kalman (SEIK) filter. The relationships between these algorithms are addressed. A simple model and WGHM are used to describe the mathematical details of the data assimilation techniques.

The observation error model, spatial resolution of observations, choice of filtering algorithm, and model ensemble size are assessed within a realistic synthetic experiment designed for the Mississippi River Basin, USA. Real GRACE products are also integrated into WGHM over this region. Investigations indicate that introducing GRACE TWSA constrains the water balance equation and corrects for insufficiently known climate forcing, in particular precipitation. Individual water states and fluxes are also adjusted but more improvements are expected by assimilating further in-situ and remotely-sensed observations. The processing choices represent important impacts on the final results. The C/DA framework is transferred to the Murray-Darling River Basin, Australia, to improve the simulation of hydrological changes under a long-term drought condition. GRACE C/DA introduces a negative trend to WGHM simulated TWSA. A validation with in-situ groundwater measurements indicates that the trend is correctly associated with the groundwater compartment. Thus, the C/DA helps to identify deficits in model simulations and improves the understanding of hydrological processes. The promising results provide a first step towards more complex C/DA applications on global scale and in conjunction with further terrestrial water storage observations.

Zusammenfassung

Zum Verständnis der physikalischen Prozesse des Wasserkreislaufes ist das Zusammenführen von Kenntnissen verschiedener Disziplinen erforderlich. Die Messungen zeitabhängiger Gravitationsfelder der Gravity Recovery And Climate Experiment (GRACE) Satellitenmission liefern einzigartige Erkenntnisse über globale Gesamtwasserspeicheränderungen (GWSA). Diese Größe repräsentiert die Summe aller einzelnen Wasserspeicherkomponenten, welche nur unzulänglich durch lokale oder andere satellitengestützte Verfahren beobachtet werden kann. Großskalige hydrologische Modelle simulieren Interaktionen zwischen terrestrischen Wasserspeicherkomponenten. Ihre Komplexität steigt heutzutage immer weiter, um alle relevanten physikalischen Prozesse im globalen Wasserkreislauf abzubilden. Sie werden durch Klimadaten angetrieben und durch Modellparameter gesteuert. Zur Gewährleistung einer realistischen Wasserbilanz in Flusseinzugsgebieten werden letztere üblicherweise gegen Durchflussmessungen kalibriert. Dennoch limitieren Unsicherheiten in den Klimadaten, in den Modellparametern und in der Modellstruktur die Zuverlässigkeit hydrologischer Prädiktionen. Um Simulationen zu verbessern ist daher die Integration von Beobachtungsdaten notwendig, welches unter dem Begriff der Datenassimilierung bekannt ist.

In dieser Arbeit wird ein neuer Kalibrierungs- und Datenassimilierungsansatz (K/DA) zur Kombination von großskalig beobachteten GWSA und hydrologischen Modellen am Beispiel des WaterGAP Global Hydrology Model (WGHM) entwickelt. WGHM ist ein konzeptionelles Wasserbilanzmodell, das tägliche Wasseränderungen auf und im Boden (inklusive Grundwasser) auf einer räumlichen Skala von $0{,}5° \times 0{,}5°$ berechnet und anthropogene Wasserentnahmen berücksichtigt. Insbesondere wird ein flexibler Ansatz zur Integration gegitterter und räumlich gemittelter GWSA eingeführt, während die Korrelationen der Beobachtungsfehler zum ersten Mal in der Assimilierung berücksichtigt werden. Eine Sensitivitätsanalyse identifiziert maßgebliche Parameter für die weltweit größten Flusseinzugsgebiete. Es wird außerdem untersucht, ob GRACE-GWSA zur Parameterkalibrierung herangezogen werden können. Um Stichprobenfehler zu reduzieren und um die rechnerische Effizienz zu steigern, wird die klassische Ensemble Kalman Filter (EnKF) Methode um das Square Root Analysis (SQRA) Schema und den Singular Evolutive Interpolated Kalman (SEIK) Filter erweitert. Die Zusammenhänge dieser Algorithmen werden dargestellt. Die mathematischen Details der Methoden werden anhand eines einfachen Modells und des WGHM beschrieben.

Das Modell der Beobachtungsfehler, die Auflösung der Beobachtungen, die Auswahl der Filteralgorithmen und die Größe des Modellensembles werden in einem realistischen synthetischen Experiment für das Flusseinzugsgebiet des Mississippis (USA) analysiert. GRACE-GWSA werden ebenfalls für dieser Region in das WGHM integriert. Untersuchungen zeigen, dass die Wasserbilanz an die Daten angepasst wird und ungenaue Klimadaten, insbesondere Niederschlag, ausgeglichen werden. Wasserspeicherkomponenten werden ebenfalls angepasst, würden aber durch die Assimilierung weiterer lokaler und satellitengestützter Daten profitieren. Der K/DA Ansatz hat einen entscheidenden Einfluss auf die Ergebnisse. Der entwickelte Ansatz wird auf das Einzugsgebiet des Murray und Darling Flusses (Australien) übertragen, um die Simulation hydrologischer Änderungen während einer Trockenperiode zu verbessern. GRACE-K/DA führt einen negativen Trend in das Modell ein. Die Validierung mit lokalen Grundwasserdaten bestätigt, dass der Trend korrekt mit dem Grundwasserspeicher assoziiert wird. Die K/DA ermöglicht somit Defizite in Modellsimulationen zu identifizieren und verbessert das Verständnis hydrologischer Prozesse. Die vielversprechenden Ergebnisse bereiten einen ersten Schritt in Richtung globaler K/DA in Verbindung mit weiteren hydrologischen Beobachtungen.

Contents

1. Introduction

1.1 Background: The Global Water Cycle

The global water cycle describes the re-distribution of water within the Earth system, including the hydrosphere, biosphere, atmosphere, and oceans (see Fig. 1.1). Water vapor is stored in the atmosphere and reaches the continents and oceans in form of precipitation. It might be stored on land in form of snow and ice, as surface water in lakes and wetlands or it infiltrates into the ground, where it is stored as soil moisture or groundwater. The continental storage is reduced by evapotranspiration, i.e. the sum of evaporation and transpiration from vegetation, or river discharge and groundwater runoff, which transport water, e.g., to seas or oceans. Due to solar radiation, water in seas and oceans also evaporates and is transported back into the atmosphere (Baumgartner and Liebscher, 1990).

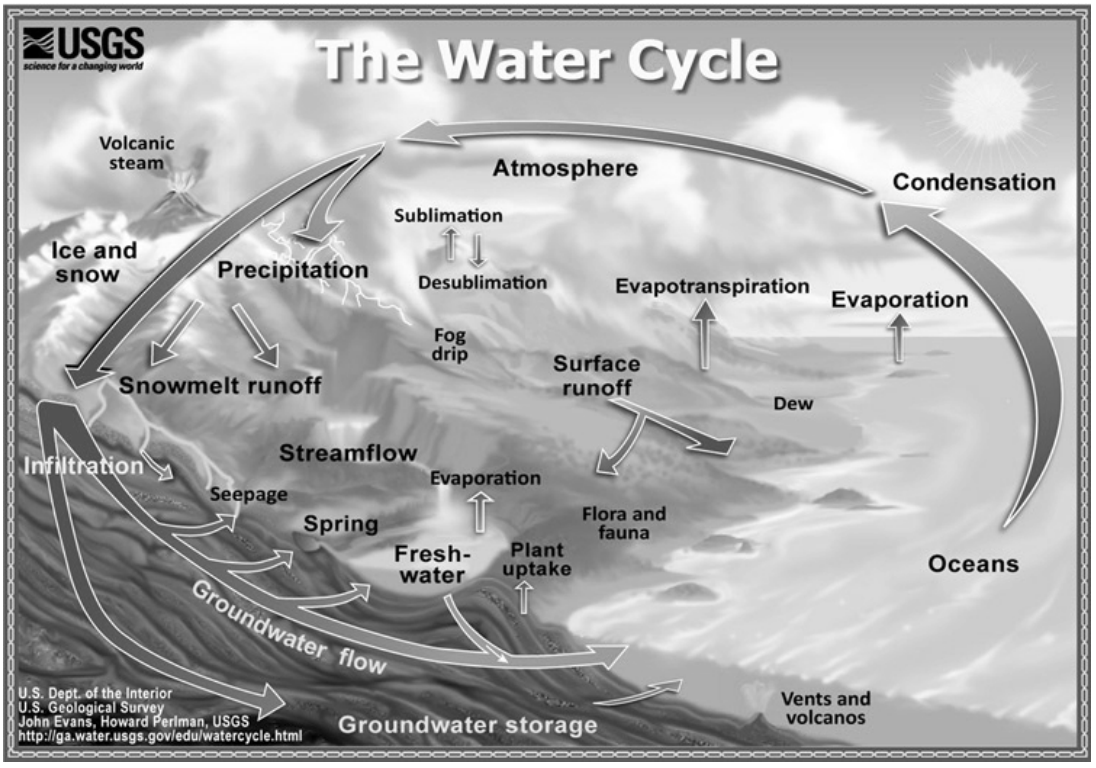

Figure 1.1: Schematic overview of the global water cycle. This figure is taken from the official website of the United States Geological Survey (USGS; http://water.usgs.gov/edu/watercycle.html).

The availability of freshwater, as well as the occurrence of natural hazards like droughts and floods impact human's life. On the other hand, changes in climate, global warming, and anthropogenic influences modify the energy and water cycle (Stocker et al., 2013). Therefore, monitoring the hydrological water cycle is an important task that requires knowledge about the underlying physical processes, as well as accurate temporally and spatially high resolved observations of water states and fluxes.

Due to the complexity of hydrological processes as well as due to their temporal and spatial variability, a large number of models has been developed to simulate the water cycle on a small, medium or large (global) scale with the aim to improve the understanding and quantification of the global water cycle (Sood and Smakhtin, 2015). These models are typically classified as conceptual (or water balance) models if empirical equations and parameters are used to represent the water dynamics, or as physics-based, e.g., land surface models, if the model equations are based on physical principles. Even though the models aim in an adequate representation of the real world, uncertainties exist due to insufficient model realism, imperfect climate input data and imperfect empirical model parameters.

Since 2002, the Gravity Recovery And Climate Experiment (GRACE) satellite mission, which consists of two satellites following each other in a polar orbit, continuously monitors the Earth system (Tapley et al., 2004). GRACE is the first geodetic tool that enables a global observation of the Earth's time-variable gravity field, which is represented in terms of potential spherical harmonic coefficients, the so-called GRACE level-2 product. The gravity field changes can be related to total water storage anomalies (TWSA) that mainly represents mass changes within the hydrosphere, biosphere, atmosphere, and oceans (Wahr et al., 1998). The temporal resolution of GRACE TWSA can vary between one month (by the official providers) to even one day (Kurtenbach et al., 2009) depending on the analysis technique, and the spatial resolution is of down to a few hundred kilometers (Schmidt et al., 2008). GRACE TWSA have been used in several hydrological, climatological and geophysical applications (see Kusche et al., 2012, Famiglietti and Rodell, 2013, Wouters et al., 2014, Famiglietti et al., 2015, Chen et al., 2016, and references therein). Since GRACE is not able to distinguish the individual sources of TWSA, e.g., soil water or groundwater, various approaches have been developed to separate GRACE TWSA into individual water compartments. Schmeer et al. (2012), Rietbroek (2014) and Forootan (2014) applied statistical decomposition and inversion techniques to separate TWSA derived by GRACE into storage compartments while using complementary data.

Merging GRACE TWSA and hydrological models calls for the application of data assimilation, which provides a twofold advantage. From the geodetic point of view, model-derived TWSA simulations that are consistent with time-variable mass estimations derived from GRACE could be very beneficial for applications that require the reduction of short-term gravity changes, e.g., de-aliasing of GRACE level-2 products (Forootan et al., 2014a) or the computation of loading effects in geometrical techniques (e.g., Collilieux et al., 2011, Fritsche et al., 2012). From the hydrological point of view, adjusting model-derived water states to GRACE observations helps improving limited simulation skills of models caused by uncertainties of input data (in particular climate forcing), errors in the model structure and uncertainties of model parameters. In this case, GRACE would act as a constraint that directly improves storage and indirectly flux simulations of hydrological models. Therefore, besides the traditional calibration of hydrological models against river discharge measurements (Gupta et al., 1998), multi-criteria calibration against river discharge and GRACE TWSA was performed by adjusting sensitive model parameters of a global model over large river basins (Werth and Güntner, 2010). Recently, a number of studies have suggested to assimilate GRACE TWSA into hydrological models (e.g., Zaitchik et al., 2008, Su et al., 2010, Forman et al., 2012, Houborg et al., 2012, van Dijk et al.,

2014, Eicker et al., 2014, Tangdamrongsub et al., 2015, Reager et al., 2015, Schumacher et al., 2016b).

1.2 Motivation to Develop a Calibration/Data Assimilation (C/DA) Framework

Geophysical and hydrological models are not able to perfectly describe processes within the Earth system. This is due to several assumptions and simplifications associated with the mathematical equations used in these models. In case of hydrological models, a major source of uncertainty are input data, in particular climate forcing such as precipitation, temperature and solar radiation (e.g., Döll et al., 2016). Empirical model parameters are defined in order to steer the model equations. Often, these parameters can not directly be measured, and therefore the estimation of their values represents another source of uncertainty in model simulations. Moreover, errors are introduced due to spatial and temporal discretization, as well as due to background information such as hydro-geology maps. Therefore, it is reasonable to merge the model outputs with measured data sets. This integration should improve the quality of model simulations by producing outputs that are closer to "reality" that is sampled by direct observations. Furthermore, assimilation of GRACE into hydrological models provides an alternative approach to separate TWSA into its individual water compartments and to downscale its relatively coarse spatial resolution. Relying on physical processes, introduced through model equations, to separate GRACE TWSA can be considered as a further benefit of data assimilation in contrast to relying on mathematical modes of storage variability as in the inversion techniques.

1.2.1 Model Parameter Calibration

The selection of appropriate values for hydrological model parameters is crucial in order to produce a reliable simulation of the water cycle. The term "parameter calibration" is referred to as adjusting the parameters of models in such a way that the "input-output behavior of the model approximates, as closely and consistently as possible, the underlying hydrological system over some historical period of time" (Vrugt, 2004, p. 10). In order to identify the "best" fit between model simulation and measurements, a so-called cost function is required. This function describes how similar model and observation values are based on an evaluation metric, such as the root mean square error (RMSE) between simulation and measurements or the Nash-Sutcliffe coefficient (NSC; Nash and Sutcliffe, 1970), that show how accurate the model represents the observations. By simulating the model outputs considering a large range of possible model parameter values, the parameter value that results in the smallest RMSE or highest NSC, respectively, can be identified as the optimal calibrated parameter.

Traditionally, hydrological models have been calibrated against river discharge measurements (Gupta et al., 1998). In case of the WaterGAP (Water Global Assessment and Prognosis) Global Hydrology Model (WGHM; Döll et al., 2003), which will be investigated in more detail in this thesis, the runoff coefficient (a factor that regulates the amount of land

runoff from soil, see, e.g., Müller Schmied et al., 2014) is calibrated against mean annual river discharge at 1319 gauge stations worldwide. In Werth (2010), GRACE TWSA products were successfully used to calibrate the six most sensitive parameters of WGHM over the 33 largest river basins worldwide. As the GRACE calibrated WGHM version provides improved storage simulations, it is adopted and further developed in this thesis.

1.2.2 Assimilation of GRACE Total Water Storage Anomaly (TWSA)

The combination of numerical model simulations with observations is called "data-model fusion" or "data assimilation". Measurements are introduced into the model in order to improve the model states that exhibit uncertainties due to imperfect input data, parameters and errors in the model structure. Data assimilation can be defined in different ways, e.g., as weighted mean between model outputs and observations based on their accuracy, or active integration of data into the model, i.e. modification of model states, which are used in the next model forward integration step.

In order to define the "best" combination of model simulations and measurements, a cost function must be formulated. The function defines how model simulations and observed values are weighted by their accuracy while averaging within the data assimilation process. In the case that an observation is very accurate and large uncertainties for the model simulation exist, the weight for the observation should be high, while the weight for the model should be small. In contrast, when the model simulation is accurate but large uncertainties are associated with observations, the model should be weighted higher and the merged value should be closer to the model value.

Data assimilation is a complex process, whereas the sources of model uncertainties have to be defined, the observation errors can either be treated as uncorrelated or correlated, the spatial resolution has to be specified, the assimilation method has to be chosen, just to name a few options (see, e.g., Liu et al., 2012). Each of the approaches might have benefits but at the same time deficiencies and difficulties of implementation. An overview of studies that address the assimilation of GRACE TWSA into hydrological models is given in Tab. 1.1.

An ensemble Kalman smoother (EnKS; Evensen and van Leeuwen, 2000) approach has been developed at the Global Modeling and Assimilation Office (GMAO) of the National Aeronautics and Space Administration (NASA, USA) to assimilate (sub-) basin averaged (Zaitchik et al., 2008, Forman et al., 2012, Houborg et al., 2012, Li et al., 2012, Reager et al., 2015) or gridded (Girotto et al., 2016, Kumar et al., 2016) GRACE TWSA products into NASA's catchment land surface model (CLSM), while assuming uncorrelated errors for GRACE TWSA observations. Forman and Reichle (2013) investigated the effect of the spatial aggregation of GRACE TWSA on the final assimilation results using NASA's framework. They concluded that TWSA observations should be assimilated at the smallest spatial scale, for which the observation errors can be considered uncorrelated.

At the Australian National University (ANU, Australia), an off-line data blending approach based on the ensemble Kalman filter (EnKF; Evensen, 1994) has been developed

to provide a Global Water Cycle Reanalysis (GWCR) for 2003-2012 (van Dijk et al., 2014). In their study, five global land surface models, i.e. four variants of NASA's Global Land Data Assimilation System (GLDAS) and the World-Wide Water Resources Assessment (W3RA) model, as well as an ensemble of TWSA estimates from GRACE time-variable products were merged. Errors of the different model estimates and GRACE products were derived through a triple collocation error estimation approach (Tavella and Premoli, 1994).

Table 1.1: Calibration (C) and sequential data assimilation (DA) studies on merging GRACE TWSA and hydrological model simulations.

Method	Model	GRACE error	Resolution (10^6km^2)	Region	Reference
EnKS (DA)	CLSM	white	0.5-1.2	Mississippi	Zaitchik et al. (2008)
”	”	”	0.3-1.6	Mackenzie	Forman et al. (2012)
”	”	”	0.3-1.2	North America	Houborg et al. (2012)
”	”	”	0.3-0.8	Europe	Li et al. (2012)
”	”	”	0.2-1.6	Mackenzie	Forman and Reichle (2013)
”	”	”	1.2	Missouri	Reager et al. (2015)
”	”	”	1°×1°	USA	Girotto et al. (2016)
”	”	”	1°×1°	USA	Kumar et al. (2016)
EnKF (DA)	GLDAS W3RA	white	1°×1°	global	van Dijk et al. (2014)
EnKF (DA)	CLM	white	4°×4°	North America	Su et al. (2010)
EnKS (DA)	”	”	”	”	
EnKF (DA)	HBV-96	white	1 km×1 km	Rhine	Tangdamrongsub et al. (2015)
EnKF (C/DA)	WGHM	correlated	3.0 &	Mississippi	Eicker et al. (2014)
”	”	”	5°×5°	”	”
EnKF (C/DA)	WGHM	white &	0.1-1.2 &	Mississippi	Schumacher et al. (2016b)
SQRA (C/DA)	”	correlated	5°×5°	”	”
SEIK (C/DA)	”	”	”	”	”
EnKF (C/DA)	WGHM	correlated	0.2-0.4	Murray-Darling	Schumacher et al. (2016c)

Other experiments based on the EnKF/EnKS were reported in Su et al. (2010) who applied data assimilation to snow dominated river basins, as well as in Tangdamrongsub et al. (2015) who assimilated GRACE TWSA into a variant of the HBV-96 (Hydrologiska Byråns Vattenbalansavdelning) model for the Rhine River Basin. These studies have mainly focused on the application of data assimilation rather than addressing the technical details of the algorithms. So far, applied hydrological data assimilation techniques were limited to the EnKF/EnKS methods.

Data assimilation has a longer history in the atmosphere and ocean sciences. In these fields, the dynamics are considered to be chaotic. As a result, small errors in the initial conditions result in large differences in the temporal evolution of model states. This is critical and a limiting factor, e.g., in numerical weather prediction. Data assimilation methods have been developed in order to estimate optimal initial conditions (e.g., Daley, 1991, Kalnay, 2003). For example, three and four dimensional variational methods (3D-Var and 4D-Var) are used for global weather predictions at the European Centre for Medium-Range Weather Forecasts (ECMWF; Courtier et al., 1994, Klinker et al., 2000).

In recent years, sequential data assimilation methods have found an increasing interest in atmospheric and oceanic applications (e.g., Houtekamer et al., 2005, Losa et al., 2012). Saynisch et al. (2015), for instance, presented an assimilation of GRACE data into an ocean model using the EnKF.

Further developments of alternative ensemble methods were also undertaken in atmospheric and oceanic sciences. For example, the square root analysis scheme (SQRA) approach (Evensen, 2004) and the singular evolutive interpolated Kalman (SEIK) filter (Pham et al., 1998) have been developed to assimilate observations into ocean models. In the context of GRACE data assimilation, Schumacher et al. (2016b) recently investigated for the first time the application of these alternative ensemble filtering approaches in hydrology.

In contrast to atmospheric and oceanic dynamics, the accuracy of terrestrial water storage simulations is predominantly influenced by uncertain meteorological forcing conditions and model parameterization (Reichle, 2008). Thus, the simultaneous estimation of water states and model parameter values is beneficial for accurate hydrological model simulations. Several studies applied ensemble filter methods to jointly estimate model states and parameters, e.g., in order to improve the simulation of conceptual rainfall-runoff models (e.g., Moradkhani et al., 2005), of land surface models (e.g., Vrugt et al., 2005, Zhang et al., 2016) or of groundwater changes in hydraulic and groundwater flow models (Chen and Zhang, 2006, Drécourt et al., 2006, Hendricks Franssen and Kinzelbach, 2008). For the first time, Schumacher (2012) investigated the potential of assimilating GRACE TWSA with their full error information into WGHM in a combined calibration and data assimilation (C/DA) framework, which was built based on the EnKF technique. This view was followed by Eicker et al. (2014) and Schumacher et al. (2016b) who assimilated sub-basin averaged and gridded GRACE TWSA ($5°\times5°$ grids) into WGHM for the Mississippi River Basin and simultaneously calibrated model parameters. Before, GRACE data assimilation studies performed only the assimilation part and no attempt was undertaken to calibrate model parameters. A transfer of the C/DA framework to the Murray-Darling River Basin was reported in Schumacher et al. (2016c).

1.2.3 Challenges of GRACE Data Assimilation for Hydrology

As summarized in Schumacher et al. (2016b), assimilating GRACE TWSA into hydrological and land surface models is challenging due to (i) the temporal and spatial resolution mismatch between model-derived simulated water states and GRACE TWSA maps, (ii) the difficulty in describing model errors due to forcing, model parameters and model structure (e.g., Reichle and Koster, 2003, Liu et al., 2012), and finally (iii) the problem to appropriately determine errors of GRACE TWSA. In particular, GRACE level-2 products contain correlated errors, which result from instrumental noise (Pierce et al., 2008), anisotropic spatial sampling of the mission (Schrama et al., 2007), and temporal aliasing caused by incomplete reduction of short-term mass variations by models (Forootan et al., 2014a). These errors manifest themselves as "striping" patterns in GRACE TWSA maps (Kusche, 2007). Although striping is reduced after applying de-correlation filters (Swenson and Wahr, 2006, Kusche et al., 2009), correlated errors still exist even after spatial

aggregation (Longuevergne et al., 2010, Sakumura et al., 2014). The choice of the data assimilation method, of the simulation and observation error models, and of the spatial discretization of observations (gridded or basin averaged) will have significant impacts on the final results of data-model fusion. Various assessments should be performed to explore the benefit of each decision, which will be discussed in this thesis.

1.3 Research Gaps

The calibration of hydrological model parameters against GRACE observations and the assimilation of GRACE data into hydrological models represent a relatively new research subject. Therefore, in the following, the major research gaps are summarized:

- The errors of GRACE TWSA are spatially correlated. GRACE data assimilation into hydrological models does not consider a realistic characterization of these uncertainties so far. Most of previous studies disregarded the correlation of observation errors by assuming white noise, i.e. spatially uncorrelated errors, for GRACE TWSA errors. Therefore, the technical aspect of introducing a full error covariance matrix of GRACE observations for data assimilation, as well as the interpretation how the correlated GRACE errors affect the data assimilation results are not fully addressed. The distribution of GRACE errors will change after applying spatial averaging. Systematic investigations to address various decisions about the spatial resolution of GRACE TWSA and the errors in data assimilation frameworks are missing.

- Different strategies have been proposed in previous studies to handle the temporal and spatial resolution mismatch of GRACE TWSA and hydrological models. Further investigations and comparisons of these approaches are required to find a "best" strategy.

- Improvements for the description of model uncertainties are desirable to better characterize and identify the water compartments that GRACE TWSA assimilation should in particular contribute to. For this, investigations on selecting appropriate a priori probability density functions for model parameters, as well as on producing representative correlated error models for climate input fields such as precipitation, temperature, and solar radiation are required.

- While GRACE TWSA assimilation improves the representation of modeled TWSA, this does not necessarily hold for individual water compartments and other state variables and fluxes. Therefore, validation with independent data sets, such as soil moisture or groundwater observations, is required to assess the performance of assimilation results. Validation is, however, not trivial due to limited available ground-based observations, limited spatial and temporal coverage, as well as resolution mismatch between model outputs and ground- or satellite-based observations. Further investigations in the direction of a multi-criteria validation is necessary for better understanding the influence of GRACE TWSA on the individual compartments and fluxes.

- For validating the assimilation results with independent measurements, adequate metrics have to be defined. These metrics, however, might not lead to the same interpretation and conclusions. For example, the RMSE measures the differences between model simulations and observations. Small RMSE values suggest a small bias and variance between model and observations. On the other hand, the correlation between the modeled and observed water state time series reveals their timing (or phase shift). It might occur that the correlation is improved after data assimilation but not the RMSE. Multi-objective metrics should be introduced to help a more extensive evaluation of data assimilation results.

- Previous GRACE data assimilation studies showed that an individual adaption of the data assimilation approach to a specific region might be required. For instance, in some regions a specific water storage compartment might represent a dominant influence on TWSA, while water fluxes such as evapotranspiration might be highly important in other areas. Finding the "best" strategy for data assimilation is however not trivial, since the performance depends on various aspects, and therefore is missing so far.

- To guarantee an improvement of the simulation of individual water storage compartments or fluxes, a multi-criteria data assimilation is desirable. A combined assimilation framework to integrate various observations of, e.g., lake and river variations from altimetry, remotely-sensed soil moisture, in-situ groundwater, and river discharge, along with GRACE TWSA might help in further improving hydrological models and therefore our understanding of the global water cycle. Such comprehensive frameworks are still missing in the literature.

1.4 Objectives of the Thesis

This PhD study aims to develop a C/DA framework that merges predicted water states from hydrological models and GRACE TWSA while simultaneously calibrating model parameters. In this section, the main objectives of the PhD research are summarized. Furthermore, hypotheses are formulated and strategies on how to test these hypotheses are proposed.

Objective 1: Establishing a C/DA framework to integrate (sub-)basin averaged and gridded GRACE TWSA into the WGHM using an ensemble Kalman filter (EnKF).

Hypothesis I of this thesis states that the calibration of model parameters using GRACE TWSA via parameter-state correlations is possible. To test this hypothesis, a sensitivity and model covariance analysis will be performed. It will be investigated if the assimilation of GRACE TWSA enables a more realistic representation of TWSA in hydrological models (hypothesis II a), as well as improvements of individual water compartments and fluxes (hypothesis II b). First, a simulation study will be set up to test the performance of the framework and validate the results with a simulated truth (twin experiment). Then observed GRACE TWSA will be assimilated into WGHM. Independent measurement data sets of individual water states

and fluxes will be used to validate the C/DA results, and thereby testing hypothesis II a and b.

Objective 2: Accounting for observation error correlations and for a flexible discretization of GRACE TWSA.

In this thesis, the proposed C/DA strategy is based on the assumption that an improved spatial resolution of TWSA enables a better adjustment to the observations ("reality") than using basin averaged observations (hypothesis III), and that considering or neglecting the correlated errors of GRACE TWSA affects the C/DA results (hypothesis IV). Thus, simulation and real case studies will be used to test both hypotheses.

Objective 3: Extending the framework by variants of the ensemble filter algorithm.

It will be investigated if introducing ensemble filter variants, i.e. alternative filtering algorithms and tuning techniques, helps to improve the performance of the C/DA framework (hypothesis V). The assimilation techniques will be described using a simple hydrological model. Then, hypothesis V will be tested in a simulation study.

Objective 4: Transferring the framework to river basins with different climatic and hydrological characteristics.

Finally, the C/DA framework will be transferred to another region with different climatic and anthropogenic conditions to show that the proposed assimilation strategy does not only work in the test region (hypothesis VI). Observed GRACE TWSA will be assimilated into WGHM and independent measurements will be used for the validation of the results.

Parts of the results of a regional and global sensitivity and covariance analysis were published in Schumacher et al. (2016a). In this thesis, additionally, an analysis of the distributions of the calibration parameters is provided (chapter 6). A synthetic experiment was proposed in Schumacher et al. (2016b) to investigate the impact of C/DA strategies on the updated water states, including the spatial resolution of GRACE TWSA and the observation error model. In addition, alternative filter methods have been introduced and their influence on the C/DA results was assessed. Furthermore, in this thesis, the mathematical foundations of data assimilation are described in details for linear and non-linear model equations (chapters 3 and 4). To illustrate the functionality of the ensemble filter algorithms, a simple model is introduced. In addition, the practical implementation of the C/DA framework to assimilate GRACE TWSA into WGHM is addressed (chapter 5). The simulation study for the complex hydrological model WGHM (Schumacher et al., 2016b) has also been extended in this thesis (chapter 7). First results of a one year assimilation of observed GRACE TWSA into WGHM for the Mississippi River Basin were presented in Eicker et al. (2014), while an extensive validation of individual water compartments and fluxes is provided in this thesis (chapter 8). The application of the C/DA framework to the Murray-Darling River Basin was presented in Schumacher et al. (2016c). In the second part of chapter 8 of this thesis, the details of this transfer are provided.

1.5 Outline of the Thesis

This PhD thesis provides technical details for integrating GRACE TWSA into hydrological model simulations within a C/DA framework. It also discusses the influence of the C/DA strategies on the model simulations. In chapter 2, observations and models of the terrestrial water cycle are described. An overview of the GRACE satellite mission is given in section 2.1. The mathematical representation of the observed gravity field changes, as well as its conversion to TWSA, are introduced. Aspects of data post-processing and the propagation of the error information are also addressed. This is followed by the description of a simple hydrological model and a detailed discussion of WGHM used for applying the proposed C/DA framework (section 2.2). In section 2.3, an overview of the hydrological data used for performing validation is provided.

The mathematical foundations of data assimilation are discussed in chapter 3. This includes an introduction of statistical foundations (section 3.1) and of the Bayes' theorem (section 3.2). The basic principles of variational and sequential data assimilation are presented in sections 3.3 and 3.4 for linear and non-linear models, respectively. State augmentation to estimate model parameter values in data assimilation is also shown. Finally, the choice of sequential data assimilation and the application of the EnKF in particular are explained.

In chapter 4, the data assimilation procedure is described. After presenting the model prediction step (section 4.1), first the classical EnKF algorithm is discussed (section 4.2.1). Subsequently, the SQRA scheme (section 4.2.2), the SEIK filter (section 4.2.3), and the EnKS (section 4.2.4) are treated, and their similarities and differences with respect to the classical EnKF are presented. In section 4.3, tuning techniques to improve the performance of the filter update are discussed, including improved initial sampling schemes, covariance inflation, and localization. The simple hydrological model (see section 2.2) is applied throughout the chapter to illustrate the ensemble filter algorithms.

The implementation of the C/DA framework to integrate GRACE TWSA into WGHM is described in chapter 5. After giving an overview of the C/DA procedure in section 5.1, the strategies to deal with the temporal and spatial resolution mismatch of GRACE TWSA and WGHM are proposed in section 5.2, followed by a summary of the model and observation error information. At the end of the chapter, pseudo codes for the EnKF, SQRA, and SEIK algorithms are presented.

In chapter 6, a sensitivity and model covariance analysis is performed to investigate whether TWSA and the individual water compartments simulated by WGHM react to changes in model parameter values and to reveal which parameters exhibit dominant influences. In addition, the correlation coefficients between water storage changes and model parameters are analyzed to test whether WGHM parameters can be calibrated using GRACE TWSA.

A synthetic experiment is set up to assess the influence of C/DA strategies on the updated water states and calibrated model parameters. In chapter 7, the synthetic experiment is applied to the Mississippi River Basin (USA), which serves as the test region. In particular, the impact of the observation error model in the filter update, the spatial resolution of GRACE TWSA, and alternative filter methods are evaluated.

In chapter 8, two real case assimilation scenarios are reported. The assimilation of GRACE TWSA into WGHM for the Mississippi River Basin for one year is presented in section 8.1. The model is run for the following three years using the calibrated parameters and improved initial states to investigate whether an improved prediction of TWSA is achieved. Independent measurements of soil moisture, surface water extent, groundwater, and river discharge are taken to evaluate the performance of individual water compartments and fluxes. Afterwards, the C/DA framework is transferred to the Murray-Darling River Basin (Australia) in section 8.2. The focus is on exploring the effect of the meteorological drought during 2003-2009 that influenced the water budget of the basin. Decline in TWSA and groundwater are estimated after C/DA to assess whether merging GRACE data with WGHM improves the representation of the hydrological processes. In-situ groundwater well observations are also used as an independent validation.

In chapter 9, the major findings of this PhD thesis are summarized and an outlook for further research is provided.

2. The Terrestrial Water Cycle

The global water cycle describes the movement of water through the Earth system and was introduced in section 1.1. A fundamental principle for the system is the conservation of mass, i.e. water is not removed from and added to the water cycle. Therefore, changes in the total terrestrial water storage ΔS (e.g., mm/month) can be directly related to precipitation P (mm/month), evapotranspiration E (mm/month), and runoff R (mm/month) within a river basin and assuming (Baumgartner and Liebscher, 1990)

$$\Delta S = P - E - R. \tag{2.1}$$

In section 2.1, the Gravity Recovery And Climate Experiment (GRACE) satellite mission is introduced. GRACE data can be converted to total water storage anomaly (TWSA) fields, i.e. ΔS in Eq. (2.1). After describing the mathematical representation (section 2.1.1), smoothing of GRACE TWSA (section 2.1.2) and spatial averaging are discussed (section 2.1.3). The section is concluded by providing a description of TWSA error estimation (section 2.1.4). In section 2.2, hydrological models are introduced that can be used to simulate water storage changes (ΔS) and fluxes (E and R) based on climate forcing such as observed precipitation (P). First, a simple model is described (section 2.2.1) and then a more complex model, the WaterGAP Global Hydrology Model (WGHM), is presented (section 2.2.2). Finally, observations of individual water storage compartments and fluxes are discussed in section 2.3.

2.1 TWSA from GRACE

GRACE is a joint satellite mission of the American National Aeronautics and Space Administration (NASA) and the German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt, DLR). Since its launch on March 17, 2002, GRACE continuously monitors the Earth's time-variable gravity field (Tapley et al., 2004). Its nominal lifetime was five years, but at the time of writing this thesis, GRACE is still operating. The mission consists of two almost-identical satellites in tandem formation (GRACE twins) chasing each other in one orbital plane initially with an altitude of 500 km and with a distance of about 220 km between each other. The orbit is near-circular, i.e. the eccentricity is smaller than 0.0005, and it has an inclination of 89.5°. Due to atmospheric drag the initial altitude is decreased to about 362 km over the last 14 years (http://www.csr.utexas.edu/grace/operations/mission_status/). The orbital period is about 94 minutes.

The distance between the two satellites is measured by a highly accurate inter-satellite K-Band microwave Ranging (KBR) system in 5 second intervals with an accuracy of about 1 μm (Fig. 2.1). Variations in the Earth's gravity field occur due to mass changes on the continents, in the atmosphere and oceans including tides, ice volume changes, as well as post-glacial rebound. If the first satellite approaches a positive gravity anomaly, its attraction to the anomaly will increase, resulting in a larger distance between the

two satellites. Then, the second satellite approaches the anomaly, leading to a higher attraction and therefore to a decrease in distance between the satellites. By this principle, GRACE observes the integral sum of mass changes in the Earth system but is not able to distinguish between the different sources. Furthermore, Global Positioning System (GPS) receivers are installed on board of the satellites to enable a precise positioning. They also allow a precise time tagging of the K-Band measurements. Satellite Laser Ranging (SLR) reflectors are used as an additional orbit determination technique. A high precision accelerometer measures non-gravitational surface forces, dominated by the atmospheric drag, which have to be removed from GRACE observations (Mayer-Gürr, 2008). Finally, two star cameras measure the absolute and relative orientation of the satellites.

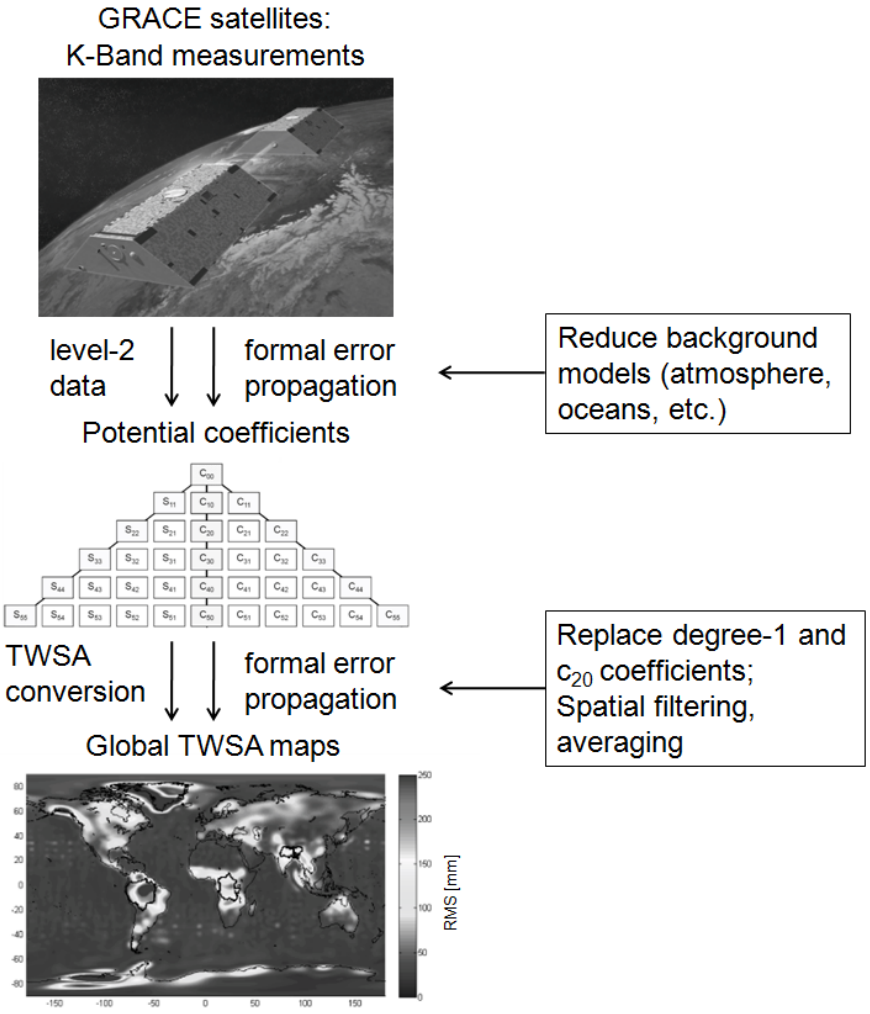

Figure 2.1: Overview of the GRACE data processing steps. The figure on the top is taken from http://photojournal.jpl.nasa.gov/catalog/PIA04235. The individual processing details are described in sections 2.1.1-2.1.4.

Three official analysis centers process the GRACE observations: the Center for Space Research (CSR, USA), the Jet Propulsion Laboratory (JPL, USA), and the GeoForschungsZentrum (GFZ, Germany). These centers provide spherical harmonic potential coefficients (level-2 data products) from the GRACE inter-satellite distance measurements. Furthermore, other research institutes estimate level-2 products, e.g., University

of Bonn (Germany), University of Graz (Austria), NASA Goddard Space Flight Center (GSFC, USA), Space Geodesy Research Group (GRGS, France), Delft University of Technology (Netherlands), and The Ohio State University (USA). During the data processing, most centers use different gravity recovery techniques and different background models to remove tidal and non-tidal high frequency mass variations, such as fast mass changes in the ocean and atmosphere, to extract hydrological mass variations from the observations. The effect of glacial isostatic adjustment (GIA) is usually not reduced but its effect is treated as post-processing in hydrological applications.

GRACE time-variable level-2 gravity products are provided with a temporal resolution of one month to even one day (Kurtenbach, 2011) depending on the analysis technique. They can be converted into TWSA (Wahr et al., 1998, Tapley et al., 2004), which have a spatial resolution of down to a few hundred kilometers (Schmidt et al., 2008).

In this thesis, GRACE level-2 products are used to estimate TWSA. Mass variations in the atmosphere and oceans have already been removed. In the following, first the mathematical representation of gravity field changes (sections 2.1.1.1 and 2.1.1.2) and their conversion to equivalent water heights (section 2.1.1.3), i.e. TWSA, are described. Due to the GRACE signal and error characteristics, the observations have to be filtered and spatially averaged (sections 2.1.2 and 2.1.3). Finally, the propagation of correlated errors of potential harmonic coefficients to correlated errors of TWSA is described (section 2.1.4). An overview of the GRACE data processing steps is given in Fig. 2.1.

2.1.1 GRACE Time-Variable Gravity Fields

2.1.1.1 Mathematical Representation of Gravity Field Changes

In this section, the description of the mathematical representation of the Earth's gravity field follows Heiskanen and Moritz (1967, chapter 1).

The Earth's gravitational potential V (J/kg) satisfies Laplace's equation outside the attracting masses

$$\Delta V = 0. \tag{2.2}$$

Therefore, its solution can be explained as a sum of harmonic basis functions $Y_n(\lambda, \theta)$

$$V(\lambda, \theta, r) = \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} Y_n(\lambda, \theta). \tag{2.3}$$

Here, (λ, θ, r) is the geocentric position vector in spherical coordinates with λ and θ denoting the geographical longitude (rad) and co-latitude (rad), while r is the distance (m) to the origin of an Earth-fixed coordinate system. Equation (2.3) is a spectral representation of the gravitational potential in the exterior of a unit sphere and $Y_n(\lambda, \theta)$ are denoted as surface spherical harmonics. It can be shown that $Y_n(\lambda, \theta)$ can be expressed as

$$Y_n(\lambda, \theta) = \sum_{m=0}^{n} [c_{nm} C_{nm}(\lambda, \theta) + s_{nm} S_{nm}(\lambda, \theta)], \tag{2.4}$$

where c_{nm} and s_{nm} are spherical harmonic coefficients, and C_{nm} and S_{nm} denote the surface spherical harmonics of degree n and order m

$$C_{nm} = \overline{P}_{nm}(\cos\theta)\cos(m\lambda), \tag{2.5}$$
$$S_{nm} = \overline{P}_{nm}(\cos\theta)\sin(m\lambda). \tag{2.6}$$

The spherical harmonics define a complete orthogonal system on the surface of a sphere. In Eq. (2.5) and (2.6), the normalized associated Legendre functions are denoted by $\overline{P}_{nm}$, which can be evaluated using a stable recursion formula. To apply Eq. (2.3) to the Earth defined as a sphere with the radius R (m) and with mass M (kg), as well as Newton's gravitational constant G ($\mathrm{m}^3/(\mathrm{kg}\cdot\mathrm{s}^2)$), the potential field can be expressed as

$$V(\lambda,\theta,r) = \frac{GM}{R}\sum_{n=0}^{\infty}\left(\frac{R}{r}\right)^{n+1}\sum_{m=0}^{n}[c_{nm}C_{nm}(\lambda,\theta) + s_{nm}S_{nm}(\lambda,\theta)]. \tag{2.7}$$

In Eq. (2.7), c_{nm} and s_{nm} are the Stokes' coefficients. Due to the choice of the factor $\frac{GM}{R}$, the coefficients are dimensionless.

Mass re-distributions in the Earth system are time-dependent. Thus, it is necessary to consider temporal changes of the gravitational potential, which can be represented in terms of changes of the Stokes' coefficients $\Delta c_{nm} = c_{nm} - c'_{nm}$ and $\Delta s_{nm} = s_{nm} - s'_{nm}$. It is common to subtract the spherical harmonic coefficients c'_{nm} and s'_{nm} from a static field or a temporal mean of monthly c_{nm} and s_{nm}, i.e. $c'_{nm} = \frac{1}{T}\sum_{t=1}^{T} c_{nm}(t)$, and s'_{nm} in the same way. Assuming global mass conservation, changes in the degree zero coefficient will vanish (see section 2.1.1.2) and Eq. (2.7) is modified as

$$\Delta V(\lambda,\theta,r) = \frac{GM}{R}\sum_{n=1}^{\infty}\left(\frac{R}{r}\right)^{n+1}\sum_{m=0}^{n}[\Delta c_{nm}C_{nm}(\lambda,\theta) + \Delta s_{nm}S_{nm}(\lambda,\theta)]. \tag{2.8}$$

2.1.1.2 Coefficients of Lower Degree

The coefficients of lower degree can be directly related to the physical shape of the Earth. The $c_{00} = 1$ coefficient can be interpreted as a scaling factor for $\frac{GM}{R}$ in Eq. (2.8). Due to the general assumption of mass conservation in the Earth system the coefficient does not change with time, thus $\Delta c_{00} = 0$. Therefore, the summation over n in Eq. (2.8) starts with $n = 1$. Changes in the degree-1 coefficients Δc_{10}, Δc_{11}, and Δs_{11} are linked to the offset $\Delta\mathbf{x}$ between the Earth's center of mass (CM) and the origin of the chosen reference system

$$\begin{pmatrix} \Delta c_{11} \\ \Delta s_{11} \\ \Delta c_{10} \end{pmatrix} = -\frac{M}{R^2}\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{pmatrix}. \tag{2.9}$$

GRACE level-2 products are evaluated with respect to a reference system with its origin located at the CM, and therefore, the degree-1 coefficients vanish. For other geodetic techniques, such as altimetry and GPS, a reference ellipsoid is used (Wahr et al., 1998) with its

origin located at the center of figure (CF). Hydrological models also simulate changes of the water states with respect to the Earth surface and therefore in a reference frame with its origin at the CF. The difference between CM and CF is known as geocenter motion. Due to the mass re-distribution within the Earth system, the offset between CF and CM varies temporally (Blewitt, 2013). GRACE observations are insensitive to these changes. These mass re-distributions represent, however, significant variations on the seasonal time scale, which was pointed out by, e.g., Chambers (2006) for ocean mass changes. Therefore, the degree-1 coefficients are usually replaced by geocenter motions estimated by analyzing physical models (http://grace.jpl.nasa.gov/data/get-data/geocenter/) or by combining GPS, GRACE, and ocean bottom pressure observations in an inversion approach (Rietbroek, 2014). In this thesis, the degree-1 time series are replaced by those estimated using the latter approach (http://www.igg.uni-bonn.de/apmg/index.php?id=geozentrum).

The influence of the replaced degree-1 coefficients on hydrological mass variations is shown in Fig. 2.2. For this, the root mean square (RMS) variability of the equivalent water height time series from degree-1 between 02/2003-06/2014 averaged over the 33 largest river basins was calculated. The smallest RMS values of 2-3 mm are found for the southern African River Basins, and the largest RMS values of 9-10 mm in the Australian River Basins, the central USA, as well as central and eastern Canada.

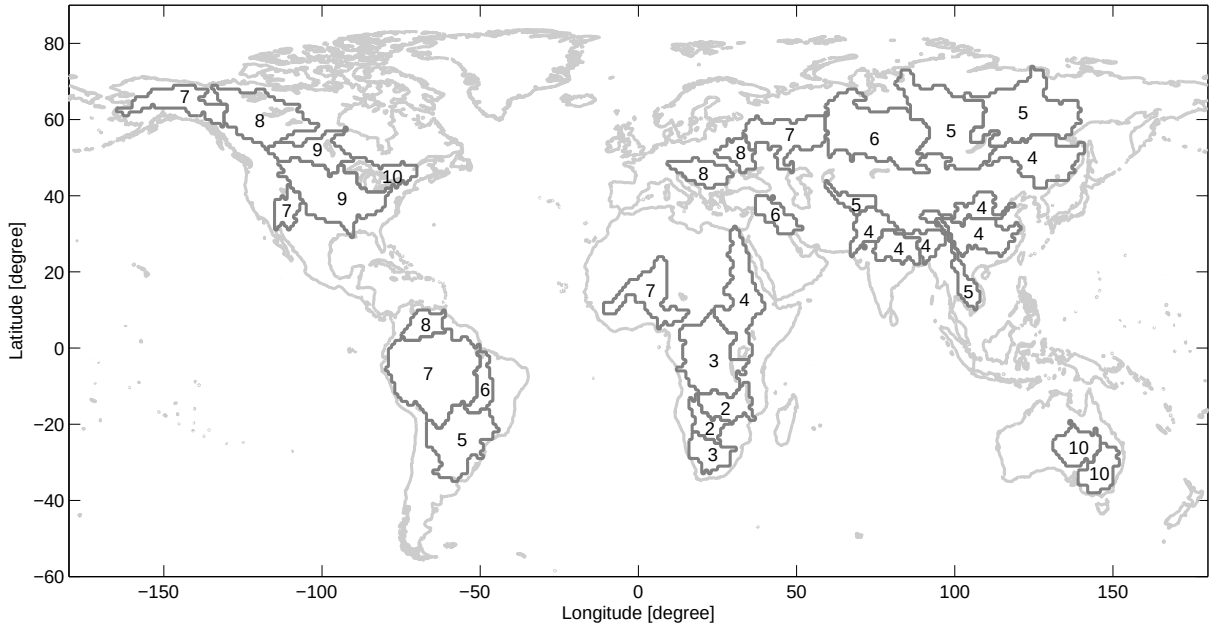

Figure 2.2: Influence of degree-1 coefficients on hydrological mass variations in terms of RMS variability (in mm) of their equivalent water height time series. Results are shown for the world's major river basins.

The degree-2 terms are related to the moments of inertia according to

$$c_{20} = -\frac{1}{R^3}\left(\theta_{33} - \frac{\theta_{11} - \theta_{22}}{2}\right). \tag{2.10}$$

Here, θ_{11}, θ_{22}, and θ_{33} denote the moments of inertia. The determination of the c_{20} coefficient (an indicator of the Earth's flattening) from GRACE is affected by various

errors, e.g., likely the S2 tidal aliasing at about one cycle in 161 days. Therefore, it is also common in hydrological applications to replace this coefficient. In this thesis, c_{20} is replaced by monthly estimates from SLR measurements (Cheng and Ries, 2012, grace.jpl.nasa.gov/data/get-data/oblateness/).

2.1.1.3 Conversion of Gravity Fields to TWSA

Wahr et al. (1998) considered a thin layer at the Earth's surface to model the loading effect of mass re-distributions from GRACE observations. Since mass changes Δm (kg) can also be expressed as surface density changes $\Delta\sigma$ (kg/m^2), the latter can be related to temporal changes of the Earth's gravity field as

$$\Delta\sigma(\lambda,\theta) = \frac{M}{4\pi R^2}\sum_{n=1}^{\infty}\sum_{m=0}^{n}\frac{(2n+1)}{(1+k_n')}[\Delta c_{nm}C_{nm}(\lambda,\theta) + \Delta s_{nm}S_{nm}(\lambda,\theta)]. \tag{2.11}$$

Considering Eq. (2.11), changes of the surface density create a direct gravitational effect due to mass changes as well as an indirect gravitational attraction through loading and deformation of the underlying solid Earth (Wahr et al., 1998). The indirect effect is represented by the degree-dependent gravitational load Love numbers $k_n^{'}$. Solutions for the load Love numbers can be estimated based on Earth models, such as the spherically symmetric, non-rotating, elastic, and isotropic Earth, so called SNREI models (e.g., Farrell, 1972, Dziewonski and Anderson, 1981).

The density changes $\Delta\sigma$ can be converted into changes of equivalent water heights ΔE (m), i.e. TWSA, by dividing the surface density by the mean density of seawater $\rho_w = 1025$ kg/m^3, i.e.

$$\Delta E(\lambda,\theta) = \frac{M}{4\pi R^2\rho_w}\sum_{n=1}^{\infty}\sum_{m=0}^{n}\frac{(2n+1)}{(1+k_n')}[\Delta c_{nm}C_{nm}(\lambda,\theta) + \Delta s_{nm}S_{nm}(\lambda,\theta)]. \tag{2.12}$$

2.1.2 Smoothing in Spectral Domain and Leakage Reduction

GRACE level-2 products, represented in terms of potential spherical harmonics, contain correlated errors, which manifest themselves as north-south "striping" patterns in GRACE TWSA in the spatial domain (Kusche, 2007). Also, the noise increases at higher degree and order of coefficients (Swenson and Wahr, 2006). To reduce the noise, smoothing (also called filtering) needs to be applied to GRACE data. In the spectral domain, this can be done by incorporating a filter matrix $\mathbf{W}$ (containing degree or degree and order dependent weights) to the spherical harmonic coefficient vector $\mathbf{x}$ to derive smoothed coefficients $\mathbf{x}^s$ as

$$\mathbf{x}^s = \mathbf{W}\mathbf{x} = \mathbf{W}\begin{pmatrix}\boldsymbol{\Delta}\mathbf{c}_{nm}\\ \boldsymbol{\Delta}\mathbf{s}_{nm}\end{pmatrix}. \tag{2.13}$$

Alternatively, the gridded GRACE TWSA can be filtered by convolving a filter kernel in the spatial domain. Generally, one distinguishes between isotropic and anisotropic filters.

Isotropic filters are only degree dependent in the spectral domain and independent of direction in the spatial domain, e.g., the Gaussian filter introduced by Jekeli (1981). Therefore, $\mathbf{W}$ in Eq. (2.13) is a diagonal matrix in the spectral domain. In contrast, anisotropic filters, e.g., those of Swenson and Wahr (2006), Kusche (2007), and Klees et al. (2008), are degree and order dependent in the spectral domain and location dependent in the spatial domain. The DDK filter (Kusche et al., 2009), for example, imitates the regularization of the GRACE level-2 normal equation matrix to construct an anisotropic smoothing kernel. In the spatial domain, the smoothing is direction-dependent, and thereby adapted to the anisotropic error correlation structure of GRACE observations. The kernel further depends on the geographical latitude. In the spectral domain, $\mathbf{W}$ in Eq. (2.13) is a full matrix, whose entries are degree and order dependent. The filtered coefficients Δc^s_{nm} and Δs^s_{nm} are then used in Eq. (2.12) to determine spatially smoothed TWSA as

$$\Delta E^s(\lambda,\theta) = \frac{M}{4\pi R^2 \rho_w} \sum_{n=1}^{\infty} \sum_{m=0}^{n} \frac{(2n+1)}{(1+k'_n)} [\Delta c^s_{nm} C_{nm}(\lambda,\theta) + \Delta s^s_{nm} S_{nm}(\lambda,\theta)]. \tag{2.14}$$

In practice, the summation over n has to be truncated at a maximum degree $n_{\max}$ according to

$$\Delta E^{s'}(\lambda,\theta) = \frac{M}{4\pi R^2 \rho_w} \sum_{n=1}^{n_{\max}} \sum_{m=0}^{n} \frac{(2n+1)}{(1+k'_n)} [\Delta c^s_{nm} C_{nm}(\lambda,\theta) + \Delta s^s_{nm} S_{nm}(\lambda,\theta)], \tag{2.15}$$

and the error due to spectral truncation can be expressed as

$$\delta\Delta E^{s'}(\lambda,\theta) = \frac{M}{4\pi R^2 \rho_w} \sum_{n=(n_{\max}+1)}^{\infty} \sum_{m=0}^{n} \frac{(2n+1)}{(1+k'_n)} [\Delta c^s_{nm} C_{nm}(\lambda,\theta) + \Delta s^s_{nm} S_{nm}(\lambda,\theta)]. \tag{2.16}$$

If Δc^s_{nm} and Δs^s_{nm} are derived from GRACE, $n_{\max} = 60$ or 90 is usually used in hydrological studies. The omission error of neglecting degrees $n > n_{\max}$ can be estimated by evaluating Eq. (2.15) up to $n_{\max} = 60$ and $n_{\max} = 90$. This results in a difference in equivalent water heights of less than 5 mm on a regular $1^\circ \times 1^\circ$ grid. The omission error of higher degrees, i.e. $n > 90$, is expected to be even smaller.

As a consequence of spatial smoothing, water mass inside a specified region might move outside of this region, as well as water from the outside might also move into this region. These effects are known as spatial leakage-out and leakage-in (Swenson and Wahr, 2002, Klees et al., 2006). Their impact on TWSA depends on the filter, area of the region, amplitude and phase of TWSA inside and outside of the region, and its location. In case of larger smoothing radii, smaller regions or along coast lines, the spatial leakage is large (e.g., Longuevergne et al., 2010).

Landerer and Swenson (2012) discussed that the differences in spatial scale between GRACE TWSA and other hydrological data such as model simulations have to be accounted for in order to consistently compare both. Otherwise differences in the data sets due to the scale mismatch might be attributed to limited skills of the observations or model simulations. To resolve this problem, filtering both observations and model simulations in the same way is suggested to harmonize their spatial scale, as applied, e.g., in Landerer et al. (2010) and Döll et al. (2014). Alternatively, Klees et al. (2006), Longuevergne et al.

(2010), as well as Landerer and Swenson (2012) suggested to estimate a damping factor of the TWSA signal (due to filtering) to reduce the effect of spatial leakage. This can be done by comparing the TWSA signal of a hydrological model before and after spatial filtering, and applying the ratios as either constant or time-dependent re-scaling factors to the filtered GRACE TWSA (Landerer and Swenson, 2012). Then, the re-scaled TWSA values can be directly compared to hydrological model outputs as, e.g., in Rodell et al. (2009) and van Dijk et al. (2011). Moreover, "mascon" (mass concentration blocks) solutions have been developed as another form of gravity field basis functions (e.g., Ramillien et al., 2012, 2015, Watkins et al., 2015). The masons act as inherent smoother on the data in such a way that spatial smoothing is not longer required in the post-processing. However, the signal amplitudes might still be attenuated and therefore scaling factors are e.g., applied to the mascon solutions determined at JPL (Watkins et al., 2015). In this thesis, re-scaling of (sub-)basin averaged and smoothed GRACE TWSA is selected to account for spatial leakage and to reduce differences in the spatial scale of observations and model simulation outputs. The reason is to save computation time and costs during the C/DA, since filtering of WGHM outputs is not required. Further investigations should be performed to find a "best" post-processing strategy for C/DA purposes.

2.1.3 Spatial Averaging of TWSA

For hydrological applications, usually GRACE TWSA are averaged over river basins. In the spectral domain, spatial averaging is performed by multiplying the filtered spherical harmonic coefficients (Δc^s_{nm} and Δs^s_{nm}) with the spherical harmonic coefficients (Δc^f_{nm} and Δs^f_{nm}) of a basin function f as

$$\Delta E^s_f = \frac{M}{4\pi R^2 \rho_w} \sum_{n=1}^{\infty} \sum_{m=0}^{n} \frac{(2n+1)}{(1+k'_n)} [\Delta c^f_{nm} \Delta c^s_{nm} + \Delta s^f_{nm} \Delta s^s_{nm}]. \tag{2.17}$$

For this purpose, harmonic coefficients of the basin function can be determined by defining a global grid with ones inside of the basin and zeros outside

$$f(\lambda, \theta) = \begin{cases} 1 & \text{within river basin} \\ 0 & \text{outside} \end{cases}, \tag{2.18}$$

and expanding the so-called basin function f into spherical harmonic functions

$$\Delta c^f_{nm} = \frac{1}{A_f} \int_{\Omega} f(\lambda, \theta) C_{nm}(\lambda, \theta) d\omega, \tag{2.19}$$

$$\Delta s^f_{nm} = \frac{1}{A_f} \int_{\Omega} f(\lambda, \theta) S_{nm}(\lambda, \theta) d\omega. \tag{2.20}$$

In Eqs. (2.19) and (2.20), the integral is calculated over the entire sphere Ω and scaled by the area of the region of interest A_f. In practice, the summation over n in Eq. (2.17) is truncated at a maximum degree $n_{\max}$ as

$$\Delta E^{s'}_f = \frac{M}{4\pi R^2 \rho_w} \sum_{n=1}^{n_{\max}} \sum_{m=0}^{n} \frac{(2n+1)}{(1+k'_n)} [\Delta c^f_{nm} \Delta c^s_{nm} + \Delta s^f_{nm} \Delta s^s_{nm}], \tag{2.21}$$

with the corresponding error due to spectral truncation $\delta\Delta E_f^{s'}$.

Alternatively, first the filtered spherical harmonic coefficients can be used to determine TWSA following Eq. (2.14), and then the equivalent water heights are integrated over the entire sphere Ω while taking the basin function in Eq. (2.18) into account

$$\Delta E_f^s = \frac{1}{A_f} \int_\Omega f(\lambda,\theta)\Delta E^s(\lambda,\theta) d\omega. \tag{2.22}$$

In practice, ΔE^s (see Eq. (2.14)) is replaced by $\Delta E^{s'}$ (see Eq. (2.15)) and a discretization of the surface of the sphere is defined, e.g., in form of a regular grid with mid points λ_j and θ_j, according to

$$\Delta E_f^{s'} = \frac{1}{A_f} \sum_{j=1}^{J} a_j \Delta E^{s'}(\lambda_j,\theta_j), \tag{2.23}$$

with the corresponding discretization error $\delta\Delta E_f^{s'}$. In Eq. (2.23), a_j is the area fraction of each grid cell j and $\sum_{j=1}^{J} a_j = A_f$. The definition of, e.g., a regular grid can be made under different assumptions. Usually, one assumes that the equivalent water height within one grid cell is constant. It would be also possible to define polynomials or finite elements to describe the equivalent water heights within one grid cell. The smoother the equivalent water heights are mathematically represented within one grid cell, the smaller is the discretization error, i.e. the comission error.

Although striping is reduced after applying de-correlation filters (Swenson and Wahr, 2006, Klees et al., 2008, Kusche et al., 2009), correlated errors still exist even after spatial aggregation (Longuevergne et al., 2010, Sakumura et al., 2014). In the spectral domain the Gibb's effect occurs and should be taken care of, while a mismatch between the basin boundaries and the resolution of the gridded observations can be a source of errors in the spatial domain.

2.1.4 Error Estimation of TWSA

To estimate the errors of GRACE TWSA, a formal error propagation is applied to convert the uncertainties of the monthly potential coefficients and the static field (or temporal mean field), which are stored in their corresponding covariance matrices $\mathbf{\Sigma}_{xx}$ and $\mathbf{\Sigma}_{x'x'}$, to the covariance matrix of equivalent water heights $\mathbf{\Sigma}_{ll}$

$$\begin{aligned}
\mathbf{\Sigma}_{ll} &= \mathbf{FW}(\mathbf{\Sigma}_{xx} + \mathbf{\Sigma}_{x'x'})\mathbf{W}^T\mathbf{F}^T \\
&= \mathbf{FW} \begin{pmatrix} \sigma^2_{c_{10}} & \sigma_{c_{10}c_{11}} & \cdots & \sigma_{c_{10}c_{n_{\max}n_{\max}}} \\ & \sigma^2_{c_{11}} & \cdots & \sigma_{c_{11}c_{n_{\max}n_{\max}}} \\ & & \ddots & \vdots \\ \text{symm.} & & & \sigma^2_{c_{n_{\max}n_{\max}}} \end{pmatrix} \mathbf{W}^T\mathbf{F}^T \\
&= \begin{pmatrix} \sigma^2_{\Delta E_1} & \sigma_{\Delta E_1\Delta E_2} & \cdots & \sigma_{\Delta E_1\Delta E_n} \\ & \sigma^2_{\Delta E_2} & \cdots & \sigma_{\Delta E_2\Delta E_n} \\ & & \ddots & \vdots \\ \text{symm.} & & & \sigma^2_{\Delta E_n} \end{pmatrix}.
\end{aligned} \tag{2.24}$$

Here, each row of the Jacobian matrix $\mathbf{F}_i = \left(\frac{\partial \Delta E_i}{\partial c_{10}} \;\; \frac{\partial \Delta E_i}{\partial c_{11}} \;\; \dots \;\; \frac{\partial \Delta E_i}{\partial c_{n_{\max} n_{\max}}} \right)$ contains the partial derivatives of Eq. (2.12) with respect to each potential coefficient

$$\frac{\partial \Delta E_i}{\partial c_{nm}} = \frac{M}{4\pi R^2 \rho_w} \frac{(2n+1)}{(1+k'_n)} C_{nm}(\lambda_i, \theta_i), \tag{2.25}$$

$$\frac{\partial \Delta E_i}{\partial s_{nm}} = \frac{M}{4\pi R^2 \rho_w} \frac{(2n+1)}{(1+k'_n)} S_{nm}(\lambda_i, \theta_i). \tag{2.26}$$

The filter matrix $\mathbf{W}$ was introduced in Eq. (2.13) to reduce the noise in GRACE data. The actual computation of $\mathbf{W}$ is addressed in Jekeli (1981) for Gaussian smoothing and in Kusche (2007) for anisotropic filter. There exists no common consensus regarding the error covariance matrix of the static field $\mathbf{\Sigma}_{x'x'}$ in Eq. (2.24). The error variances are much smaller than the error variances of the monthly gravity field coefficients. In this thesis, the error information of the static field is considered.

Section 2.1.1.2 describes how the spherical harmonic coefficients of degree-1 and c_{20} from GRACE are usually replaced by external time series. Rietbroek et al. (2009) and Rietbroek et al. (2012) estimated a full error covariance matrix for the degree-1 coefficients based on measurements of GPS, GRACE and ocean bottom pressure. Furthermore, an estimation of the error variance of c_{20} is available (grace.jpl.nasa.gov/data/get-data/oblateness/). Therefore, the 3×3 block of the degree-1 error information in Eq. (2.24) is replaced by the external data, as well as the error variance of c_{20}. The covariances between the replaced lower degree coefficients and the spherical harmonic coefficients from GRACE are assumed to be zero.

During the data processing, background models are applied to reduce high frequency mass variations within other Earth system components. Errors in these models, e.g., in the tidal and non-tidal ocean-atmosphere de-aliasing products (Forootan et al., 2013, 2014a), are not included in the formal error propagation. Since such errors are neglected in this thesis, the total error estimation might still be underestimated.

2.2 Hydrological Model Simulations

To simulate large-scale continental hydrology, several land surface and hydrological water balance models have been developed. Land surface models basically represent the land-atmosphere interface in climate models and numerical weather prediction, and aim to represent the energy and water fluxes by implementing surface energy and water balance equations. In contrast, hydrological water balance models have mainly been designed to simulate river discharge, and are often used for water resources assessments and flooding predictions. They are of conceptual structure, i.e. even though the complex physical processes are often known, the model equations are reasonably simplified due to the lack of adequate forcing data sets (Kaspar, 2004). In the following, two models that both belong to the class of hydrological water balance models are described: a simple one-bucket model and the complex WGHM.

2.2.1 Simple One-Bucket Model

In this section, a simple linear model is introduced. The model consists of only one water storage S (the so-called "bucket" or total water storage TWS), which is filled by net precipitation P-E (precipitation minus evapotranspiration) and reduced by runoff (or discharge) R (see Fig. 2.3 a). The state of the storage S might be expressed in volume, e.g., m^3, and the fluxes P, E, and R as volume per time, e.g., m^3/day. Here, net precipitation and runoff are accumulated over one day and introduced as volume to the water storage S. The hydrological processes are described by two simple model equations (de Zeeuw, 1973). The storage of time step k is determined by adding net precipitation to the storage of the previous time step $k-1$ and reducing the runoff

$$S_k = S_{k-1} + (P-E)_k - R_k. \tag{2.27}$$

The runoff R_k depends linearly on the storage value S_{k-1}

$$R_k = KS_{k-1}, \tag{2.28}$$

which is shown in Fig. 2.3 b. The amount of runoff is controlled by the outflow coefficient K, which is dimensionless. The model might be applied to a river basin or to (sub-) catchments. By introducing the daily accumulated net precipitation P-E time series of

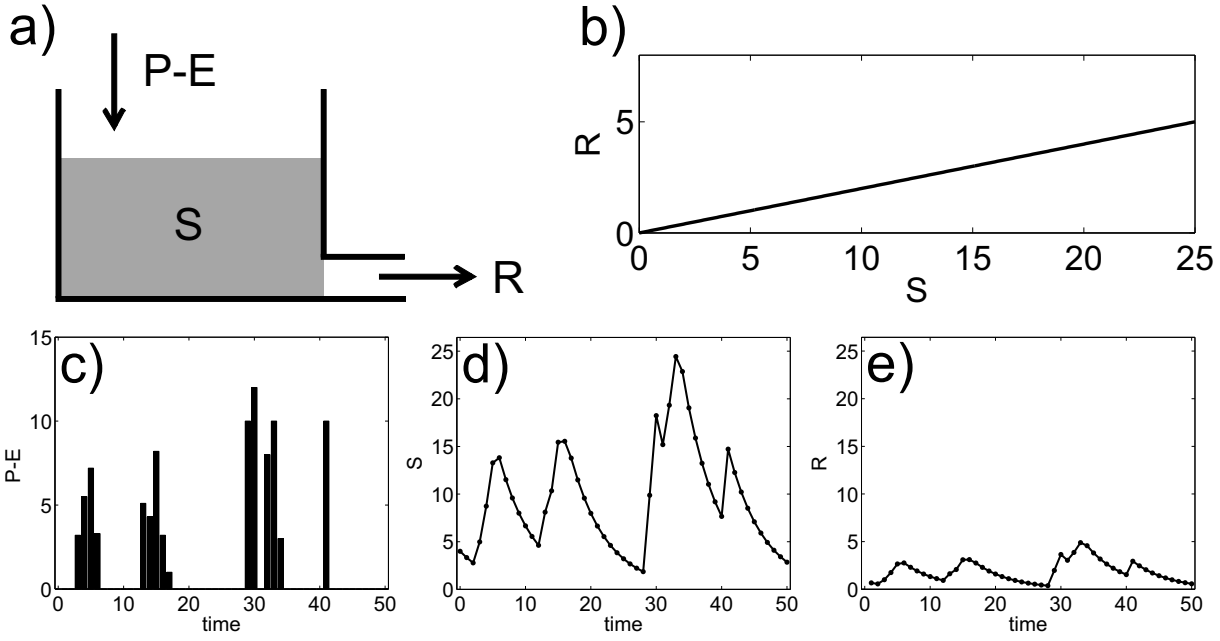

Figure 2.3: a) Scheme of a simple hydrological model, which consists of a one-bucket storage S. b) The relation between the water storage compartment S (in m^3) and daily accumulated outflow R (in m^3) is shown, which is linear. By defining the model input, i.e. daily accumulated net precipitation $P-E$ (in m^3) as in c), the storage S (in m^3) and daily accumulated outflow R (in m^3) change over time as shown in d) and e), respectively.

Fig. 2.3 c, the time evolution of the water storage S (Fig. 2.3 d) and the daily accumulated runoff R (Fig. 2.3 e) are determined. Thus, by definition and as it can be seen in Fig. 2.3 d and e, it is clear that S and R are correlated with a correlation coefficient of 1.0. The

motivation to select such a simple model is to better illustrate the steps of the assimilation methods of this thesis (see chapter 4). The time step and units for S, P-E, and R have been chosen arbitrarily, since the simple model is used to describe the mathematical details.

2.2.2 The WaterGAP Global Hydrology Model (WGHM)

The global model of water availability and water use WaterGAP (Water-Global Assessment and Prognosis) consists of WGHM and the groundwater and surface water use model (GWSWUSE; Alcamo et al., 2003, Döll et al., 2003, Müller Schmied et al., 2014). WGHM simulates water storage changes and fluxes for all continents excluding Antarctica on a $0.5°\times0.5°$ grid, i.e. 66896 grid cells in total, with a daily time step. Water storage changes in ten individual compartments, i.e. canopy, snow, soil, groundwater, local wetlands, global wetlands, local lakes, global lakes, global reservoirs and rivers are calculated for each grid cell. Local lakes and wetlands receive only runoff from the specific cell, while global surface water bodies as well as rivers receive inflow from the upstream grid cells, too. Human water use for irrigation, livestock, households, manufacturing and cooling of thermal power plants is simulated by separate sub-models. GWSWUSE defines whether human water abstraction is taken from surface water bodies or groundwater. This is then considered in the WGHM simulations. The model is calibrated against mean annual river discharge at 1319 stations by adjusting the runoff coefficient γ for all grid cells in each calibration basin. In the case that the simulated discharge deviates more than 1 % from the observed discharge, up to two additional correction factors are applied (Hunger and Döll, 2008, Müller Schmied et al., 2014).

2.2.2.1 Overview of the WGHM Structure

An overview of the WGHM structure is given in Fig. 2.4. The vertical water balance describes the transport of water through the canopy, snow, soil, and groundwater compartment, as well as the partitioning of precipitation into evapotranspiration and runoff. Water transport as runoff from the land area is partitioned into fast surface and subsurface runoff, which flows directly into the surface water bodies, and groundwater recharge. The latter first flows into the groundwater and subsequently as groundwater outflow into surface water bodies. In addition, precipitation over surface water is added to the lake, wetland, reservoir and river compartments, while evaporation reduces the amount of water. The river compartment is the final storage of the grid cells. The outflow for each cell, and thus the inflow of the global lake, reservoir and wetland or river compartment of the next cell is directed laterally on the basis of the global Drainage Direction Map (DDM30; Döll and Lehner, 2002). Furthermore, the impact of human water use as simulated by WaterGAP water use sub-models is taken into account in WGHM. Net water use (water abstractions minus return flows) are abstracted from surface water bodies (including rivers) or groundwater (Döll et al., 2012).

WGHM is a conceptual model with significant complexity that accounts for a variety of known hydrological processes as described in this section. In the following, the climate forcing fields that drive WGHM are introduced (section 2.2.2.2). Then, the model equations

describing the vertical water balance (section 2.2.2.3) and the horizontal water balance, as well as the lateral routing are presented (section 2.2.2.4). An overview of the simulation of anthropogenic water use and the model calibration procedure are presented in sections 2.2.2.5 and 2.2.2.6, respectively. Finally, sources of uncertainties in the simulation of water storage changes and fluxes are described (section 2.2.2.7).

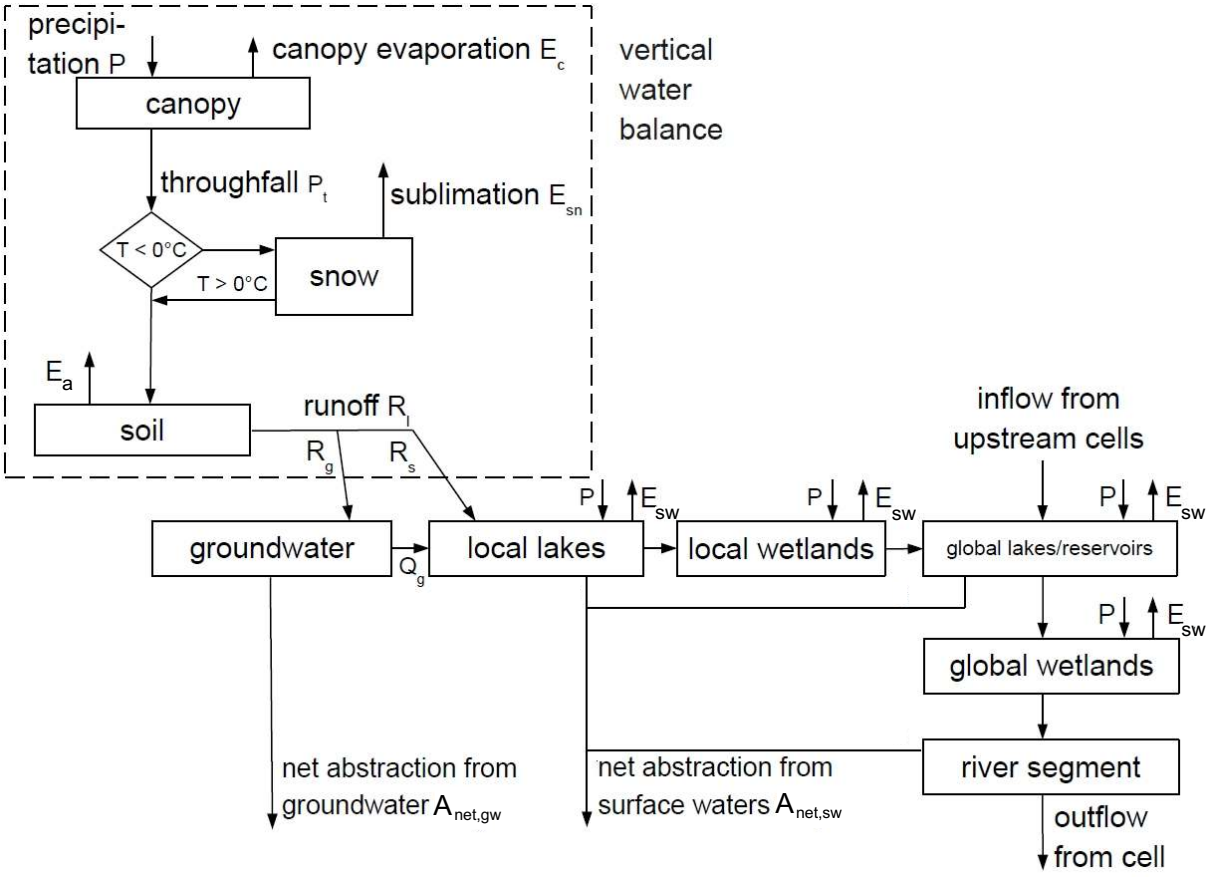

Figure 2.4: Schematic structure of WGHM water storage compartments and fluxes that are computed within each 0.5°×0.5° grid cell (Müller Schmied et al., 2014, p. 3530). Water storage compartments are represented by boxes and water fluxes by arrows.

2.2.2.2 Climate Forcing

WGHM can be forced by several climate input data sets. In this thesis, monthly time series of the number of wet days in month, temperature and cloudiness were used from the data set CRU TS 3.2 (Climate Research Unit's Time Series; Harris et al., 2013), whereas monthly precipitation input fields were taken from the GPCC (Global Precipitation Climatology Centre data, version 6) precipitation monitoring product (Schneider et al., 2014). In WGHM, precipitation values are equally partitioned into the number of wet days in month, while the wet days are distributed using a first order Markov chain algorithm. Daily short- and long-wave radiation is determined from the cloudiness information. Alternatively, daily time series of precipitation, temperature, short- and long-wave radiation from the WFDEI meteorological data set (WATCH Forcing Data methodology

applied to ERA-Interim data; Weedon et al., 2014) were used in this thesis. The impact of using these two different climate input data sets on water flows and water storage changes as computed by WaterGAP 2.2 has been reported in Müller Schmied et al. (2014, 2016).

2.2.2.3 Vertical Water Balance

The description of the vertical water balance in this section and the horizontal water balance in section 2.2.2.4 follows those in Döll et al. (2003) and Müller Schmied et al. (2014).

Canopy Water Balance

Daily changes in the canopy storage S_c (mm) over time t are determined as the difference between incoming daily precipitation P (mm/day), throughfall P_t (mm/day), and evaporation E_c (mm/day)

$$\frac{dS_c}{dt} = P - P_t - E_c. \tag{2.29}$$

In this thesis, a precipitation multiplier (parameter with identification number IN=22 in Tab. 2.1) was introduced. The maximum amount of water that can be stored in canopy $S_{c,\max}$ (mm) is defined by the maximum canopy water height m_c (mm, IN=10 in Tab. 2.1) and the leaf area index I_{LAI}

$$S_{c,\max} = m_c \cdot I_{\mathrm{LAI}}, \tag{2.30}$$

where the latter depends on a biomass multiplier and a specific leaf area multiplier (IN=11 in Tab. 2.1). By defining the difference between the maximum canopy water storage $S_{c,\max}$ and the current canopy water storage S_c as canopy deficiency $S_{c,\mathrm{def}} = S_{c,\max} - S_c$, the throughfall P_t is calculated as

$$P_t = \begin{cases} 0 & \text{if} \quad P \leq S_{c,\mathrm{def}} \\ P - S_{c,\mathrm{def}} & \text{if} \quad P > S_{c,\mathrm{def}} \end{cases}. \tag{2.31}$$

In addition, canopy evaporation E_c reduces the storage (Deardorff, 1978) as

$$E_c = E_p \left(\frac{S_c}{S_{c,\max}} \right)^{\frac{2}{3}}. \tag{2.32}$$

Herein, E_p denotes the potential evapotranspiration (mm/day), which is determined using the Priestley-Taylor equation (Priestley and Taylor, 1972) according to

$$E_p = \alpha_{\mathrm{PT}} \frac{\Delta}{\Delta + \gamma_{\mathrm{psy}}} (R_n - G). \tag{2.33}$$

In Eq. (2.33), Δ is the saturation vapor pressure, γ_{psy} denotes the psychrometer constant, G labels the soil heat flow, α_{PT} is the empirical Priestley-Taylor coefficient, and R_n characterizes the net radiation. The Priestley-Taylor coefficient is defined for humid and arid

Table 2.1: WGHM parameters and their properties that are calibrated within the calibration and data assimilation (C/DA) framework of this thesis. The identification number of each parameter is shown under "IN", while "mode" represents the value that is used in WaterGAP version 2.2, and under "limits" the spread of parameter values used for ensemble generation are summarized (see section 5.2.3). To generate ensembles of parameters, either triangular or uniform distributions were assumed, indicated in the first column by "$^{\triangle}$" and "$^{\circ}$", respectively. Units of parameters are given in the second column. Parameters, marked with "*", are not integrated in the original WaterGAP 2.2 version but are extra parameters that are calibrated in the C/DA framework in this thesis. Based on the experiences in the sensitivity and covariance study (chapter 6), the triangular distributions of parameters IN=4 and IN=19 were modified (see section 6.2.1).

IN	Calibration Parameter	Mode	Limits
$1^{*\triangle}$	root depth multiplier	1	[0.5 2.0]
$2^{*\triangle}$	river roughness coefficient multiplier	1	[0.5 2.0]
$3^{\triangle}$	lake depth (m)	5	[1 20]
$4^{\triangle}$	wetland depth (m)	2	[0.5 5]
$5^{\triangle}$	surface water outflow coefficient (day^{-1})	0.01	[0.001 0.1]
$6^{*\triangle}$	net radiation multiplier	1	[0.5 2.0]
$7^{\triangle}$	Priestley-Taylor coefficient (humid)	1.26	[0.885 1.65]
$8^{\triangle}$	Priestley-Taylor coefficient (arid)	1.74	[1.365 2.115]
$9^{\triangle}$	maximum daily potential evapotranspiration (mm/day)	15	[7.25 22.5]
$10^{\triangle}$	maximum canopy water height per leaf area (mm)	0.3	[0.1 1.4]
$11^{*\triangle}$	specific leaf area multiplier	1	[0.5 2.0]
$12^{\triangle}$	snow freeze temperature (°C)	0	[-1.0 3.0]
$13^{\triangle}$	snow melt temperature (°C)	0	[-3.75 3.75]
$14^{*\triangle}$	degree day factor multiplier	1	[0.5 2.0]
$15^{\triangle}$	temperature gradient (°C/m)	0.006	[0.004 0.01]
$16^{*\triangle}$	groundwater recharge factor multiplier	1	[0.5 2.0]
$17^{*\triangle}$	maximum groundwater recharge multiplier	1	[0.5 2.0]
18°	critical precipitation for groundwater recharge (mm/day)	10	[2.5 20.0]
$19^{\triangle}$	groundwater baseflow coefficient (day^{-1})	0.006	[0.006 0.018]
$20^{*\triangle}$	net abstraction surface water multiplier	1	[0.5 2.0]
$21^{*\triangle}$	net abstraction groundwater multiplier	1	[0.5 2.0]
$22^{*\triangle}$	precipitation multiplier	1	[0.8 1.2]

regions (IN=7 and IN=8 in Tab. 2.1). In the WaterGAP version used in this thesis, a multiplier is applied to the net radiation (IN=6 in Tab. 2.1). A detailed description of the parameters can be found in Shuttleworth (1993) and Kaspar (2004).

Snow Water Balance

The throughfall P_t is stored in the form of snow if the daily temperature (°C) is below the snow freeze temperature T_f (°C, IN=12 in Tab. 2.1). Then, daily changes of the snow water storage S_{sn} (mm) are determined as

$$\frac{dS_{\mathrm{sn}}}{dt} = P_{\mathrm{sn}} - M - E_{\mathrm{sn}}, \tag{2.34}$$

where P_{sn} is precipitation falling as snow (mm/day), M is snow melt (mm/day), and E_{sn} denotes sublimation (mm/day). If the current temperature T is above the snow melt temperature T_m (°C, IN=13 in Tab. 2.1), snow melt is determined depending on a land cover specific degree-day factor D_f (mm/(day°C))

$$M = D_f(T - T_m). \tag{2.35}$$

In the WaterGAP version used in this thesis, a degree-day factor multiplier (IN=14 in Tab. 2.1) was introduced. Furthermore, the effect of elevation on the temperature (IN=15 in Tab. 2.1) is taken into account. The snow storage decreases due to sublimation E_{sn} (mm/day), which is calculated according to Eq. (2.33) but using snow specific parameter values. For the snow algorithm, a detailed description can be found in Hunger and Döll (2008).

Soil Water Balance

Changes of the soil storage S_s (mm)

$$\frac{dS_s}{dt} = P_{\text{eff}} - E_a - R_l \tag{2.36}$$

depend on the effective precipitation $P_{\text{eff}} = P_t - P_{\text{sn}} + M$ (mm/day), the actual evaporation from soil E_a (mm/day), and the runoff from landside R_l (mm/day). The soil storage represents the water content in the effective root zone of the ground covering plants. Its maximum capacity (mm) is determined as

$$S_{s,\max} = m_{\text{droot}} d_{\text{root}} C_s, \tag{2.37}$$

based on the root depth multiplier m_{droot} (IN=1 in Tab. 2.1), as well as the root depth d_{root} (mm) and the water capacity of the root zone C_s, which are associated with a specific land-cover class. The actual evaporation of water from soil (mm/day)

$$E_a = \min((E_p - E_c), (E_{p,\max} - E_c)\frac{S_s}{S_{s,\max}}) \tag{2.38}$$

depends on the difference between the maximum daily potential evapotranspiration $E_{p,\max}$ (mm, IN=9 in Tab. 2.1), the canopy evaporation E_c, and the degree of saturation of the soil water storage. Its value is limited by the potential evapotranspiration E_p reduced by canopy evaporation. The runoff from landside R_l depends on the effective precipitation, the degree of saturation of the soil water storage and the runoff coefficient γ

$$R_l = \begin{cases} 0 & \text{if} \quad S_s = 0 \\ P_{\text{eff}}(\frac{S_s}{S_{s,\max}})^{\gamma} & \text{if} \quad 0 < S_s < S_{s,\max} \\ P_{\text{eff}} & \text{if} \quad S_s = S_{s,\max} \end{cases}. \tag{2.39}$$

The parameter γ has already been calibrated in the WaterGAP 2.2 version used in this thesis (see section 2.2.2.6). It is not considered in the calibration of WGHM parameters in the C/DA framework that is developed in this thesis.

Groundwater Balance

The groundwater storage S_g (mm) is increased by groundwater recharge R_g (mm/day), and reduced by groundwater baseflow Q_g (mm/day) and net abstractions from groundwater $A_{\text{net,gw}}$ (mm/day) according to

$$\frac{dS_g}{dt} = R_g - Q_g - A_{\text{net,gw}}. \tag{2.40}$$

Herein, the amount of groundwater recharge is a fraction of the total runoff from landside R_l in Eq. (2.39) controlled by the groundwater recharge factor f_g, which accepts values between zero and one, and is limited by the maximum groundwater recharge $R_{g,\max}$ (mm) as

$$R_g = \min(R_{g,\max}, f_g \cdot R_l). \tag{2.41}$$

Multipliers for the groundwater recharge factor (IN=16 in Tab. 2.1) and the maximum groundwater recharge (IN=17 in Tab. 2.1) were introduced to be calibrated in this thesis.

In arid regions, groundwater recharge only occurs if a critical precipitation value is exceeded (mm/day, IN=18 in Tab. 2.1). A detailed description can be found in Döll and Fiedler (2008). The groundwater baseflow Q_g (mm/day) is regulated by the groundwater baseflow coefficient k_g (1/day, IN=19 in Tab. 2.1) as

$$Q_g = k_g S_g. \tag{2.42}$$

Net abstraction from groundwater $A_{\text{net,gw}}$ is simulated in the global water use model GWSWUSE and can take positive or negative values (see section 2.2.2.5). In this thesis, a net abstraction from groundwater multiplier was introduced in the C/DA (IN=21 in Tab. 2.1). While the minimum water state of all other storage compartments is limited by zero, the water state of the groundwater storage can always be reduced. In case that the sum of groundwater baseflow and anthropogenic water use is larger than groundwater recharge, depletion of groundwater is simulated by allowing the groundwater storage to take negative values. Then, the groundwater baseflow Q_g is zero.

2.2.2.4 Horizontal Water Balance

Surface Water Balance

Changes in surface water storage S_{sw} (m^3), e.g., lake or wetland, are determined as

$$\frac{dS_{\text{sw}}}{dt} = P + I - E_{\text{sw}} - Q_{\text{sw}} - A_{\text{net,sw}}. \tag{2.43}$$

Precipitation P (m^3/day) over open surface water and inflow I (m^3/day) from other storage compartments, for example groundwater baseflow, as well as from upstream grid cells in case of global lakes and wetlands (see Fig. 2.4) increase the storage. Potential evaporation of open water surfaces E_{sw} (m^3/day) is calculated similar to Eq. (2.33), but with parameter values for open water bodies, and reduced from S_{sw}. In addition, surface

water outflow Q_{sw} (m^3/day) and net abstraction from surface water $A_{\mathrm{net,sw}}$ (m^3/day) are reduced. The maximum water storage of lakes and wetlands $S_{\mathrm{sw,max}}$ (m^3) is the product of their constant area F (m^2) and a constant maximum depth h_{sw} (m), which correspond to IN=3 and IN=4 in Tab. 2.1, respectively,

$$S_{\mathrm{sw,max}} = F h_{\mathrm{sw,max}}. \tag{2.44}$$

WaterGAP 2.2 does not consider shrinking or expansion of water body areas. Instead, a reduction parameter is introduced, which reduces evaporation if the fill level of the surface water storage is low (for details see Hunger and Döll, 2008). The outflow of local surface water bodies $Q_{\mathrm{sw,loc}}$ (m^3/day) depends on the degree of saturation of the storage and the surface water outflow coefficient k_{sw} (1/day, IN=5 in Tab. 2.1) as

$$Q_{\mathrm{sw,loc}} = k_{\mathrm{sw}} S_{\mathrm{sw}} \left(\frac{S_{\mathrm{sw}}}{S_{\mathrm{sw,max}}}\right)^{\frac{3}{2}}. \tag{2.45}$$

In case of global lakes and wetlands, $Q_{\mathrm{sw,glo}}$ is simulated as linear outflow similar to the groundwater baseflow in Eq. (2.42). Net abstraction from surface water $A_{\mathrm{net,sw}}$ is simulated in the global water use model (GWSWUSE, see section 2.2.2.5). In this thesis, a net abstraction from surface water multiplier was introduced (IN=20 in Tab. 2.1).

River Routing

The outflow from the grid cells, and thus the inflow of global lakes, reservoirs and wetlands, as well as the river compartment of the next grid cell is routed through the stream network based on the global map DDM30 (Fig. 2.5). In contrast to groundwater baseflow Q_g in Eq. (2.42), fast surface runoff R_s (mm/day), which is a fraction of the total runoff from landside $R_s = R_l - R_g$, is directed to the surface water storage without delay. The outflow Q_r (m^3/day) from the river storage S_r (m^3) is defined as

$$Q_r = \frac{v_r}{s} S_r, \tag{2.46}$$

where its magnitude depends on the river velocity v_r (m/day) and the distance s (m) between neighboring cells. The river velocity depends on the river bed roughness, for which a multiplier was introduced in this thesis (IN=2 in Tab. 2.1).

2.2.2.5 Anthropogenic Water Consumption

The global water use model simulates anthropogenic water consumption from surface water bodies and groundwater (Döll et al., 2012). For this, five separate sub-models are used to estimate water use for irrigation (Döll and Lehner, 2002, Siebert et al., 2005), livestock, households, and manufacturing (Flörke et al., 2013), as well as cooling of thermal power plants (Vassolo and Döll, 2005, Flörke et al., 2013). In the sub-models, the source of water abstraction is not specified. The values for net abstraction from groundwater $A_{\mathrm{net,gw}}$ (in Eq. (2.40)) or from surface water $A_{\mathrm{net,sw}}$ (in Eq. (2.43)) are then determined in the GWSWUSE model and are used as inputs for the WGHM simulations.

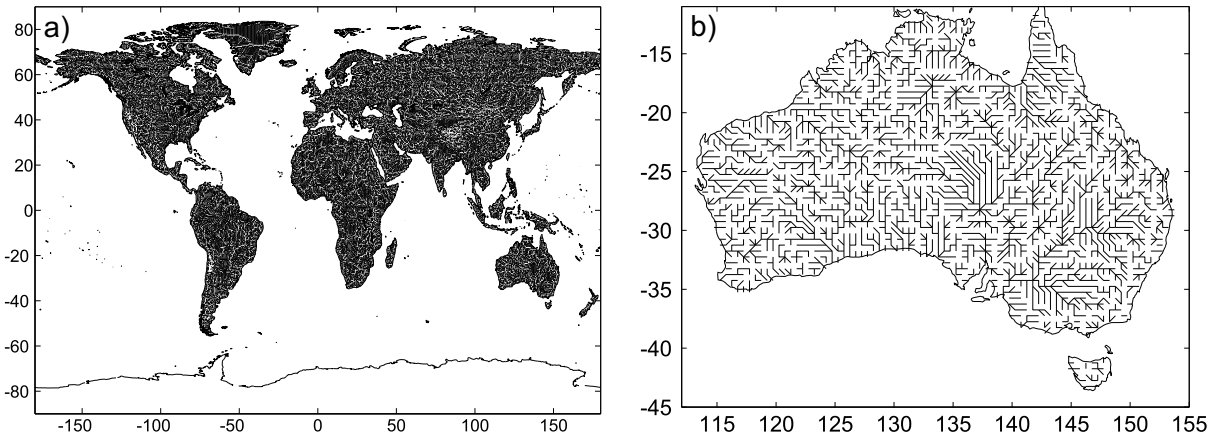

Figure 2.5: a) The Drainage Direction Map (DDM30; Döll and Lehner, 2002) is defined over the globe, from which a zoom over Australia is shown in b).

2.2.2.6 Model Calibration

WGHM is calibrated against mean annual river discharge at 1319 stations in 724 river basins (Fig. 2.6) based on measurements provided by the Global Runoff Data Centre (GRDC; Bundesanstalt für Gewässerkunde (BfG), Koblenz, Germany). The runoff coefficient γ in Eq. (2.39) was adjusted homogeneously for all grid cells in each calibration basin (Döll et al., 2003). The coefficient is allowed to take values between 0.1 and 5 to ensure realistic simulations of soil dynamics. In case that the simulated long-term averaged discharge still deviates more than 1 % from the observed discharge after calibration, up to two additional correction factors can be applied (Hunger and Döll, 2008). These differences might be caused by a strong impact of processes on the basin hydrology that are not taken into account in WGHM and by errors in climate input data (Döll et al., 2003). The calibration was performed using all available data during the 30-year period 1971-2000, for which the climate input of WGHM is most reliable (Müller Schmied et al., 2014).

2.2.2.7 Uncertainties of Model Simulations

Uncertainties of input data, imperfect parameters, errors in the model structure, discretization errors, and heterogeneity within a grid cell are the major sources of uncertainty in hydrological model simulations and might result in large differences between model simulations and observed values (Döll et al., 2003). It is not trivial to quantify these errors and determine errors of the model outputs. In this thesis, a Monte-Carlo approach is applied to empirically estimate the uncertainties of WGHM simulations (see section 5.2.3).

2.3 Supplementary Observation Data

Various in-situ and remotely-sensed observations of individual water storage compartments can be used to assess the accuracy of WGHM simulations. These include snow

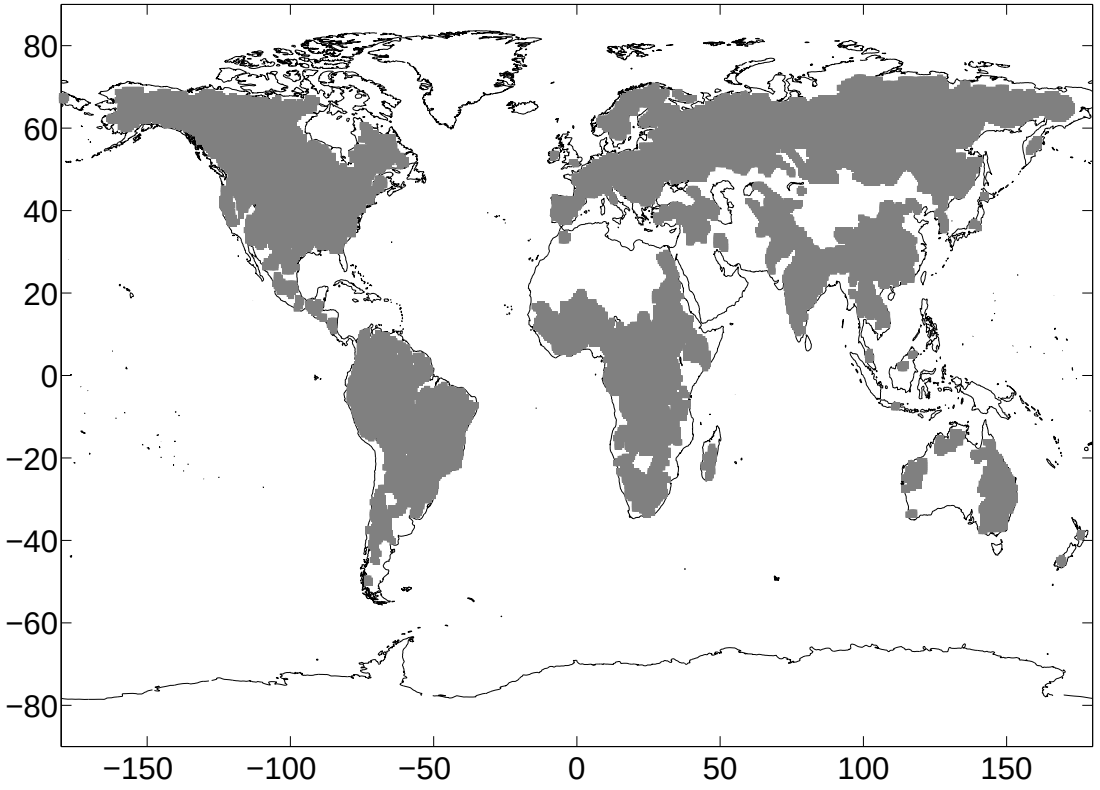

Figure 2.6: The runoff coefficient has been calibrated for the 724 gray shaded river basins.

water, soil moisture, groundwater, surface water body levels and extent, as well as water fluxes, e.g., river discharge and evapotranspiration. In the following, the data sets used for validation in this thesis are described.

In-situ Soil Moisture

Daily in-situ volumetric soil moisture data (in m^3m^{-3}) were obtained from the Soil Climate Analysis Network (SCAN; http://www.wcc.nrcs.usda.gov/scan/) over the entire United States, from which 69 observation sites are located in the Mississippi River Basin (Schaefer et al., 2007), which serves as test region in this thesis. Measurements are provided at depths of 5, 10, 20, 51 and 102 cm. A total of 29 stations are analyzed in this study after excluding stations (i) for which no information about the bottom depth is available (13 stations), (ii) that have no or little data during the study period (22 stations), and (iii) for which the time series contain large jumps (possibly due to an exchange of the sensor; 5 stations). In this thesis, remotely-sensed soil moisture was not considered, since these measurements represent changes in a thin upper soil layer (up to ∼5 cm depth), and such a layer is not defined in WGHM. The uppermost measurement is assumed to represent volumetric soil moisture from 0 to 7.5 cm, the second measurement from 7.5 to 15 cm, the third from 15 to 35.5 cm, the fourth from 35.5 to 76.5 cm, and the fifth from 76.5 cm to the bottom of the soil zone. To define the bottom depth, the deepest part of the soil horizons was used, which is provided by the National Soil Survey Center (NSSC, http://ncsslabdatamart.sc.egov.usda.gov/). The definition of the soil layers is based on the measurement depths and not necessarily due to, e.g., changes

in soil (layer) properties. To convert soil moisture into equivalent water heights, the observed values were multiplied by the height of the corresponding observation layers. The sum of all layers defines the equivalent water height of the total soil water storage.

Groundwater Well Measurements

Daily depths of groundwater for 58 wells located in the Mississippi River Basin were obtained from the US Geological Survey (http://www.usgs.gov/water/) and Illinois State Water Survey (http://il.water.usgs.gov/). The observations were processed by Rodell et al. (2007) and provided for this study. Time series of groundwater storage anomalies were derived by taking into account specific yield estimates for the well locations, i.e. factors to convert groundwater tables to equivalent water heights were applied, which depend on the soil properties. The considered measurement wells were located in unconfined or semi-confined aquifers and were not influenced by nearby pumping or injections. Monthly basin-averaged time series were computed using the Thiessen polygon areas to weight the individual well time series for the whole Mississippi River Basin and its four major sub-basins.

In addition, groundwater tables for around 15800 wells within the Murray-Darling River Basin, which serves as the second test region of this thesis, were provided by Dr. Russell Crosbie (CSIRO Land and Water, Adelaide, Australia). The measurements were averaged over $1^\circ \times 1^\circ$ grid cells, including between one to around 2680 wells per grid cell. The locations of the individual measurement wells are reported in Tregoning et al. (2012, chapter 5.1) who evaluated groundwater estimates from GRACE. It was reported that these wells might be influenced by local effects such as pumping that might cause drawdown or recharge due to irrigation. A specific yield of 0.1 as a typical value for water aquifers was applied to convert groundwater tables to equivalent water heights as proposed by Tregoning et al. (2012). In this study, a specific yield map based on geological information is additionally applied (Johnson, 1967).

River Discharge

Monthly river discharge data are obtained from GRDC, from which values at 1319 stations world-wide were previously used to calibrate the river basin-specific runoff coefficients of WGHM. A total of 84 of these stations are located within the Mississippi River Basin and eleven stations are located in the Murray-Darling River Basin. Please note that the calibration was performed against mean annual river discharge, and that monthly variations of river discharge are evaluated after applying C/DA. Therefore, the discharge data can be used for validation.

Surface Water Extent

The Global Inundation Extent from Multi-Satellite (GIEMS) data set globally describes the area that is covered by surface water, i.e. lakes, wetlands, rivers, irrigated agriculture, as well as episodic and seasonal inundation on an equal area grid (Papa et al., 2008). The spatial resolution of these maps is 773 km^2, corresponding to $0.25^\circ \times 0.25^\circ$ at the equator.

The monthly data are available from 01/1993 to 12/2007. Grid cells that are covered by snow or ice, located at the coast, or not subject to inundation are not considered in these maps. The data set is based on a combination of multi-satellite observations of surface water extent, water levels from the altimetry satellite mission Topex/Poseidon, and in-situ hydrographic measurements (for details see e.g., Papa et al., 2006, Prigent et al., 2007, Papa et al., 2010). In this thesis, spatial averages of the gridded surface water extent data set for the entire Mississippi River Basin and its four major sub-basins were determined.

3. Data Assimilation

In this chapter, the mathematical foundations of variational and sequential data assimilation techniques for both linear and non-linear models are described. Therefore, first, statistical foundations are introduced in section 3.1. Then, the Bayes' theorem that relates the probability density functions (PDF) of observations and unknown parameters is introduced in section 3.2. Section 3.3 addresses the variational data assimilation approach, including the formulation for linear (section 3.3.1) and non-linear models (section 3.3.2). This is followed by an introduction to sequential data assimilation in section 3.4. The traditional Kalman filter (KF) equations are derived for combining states from linear models with observations (section 3.4.1). Based on these equations, ensemble methods are introduced to deal with non-linear models (section 3.4.2), including the ensemble Kalman filter (EnKF) and its variants, as well as non-linear ensemble methods. Methods that are able to deal with non-linear models are of interest to this PhD thesis since the model equations of the WaterGAP Global Hydrology Model (WGHM, see section 2.2.2) are non-linear. Since many variants of EnKF have been developed to make the original approach more efficient or to adapt it to a specific application, in this chapter only a brief overview will be provided. Then, in chapter 4 the algorithms of several selected variants will be described in detail. A strategy for simultaneous parameter estimation and sequential data assimilation is introduced in section 3.4.3. Finally, alternative sequential ensemble data assimilation methods used in this thesis are explained in section 3.4.4.

3.1 Statistical Foundations for Data Assimilation

The description of the statistical foundations in this section follows Koch (2007) and Evensen (2007).

Probability Density Function

The probability distribution function F defines the likelihood that a realization of a continuous random variable $\mathcal{X}$ takes a value less than or equal to x, thus

$$\mathrm{F(x)} = \int_{-\infty}^{\mathrm{x}} \mathrm{f}(t)\, dt. \tag{3.1}$$

Herein, $\mathrm{f}(t)$ denotes the PDF that indicates the likelihood that a random variable $\mathcal{X}$ takes a particular value x, and t is an integration variable. The PDF must satisfy the conditions that the probability for the random variable $\mathcal{X}$ to take a value x is positive or zero, and that its integral over the space of real numbers is equal to one

$$\mathrm{f(x)} \geq 0 \quad \text{and} \quad \int_{-\infty}^{\infty} \mathrm{f(x)}\, d\mathrm{x} = 1. \tag{3.2}$$

Most methods in data assimilation are optimal under the assumption of Gaussian distributions. The bell shape of this distribution is isotropic and can be fully described by the

first and second order statistical moments, i.e. the mean and variance, while higher order statistical moments are equal to zero or constant. The Gaussian distribution, also called normal distribution, is defined as

$$\mathrm{f(x)} = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(\mathrm{x}-\mu)^2}{2\sigma^2}\right), \tag{3.3}$$

in which μ denotes the mean value and σ the standard deviation. This can also be expressed as $\mathcal{X} \sim \mathcal{N}(\mu, \sigma^2)$.

Second Order Statistical Moments

The first statistical moment defines the expectation value $E[\mathcal{X}]$ of the random variable $\mathcal{X}$ as

$$E[\mathcal{X}] = \int_{-\infty}^{\infty} \mathrm{x}\,\mathrm{f(x)}\,d\mathrm{x} = \mu, \tag{3.4}$$

and specifies the average value of an infinite number of samples drawn from the PDF f(x). The second central moment defines the variance as

$$E[(\mathcal{X} - E[\mathcal{X}])^2] = \int_{-\infty}^{\infty} (\mathrm{x} - E[\mathcal{X}])^2\,\mathrm{f(x)}\,d\mathrm{x} = \sigma^2. \tag{3.5}$$

It describes the expected value of the squared deviations of x from its expectation value $E[\mathcal{X}]$. In case of two random variables $\mathcal{X}$ and $\mathcal{Y}$ with their PDFs f(x) and f(y), the joint probability is defined as f(x,y)=f(x|y)f(y)=f(y|x)f(x). This is the Bayes' theorem (Koch, 2007, p.31, Eq. (2.122)) in which f(x|y) denotes the conditional PDF of a random variable $\mathcal{X}$. The covariance of the random variables $\mathcal{X}$ and $\mathcal{Y}$ is defined as

$$E[(\mathcal{X} - E[\mathcal{X}])(\mathcal{Y} - E[\mathcal{Y}])] = \int\int_{-\infty}^{\infty} (\mathrm{x} - E[\mathcal{X}])(\mathrm{y} - E[\mathcal{Y}])\,\mathrm{f(x,y)}\,d\mathrm{x}\,d\mathrm{y} = \sigma_{\mathrm{x,y}}. \tag{3.6}$$

Statistical Moments from Samples

Usually, only a limited number N_e of samples, i.e. realizations from the PDF f(x), is available to determine a "best guess" of the expectation value in Eq. (3.4) as

$$\mu = E[\mathcal{X}] \simeq \overline{\mathrm{x}} = \frac{1}{N_e}\sum_{i=1}^{N_e} \mathrm{x}^i. \tag{3.7}$$

Herein, $\overline{\mathbf{x}}$ denotes the (ensemble) mean. The sample variance s^2 can be determined as

$$\sigma^2 = E[(\mathcal{X} - E[\mathcal{X}])^2] \simeq \mathrm{s}^2 = \frac{1}{N_e - 1}\sum_{i=1}^{N_e} (\mathrm{x}^i - \overline{\mathrm{x}})^2, \tag{3.8}$$

and the sample covariance $\mathrm{s}_{\mathrm{x,y}}$ can be calculated as

$$\sigma_{\mathrm{x,y}} = E[(\mathcal{X} - E[\mathcal{X}])(\mathcal{Y} - E[\mathcal{Y}])] \simeq \mathrm{s}_{\mathrm{x,y}} = \frac{1}{N_e - 1}\sum_{i=1}^{N_e} (\mathrm{x}^i - \overline{\mathrm{x}})(\mathrm{y}^i - \overline{\mathrm{y}}). \tag{3.9}$$

Sampling from Gaussian Distributions

In the following, vectors will be indicated by small letters and bold, and matrices will be indicated by capital letters and bold. If the mean vector $\boldsymbol{\mu}$ and the error covariance matrix $\boldsymbol{\Sigma}$ of a (multi-variate) Gaussian distribution are known, samples can be generated by, e.g., using random Monte Carlo sampling (Fishman, 1996). For this, a Cholesky decomposition of the error covariance matrix is performed according to $\boldsymbol{\Sigma} = \mathbf{G}\mathbf{G}^T$, where $\mathbf{G}$ denotes a regular lower triangular matrix. Following Koch (2007, p. 197, Eq. (6.13)), realizations $\mathbf{e}$ of a standard normal distributed random variable $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ can be generated and transformed to realizations $\mathbf{x}$ of the random variable $\mathcal{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ as

$$\mathbf{x} = \boldsymbol{\mu} + \mathbf{G}\mathbf{e}. \tag{3.10}$$

Alternatively, the minimum second order exact sampling (Appendix in Pham, 2001) can be applied to generate realizations of the random variable from an eigenvalue-decomposed error covariance matrix. It is used in the singular evolutive interpolated Kalman (SEIK) filter (Pham et al., 1998) and described in detail along with the filter equations in section 4.2.3.

In the following, the random variable $\mathcal{X}$ and its realization x will not be distinguished in notation, since in ensemble data assimilation the random variable is replaced by samples.

3.2 Bayes' Theorem for Data Assimilation

In this section a function that specifies how to combine model simulations and measurements, the so-called "cost function", is introduced. It is assumed that error information about model and observations is available in form of error covariance matrices and thus leads to a weighting in the cost function. Based on the uncertainty information the cost function is derived using the Bayes' theorem (see e.g., chapter 3 in Evensen, 2007, Nichols, 2010).

One can consider the prediction $\mathbf{x}^-$, e.g., of a model simulation, and the measurements $\mathbf{y}$ as two realizations of the true states stored in the $m \times 1$ vector $\mathbf{x}^{\text{true}}$, where m indicates the number of states:

$$\mathbf{x}^- = \mathbf{x}^{\text{true}} + \mathbf{q}^-, \tag{3.11}$$

$$\mathbf{y} = \mathbf{A}\mathbf{x}^{\text{true}} + \boldsymbol{\epsilon}, \tag{3.12}$$

In Eq. (3.12), $\mathbf{q}^-$ denotes the errors of the model prediction and $\boldsymbol{\epsilon}$ the measurement errors, as well as $\mathbf{A}$ is the $n \times m$ design matrix, which defines the relation between $\mathbf{x}^{\text{true}}$ and $\mathbf{y}$, and n is the number of measurements.

It is of interest to find an improved state estimate $\mathbf{x}^+$ based on the error information of model prediction and observations. For the time being, Gaussian distributions are assumed for the error terms, i.e. the mean values $\overline{\mathbf{q}^-}$ and $\overline{\boldsymbol{\epsilon}}$, denoted by the overline, of the model prediction and the observation error vectors are zero, while the error covariance matrices

$\mathbf{C}(\mathbf{x}^-) = \overline{\mathbf{q}^-(\mathbf{q}^-)^T}$ and $\mathbf{\Sigma}_{\text{yy}} = \overline{\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T}$ hold the second order statistical moments (variances and co-variances of predicted and observed states, respectively). Furthermore, the model states and observation errors are assumed to be uncorrelated, i.e. the expectation of their inner product is zero $\overline{\mathbf{q}^-\boldsymbol{\epsilon}^T} = \mathbf{0}$.

Based on the assumption of Gaussian distributions, the prior PDF $p(\mathbf{x})$ (e.g., Koch, 2007, p. 80) of the optimal model states $\mathbf{x}$ is given by

$$p(\mathbf{x}) \propto \exp[-\frac{1}{2}(\mathbf{x}-\mathbf{x}^-)^T\mathbf{C}(\mathbf{x}^-)^{-1}(\mathbf{x}-\mathbf{x}^-)], \tag{3.13}$$

and the likelihood function $p(\mathbf{y}|\mathbf{x})$ of the observations depending on the optimal model states (Koch, 2007, p. 86, Eq. (4.4)) is

$$p(\mathbf{y}|\mathbf{x}) \propto \exp[-\frac{1}{2}(\mathbf{A}\mathbf{x}-\mathbf{y})^T\mathbf{\Sigma}_{\text{yy}}^{-1}(\mathbf{A}\mathbf{x}-\mathbf{y})]. \tag{3.14}$$

Using the Bayes' theorem $p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{x})p(\mathbf{y}|\mathbf{x})$ (Koch, 2007, p.31, Eq. (2.122)), the posterior PDF $p(\mathbf{x}|\mathbf{y})$ of the optimal model states depending on the observations (Koch, 2007, p.81, Eq. (3.78)) results in

$$p(\mathbf{x}|\mathbf{y}) \propto \exp[-J(\mathbf{x})], \tag{3.15}$$

$$\text{with} \quad J(\mathbf{x}) = \underbrace{\frac{1}{2}(\mathbf{x}-\mathbf{x}^-)^T\mathbf{C}(\mathbf{x}^-)^{-1}(\mathbf{x}-\mathbf{x}^-)}_{J_1} + \underbrace{\frac{1}{2}(\mathbf{A}\mathbf{x}-\mathbf{y})^T\mathbf{\Sigma}_{\text{yy}}^{-1}(\mathbf{A}\mathbf{x}-\mathbf{y})}_{J_2}. \tag{3.16}$$

The cost function $J(\mathbf{x})$ is a linear combination of the residuals of the model forecast $\mathbf{x}^-$ (J_1) and the residuals of the measurements $\mathbf{y}$ (J_2) weighted by their error covariance matrices $\mathbf{C}(\mathbf{x}^-)$ and $\mathbf{\Sigma}_{\text{yy}}$. The solution $\mathbf{x}^+$ that maximizes the likelihood estimate in Eq. (3.15) for $\mathbf{x}$, i.e. that gives the maximum of the posterior PDF, is equivalent to the least squares solution that minimizes $J(\mathbf{x})$ in Eq. (3.16). The solution is the best linear unbiased estimator (BLUE), in which "best" is defined as the estimator with minimum variance. The cost function is used in the following sections to derive a formulation for the BLUE.

3.3 Variational Data Assimilation

It is assumed that a model simulation $\mathbf{x}^-$ is available with initial values $\mathbf{x}_0$ and predictions $\mathbf{x}_k^-$ at time steps t_k with $k = 1, \ldots, K$ (black points in Fig. 3.1). Additionally, for each time step observations $\mathbf{y}_k$ are available (black triangles in Fig. 3.1). In variational data assimilation, all available measurements are used simultaneously to correct the initial conditions of the model $\mathbf{x}_0^+$. The idea is to obtain the best overall fit of the corrected model values $\mathbf{x}_k^+$ (white points in Fig. 3.1) and the observations. For this, it is necessary to integrate the model forward and backward in time. In addition, adjoint code of the model is required to determine gradients for error estimation. In practice, time windows are defined and only observations within these windows are taken into account to estimate corrected initial model conditions.

Variational data assimilation is of particular interest when the behavior of a model system is very much under influence of the accuracy of initial conditions, e.g., in atmospheric modeling. Following this section, formulations of the cost functions for variational data assimilation when considering linear (section 3.3.1) and non-linear (section 3.3.2) model equations with and without model errors are presented. A detailed description on formulating the cost functions can be found in, e.g., Evensen (2007) or Talagrand (2010).

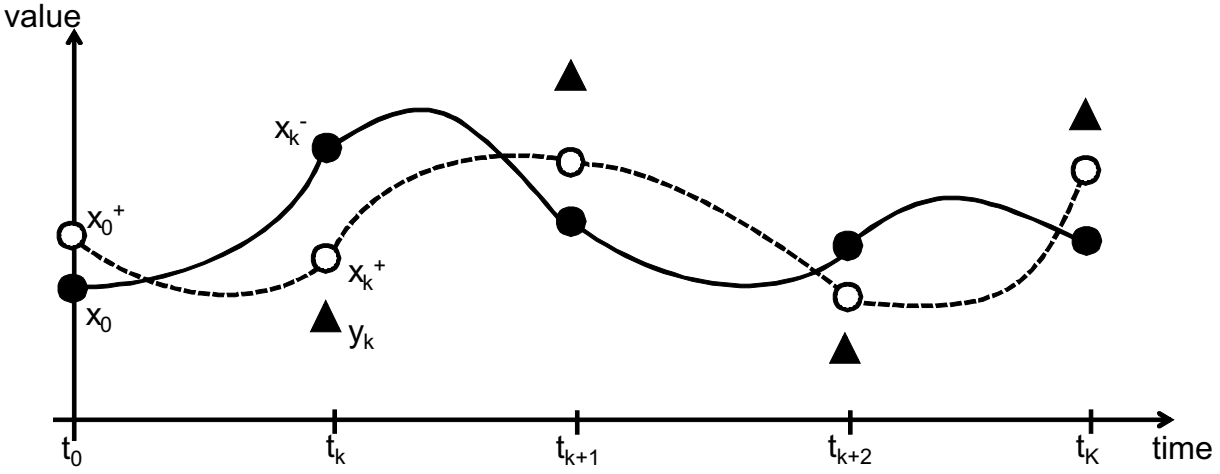

Figure 3.1: Scheme of variational data assimilation: All available data $\mathbf{y}_k$ (black triangles) are used simultaneously to improve the initial model conditions $\mathbf{x}_0$ and original model states $\mathbf{x}_k^-$ (black points and solid line). The corrected initial conditions $\mathbf{x}_0^+$ yield the best overall fit of model states $\mathbf{x}_k^+$ (white points and dashed line) and observations.

3.3.1 Linear Variational Problems

Sasaki (1970a) introduced two main approaches to implement variational data assimilation for linear models. These include the "strong" and "weak" constraint formulations. For the first, a perfect model is assumed, while the latter considers model errors.

Strong Constraint Formulation

First, it is assumed that the temporal evolution of model states can be described by a perfect linear discretized model, i.e. the model states $\mathbf{x}_k$ of time step t_k can be obtained by applying the linear model operator $\mathbf{F}_{k-1}$ to the model states $\mathbf{x}_{k-1}$ of the previous time step $k-1$, i.e.

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1}, \quad k = 1, \ldots, K. \tag{3.17}$$

Thus, in Eq. (3.17) model errors are ignored. Additionally, it is assumed that initial states $\mathbf{x}_0^-$ at time step $k = 0$ along with their error covariance matrix $\mathbf{C}(\mathbf{x}_0^-)$ are given and

that these errors are uncorrelated with the observation errors $\mathbf{\Sigma}_{\mathrm{yy}}$. The cost function in Eq. (3.16) is modified to

$$J(\mathbf{x}_0) = \underbrace{\frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^-)^T \mathbf{C}(\mathbf{x}_0^-)^{-1}(\mathbf{x}_0 - \mathbf{x}_0^-)}_{\text{Part 1}} + \underbrace{\frac{1}{2}\sum_{k=1}^{K}(\mathbf{A}_k\mathbf{x}_k - \mathbf{y}_k)^T \mathbf{\Sigma}_{\mathrm{yy}}^{-1}(\mathbf{A}_k\mathbf{x}_k - \mathbf{y}_k)}_{\text{Part 2}}, \tag{3.18}$$

in which Part 1 minimizes the residuals of the initial model conditions $\mathbf{x}_0^-$, since it is of interest to find their optimal values $\mathbf{x}_0^+$. This automatically leads to the best fit between the model states $\mathbf{A}_k\mathbf{x}_k$ and the observations $\mathbf{y}_k$ of each time step t_k (Part 2), because the model forward integration is error-free. Herein, the design matrix $\mathbf{A}_k$ relates the modeled states to the observations. This form of data assimilation is the so-called strong constraint variational problem (for details see Sasaki, 1970a, Talagrand, 2010).

Weak Constraint Formulation

In this version, it is still assumed that the temporal evolution of model states can be described by a linear discrete relation, but a random model error term $\mathbf{q}_{k-1}$ with covariance matrix $\mathbf{Q}_{k-1}$ is added to the equation

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{q}_{k-1}, \quad k = 1, \ldots, K. \tag{3.19}$$

This means that the model does not perfectly describe the relation between $\mathbf{x}_k$ and $\mathbf{x}_{k-1}$. In Eq. (3.19), the model errors $\mathbf{q}_{k-1}$ are assumed to be uncorrelated in time and uncorrelated with the observation errors and the errors of the inital values. When formulating the cost function, it is not only of interest to optimize the initial model states $\mathbf{x}_0$ (like in Eq. (3.18)) but it is also of interest to improve the sequence of model states $\{\mathbf{x}_k, k = 1, \ldots, K\}$

$$\begin{aligned} J(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_K) = {} & \underbrace{\frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^-)^T \mathbf{C}(\mathbf{x}_0^-)^{-1}(\mathbf{x}_0 - \mathbf{x}_0^-)}_{\text{Part 1}} + \underbrace{\frac{1}{2}\sum_{k=1}^{K}(\mathbf{A}_k\mathbf{x}_k - \mathbf{y}_k)^T \mathbf{\Sigma}_{\mathrm{yy}}^{-1}(\mathbf{A}_k\mathbf{x}_k - \mathbf{y}_k)}_{\text{Part 2}} \\ & + \underbrace{\frac{1}{2}\sum_{k=1}^{K}(\mathbf{x}_k - \mathbf{F}_{k-1}\mathbf{x}_{k-1})^T \mathbf{Q}_{k-1}^{-1}(\mathbf{x}_k - \mathbf{F}_{k-1}\mathbf{x}_{k-1})}_{\text{Part 3}}. \end{aligned} \tag{3.20}$$

Part 3 appears here in addition to Part 1 and 2 in Eq. (3.18), which minimizes the model errors $\mathbf{q}_{k-1}$. This form of data assimilation is the so-called weak constraint variational problem (Sasaki, 1970a).

A large number of methods exists to minimize the cost function of variational problems. Commonly used approaches to minimize $J(.)$ are the representer method (Bennett, 1992, 2002) and adjoint methods (Talagrand and Courtier, 1987, Courtier and Talagrand, 1987), including the three- or four-dimensional variational assimilation (3D/4D-Var; Sasaki, 1970a,b,c). Since this thesis focuses on sequential data assimilation algorithms, the reader is referred to the mentioned literature for more details.

3.3.2 Non-Linear Variational Problems

In the following, it is assumed that the time evolution of the model states is given by a non-linear discrete model, denoted by $f(.)$,

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{q}_{k-1}, \quad k = 1, \ldots, K, \tag{3.21}$$

while accounting for model errors $\mathbf{q}_{k-1}$. When minimizing the cost function

$$J(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_K) = \underbrace{\frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^-)^T \mathbf{C}(\mathbf{x}_0^-)^{-1}(\mathbf{x}_0 - \mathbf{x}_0^-)}_{\text{Part 1}} + \underbrace{\frac{1}{2}\sum_{k=1}^{K}(\mathbf{A}_k\mathbf{x}_k - \mathbf{y}_k)^T \mathbf{\Sigma}_{yy}^{-1}(\mathbf{A}_k\mathbf{x}_k - \mathbf{y}_k)}_{\text{Part 2}}$$
$$+ \underbrace{\frac{1}{2}\sum_{k=1}^{K}(\mathbf{x}_k - f(\mathbf{x}_{k-1}))^T \mathbf{Q}_{k-1}^{-1}(\mathbf{x}_k - f(\mathbf{x}_{k-1}))}_{\text{Part 3}}, \tag{3.22}$$

basically only Part 3 of Eq. (3.22) differs from the definition of the cost function for linear models in Eq. (3.20). Various methods are used to solve Eq. (3.22), e.g., the four-dimensional variational method (Rabier et al., 2000), the iterative representer method (Bennett et al., 1996, Evensen, 2007, p. 67f.), the linearized model operators for adjoint/backward equation (Evensen, 2007, p.71), genetic algorithms such as the adapted Metropolis algorithm (Metropolis et al., 1953), and the Hybrid Monte Carlo algorithm (Duane et al., 1987, Bennett and Chua, 1994). The solution of the listed methods is only optimal (in sense of the BLUE) when the model equations are linear. A more detailed description of variational data assimilation for non-linear models is given in Evensen (2007), Talagrand (2010), and references therein.

3.4 Sequential Data Assimilation

In sequential data assimilation, observations are used to correct the present states of a model as soon as they become available. In Fig. 3.2, the concept of sequential data assimilation is illustrated. The model run starts at time t_0 with initial conditions $\mathbf{x}_0$ and is integrated forward until time t_k, for which observations $\mathbf{y}_k$ are available (black triangles in Fig. 3.2). The measurements are directly incorporated to correct the model prediction $\mathbf{x}_k^-$ (black points in Fig. 3.2). The corrected values $\mathbf{x}_k^+$ (white points in Fig. 3.2) are subsequently used to start the next model forward integration. This procedure is repeated sequentially for each time step, at which observations are available. The approach is suitable when systems are driven by forcing fields, e.g., precipitation, temperature and other climate variables in hydrological models, and when initial conditions are less relevant.

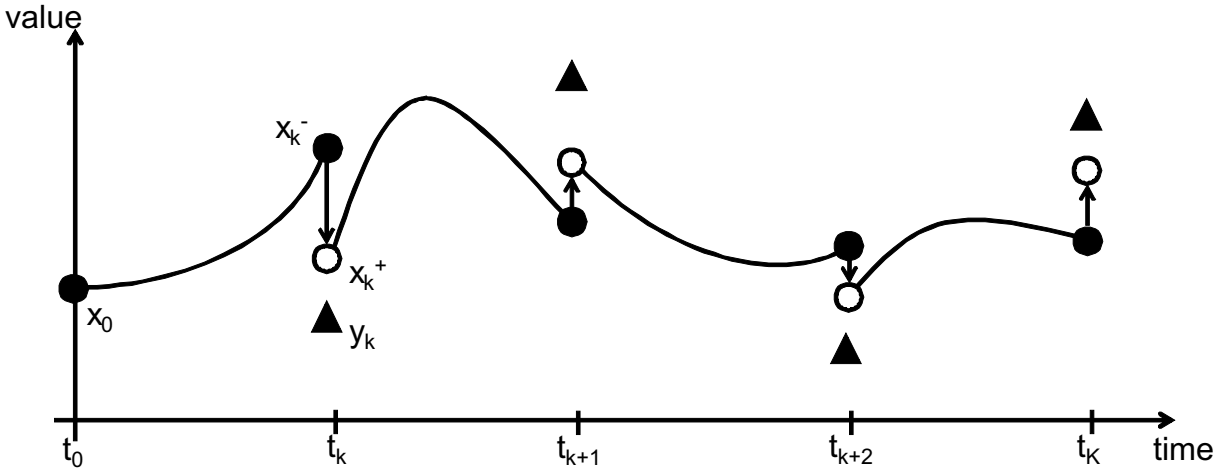

Figure 3.2: Scheme of sequential data assimilation: Observations (black triangles) are used as soon as they are available to improve the current model states $\mathbf{x}_k^-$ (black points and solid line). The updated model states $\mathbf{x}_k^+$ (white points) give the best fit to the observations at the current time step k.

3.4.1 Linear Problems: The Kalman Filter

The Kalman filter (KF; Kalman, 1960) provides a formulation to find the BLUE for the predicted model states for each time step at which observations are available. Here, the cost function in Eq. (3.16) is applied to one specific time step k

$$J(\mathbf{x}_k) = \underbrace{\frac{1}{2}(\mathbf{x}_k - \mathbf{x}_k^-)^T \mathbf{C}(\mathbf{x}_k^-)^{-1}(\mathbf{x}_k - \mathbf{x}_k^-)}_{\text{Part 1}} + \underbrace{\frac{1}{2}(\mathbf{A}_k\mathbf{x}_k - \mathbf{y}_k)^T \boldsymbol{\Sigma}_{\text{yy}}^{-1}(\mathbf{A}_k\mathbf{x}_k - \mathbf{y}_k)}_{\text{Part 2}}, \tag{3.23}$$

where Part 1 gives the optimal model condition at time step k that guarantees the best fit between model and observations at this specific time step (Part 2). Part 2 of Eq. (3.23) is similar to Part 2 of Eq. (3.18) and (3.20) with the small difference that it only focuses on the current time step. The optimal guess of the model states can be determined by minimizing the derivative of the cost function $J(\mathbf{x}_k)$ in Eq. (3.23) as

$$\frac{J(\mathbf{x}_k)}{d\mathbf{x}_k} = 2(\mathbf{x}_k - \mathbf{x}_k^-)^T \mathbf{C}(\mathbf{x}_k^-)^{-1} + 2(\mathbf{A}_k\mathbf{x}_k - \mathbf{y}_k)^T \boldsymbol{\Sigma}_{\text{yy}}^{-1} = 0. \tag{3.24}$$

In the update step of the KF, the optimal estimate $\mathbf{x}_k^+$ is given by the model prediction $\mathbf{x}_k^-$ plus a weighted linear combination of the model prediction and the measurements $\mathbf{y}_k$

$$\mathbf{x}_k^+ = \mathbf{x}_k^- + \mathbf{K}_k(\mathbf{y}_k - \mathbf{A}_k\mathbf{x}_k^-), \tag{3.25}$$

$$\text{with} \quad \mathbf{K}_k = \mathbf{C}(\mathbf{x}_k^-)\mathbf{A}_k^T(\mathbf{A}_k\mathbf{C}(\mathbf{x}_k^-)\mathbf{A}_k^T + \boldsymbol{\Sigma}_{\text{yy}})^{-1}, \tag{3.26}$$

where $\mathbf{K}_k$ is the Kalman gain matrix. The time evolution of the covariance matrix of the model prediction at time step $k-1$ to time step k is calculated by a formal variance error propagation as

$$\mathbf{C}(\mathbf{x}_k^-) = \mathbf{F}^T\mathbf{C}(\mathbf{x}_{k-1}^-)\mathbf{F} + \mathbf{Q}_{k-1}. \tag{3.27}$$

Herein, the transition matrix $\mathbf{F}$ relates the covariance matrices of time steps $k-1$ and k. In case of a perfect model as in Eq. (3.17), the term $\mathbf{Q}_{k-1}$ will be neglected. For an imperfect model as in Eq. (3.19), the model errors are reflected in the covariance matrix of the model prediction $\mathbf{Q}_{k-1}$. Considering Eq. (3.25), the covariance matrix of the updated model prediction $\mathbf{C}(\mathbf{x}_k^+)$ is determined by a formal variance error propagation, i.e.

$$\mathbf{C}(\mathbf{x}_k^+) = (\mathbf{I} - \mathbf{K}_k\mathbf{A}_k)\mathbf{C}(\mathbf{x}_k^-). \tag{3.28}$$

The model update $\mathbf{x}_k^+$ in Eq. (3.25) is the least squares solution of Eq. (3.24) and it is only optimal for linear models with Gaussian distributed model prediction and observation errors since the estimation considers only second order moments. For linear problems, the KF and the weak constraint variational method provide identical solutions at the final time step, when introducing prior error statistics and formulations consistently (Evensen, 2007, p. 50).

3.4.2 Non-linear Sequential Problems

In case of the KF, where linear model equations are evaluated and Gaussian error statistics are assumed, the time evolution of the PDF is fully described by the variance error propagation of the model covariance matrix $\mathbf{C}(\mathbf{x}_{k-1})$ to $\mathbf{C}(\mathbf{x}_k)$ (see Eq. (3.27)). In case of non-linear model equations, the full PDF needs to be evolved in time, which is described by the Fokker-Planck equation when assuming additive Gaussian model errors (e.g., Risken and Frank, 1996). The explicit solution of the Fokker-Planck equation is not possible for high dimensions. However, it is possible to derive statistical moments of the PDF. In case of linear model equations and Gaussian distributions as assumed in the KF, mean and covariance represent the full PDF of the model prediction. For non-linear models, the PDF is not fully characterized by the first and second order statistical moments. However, they are often used to represent valuable information on errors. Therefore, for non-linear model equations, the traditional KF approach is replaced by the extended Kalman filter (EKF; Kalman and Bucy, 1961) that uses an approximate linear equation for predicting the error covariance matrix (Evensen, 2007, p. 32) and that allows to handle non-linear model equations. The disadvantage of this filter choice is that it uses an approximation of the error statistics due to linearization and leads to large computational costs for the forward integration of the error covariance matrix. To overcome these problems, the EnKF has been developed (Evensen, 1994), which applies a Markov Chain Monte Carlo (MCMC) method (e.g., Gamerman and Lopes, 1997) to solve the Fokker-Planck equation (Evensen, 2009, p. 40). In this section, the focus is on ensemble filter methods, which can be interpreted as suboptimal KF in the sense that higher order statistical moments are neglected in the filter update.

Ensemble Kalman Filter

In ensemble methods, the PDF is estimated by an ensemble of N_e model realizations and evolves in time by evaluating the non-linear model equations of Eq. (3.21) for each ensemble member. Thus, the ensemble mean statistically represents the best estimate of $\mathbf{x}_k^-$ in Eq. (3.25) and the ensemble spread defines the error in the ensemble, i.e. it

approximates $\mathbf{C}(\mathbf{x}_k^-)$ in Eq. (3.27). Each of the model states represents one point in the n-dimensional state space. The cloud of points in the state space can be described using a PDF. The full state space and thus the full PDF $p(\mathbf{x})$ would be described with an infinite ensemble size, while in practice, it is approximated by a finite ensemble

$$\lim_{N_e \to \infty} \frac{dN_e}{N_e} = p(\mathbf{x}), \tag{3.29}$$

in which dN_e denotes the number of points in a small unit volume and N_e is the total number of points. In a one-dimensional state space, Eq. (3.29) can be visualized by a histogram. With knowledge about the PDF or the ensemble representation, statistical moments can be determined (Evensen, 2007, p. 39). In the update step of the EnKF, which implies that the PDF of model and observations are Gaussian, the model mean and error covariance matrix are approximated by the finite ensemble. The EnKF does not use higher order statistical moments. The update step of the EnKF is very close to that of the KF. However, the expression for the analytical error covariance matrix of the model prediction in Eq. (3.26) is replaced by an empirical approximation based on the model ensemble. The filter algorithm is described in detail in section 4.2.1.

Variants of the Ensemble Kalman Filter

In the following, a selection of various variants of the EnKF algorithm is described in order to give an overview of the large number of methods that have been developed to make the original approach more efficient (Tab. 3.1). Each of these modifications focused on a specific characteristic of the EnKF. Burgers et al. (1998) showed that an ensemble of observation errors needs to be introduced to keep the error statistics of the ensemble update unbiased (see section 4.2.1). To mitigate additional sampling errors that occur when generating an observation ensemble, several square root analysis schemes (SQRA) have been developed that update the ensemble mean instead of each ensemble member (e.g., Whitaker and Hamill, 2002, Tippett et al., 2003, and references therein). Evensen (2004) developed a straight forward square root variant of the classical EnKF, which is implemented in this thesis (see section 4.2.2). Besides a reduced rank SQRA Kalman filter was developed, in which the ensemble is forced to represent the most dominant modes of the error description (see, e.g., Verlaan and Heemink, 2001). Anderson (2001) noticed that the covariance structure of model states is destroyed when re-sampling of the update is required, such as in methods based on SQRA. Therefore, the ensemble adjustment Kalman filter (EAKF) was developed to preserve information about prior covariances between model states. This is also guaranteed when applying the classical EnKF. Hybrid variants of the EnKF with SQRA approaches have been applied, e.g., in Heemink et al. (2001) who introduced the partially orthogonal ensemble Kalman filter (POEnKF) and the complementary orthogonal subspace filter for efficient ensembles (COFFEE). Furthermore, ensemble transform Kalman filter (ETKF; Bishop et al., 2001) and error subspace statistical estimation (ESSE; Lermusiaux and Robinson, 1999) exist. The SEIK filter propagates the ensemble members in the same way as the original EnKF does (Pham et al., 1998). It differs in the generation of the initial ensemble, which is not completely random but relies on the main orthogonal directions of the model error covariance matrix (Drécourt, 2004). By applying the minimum second order exact sampling (see section 2.3) the

exact representation of the first and second order error statistics by the finite ensemble is guaranteed. In addition, the SEIK reformulates the Kalman gain matrix in such a way that the update is performed in the ensemble space instead of the observation space (see section 4.2.3). A detailed overview of EnKF variants and EnKF applications is given in Evensen (2004).

Table 3.1: Selection of ensemble filter and smoother approaches that are used for data-model fusion. Filters that have been implemented in this thesis are shown in bold.

Abbreviation	Ensemble Filter	Reference
EnKF	ensemble Kalman filter	Evensen (1994), Burgers et al. (1998)
SQRA	square root analysis scheme	e.g., Evensen (2004)
SQRA KF	reduced rank square root Kalman filter	e.g., Verlaan and Heemink (2001)
EAKF	ensemble adjustment Kalman filter	Anderson (2001)
POEnKF	partially orthogonal ensemble Kalman filter	e.g., Heemink et al. (2001)
COFFEE	complementary orthogonal subspace filter for efficient ensembles	e.g., Heemink et al. (2001)
ETKF	ensemble transform Kalman filter	Bishop et al. (2001)
ESSE	error subspace statistical estimation	Lermusiaux and Robinson (1999)
SEIK	singular evolutive interpolated Kalman filter	Pham et al. (1998)
EnKS	ensemble Kalman smoother	Evensen and van Leeuwen (2000)
PF	particle filter	Pham (2001)

Ensemble Kalman Smoother

The ensemble Kalman smoother (EnKS) was introduced by Evensen and van Leeuwen (2000) and can be interpreted as an extension of the EnKF (Evensen, 2007, p.130). In contrast to the EnKF, the EnKS uses the current observations not only to update the current model states but also to update model states backward in time. For this, the model prediction vector of the EnKF is extended with previous model states (see section 4.2.4). Therefore, the empirical error covariance matrix of the model states describes the covariances in space and time, while the error covariance matrix of the EnKF only holds the covariances in space. The EnKS might be limited to a maximum number of updates backward in time (Cohn et al., 1994).

Non-linear Ensemble Methods

The EnKF, EnKS, and their variants perform a linear combination of model predictions and observations in the update step, while considering the first and second order moments of the uncertainty information. In case of non-linear model equations, the consideration of the full PDF of model and observations in the update step is expected to more realistically

weight model predictions and measurements, and therefore to improve the data assimilation results (Evensen, 2007, p. 265). An example for a non-linear ensemble method is the particle filter (PF; Pham, 2001) that is already successfully applied to low-dimensional systems. The model forecast step of the PF is performed by evaluating the non-linear model equations with an ensemble of model states, such as in the EnKF/EnKS prediction step. In the update step of the PF, the full PDF of the model predictions and observations are considered to obtain corrected model updates. In contrast to EnKF/EnKS, the PDF of each ensemble member (the so-called "particle") is updated in the PF approach rather than the states of the particle. Thus, for non-linear problems the application of PF might be more desirable. However, so far the PF or other non-linear ensemble methods (see e.g., references in Evensen, 2007, p. 265) could not be practically used for high-dimensional systems as it is the case for this thesis.

3.4.3 Simultaneous C/DA

In hydrological modeling, it is common to calibrate basin-specific empirical model parameters that are usually assumed to be temporally constant. Some of these parameters describe physio-geographic characteristics, e.g., average lake depth, while other parameters appear as conceptual such as the groundwater outflow coefficient in WGHM. In data assimilation, the model ensemble prediction vector can be augmented by model parameters for a simultaneous calibration and data assimilation (C/DA) in the EnKF analysis step. Therefore, here the prediction vector $\mathbf{x}_k^-$ is composed of two parts

$$\mathbf{x}_k^- = \begin{pmatrix} \mathbf{w}_k^- \\ \mathbf{p}_k^- \end{pmatrix}, \tag{3.30}$$

in which $\mathbf{w}_k^-$ contains the model state values and $\mathbf{p}_k^-$ comprises the model calibration parameters. The latter typically cannot be observed, and they are therefore updated via the cross-correlations of model states and parameters. The update of model parameters in the EnKF is also called parameter estimation. In contrast to traditional parameter calibration (Gupta et al., 1998), the parameters are updated in each EnKF analysis step, and therefore, their values vary temporally. Schumacher et al. (2016a), for instance, showed how GRACE observations contribute in calibrating WGHM parameters (see chapter 6). This is especially the case whenever large correlations between model states and calibration parameters exist, as expected.

3.4.4 Discussion of the Choice of Sequential Data Assimilation Methods

In this thesis, the EnKF has been chosen to develop a new C/DA framework for integrating GRACE data into WGHM. The reason was threefold: (1) Variational methods require the integration of the model backward in time. In addition, these methods demand the implementation of adjoint models for uncertainty estimation, i.e. models that are used to obtain the gradient of the cost function $J(.)$ with respect to the initial conditions. These methods are usually chosen for atmospheric and oceanic data assimilation to estimate

optimal initial conditions. The dynamic processes are described using physical equations, i.e. nonlinear partial differential equations, which are numerically solved by applying, e.g., spectral methods or finite element methods. The stochastics of the system are represented by the physical and energetic equations which are, e.g., described by stochastically perturbed physics tendencies and stochastic kinetic energy backscatter schemes in ECMWF's integrated forecasting system (IFS; Shutts et al., 2007, Palmer et al., 2009). (2) In contrast, hydrological models are driven by forcing input fields including precipitation, temperature, long- and short-wavelength radiation. These boundary conditions are not perfectly known but subject to measurement and interpolation errors. In addition, uncertain model parameters highly influence the simulation accuracy. The stochastics lie in the forcing data and background information such as topography and land cover. Therefore, it is necessary to ingest observation data into the model as soon as they are available to constrain the current state of the system. Hence, sequential data assimilation is a suitable method in hydrological science. (3) Due to the simple form of the EnKF, ease of implementation, and affordable computational requirements (Evensen, 2007, p. 38), as well as already successful applications to oceanography, atmosphere, and recently to hydrology, the EnKF and two of its extensions, the SQRA according to Evensen (2004) and the SEIK filter (Pham et al., 1998), are implemented in this thesis and applied to a simple hydrological model in chapter 4 and to WGHM in chapters 7 and 8. In addition, the EnKS is implemented and applied to the simple hydrological model.

4. Ensemble Kalman Filter Methods

In this thesis, a new calibration and data assimilation (C/DA) framework is introduced. This framework has been designed based on the classical ensemble Kalman filter (EnKF; Evensen, 1994, Burgers et al., 1998, Evensen, 2003). After first successful applications (Schumacher, 2012, Eicker et al., 2014), the framework has been extended in Schumacher et al. (2016b) by the square root analysis scheme (SQRA; Evensen, 2004) and the singular evolutive interpolated Kalman filter (SEIK; Pham et al., 1998). The first motivation for considering variants of the filter algorithm is that the classical EnKF approach uses an observation ensemble that introduces an additional source of sampling errors to the algorithm (Evensen, 2004). Whitaker and Hamill (2002) showed that for small ensemble sizes the sampling errors decrease by applying SQRA (Tippett et al., 2003, and references therein) as will be shown later in this chapter. The second motivation is to reduce computation time in the update step. The update step of the SEIK filter is performed in the ensemble space instead of the observation space, unlike for the EnKF and SQRA methods. Therefore, especially the assimilation of large numbers of observations, i.e. much larger than the ensemble size, is usually better handled by the SEIK filter. Thus, the SEIK filter allows to establish a flexible framework for larger river basins and for integrating additional observations of, e.g., river discharge, lake level, soil moisture or snow water equivalent in future work. Additionally, the ensemble Kalman smoother (EnKS; Evensen and van Leeuwen, 2000) is introduced for smoothing possibly occurring jumps in the EnKF update due to introduced observations.

In this chapter, the two-step procedure of the C/DA is described: (i) the ensemble prediction step (section 4.1), i.e. the forward integration of the model for each ensemble member, which is basically independent of the applied filter algorithm, and (ii) the update (or analysis) step that merges model states and observations (section 4.2). The algorithms of EnKF (section 4.2.1) as well as SQRA (section 4.2.2), SEIK (section 4.2.3) and EnKS (section 4.2.4) are discussed and related to the EnKF. A range of tuning techniques for optimizing the C/DA results, e.g., improved initial sampling (Evensen, 2004), variance inflation factors (Hamill and Snyder, 2002) and localization (Houtekamer and Mitchell, 2001, Hamill et al., 2001) are addressed in section 4.3. Furthermore, throughout the chapter a simple example is used to illustrate the filtering procedure and to highlight selected issues. Therefore, a simple hydrological model was introduced in section 2.2.1.

4.1 Model Prediction

The model forward integration is implemented by evaluating the non-linear dynamical model equations, denoted by $f(.)$,

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{p}) + \mathbf{q}_{k-1}. \tag{4.1}$$

The model states $\mathbf{x}_k$ of the current time step k depend non-linearly on the model states $\mathbf{x}_{k-1}$ of the previous time step $(k-1)$, time dependent input forcing fields $\mathbf{u}_k$ and constant

model parameters $\mathbf{p}$, as well as on unknown model errors $\mathbf{q}_{k-1}$ of any probability distribution, i.e. not necessarily of the Gaussian distribution. Note that in addition to Eq. (3.21) forcing fields ($\mathbf{u}_k$) and model parameters ($\mathbf{p}$) are considered. In linear approaches, the error covariance matrix of the model is obtained from error propagation of the previous model state covariance matrix $\mathbf{C}(\mathbf{x}_{k-1})$ to the current time step, as given in Eq. (3.27), for which the model error covariance matrix $\mathbf{Q}_{k-1} = E(\mathbf{q}_{k-1}\mathbf{q}_{k-1}^T)$ should be given.

In ensemble based data assimilation, the model equations are evaluated for each of the $i = 1, \ldots, N_e$ ensemble members (e.g., Evensen, 2007), i.e.

$$\mathbf{x}_k^{(i)-} = f(\mathbf{x}_{k-1}^{(i)}, \mathbf{u}_k^{(i)}, \mathbf{p}^{(i)}). \tag{4.2}$$

The model states $\mathbf{x}_k^{(i)-}$ of the current time step k, referred to as model predictions, are denoted with the superscript "$-$". In this work, $\mathbf{q}_{k-1}$ is neglected, i.e. no explicit realizations of the model errors are generated, due to the difficulty in specifying the matrix $\mathbf{Q}_{k-1}$. However, an alternative strategy to consider these errors is introduced in section 4.3. Uncertainties in model states $\mathbf{x}_{k-1}$, input forcing fields $\mathbf{u}_k$, and model parameters $\mathbf{p}$ are introduced by generating N_e ensemble members. In Eq. (4.2), only model states are written to the model prediction vector, i.e. $\mathbf{x}_k^{(i)-} = \mathbf{w}_k^{(i)-}$ (see Eq. (3.30)).

For a simultaneous C/DA, Eq. (4.2) is reformulated to

$$\mathbf{x}_k^{(i)-} = \begin{pmatrix} \mathbf{w}_k^{(i)-} \\ \mathbf{p}_k^{(i)-} \end{pmatrix} = \begin{pmatrix} f(\mathbf{w}_{k-1}^{(i)}, \mathbf{u}_k^{(i)}, \mathbf{p}_{k-1}^{(i)}) \\ \mathbf{p}_{k-1}^{(i)} \end{pmatrix}. \tag{4.3}$$

Here, the time indices $k-1$ and k are added to the model calibration parameters $\mathbf{p}$. In the model prediction step, the values of the calibration parameters are constant. Then, in the ensemble filter update the calibration parameters are simultaneously updated along with the model states $\mathbf{w}$.

The error statistics of the model prediction are represented by the ensemble mean $\overline{\mathbf{x}}_k = \frac{1}{N_e}\sum_{i=1}^{N_e}\mathbf{x}_k^{(i)-}$ and the empirical error covariance matrix (e.g., Ripley, 1987)

$$\mathbf{C}^e(\mathbf{x}_k^-) = \frac{1}{N_e - 1}\Delta\mathbf{X}_k^-(\Delta\mathbf{X}_k^-)^T, \tag{4.4}$$

determined from the ensemble spread. Here, the matrix $\Delta\mathbf{X}_k^-$ stores the ensemble perturbations $\Delta\mathbf{x}_k^{(i)-} = \mathbf{x}_k^{(i)-} - \overline{\mathbf{x}}_k$ in its columns. It is possible to define $\Delta\mathbf{X}_k^- = \mathbf{X}_k^-\mathbf{W}$ with $\mathbf{X}_k^- = (\mathbf{x}_k^{(1)-}, \ldots, \mathbf{x}_k^{(N_e)-})$ and the idempotent $(N_e \times N_e)$-projection matrix $\mathbf{W}$ with elements equal to $1 - N_e^{-1}$ on its diagonal and $-N_e^{-1}$ as off-diagonal entries. By introducing $\mathbf{W}$ in the mentioned way, the rank of the matrix is $(N_e - 1)$ and the formulation of the model covariance matrix results in

$$\mathbf{C}^e(\mathbf{x}_k^-) = \frac{1}{N_e - 1}\mathbf{X}_k^-\mathbf{W}(\mathbf{X}_k^-)^T. \tag{4.5}$$

In the following, five examples are designed to describe the details of the sequential ensemble data assimilation methods (box 1-5).

Box 1: Prediction with a Simple Model and WGHM

Let us assume that we aim to improve the simulation of the total water storage (TWS), denoted by S, of the simple model presented in section 2.2.1, and simultaneously calibrating the model parameter K by introducing observations Y of TWS. In the model prediction step, an ensemble of model states is first initialized. In this example, $N_e = 30$ realizations of K_0 (dimensionless), of the initial water state S_0 (m^3) and of the precipitation minus evapotranspiration time series $(P-E)_k$ (m^3) are generated based on arbitrarily chosen PDFs (Tab. 4.1).

Table 4.1: Details to generate an initial ensemble of model runs by introducing uncertainties of the initial water state S_0, the model parameter K, and the input forcing field $(P-E)$, i.e. net precipitation. The model prediction is then performed using the equations of the simple model which was presented in section 2.2.1.

Ensemble of	Variable name in (Eq. (4.3))	Simple Model in Eqs. (2.27) and (2.28)	Initial Values
model state	$\mathbf{w}$	S	$S_0^{(i)} \in [2, 8]$
calibration parameter	$\mathbf{p}$	K	$K_0^{(i)} \in [0.01\ 0.99]$
forcing field	$\mathbf{u}$	$(P-E)$	$(P-E)_k^{(i)} = m^{(i)} \cdot (P-E)_k$ with $m^{(i)} \in [0.8, 1.2]$

For K_0, a minimum and maximum limit of 0.01 and 0.99 was defined, respectively. Subsequently, 30 uniformly distributed ensemble members $K_0^{(i)}$ were generated within these limits. The same procedure was applied to S_0 while considering 2 m^3 and 8 m^3 as limits. In order to account for uncertainties in the forcing field, i.e. net precipitation in this example, a multiplicative error centered at one and with maximum limits of 0.8 and 1.2 was assumed for $(P-E)_k$. Thus, larger uncertainties were associated with higher net precipitation values. It is worth mentioning that the ensemble of $(P-E)_k$ is only used to represent uncertainties of the forcing field and, therefore, to realistically represent uncertainties of the model simulation. It is not updated in the ensemble filter. The model in Eq. (4.3) depends on $\mathbf{w}_{k-1}^{(i)}$, which is the water storage $S_{k-1}^{(i)}$ of the previous time step $k-1$ for sample i, $\mathbf{u}_k^{(i)}$ (the i-th realization of $(P-E)_k^{(i)}$ at time step k), and $\mathbf{p}_{k-1}^{(i)}$ (associated with the parameter $K_{k-1}^{(i)}$). The initial ensemble mean and error covariance matrix of S_0 and K_0, which both should be improved, are calculated using their respective ensembles and given by

$$\overline{\mathbf{x}_0} = \begin{bmatrix} \overline{S_0} \\ \overline{K_0} \end{bmatrix} = \begin{bmatrix} 4.88 \\ 0.47 \end{bmatrix}, \text{ and } \mathbf{C}^e(\mathbf{x}_0) = \begin{bmatrix} \sigma_{S_0}^2 & \sigma_{S_0K_0} \\ \sigma_{K_0S_0} & \sigma_{K_0}^2 \end{bmatrix} = \begin{bmatrix} 3.28 & -0.05 \\ -0.05 & 0.09 \end{bmatrix}.$$

By evaluating Eqs. (2.27) and (2.28) for each ensemble member over 24 time steps, the open loop (OL) simulations, i.e. no observations of S are assimilated, in Fig. 4.1 are determined.

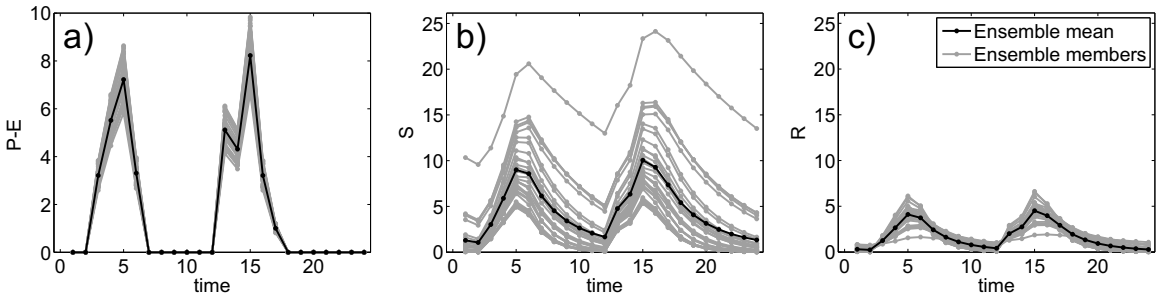

Figure 4.1: An overview of $N_e = 30$ ensemble members of the generated daily a) accumulated net precipitation $P - E$ (in m^3), b) the open loop simulations of storage S (in m^3), and c) accumulated runoff R (in m^3).

For data assimilation, the model prediction is performed until time step k, at which observations are available. The model simulations of each ensemble member are written into a prediction vector

$$\mathbf{x}_k^{(i)-} = \begin{bmatrix} \mathbf{w}_k^{(i)-} \\ \mathbf{p}_k^{(i)-} \end{bmatrix} = \begin{bmatrix} S_k^{(i)} \\ K_k^{(i)} \end{bmatrix} = \begin{bmatrix} (1 - K_{k-1}^{(i)}) S_{k-1}^{(i)} + (P - E)_{k-1}^{(i)} \\ K_{k-1}^{(i)} \end{bmatrix},$$

with m=2 rows, which is the number of states and parameters that will be updated, and one column. These vectors are collected in the model prediction matrix

$$\mathbf{X}_k^- = \begin{bmatrix} S_k^{(1)} & S_k^{(2)} & \dots & S_k^{(N_e)} \\ K_k^{(1)} & K_k^{(2)} & \dots & K_k^{(N_e)} \end{bmatrix},$$

with m=2 rows and N_e columns. The second order error statistics of the model prediction are represented by the ensemble mean $\overline{\mathbf{x}_k^-}$ and the empirical model error covariance matrix $\mathbf{C}^e(\mathbf{x}_k^-)$, which are given at time step k =1 by

$$\overline{\mathbf{x}_1^-} = \begin{bmatrix} \overline{S_1} \\ \overline{K_1} \end{bmatrix} = \begin{bmatrix} 3.47 \\ 0.47 \end{bmatrix}, \text{ and } \mathbf{C}^e(\mathbf{x}_1^-) = \begin{bmatrix} \sigma_{S_1}^2 & \sigma_{S_1 K_1} \\ \sigma_{K_1 S_1} & \sigma_{K_1}^2 \end{bmatrix} = \begin{bmatrix} 2.46 & -0.24 \\ -0.24 & 0.09 \end{bmatrix}.$$

The variance of the TWS S is ∼27 times larger than the variance of the model parameter K, since the minimum and maximum limits form a much larger interval for TWS. This depends obviously on the selected units for S. Since the parameter is not observed, the variance of K or its ratio to the variance of S will not affect the C/DA results. Instead, the parameter K will be calibrated via the dimensionless correlation to the state S. The model state and parameter have a covariance of -0.24 and a correlation coefficient of -0.52. The correlation coefficient does not depend on the unit of S and, thus, the calibration of K is independent of the selected unit.

To transfer this example to the more complex WaterGAP Global Hydrology Model (WGHM), the model prediction vector is extended in three ways: (1) In WGHM, not only one TWS but ten individual water storage compartments are defined (see section 2.2.2). Therefore, the simulated values of each compartment are written to the model prediction vector. (2) WGHM does not only consider one storage for a

river basin but simulates water changes on a 0.5°×0.5° grid. Thus, the water storage values in all grid cells within the river basin of interest are included in the prediction vector. (3) It is of interest to calibrate not only one but 22 model parameters, which are considered as constant for all grid cells within one river basin. Therefore, all of these parameter values are added to the prediction vector. For each model forward integration i=1,..., N_e, the prediction vector is determined as

$$\mathbf{x}_k^{(i)-} = \begin{pmatrix} \text{storage compartments in cell } 1^{(i)} \\ \vdots \\ \text{storage compartments in cell } n^{(i)} \\ \text{WGHM calibration parameters}^{(i)} \end{pmatrix}.$$

The vector includes $10 \cdot n + p$ considering ten water storage compartments in n grid cells and p model parameters. The Mississippi River Basin, the test region for the C/DA framework, comprises n=1382 grid cells, which results in 13842 rows of the prediction vector.

4.2 Filter Update

4.2.1 Ensemble Kalman Filter (EnKF)

In the update (or analysis) step of the EnKF (Evensen, 2003), each model prediction sample $\mathbf{x}_k^{(i)-}$ is informed by a perturbed version $\mathbf{y}_k + \delta\mathbf{y}_k^{(i)}$ of the observation data. By introducing the perturbations $\delta\mathbf{y}_k^{(i)}$, the observation vector is treated as a random variable in such a way to keep the update error covariance matrix within the ensemble unbiased. Burgers et al. (1998) showed that, when neglecting these observation perturbations, the variance of the updated ensemble is too low (see example in box 2). The ensemble of EnKF updated states $\mathbf{X}_k^+ = (\mathbf{x}_k^{(1)+}, \ldots, \mathbf{x}_k^{(N_e)+})$ is denoted with superscript "+", and obtained from

$$\mathbf{X}_k^+ = \mathbf{X}_k^- + \mathbf{K}_k((\mathbf{Y}_k + \Delta\mathbf{Y}_k) - \mathbf{A}\mathbf{X}_k^-), \tag{4.6}$$

with

$$\mathbf{K}_k = \mathbf{C}^e(\mathbf{x}_k^-)\mathbf{A}^T(\mathbf{A}\mathbf{C}^e(\mathbf{x}_k^-)\mathbf{A}^T + \mathbf{\Sigma}_{\mathrm{yy}})^{-1}. \tag{4.7}$$

Herein, $\mathbf{Y}_k$ contains the observation vector $\mathbf{y}_k$ in each of its columns, while $\Delta\mathbf{Y}_k$ stores the realizations of the observation perturbations $\delta\mathbf{y}_k^{(i)}$. The difference between the measured (and perturbed) and the predicted observations $((\mathbf{Y}_k + \Delta\mathbf{Y}_k) - \mathbf{A}\mathbf{X}_k^-)$ is weighted and used to correct the predicted model ensemble $\mathbf{X}_k^-$. In Eq. (4.7), $\mathbf{A}$ is the design matrix that relates model states to observations and, here, is assumed to be constant over time. The gain matrix $\mathbf{K}_k$ weights the empirical ensemble covariance matrix of the model prediction

$\mathbf{C}^e(\mathbf{x}_k^-)$ and the observation error covariance matrix $\mathbf{\Sigma}_{\text{yy}} = E(\delta\mathbf{y}_k\delta\mathbf{y}_k^T)$. The update error covariance matrix $\mathbf{C}^e(\mathbf{x}_k^+)$ is given by

$$\mathbf{C}^e(\mathbf{x}_k^+) = (\mathbf{I} - \mathbf{K}_k\mathbf{A})\mathbf{C}^e(\mathbf{x}_k^-), \tag{4.8}$$

in which $\mathbf{I}$ denotes the identity matrix.

Box 2: EnKF Update with Simple Model and WGHM

In general, three extreme cases might happen when the EnKF update step adjusts the model prediction. Assume that the difference between a TWS observation Y and a model prediction X^- is 10 mm (Fig. 4.2).

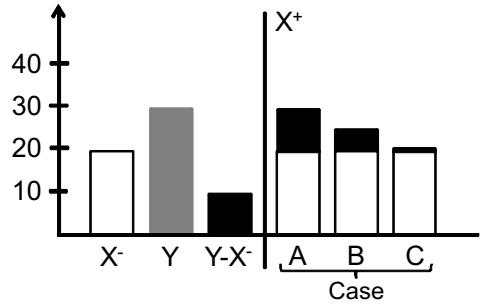

Figure 4.2: Scheme of three extreme cases that might happen when the EnKF update adjusts the model prediction.

In case (A) the observation error is small, while the model prediction error is large. Therefore, nearly 10 mm are added to the model prediction to determine the model update X^+. In case (B), both, model prediction and observation are equally accurate. Thus, 5 mm are added to the model prediction. Finally, in case (C), the model prediction has much smaller uncertainties than the observation, and therefore, its value is rarely influenced.

In the following, the model update by the EnKF is calculated for the model prediction vector in the example described in box 1. Assuming the synthetic true model parameter $K^{\text{true}} = 0.3$ and evaluating the simple one-bucket model over 24 time steps gives the true TWS states $\mathbf{S}^{\text{true}}$ (Fig. 4.3 a). For obtaining synthetic observations $\mathbf{y}$, a multiplicative error is added to the true TWS values, i.e.

$$\mathbf{y} = \mathbf{S}^{\text{true}} + 0.3 \cdot \mathbf{S}^{\text{true}} \cdot \mathbf{noise}^{\text{uniform}}.$$

Using a multiplicative observation error model leads to observation errors depending on the fill level of the storage S. An ensemble of $N_e = 30$ observations is generated by adding realizations of the same noise type. The observation ensemble matrix $(\mathbf{Y}_k + \Delta\mathbf{Y}_k)$ has $N_e = 30$ entries at each time step k

$$(\mathbf{Y}_k + \Delta\mathbf{Y}_k) = \left[\begin{array}{cccc} \mathbf{y}_k^{(1)} & \mathbf{y}_k^{(2)} & \dots & \mathbf{y}_k^{(N_e)} \end{array} \right].$$

The observation ensemble is used at each time step k to update the model prediction ensemble. At time step $k = 1$ the observation ensemble mean $\overline{\mathbf{y}_1}$ is 5.34 m^3 and its error variance $\mathbf{\Sigma}_{\text{yy}}$ is 4.10 $(\text{m}^3)^2$. The design matrix $\mathbf{A} = [1\ 0]$ relates the model prediction vector to the observations. The Kalman gain $\mathbf{K}$ is applied to the difference Δ between the observed and the model predicted water storage S resulting in N_e

innovation vectors $\mathbf{d}$, which are added to the N_e model prediction vectors. Here, the ensemble mean of Δ and the ensemble mean $\mathbf{d}_1$ of the innovation vectors are provided

$$\mathbf{K}_1 = \begin{bmatrix} \sigma_{S_1}^2(\sigma_{S_1}^2 + \sigma_{Y_1}^2)^{-1} \\ \sigma_{K_1S_1}(\sigma_{S_1}^2 + \sigma_{Y_1}^2)^{-1} \end{bmatrix} = \begin{bmatrix} 0.37 \\ -0.04 \end{bmatrix}, \quad \Delta = 1.88, \text{ and } \mathbf{d}_1 = \begin{bmatrix} 0.70 \\ -0.07 \end{bmatrix}.$$

The updated model ensemble mean and the corresponding error covariance matrix result in

$$\overline{\mathbf{x}_1^+} = \begin{bmatrix} 4.17 \\ 0.41 \end{bmatrix}, \text{ and } \mathbf{C}^e(\mathbf{x}_1^-) = \begin{bmatrix} 1.57 & -0.15 \\ -0.15 & 0.08 \end{bmatrix}.$$

In this example, the model prediction of S is more accurate than the observations, and therefore, the model update is closer to the model prediction value than the observed value.

The time series of the ensemble means of the model prediction and the update of S, as well as the synthetic observations and the truth are shown in Fig. 4.3 a.

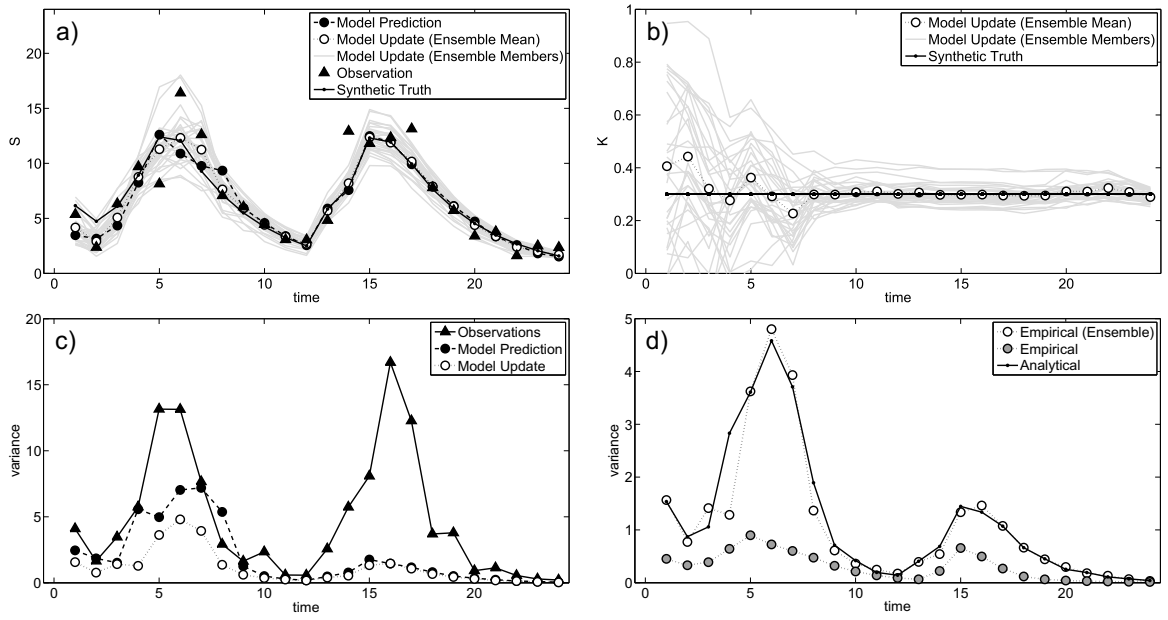

Figure 4.3: Results of the EnKF update for the simple model in section 2.2.1 over 24 time steps: the time series are shown for a) the ensemble mean of model TWS prediction $\overline{S_1^-}$, the ensemble mean $\overline{S_1^+}$ and ensemble members of the model update, observations $\mathbf{Y}_1$, and synthetic truth (in m^3); b) ensemble mean $\overline{K_1^+}$ and ensemble members of the updated model parameter values, and the true parameter value; c) variances of observations as well as of model prediction and update (in m^3), which are shown in a). d) Comparison of empirically estimated variances (in $(m^3)^2$) of the filter updates when using an observation ensemble (Empirical (Ensemble) in d)) or neglecting the observation ensemble (Empirical in d)) and the analytical variances (determined by applying formal variance error propagation) shows that the observation perturbations keep the update ensemble unbiased.

During the first twelve update steps the innovation, i.e. the correction of the predicted value, is larger than during the second twelve updates, in which the model prediction is already very close to the truth. Fig. 4.3 b shows the ensemble mean and ensemble members of the updated model parameter K. After eight updates the calibrated parameter converges to the true parameter value. This results in a much more precise model prediction during the following model forward integrations. In the present case, the model parameter should be calibrated for around ten update steps until it sufficiently represents the true value. Afterwards, the assimilation of TWS observations has only a marginal influence on the model predictions. This is also reflected in the variances of model predictions and observations (Fig. 4.3 c). By assimilating TWS observations, the variances of the model update $\overline{\mathbf{x}^+}$ is reduced compared to the variances of the model prediction $\overline{\mathbf{x}^-}$, i.e. that the spread of the model ensemble decreases with each update. Depending on the model input $P - E$ and the parameter sample K_i, the ensemble spread increases in the next model forward integration. During the second half of the update phase the variance of the model prediction is small compared to the standard deviation of the observations. The gain matrix therefore has entries close to zero, i.e. the influence of the observations is nearly negligible.

The updated model states were determined by evaluating Eq. (4.6) while (i) considering the observation perturbation matrix $\Delta\mathbf{Y}_k$ in one case and (ii) neglecting it in the other case. The results in Fig. 4.3 a, b and c are shown for case (i). The analytical covariance matrix of the update is determined by evaluating Eq. (4.8). For calculating the empirical model covariance matrix using the ensemble members of the filter update in case (i), the main diagonal elements represent the variances determined by formal error propagation well (Fig. 4.3 d). However, by neglecting the perturbation of observations, the variances of the filter updates are considerably underestimated (Fig. 4.3 d). This justifies the introduction of the observation perturbations $\Delta\mathbf{Y}_k$ in Eq. (4.6) as proposed by Burgers et al. (1998).

As indicated in the example of box 1, the prediction vector of WGHM does not contain TWS but values of individual water compartments. Here, it is assumed that simulated TWS is the sum of a soil water and a groundwater compartment to illustrate the vertical disaggregation of TWS (Fig. 4.4). The difference between observation Y and model prediction X^- is 10 mm as in Fig. 4.2, and the observation is more accurate than the model prediction (case A in Fig. 4.2). In case (A) of Fig. 4.4, the simulation of groundwater is very accurate but the simulation of soil water exhibits large uncertainties. Thus, a large portion of $Y - X^-$ is associated with the soil water storage, i.e. 9 mm in this example, and only 1 mm is added to the groundwater storage. In case (B), both storage compartments have equal uncertainties, and 5 mm are added to each storage. Finally, in case (C), the simulation of groundwater is uncertain but the soil water compartment is precisely known (in a real case study, this might have been checked with independent in-situ measurements). Here, nearly 10 mm are associated with the groundwater compartment. Case (A), (B) and (C) are determined for uncorrelated soil water and groundwater compartments. In case (D), a moderate correlation of 0.5 and uncertainties as in case (C) are assumed. This results in a slightly larger update of the soil water compartment compared to case (C). Then, in case (E), a high

correlation of 0.99 is introduced, for which the update increment of soil water is found to be even higher.

In addition to the vertical disaggregation of TWS into its individual compartments, the spatially coarser GRACE TWS anomalies are horizontally disaggregated to the finer resolved WGHM grid cells within the EnKF update (Fig. 4.5). Here, the spatial average of TWS in two grid cells (X^-) is compared to the observed TWS Y. It is assumed that the difference of 10 mm is completely introduced into the model. If the left grid cell is more accurate than the right grid cell, the update will be larger for the right grid cell (case A), and vice versa (case C). For equally accurate grid cells, both cells receive an update of 5 mm (case B). Spatial correlations between storage compartments might exist, e.g., if neighboring grid cells belong to one surface water body. In this case, the water level will not change strongly from one to the next grid cell. Assuming uncertainties as in case C but introducing a moderate correlation between the grid cells of 0.5 (case D) or a high correlation of 0.99 (case E), the updated grid cell values get closer together.

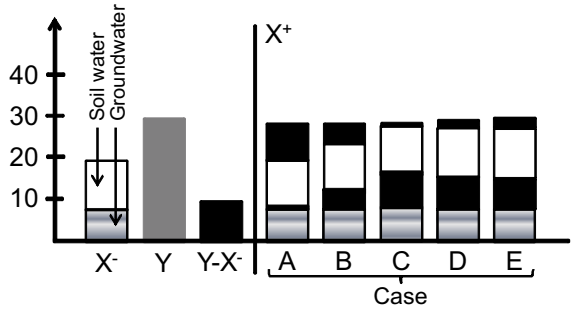

Figure 4.4: Vertical disaggregation of the EnKF update.

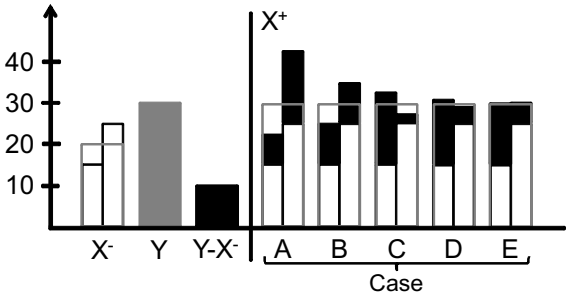

Figure 4.5: Horizontal disaggregation of the EnKF update.

In case of WGHM, vertical and horizontal disaggregation is performed simultaneously, while additionally model parameter values are adjusted. Thus, the update results might not be easily traced back to their cause.

4.2.2 Square Root Analysis Scheme for EnKF (SQRA)

The idea of the SQRA method is to avoid the generation of observation perturbations, which are required in the classical EnKF. The SQRA update (Evensen, 2004, 2007) consists of two parts: (1) the update of the ensemble mean and (2) the update of the ensemble perturbations. In contrast to the EnKF, SQRA does not perform the update for each sample individually (see Eq. (4.6)) but for the ensemble mean of the model predictions, i.e.

$$\overline{\mathbf{x}_k^+} = \overline{\mathbf{x}_k^-} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{A}\overline{\mathbf{x}_k^-}). \tag{4.9}$$

Here, the original observation vector $\mathbf{y}_k$ is used for correcting the predicted ensemble mean $\overline{\mathbf{x}_k^-}$. In this way, the generation of observation perturbations is avoided, which is required

in the EnKF but provides an additional source of sampling errors (Whitaker and Hamill, 2002, see example in box 3).

Since an unbiased ensemble of updated model states $\mathbf{X}_k^+$ is needed for the next model forward integration, re-sampling of the ensemble perturbations is required. In this PhD thesis, the simple and straightforward version of the SQRA introduced by Evensen (2004) is implemented. First, Eq. (4.4) is used to define the error covariance matrix of the model update as $\mathbf{C}^e(\mathbf{x}_k^+) = \frac{\Delta\mathbf{X}_k^+(\Delta\mathbf{X}_k^+)^T}{N_e-1}$. Then, the ensemble versions of $\mathbf{C}^e(\mathbf{x}_k^-)$ and $\mathbf{C}^e(\mathbf{x}_k^+)$ are inserted in Eq. (4.8) to calculate $\Delta\mathbf{X}_k^+$ as depending on the ensemble perturbations of the predictions

$$\begin{aligned}&\Delta\mathbf{X}_k^+(\Delta\mathbf{X}_k^+)^T = \\ &\Delta\mathbf{X}_k^-(\mathbf{I}-(\Delta\mathbf{X}_k^-)^T\mathbf{A}^T(\mathbf{A}\Delta\mathbf{X}_k^-(\Delta\mathbf{X}_k^-)^T\mathbf{A}^T+(N_e-1)\mathbf{\Sigma}_{\mathrm{yy}})^{-1}\mathbf{A}\Delta\mathbf{X}_k^-)(\Delta\mathbf{X}_k^-)^T.\end{aligned} \tag{4.10}$$

An eigenvalue decomposition is applied to

$$(\mathbf{A}\Delta\mathbf{X}_k^-(\Delta\mathbf{X}_k^-)^T\mathbf{A}^T+(N_e-1)\mathbf{\Sigma}_{\mathrm{yy}})^{-1} = \mathbf{Z}\mathbf{\Lambda}^{-1}\mathbf{Z}^T, \tag{4.11}$$

and Eq. (4.10) is then reorganized to

$$\Delta\mathbf{X}_k^+(\Delta\mathbf{X}_k^+)^T = \Delta\mathbf{X}_k^-(\mathbf{I}-\underbrace{(\mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{Z}^T\mathbf{A}\Delta\mathbf{X}_k^-)^T}_{\mathbf{D}^T}(\underbrace{\mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{Z}^T\mathbf{A}\Delta\mathbf{X}_k^-}_{\mathbf{D}}))(\Delta\mathbf{X}_k^-)^T. \tag{4.12}$$

The singular value decomposition of $\mathbf{D} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ is inserted into Eq. (4.12), which results in

$$\begin{aligned}\Delta\mathbf{X}_k^+(\Delta\mathbf{X}_k^+)^T &= \Delta\mathbf{X}_k^-(\mathbf{I}-(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T)^T(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T))(\Delta\mathbf{X}_k^-)^T \\ &= \Delta\mathbf{X}_k^-\mathbf{V}(\mathbf{I}-\mathbf{\Sigma}^T\mathbf{\Sigma})\mathbf{V}^T(\Delta\mathbf{X}_k^-)^T.\end{aligned} \tag{4.13}$$

Using the square root of the diagonal matrix $(\mathbf{I}-\mathbf{\Sigma}^T\mathbf{\Sigma})$, Eq. (4.13) is written as

$$\Delta\mathbf{X}_k^+(\Delta\mathbf{X}_k^+)^T = (\Delta\mathbf{X}_k^-\mathbf{V}\sqrt{\mathbf{I}-\mathbf{\Sigma}^T\mathbf{\Sigma}})(\Delta\mathbf{X}_k^-\mathbf{V}\sqrt{\mathbf{I}-\mathbf{\Sigma}^T\mathbf{\Sigma}})^T, \tag{4.14}$$

which represents a symmetric expression that can be used to generate normally distributed perturbation vectors with zero mean and covariance matrix $\mathbf{C}^e(\mathbf{x}_k^+)$ (see Eq. (3.10)). Finally, the updated ensemble perturbations are added to the updated ensemble mean

$$\mathbf{X}_k^+ = \overline{\mathbf{X}_k^+} + \underbrace{\Delta\mathbf{X}_k^-\mathbf{V}\sqrt{\mathbf{I}-\mathbf{\Sigma}^T\mathbf{\Sigma}}}_{\Delta\mathbf{X}_k^+}\mathbf{\Theta}^T. \tag{4.15}$$

In Eq. (4.15), $\mathbf{\Theta^T}$ holds standard normal distributed random numbers. It contains the right hand side eigenvectors of a matrix $\mathbf{M} = \mathbf{\Theta}\mathbf{\Lambda}\mathbf{\Theta^T}$ that holds uniformly distributed random numbers $\mathbf{M}_{ij}$ in its rows i and columns j. Using this random orthonormal matrix guarantees that the vectors in each column of $\mathbf{\Theta^T}$ are independent from each other. By multiplying $\Delta\mathbf{X}_k^+$ with $\mathbf{\Theta^T}$, realizations of ensemble perturbations are generated from the update error covariance matrix $\mathbf{C}^e(\mathbf{x}_k^+)$ by Monte Carlo sampling (e.g., Kusche, 2003). A detailed derivation of the algorithm and a comparison to the classical EnKF can be found in Evensen (2004) and Evensen (2007).

Box 3: SQRA Update with Simple Model

The examples in box 1 and box 2 are extended in the following, by performing the filter update applying the SQRA approach. Here, the focus is on the representation of the uncertainties of the model update. The differences of the main diagonal elements of the empirical ensemble update covariance matrix to the analytical solution, which was determined by formal error propagation, are shown in Fig. 4.6. The variances of the updated model values are represented much more accurate when applying the SQRA compared to the EnKF that relies on an ensemble of observation perturbations.

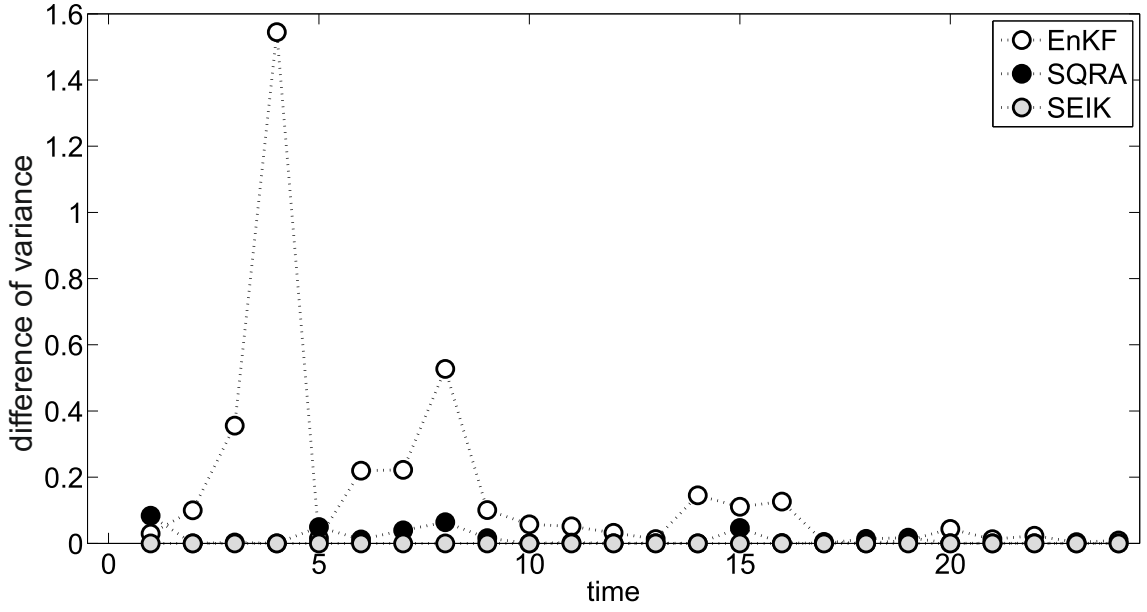

Figure 4.6: Absolute differences between the variances of the model update determined by formal error propagation (in $(m^3)^2$) and the empirical variances when applying the EnKF, SQRA or SEIK method.

In this and the following examples, it is of interest to provide insights about the ensemble filter methods. The simple model is used for this. The underlying concept is identical for more complex models such as WGHM.

4.2.3 Singular Evolutive Interpolated Kalman Filter (SEIK)

The idea of the SEIK algorithm (Pham et al., 1998) is to reformulate the equation of the Kalman gain matrix to reduce computational time in the filter update step. In the SEIK filter, the ensemble representation of the model prediction error covariance matrix is given in form of

$$\mathbf{C}^e_{\mathrm{SEIK}}(\mathbf{x}_k^-) = \mathbf{L}^e_k \mathbf{G}^e \mathbf{L}^{e^T}_k , \tag{4.16}$$

where the matrix $\mathbf{L}_k^e = \mathbf{X}_k^- \mathbf{T}$ is of dimension $m \times (N_e - 1)$, m is the number of entries in the model prediction vectors $\mathbf{x}_k^{(i)-}$, and N_e is the ensemble size. Here, $\mathbf{T}$ is a full rank matrix with zero column sums, which consists of the first $(N_e - 1)$ columns of the matrix $\mathbf{W}$ in Eq. (4.5): $\mathbf{W} = [\mathbf{T}|\mathbf{t}]$ with $\mathbf{t}$ representing the last column of $\mathbf{W}$. $\mathbf{G}^e = \frac{1}{N_e}(\mathbf{T}^T\mathbf{T})^{-1}$ is normalized by the ensemble size N_e. Using Eq. (4.16), the model prediction errors are represented in the space that is spanned by the columns of $\mathbf{L}_k^e$.

As for the EnKF, the formulation of the SEIK filter update can be derived from the KF equations. Here, however, the model prediction error covariance matrix in Eq. (4.7) is replaced by the ensemble representation defined in Eq. (4.16)

$$\mathbf{K}_k = \mathbf{L}_k^e \mathbf{G}^e \mathbf{L}_k^{e^T} \mathbf{A}^T (\mathbf{A}\mathbf{L}_k^e \mathbf{G}^e \mathbf{L}_k^{e^T} \mathbf{A}^T + \mathbf{\Sigma}_{\mathrm{yy}})^{-1}. \tag{4.17}$$

By applying the matrix identity $\mathbf{QW}(\mathbf{Z} + \mathbf{VQW})^{-1} = (\mathbf{Q}^{-1} + \mathbf{WZ}^{-1}\mathbf{V})^{-1}\mathbf{WZ}^{-1}$ (Koch, 1999, p. 34, Eq. (1.115)) for invertible matrices $\mathbf{Q}$ and $\mathbf{Z}$ and arbitrary matrices $\mathbf{V}$ and $\mathbf{W}$ to Eq. (4.17), the formulation of the gain matrix becomes

$$\mathbf{K}_k = \mathbf{L}_k^e \underbrace{[(\mathbf{G}^e)^{-1} + \mathbf{L}_k^{e^T} \mathbf{A}^T \mathbf{\Sigma}_{\mathrm{yy}}^{-1} \mathbf{A}\mathbf{L}_k^e]^{-1}}_{N_e \times N_e} \mathbf{L}_k^{e^T} \mathbf{A}^T \mathbf{\Sigma}_{\mathrm{yy}}^{-1}. \tag{4.18}$$

This is the SEIK filter formulation implemented in this thesis. Here, the observation error covariance matrix $\mathbf{\Sigma}_{\mathrm{yy}}$ is transformed to the ensemble space by applying $\mathbf{A}\mathbf{L}_k^e$ to $\mathbf{\Sigma}_{\mathrm{yy}}^{-1}$: It becomes obvious that the size of the matrix to be inverted depends on the model ensemble size N_e. The update is performed in the ensemble space, and if the number of observations is much larger than the ensemble size, the application of the SEIK filter is efficient. By defining

$$\mathbf{U}_k = ((\mathbf{G}^e)^{-1} + (\mathbf{A}\mathbf{L}_k^e)^T \mathbf{\Sigma}_{\mathrm{yy}}^{-1} \mathbf{A}\mathbf{L}_k^e)^{-1}, \tag{4.19}$$

as well as $\mathbf{a}_k = \mathbf{U}_k (\mathbf{A}\mathbf{L}_k^e)^T \mathbf{\Sigma}_{\mathrm{yy}}^{-1} (\mathbf{y}_k - \mathbf{A}\overline{\mathbf{x}_k^-})$, and inserting these together with Eq. (4.18) into Eq. (4.9), the formulation of the model update is finally converted to the common notation of the SEIK filter

$$\overline{\mathbf{x}_k^+} = \overline{\mathbf{x}_k^-} + \mathbf{L}_k^e \mathbf{a}_k. \tag{4.20}$$

Basically, one projects the errors of the updated states onto the space that is spanned by the columns of $\mathbf{L}_k^e$, which results in the formulation of the model update covariance matrix $\mathbf{C}^e(\mathbf{x}_k^+)$ as

$$\mathbf{C}^e(\mathbf{x}_k^+) = \mathbf{L}_k^e \mathbf{U}_k \mathbf{L}_k^{e^T}. \tag{4.21}$$

A detailed derivation of Eq. (4.21) can be found in Pham et al. (1998).

Finally, the update of the ensemble perturbations is performed. To this end, the minimum second order exact sampling is used (Appendix in Pham, 2001). Ensemble perturbations are generated from the eigenvalue-decomposed error covariance matrix of the filter update. The ensemble mean and the ensemble covariance matrix need to match exactly the updated ensemble mean $\overline{\mathbf{x}_k^+}$ and the updated error covariance matrix $\mathbf{C}(\mathbf{x}_k^+)$, i.e.

$$\frac{1}{N_e} \sum_{i=1}^{N_e} \mathbf{x}_k^{(i)} = \overline{\mathbf{x}_k} \equiv \overline{\mathbf{x}_k^+}, \tag{4.22}$$

$$\mathbf{L}_0 \mathbf{C}_0^T \mathbf{\Omega}_0^T \mathbf{\Omega}_0 \mathbf{C}_0 \mathbf{L}_0^T = \mathbf{S}_0 \equiv \mathbf{C}(\mathbf{x}_k^+). \tag{4.23}$$

This is realized by determining a low (N_e-1)-rank approximation of the covariance matrix, using the leading eigenvalues and eigenvectors (or dominant orthogonal modes) of the ensemble update error covariance matrix $\mathbf{C}^e(\mathbf{x}_k^+)$, whose eigenvectors and eigenvalues are stored in $\mathbf{L}_0$ and $\mathbf{U}_0 = \mathbf{C}_0^T\mathbf{C}_0$, respectively. In Eq. (4.23), $\mathbf{\Omega}_0$ is an orthonormal matrix. Its columns are orthogonal to a vector that contains only the numbers one. This matrix can, for example, be determined by Householder transformation (Hoteit et al., 2002, Appendix, p. 125-126). The update ensemble $\mathbf{X}_k^+$ is determined by adding the generated perturbations to the updated ensemble mean, which is stored in each column of $\overline{\mathbf{X}_k^+}$:

$$\mathbf{X}_k^+ = \overline{\mathbf{X}_k^+} + \sqrt{N_e}\mathbf{L}_0\mathbf{C}_0^T\mathbf{\Omega}_0^T. \tag{4.24}$$

A comparison of the classical EnKF and SEIK filter can also be found e.g., in Nerger (2003). The formulation of the Kalman gain matrix based on the EnKF ensemble representation $\mathbf{C}^e(\mathbf{x}_k^-)$ in Eq. (4.7) and on the SEIK ensemble representation $\mathbf{C}^e_{\mathrm{SEIK}}(\mathbf{x}_k^-)$ in Eq. (4.18) of the model prediction error covariance matrix is only identical during the first update (identical model configuration and initial state estimate and covariance matrix implied). However, the EnKF and SEIK updated model state vectors differ from each other, since the EnKF relies on an observation ensemble but the SEIK considers an update of the ensemble mean of the model prediction vector similar to the SQRA method. Therefore, the sequence of updates will numerically differ in both approaches. In the limit $N_e \to \infty$, assuming ergodicity, the two ensemble representations fall back to the conventional Kalman filter and thus would lead to identical data assimilation results.

Box 4: SEIK Update with Simple Model

Here, the examples in boxes 1-3 are solved by applying the SEIK filter. Due to the fact that the minimum second order exact sampling is used for generating the perturbations of the updated model values, the variances derived from the SEIK filter ensemble exactly equal the analytical solution derived by formal variance error propagation (Fig. 4.6). Even though the three filter methods that are reviewed in this thesis can all be derived from the original Kalman filter equations, the calibration and data assimilation results are not the same. This is due to the different approaches of generating perturbations of the filter update (and of the observations in case of the EnKF).

4.2.4 Ensemble Kalman Smoother (EnKS)

Similar to the EnKF, the EnKS (Evensen and van Leeuwen, 2000) processes the measurements sequentially in time. The essential difference is that the observations are not only used to correct the ensemble members of the model prediction at the current time step but also backward in time. For describing the algorithm, a second time index is introduced. The index k still refers to the time steps when measurements are available. The index $l(k)$ stands for the number of model predictions until time step k. Each model prediction

vector $\mathbf{x}_{l(k)}^{(i)-}$ of the current k-th available measurements is extended by the corrected ensemble members of the previous update $\widetilde{\mathbf{x}}_{l(k-1)}^{(i)+}$ and the predicted model states starting from that last update $\mathbf{x}_{[l(k-1)]+1}^{(i)-}, \ldots, \mathbf{x}_{l(k)}^{(i)-}$

$$\widetilde{\mathbf{x}}_k^{(i)-} = \begin{bmatrix} \widetilde{\mathbf{x}}_{l(k-1)}^{(i)+} \\ \mathbf{x}_{[l(k-1)]+1}^{(i)-} \\ \vdots \\ \mathbf{x}_{l(k)}^{(i)-} \end{bmatrix}, \tag{4.25}$$

which results in $N_{m(k)} = l(k) \cdot (n+p)$ rows of the model prediction vectors, with n being the number of model states and p the number of calibration parameters. $\widetilde{\mathbf{x}}_{l(k-1)}^{(i)+}$ holds the EnKS updates (model states and parameters) until the starting point of the data assimilation and calibration phase. The associated empirical model prediction error covariance matrix, i.e.

$$\mathbf{C}^e(\widetilde{\mathbf{x}}_k^-) = \begin{bmatrix} \mathbf{C}^e(\mathbf{x}_0) & \mathbf{C}^e(\mathbf{x}_0, \mathbf{x}_1^-) & \ldots & \mathbf{C}^e(\mathbf{x}_0, \mathbf{x}_{l(k)}^-) \\ \vdots & \mathbf{C}^e(\mathbf{x}_1^-) & & \mathbf{C}^e(\mathbf{x}_1^-, \mathbf{x}_{l(k)}^-) \\ & & \ddots & \vdots \\ \text{symm.} & \ldots & & \mathbf{C}^e(\mathbf{x}_{l(k)}^-) \end{bmatrix}, \tag{4.26}$$

contains the covariances of the model states and parameters in space and time.

The ensemble of the EnKS update is

$$\widetilde{\mathbf{X}}_k^+ = \widetilde{\mathbf{X}}_k^- + \underbrace{\mathbf{C}^e(\widetilde{\mathbf{x}}_k^-)\mathbf{A}_k^T}_{\text{Part 1}} \underbrace{(\mathbf{A}_k \mathbf{C}^e(\widetilde{\mathbf{x}}_k^-)\mathbf{A}_k^T + \boldsymbol{\Sigma}_{\text{yy}})^{-1}((\mathbf{Y}_k + \Delta\mathbf{Y}_k) - \mathbf{A}_k\widetilde{\mathbf{X}}_k^-)}_{\text{Part 2}}. \tag{4.27}$$

The time index k is now written to the design matrix $\mathbf{A}$, since the size of the model prediction matrix and therefore the size of the design matrix increase with each EnKS update. Part 2 in Eq. (4.27) only depends on the model predictions and observations at time step k, while Part 1 distributes the current update increment to the current model predictions and the model states backward in time.

By inserting Eq. (4.4) and the definition $\mathbf{D}_k = ((\mathbf{Y}_k + \Delta\mathbf{Y}_k) - \mathbf{A}_k\widetilde{\mathbf{X}}_k^-)$ into Eq. (4.27), the update equation can be reformulated (see e.g., Evensen, 2007, p. 127) as

$$\begin{aligned} \widetilde{\mathbf{X}}_k^+ &= \widetilde{\mathbf{X}}_k^- + \frac{\Delta\widetilde{\mathbf{X}}_k^-(\Delta\widetilde{\mathbf{X}}_k^-)^T}{N_e-1}\mathbf{A}_k^T(\mathbf{A}_k\frac{\Delta\widetilde{\mathbf{X}}_k^-(\Delta\widetilde{\mathbf{X}}_k^-)^T}{N_e-1}\mathbf{A}_k^T + \boldsymbol{\Sigma}_{\text{yy}})^{-1}\mathbf{D}_\mathbf{k} \\ &= \widetilde{\mathbf{X}}_k^- + \Delta\widetilde{\mathbf{X}}_k^-\mathbf{S}_k^T\mathbf{C}_k^{-1}\mathbf{D}_k \\ &\text{with} \quad \mathbf{S}_k = \mathbf{A}_k\Delta\widetilde{\mathbf{X}}_k^- \quad \text{and} \quad \mathbf{C}_k = \mathbf{S}_k\mathbf{S}_k^T + (N_e-1)\boldsymbol{\Sigma}_{\text{yy}}. \end{aligned}$$

By additionally using the definition of the state perturbations $\Delta\widetilde{\mathbf{X}}_k^- = (\mathbf{I} - \mathbf{I}_N)\widetilde{\mathbf{X}}_k^-$, in which $\mathbf{I}_N$ is a quadratic matrix that contains $\frac{1}{N_e}$ for each element, a compact formulation is obtained

$$\begin{aligned}\widetilde{\mathbf{X}}_k^+ &= \widetilde{\mathbf{X}}_k^- + (\mathbf{I} - \mathbf{I}_N)\widetilde{\mathbf{X}}_k^- \mathbf{S}_k^T \mathbf{C}_k^{-1} \mathbf{D}_k \\ &= \widetilde{\mathbf{X}}_k^- (\mathbf{I} + (\mathbf{I} - \mathbf{I}_N)\mathbf{S}_k^T \mathbf{C}_k^{-1} \mathbf{D}_k) \\ &= \underbrace{\widetilde{\mathbf{X}}_k^-}_{\text{Part 1}} \underbrace{(\mathbf{I} + \mathbf{S}_k^T \mathbf{C}_k^{-1} \mathbf{D}_k)}_{\text{Part 2}}. \end{aligned} \tag{4.28}$$

Herein, $\mathbf{I}_N \mathbf{S}_k^T = 0$ is used. The EnKS update equation still consists of two parts, from which the second part only depends on the current update time step k and therefore adds the incremental information of the current measurements. The first part includes information of the current and all time steps backward in time that have already been corrected by the observations at time steps 1 to $k-1$. According to Evensen (2007, p. 129), the update is a weak non-linear combination of the prior ensemble. In practice, time windows are defined and only observations within these windows are taken into account in the EnKS update (see lagged EnKS in Eq. (4.30)).

EnKS using EnKF as first guess

Since the EnKS can be interpreted as an extension of the EnKF (Evensen, 2007, p. 130), here the EnKS update is determined by interpreting the EnKF update as first guess of the EnKS solution. Thus, the update equation can be written as

$$\widetilde{\mathbf{X}}_{\text{EnKS}}^+ = \mathbf{X}_{\text{EnKF}}^+ \prod_{j=t}^{k} (\mathbf{I} + \mathbf{S}_j^T \mathbf{C}_j^{-1} \mathbf{D}_j). \tag{4.29}$$

The observations of all time steps until time step t are already included in the EnKF update. The observations after time step t until the current update time step k are used to determine the EnKS solution, which is therefore informed by all available measurements until the present time step k.

Lagged EnKS

The lagged EnKS exactly works like the EnKS but the propagation of update increments backward in time is limited to a constant lag (Evensen, 2007, p. 136). This is motivated by the assumption that the influence of the observations strongly decreases with larger distances in time and is negligible after some time. For a fixed time lag of b measurement time steps, Eq. (4.29) becomes

$$\widetilde{\mathbf{X}}_{\text{EnKS}}^+ = \mathbf{X}_{\text{EnKF}}^+ \prod_{j=k-b}^{k} (\mathbf{I} + \mathbf{S}_j^T \mathbf{C}_j^{-1} \mathbf{D}_j). \tag{4.30}$$

The lag might be chosen with respect to the temporal correlation length of (geophysical or hydrological) time series.

Box 5: EnKS Update with Simple Model

In this example, the EnKS and lagged EnKS with a lag of two time steps are used to derive the EnKS solution of the simple model in section 2.2.1. The lag of only two time steps was chosen to show the extremes of using all available information in the EnKS and only a very limited number of observations in the lagged EnKS. The EnKS has a large positive impact on the state estimation S, where the EnKF update showed large discrepancies to the true state (Fig. 4.7 a, e.g., at time step $k = 2$, 5 and 7). This was the case when the observation noise led to a poor representation of the true states (see Fig. 4.3). The performance of the EnKS was however not always closer to the true value than the lagged EnKS. This shows that the influence of observations decreases with increasing time difference.

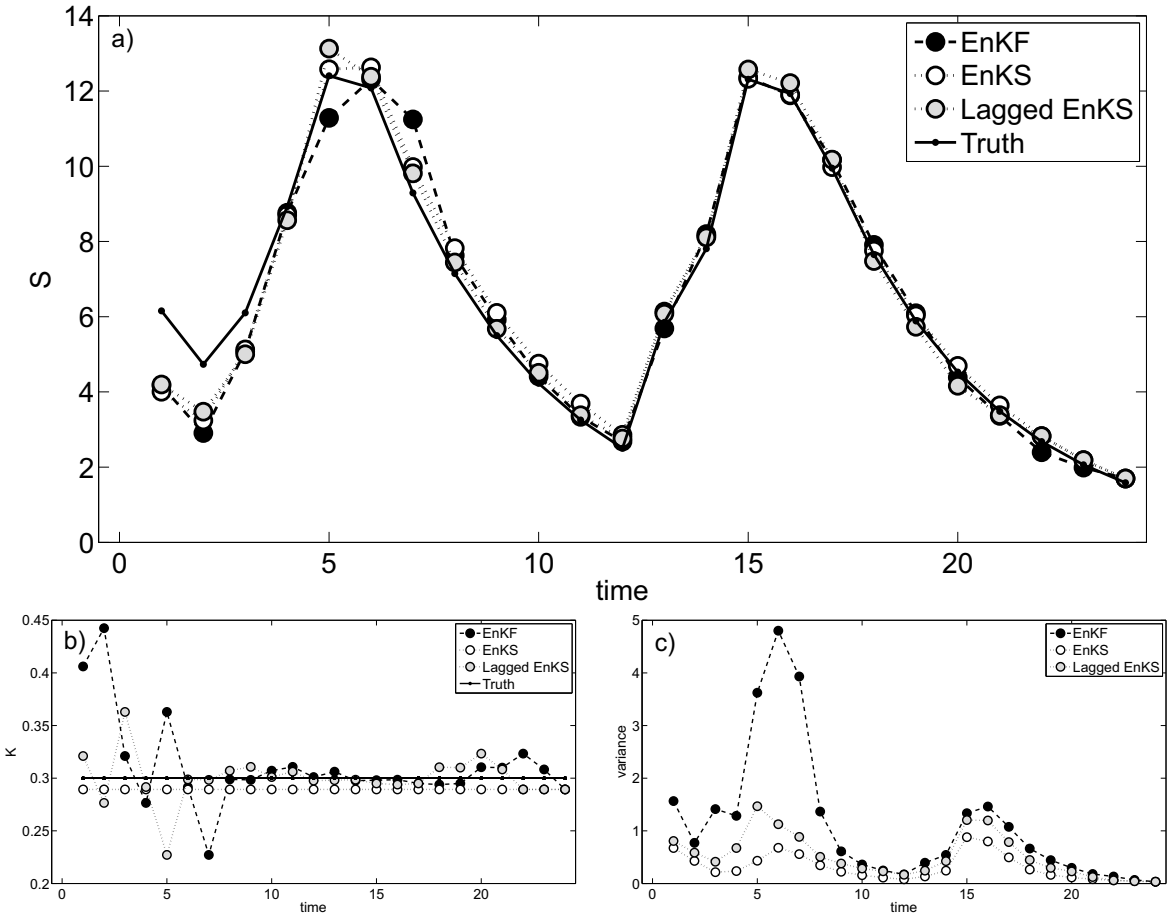

Figure 4.7: Results of the EnKF, EnKS and lagged EnKS update for the simple model in section 2.2.1 over 24 time steps: the time series are shown for a) the ensemble means of model TWS updates and synthetic truth (in m^3); b) the ensemble mean of the updated model parameter values K and the true parameter value; c) the empirical variances (in $(m^3)^2$) of model updates of S that are shown in a).

Regarding the calibration parameter, the EnKS update at time step k-lag is exactly the same as the EnKF update at time step k (Fig. 4.7 b). This shows that considering

the same observations leads to identical results for the calibrated parameter. Since usually the parameter value of the last update step is defined as calibrated model parameter and used in future model simulations, the choice of using the EnKF or (lagged) EnKS does not influence the calibrated parameter value, which is identical for all three variants at the last update time step. Obviously, this holds only if the initial model configuration is introduced consistently among the approaches.

Applying the EnKS also affects the variances, i.e. the spread in the ensemble of the updated water states S (shown in Fig. 4.7 c) and parameter K. At time step $k = 5$, 6 and 7, the ensemble members are quite different from each other after applying the EnKF and therefore the values of their variances are relatively high. At these time steps, the uncertainty of the observations is large compared to the uncertainties in the model predictions (shown in Fig. 4.3). Thus, the observations have only a small influence on the water state S and the parameter K in the EnKF update due to small weights in the gain matrix. In contrast, by applying the lagged EnKS, the model ensemble members are closer together since more observation information is used to constrain the model simulation. Thus, the variances of S and K are smaller. As a result, the variances of the updated water state S and parameter K are further reduced. The reduction of the variance of S by applying the EnKS using all available data instead of the lagged EnKS is however not as strong as the reduction derived by applying the lagged EnKS instead of the EnKF. Therefore, and due to saving computational costs, the lagged EnKS is preferred here.

4.3 Tuning Techniques to Improve the Filter Performance

Applying EnKF/EnKS techniques involves few mathematical issues, which might potentially limit their performance including (1) small ensemble size, (2) fast filter convergence, and (3) artificial state correlations: (1) Due to the high dimensions of hydrological or geophysical models, it is desirable to keep the number of model forward integrations small to reduce computational loads. Therefore, an ensemble of limited size is used in practice that should be sufficiently large to represent the error statistics of the model states. By choosing the ensemble members wisely, a smaller ensemble might represent the error statistics as accurate as a randomly generated larger ensemble (Evensen, 2004). (2) In each update step, the model prediction ensemble members are pulled towards the observations. Thus, the model ensemble spread is even more reduced with increasing the number of update steps. Ensemble inbreeding might be the consequence, which unrealistically reduces the influence of observations on the model prediction in the update step (see Fig. 4.3 a, b, c in the second update phase). This is less critical for data assimilation only, since an ensemble of constant parameters usually keeps the model ensemble spread large. However, for a simultaneous C/DA, fast ensemble convergence is challenging. To inflate the ensemble, a multiplication factor can be introduced (Hamill and Snyder, 2002), which can be interpreted as a strategy to account for model errors. (3) The limited ensemble size can

cause artificial model state correlations (e.g., Liu et al., 2012). Localization methods have been developed to reduce correlations that are likely caused by numerical problems rather than physical reasons. Strategies to address these issues are presented in the following.

4.3.1 Improved Initial Sampling

The improved sampling scheme described here follows the one introduced by Evensen (2004). To choose initial ensemble members that represent the most dominant directions of the error space, firstly, a large ensemble of size $z \times N_e$ is generated, e.g., with $z = 10$ (Evensen, 2004) and $N_e = 30$. The perturbations of the ensemble members around their mean are collected in the perturbation matrix $\Delta\mathbf{X}_{\text{init}}$. The number of its rows is equal to the number of model states m (model parameters are excluded here) and the number of its columns is equal to the ensemble size zN_e. A singular value decomposition (SVD) gives

$$\Delta\mathbf{X}_{\text{init}} = \mathbf{U}_{\text{init}}\mathbf{\Lambda}_{\text{init}}\mathbf{V}_{\text{init}}^T, \tag{4.31}$$

where the singular values are the main diagonal elements of $\mathbf{\Lambda}_{\text{init}}$, as well as $\mathbf{U}_{\text{init}}$ and $\mathbf{V}_{\text{init}}^T$ store the right- and left-side singular vectors in their columns, respectively. The first N_e dominant singular values are stored in the full rank $(N_e \times N_e)$ matrix $\mathbf{\Lambda}$ and their corresponding first N_e left-side singular vectors are stored in the $(m \times N_e)$ matrix $\mathbf{U}$ as

$$\mathbf{\Lambda}_{\text{init}} = \begin{bmatrix} \mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \ddots \end{bmatrix}, \quad \text{and} \quad \mathbf{U}_{\text{init}} = [\mathbf{U}|\ldots]\,. \tag{4.32}$$

Further, a random orthogonal matrix $\mathbf{V}_{\text{o}}^T$ is determined by computing the SVD of a quadratic random matrix $\mathbf{R} = \mathbf{U}_{\text{o}}\mathbf{\Lambda}_{\text{o}}\mathbf{V}_{\text{o}}^T$ with N_e rows and columns. The improved initial ensemble is obtained by

$$\Delta\mathbf{X}_0 = \mathbf{U}\mathbf{\Lambda}\mathbf{V}_{\text{o}}^T, \tag{4.33}$$

where the dominant singular values are scaled by z to retain the correct variance of the new ensemble perturbations $\Delta\mathbf{X}_0$. Finally, the ensemble mean of the new ensemble perturbations is subtracted from each new ensemble member to ensure that the mean is zero. The described procedure is similar to the minimum second order exact sampling (see Appendix in Pham, 2001, and section 4.2.3).

4.3.2 Covariance Inflation

Estimation of the empirical model covariance matrix $\mathbf{C}^e(\mathbf{x}_k^-)$ might be too optimistic, i.e. the variances might be too small due to neglecting errors in the model structure ($\mathbf{q}_{k-1}$ in Eq. (4.1)). In the absence of reliable information about these errors, alternative strategies to enlarge the ensemble spread have been developed: Hamill and Snyder (2002) introduced the so-called inflation factor. Here, the ensemble perturbations are multiplied by a constant inflation factor m_c

$$\mathbf{X}_k'^- = m_c(\mathbf{X}_k^- - \overline{\mathbf{X}_k^-}) + \overline{\mathbf{X}_k^-}, \tag{4.34}$$

prior to the introduction of the predicted model states into the classical EnKF or SQRA. As a result, $\mathbf{X}_k^{'-}$ appears as the predicted ensemble with increased perturbations. This factor avoids fast ensemble convergence due to the reduction of the variances, i.e. the ensemble spread after each filter update. In the SEIK filter, the inverse matrix $(\mathbf{G}^e)^{-1}$ in Eq. (4.19) is replaced by $\frac{1}{m_c}\mathbf{G}^{-1}$, where $\frac{1}{m_c}$ is denoted as forgetting factor in Pham et al. (1998). Furthermore, adaptive inflation factors have been developed that replace manual optimization (e.g., Anderson, 2007, 2009, Li et al., 2009, Miyoshi, 2011). These factors might vary in space and time. Hendricks Franssen and Kinzelbach (2008), for example, discussed the fast ensemble convergence in the joint estimation of states and parameters of a groundwater flow model. They compared adaptive inflation factors and efficient parameter space sampling and found that a combination of both strategies yielded the best result, i.e. the strongest reduction of the fast convergence problem. Nerger et al. (2007) assimilated synthetic sea surface heights into the Finite Element Ocean Model (FEOM) while using a constant forgetting factor to inflate the underestimated variances and to stabilize the SEIK filter process. Applying adaptive inflation factors increases the computational loads of the ensemble filter process, since their estimation requires information about observation and model residuals. Therefore, in this thesis, the constant inflation factor in Eq. (4.34) is implemented. The estimation of variance components might be investigated in future work.

4.3.3 Localization

Covariance Localization

A weighting matrix $\mathbf{M}$, denoted as covariance localization, can be applied to the model prediction covariance matrix $\mathbf{C}^e(\mathbf{x}_k^-)$ to reduce the correlations between model states with large spatial distances (Houtekamer and Mitchell, 2001, Hamill et al., 2001). The formulation of the Kalman gain matrix in Eq. (4.7) is extended to

$$\mathbf{K}_k = (\mathbf{M} \circ \mathbf{C}^e(\mathbf{x}_k^-))\mathbf{A}^T(\mathbf{A}(\mathbf{M} \circ \mathbf{C}^e(\mathbf{x}_k^-))\mathbf{A}^T + \mathbf{\Sigma}_{\mathrm{yy}})^{-1}, \tag{4.35}$$

in which $\circ$ denotes the Schur or Hadamard product that is an element-wise multiplication of two matrices. The weight matrix $\mathbf{M}$ is usually chosen to be a correlation matrix of compact support drawn from a correlation function with local support. Several examples of localization functions with local support can be found, e.g., in Gaspari and Cohn (1999) including the widely used fifth-order polynomial function, which might be interpreted as Gaussian function with local support. Covariance localization is applicable for all filter methods that compute the covariance matrix or its projection on the observation space explicitly, such as the EnKF (section 4.2.1) and SQRA (section 4.2.2). The challenge is to select a reasonable localization distance (e.g., Kirchgessner et al., 2014).

Usually, covariance localization has been applied to pixel data. Its application has never been discussed in the context of GRACE TWSA assimilation. Since GRACE provides the integral of water mass over a specific region, its horizontal distribution into model grid cells and its vertical distribution into individual water compartments needs to be specified by variances and covariances of the model states. For this, the spatial covariances of simulated water states have to be estimated, as well as covariances between different

water compartments, i.e. multi-variate covariances. A data-adaptive approximation of analytically positive definite covariance functions would probably allow a more realistic description of the model error covariance matrix and should be investigated in future work.

Domain Localization

For ensemble filter methods that do not explicitly determine the model prediction error covariance matrix, such as the SEIK filter (section 4.2.3), domain localization can be applied. It is also applicable for the other filters. Therefore, the assimilation process is split up into local regions, in which independent assimilation updates are performed. For each of the local analysis, measurements are considered within a predefined cut-off radius, which is similar to the localization length in covariance localization. Often, domain localization is combined with additional observation localization. The influence of observations with increasing distances is reduced. Details on these methods are given in Ott et al. (2004), Nerger et al. (2006) and Hunt et al. (2007). Comparisons between covariance and domain localization have shown that the performance is similar, where domain localization generally results in weaker localization. However, the optimal correlation length in covariance localization is wider (see e.g., Janjić et al., 2011, Greybush et al., 2011).

5. Implementing C/DA to Merge GRACE and WGHM

In chapters 3 and 4, a calibration and data assimilation (C/DA) framework was introduced. In this chapter, details of implementing the merging process of monthly GRACE total water storage anomalies (TWSA) with the WaterGAP Global Hydrology Model (WGHM) are discussed. To implement the C/DA framework, a modular structure is chosen that separates the model forward integration, assembling the model prediction vectors and observations, as well as the filter algorithms in order to allow a simple exchange of individual program parts (e.g., the filter update method). In this way, only minimal changes of the WGHM source code are required for implementation. This strategy makes the transfer of the proposed C/DA to other hydrological models easier.

In this chapter, first an overview of the developed C/DA procedure is provided (section 5.1). Then, approaches to address the temporal and spatial resolution mismatch between WGHM and GRACE (section 5.2.1 and 5.2.2), as well as the description of model and observation uncertainties are described (section 5.2.3 and 5.2.4). Finally, pseudo codes to implement the ensemble Kalman filter (EnKF), the square root analysis scheme (SQRA), and the singular evolutive interpolated Kalman (SEIK) filter algorithms are presented (section 5.3).

5.1 Overview of the C/DA Procedure

In the following, the C/DA framework developed for assimilating GRACE TWSA data into WGHM is described (see Fig. 5.1). For the model forward integration phase, an ensemble of N_e WGHM runs with different model parameters, climate forcing, and initial water states is performed. A proper selection of probability distributions to generate the required ensembles is discussed in section 5.2.3. All model forward integrations are evaluated in parallel, since the runs are completely independent from each other.

In the filter update, it is intended to correct the monthly means of water states, which are stored as model outputs, in a single river basin. Since WGHM always runs globally, the grid cells are masked out for the basin of interest. This is possible, since the calculation of the water balance equation for each river basin is independent. However, a simultaneous C/DA of multiple river basins can also be performed. In the presented set-up, the usage of a single river basin results in a model prediction state vector $\mathbf{x}_k^{(i)-}$ for each ensemble member, which contains the WGHM water storage states in ten individual model compartments (canopy, snow, soil, local lake, global lake, local wetland, global wetland, reservoir, river and groundwater) for each grid cell of the basin. In order to jointly calibrate model parameters within the assimilation procedure, the state vector is augmented by the WGHM calibration parameters. The design matrix $\mathbf{A}$ establishes the relationship between the model states and the TWSA observed by GRACE. In the following, $\mathbf{A}$ will

be split into a measurement operator **H** that accumulates the ten individual storage compartments for each model grid cell and a mapping operator **B** that spatially averages the gridded model predicted TWSA to match the resolution of GRACE observations stored in the vector **y**.

Either the EnKF, the SQRA or the SEIK filter algorithm can be chosen to perform the update of model water states and parameters. In case of EnKF, an ensemble of observations needs to be generated. Finally, the updated state vector is prepared for the next model forward integration. This procedure is repeated sequentially as long as observations are available.

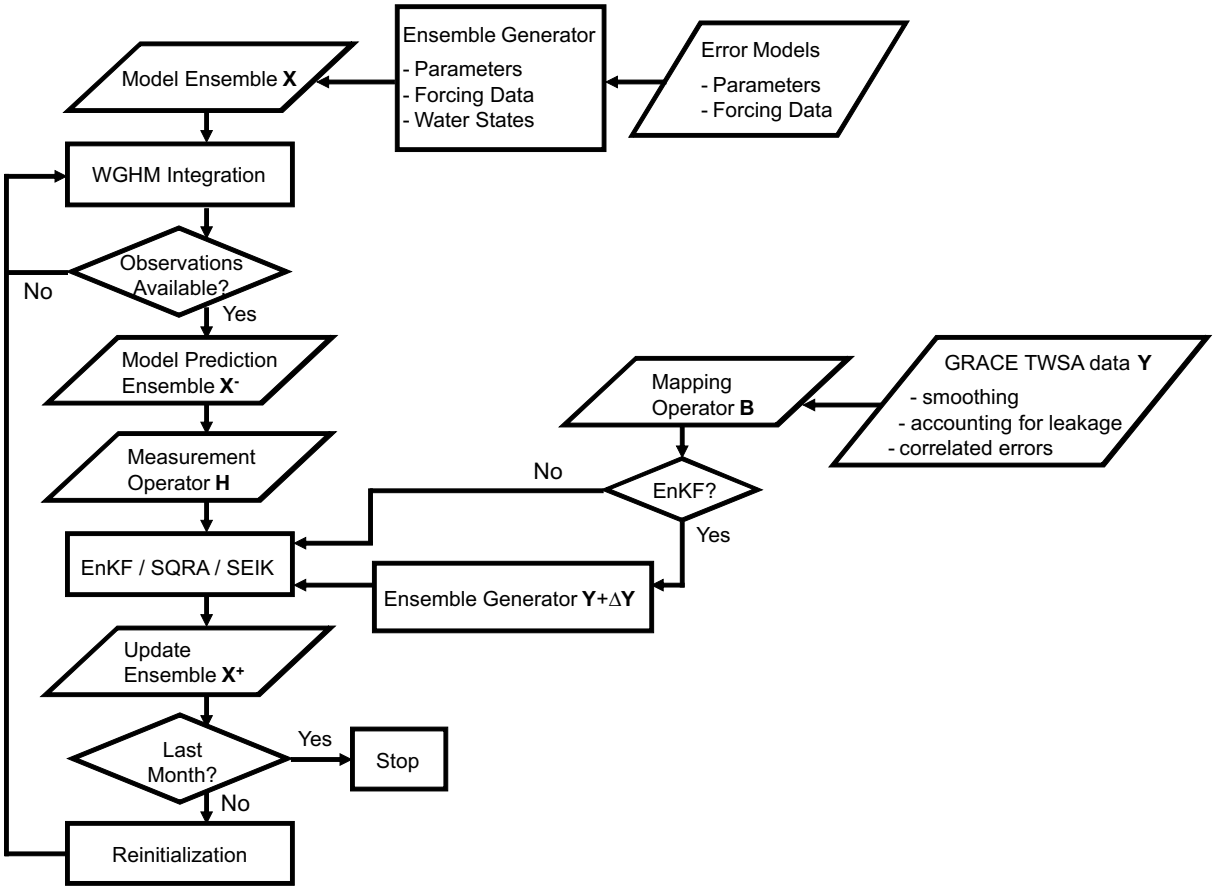

Figure 5.1: Flowchart of the calibration and data assimilation (C/DA) procedure.

5.2 Addressing the Challenges of Merging GRACE TWSA and Models

It has been discussed in section 1.2 that several challenges occur when combining model-predicted water states and TWSA observations from GRACE. In this section, approaches are discussed that deal with the temporal and spatial resolution mismatch between the hydrological model and GRACE TWSA products (section 5.2.1 and section 5.2.2), as well as methods that address the uncertainty estimation for the model forward prediction and for GRACE TWSA observations (section 5.2.3 and section 5.2.4).

5.2.1 Temporal Resolution Mismatch

The simulation step of WGHM is one day (black points in Fig. 5.2 a), whereas only monthly means of TWSA from GRACE are available to inform the model (dashed gray lines). To deal with this resolution mismatch, first, the monthly mean of the daily predicted model states is determined (solid black line). The model predicted monthly TWSA mean is informed by GRACE (dashed black line) using one of the proposed ensemble filter algorithms (section 4.2.1, 4.2.2, and 4.2.3). Since GRACE does not provide information on sub-monthly dependencies, it is assumed that the disagreement between model prediction and observations ΔX can be applied in the same way to each day. Therefore, each of the daily values is shifted by ΔX (gray arrows), leading to daily updates (black stars). The temporal mean of the updated daily values therefore equals exactly the updated monthly mean. To start the next model forward integration, the updated last day of the current month k (black arrow) is used and the model is started for the next month $k+1$. Since at the time being a more sophisticated approach for correcting daily values is not applied, the monthly updated values after completing the C/DA procedure are evaluated. It is important to realize that no daily updates are calculated but only the last day of a month is updated to enable the model forward integration for the next month.

In previous studies, different strategies were applied to deal with the temporal resolution mismatch. Zaitchik et al. (2008) proposed to estimate the monthly mean of the model forward prediction using the 5-th, 15-th and 25-th day of a month to mimic the approximately three overpasses of the GRACE satellites over one river basin (larger black points in Fig. 5.2 b). An ensemble Kalman smoother (EnKS) approach was used to update the model state at the first day of a month based on the covariances between the model predicted first day and the model predicted monthly mean (solid black lines). Subsequently, the update is equally distributed over the days of each month. For this, the model is once again integrated for the current month, and the fraction of the update (ΔX divided by the number of days of the current month) is added to the model for each day. This approach ensures that the discontinuity between the daily values of two successive months (introduced by the update) is small, whereas it is likely larger in the approach applied in this thesis (black dashed arrow in Fig. 5.2 a). A disadvantage of the approach suggested by Zaitchik et al. (2008) is that the model forward simulation needs to be performed twice for the same month. The approach was also applied by Forman et al. (2012), Houborg et al. (2012), Li et al. (2012), as well as Forman and Reichle (2013). Girotto et al. (2016) investigated how introducing the entire update increment at the beginning or at the end of a month influences the assimilation performance and found only a small differences for groundwater and soil moisture simulations. Alternatively, Tangdamrongsub et al. (2015) assumed that the monthly mean of GRACE data corresponds to the middle of a month and they suggested to interpolate the time series at five-day intervals (Fig. 5.2 c). Then, the interpolated GRACE TWSA are assimilated every five days. However, interpolation of monthly to daily observations introduces interpolation errors, which should be considered as an additional source of uncertainties for GRACE observations. A different model-data blending approach is introduced by van Dijk et al. (2014) that only merged monthly values of model and GRACE observations avoiding disaggregation to daily values. This however restricts the evaluation of the updated hydrological model states to monthly resolution, which poses no problem when only the seasonal time-scale is of interest of the study.

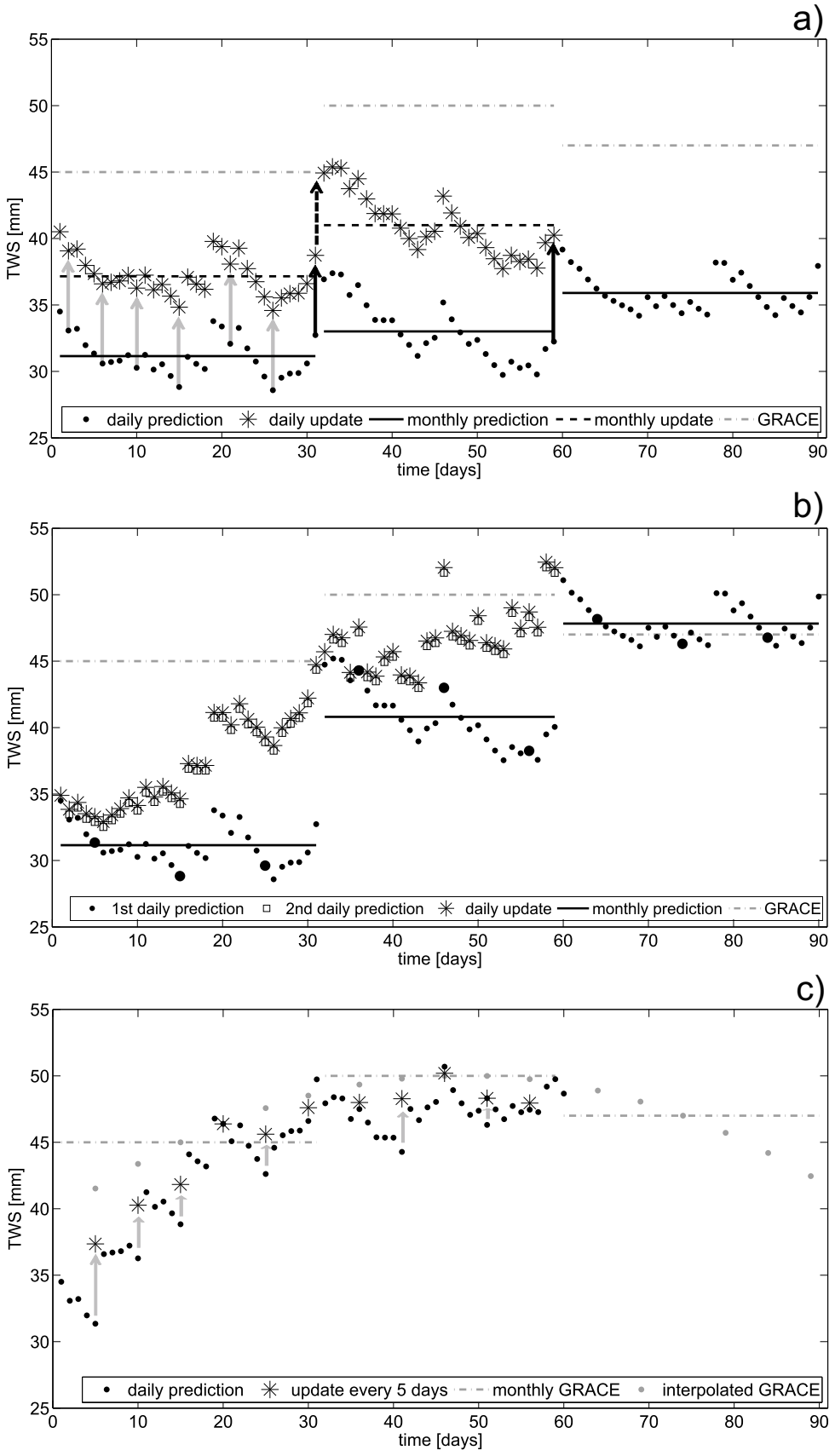

Figure 5.2: Approaches to deal with the temporal resolution mismatch between model predicted and GRACE TWSA, a) which is implemented in this thesis, b) was proposed by Zaitchik et al. (2008), and c) was suggested by Tangdamrongsub et al. (2015).

5.2.2 Spatial Resolution Mismatch

5.2.2.1 Resolution of GRACE TWSA Observations

As mentioned in section 2.2.2, the spatial resolution of WGHM is 0.5°×0.5°. However, an accurate estimation of GRACE TWSA is limited to larger grid size or to (sub-) basin average estimates (see section 2.1). In order to assign a reasonable choice for the spatial resolution of GRACE TWSA observations, the rank and condition number of the observation error covariance matrix $\mathbf{\Sigma}_{yy}$ smoothed with a 500 km Gaussian filter are calculated for the Mississippi River Basin (Eicker et al., 2014), which serves as the main test region of this thesis (see chapter 7 and section 8.1). Table 5.1 reports the matrix dimension, rank (determined by singular value decomposition), and condition number for different spatial grid resolutions. For grid sizes smaller than 2°×2°, the error covariance matrix clearly shows a rank deficiency. Starting from 2°×2°, the matrix has numerically full rank, however, the condition numbers are very poor for small grid sizes, making stable computations very difficult. Therefore, observing the trade-off between spatial resolution and stability of the matrices, an averaging grid size of 5°×5° is proposed (Eicker et al., 2014). This also matches well to the 500 km Gaussian smoothing, which was applied to the GRACE solution for this investigation. The optimal spatial resolution of GRACE data might change when considering other large-scale river basins or when improved GRACE level-2 products with better post-processing (e.g., anisotropic smoothing approaches; see section 2.1.2) are available.

Table 5.1: Matrix dimension, rank, and condition number of the GRACE observation error covariance matrix depending on the selected grid cell size. The table is taken from Eicker et al. (2014).

cell size	matrix dimension	rank	condition number
0.5°×0.5°	1382×1382	128	$4.6 \cdot 10^{20}$
1°×1°	357×357	128	$3.6 \cdot 10^{19}$
2°×2°	91×91	91	$1.3 \cdot 10^{14}$
3°×3°	38×38	38	$2.2 \cdot 10^{07}$
4°×4°	25×25	25	$2.5 \cdot 10^{05}$
5°×5°	17×17	17	$6.4 \cdot 10^{03}$
10°×10°	5×5	5	$4.6 \cdot 10^{01}$

In previous studies either (sub-)basin-averaged (e.g., Zaitchik et al., 2008, Forman et al., 2012, Houborg et al., 2012, Li et al., 2012), 1°×1° gridded GRACE observations (van Dijk et al., 2014, Girotto et al., 2016, Kumar et al., 2016) or 1 km × 1 km gridded observations respectively (Tangdamrongsub et al., 2015) were assimilated into hydrological models. It should be emphasized that neighboring TWSA values on fine grids, i.e. smaller than ∼ 200000 km^2 and therefore below the spatial resolution that can be achieved with GRACE observations, are highly correlated. Forman and Reichle (2013) systematically investigated the effect of spatial aggregation of GRACE TWSA in a data assimilation framework by conducting a synthetic experiment. However, these authors considered only

white noise for simulated TWSA. They investigated the assimilation of TWSA aggregated to one, two, four, and six sub-basins in the Mackenzie River Basin, and concluded that the observations should be assimilated at the smallest spatial scale for which the observations can be considered as uncorrelated, i.e. around 200000-360000 km^2. In the following, similar to Eicker et al. (2014) and Schumacher et al. (2016b), the design matrix $\mathbf{A}$ is implemented to merge e.g., 5°×5° or (sub-)basin average GRACE TWSA with 0.5°×0.5° WGHM simulations.

5.2.2.2 Aggregation and Mapping Operator

First, the water states and model parameters are assembled in the model prediction vector $\mathbf{x}_k^{(i)-}$ for each sample i=1,...,N_e (see Eq. 4.3). The first part of the vector $\mathbf{w}_k^{(i)-}$ holds the individual water compartment states in each grid cell within the river basin of interest, and $\mathbf{p}_k^{(i)-}$ contains the calibration parameters (see Eq. 3.30). Then, the design matrix (e.g., in Eqs. (4.6) to (4.8) of the EnKF, in Eq. (4.9) of the SQRA, and in Eqs. (4.17) to (4.19) of the SEIK filter) is split according to $\mathbf{A} = \mathbf{BH}$, consisting of the vertical aggregation operator $\mathbf{H}$ and the horizontal mapping operator $\mathbf{B}$ (Fig. 5.3). The vertical sum of all modeled storage compartments (canopy, snow, soil, local lake, global lake, local wetland, global wetland, reservoir, river, and groundwater) is determined for each grid cell by incorporating $\mathbf{H}$ that is of dimension $n \times (10 \cdot n + p)$, i.e.

$$\begin{pmatrix} \text{TWS in cell } 1^{(i)} \\ \vdots \\ \text{TWS in cell } n^{(i)} \end{pmatrix}_{n\times 1} = \mathbf{H}_{n\times(10\cdot n+p)}(\mathbf{x}_k^{(i)-})_{(10\cdot n+p)\times 1} \tag{5.1}$$

$$= \mathbf{H}_{n\times(10\cdot n+p)} \begin{pmatrix} (\mathbf{w}_k^{(i)-})_{(10\cdot n)\times 1} \\ (\mathbf{p}_k^{(i)-})_{p\times 1} \end{pmatrix}$$

$$= \begin{pmatrix} 1\ldots1 & 0\ldots0 & \ldots & 0\ldots0 \\ 0\ldots0 & 1\ldots1 & \ldots & 0\ldots0 \\ \vdots & \vdots & \ddots & \vdots \\ \underbrace{0\ldots0}_{10} & \underbrace{0\ldots0}_{10} & \ldots & \underbrace{0\ldots0}_{p} \end{pmatrix} \begin{pmatrix} \text{storage compartments in cell } 1^{(i)} \\ \vdots \\ \text{storage compartments in cell } n^{(i)} \\ \text{WGHM calibration parameters}^{(i)} \end{pmatrix}.$$

The mapping operator $\mathbf{B}$ of dimension $J \times n$ is introduced to upscale the model simulated TWS (here on a 0.5°×0.5° grid) to coarser resolved grid cells or to (sub-)basin averages

$$\begin{pmatrix} \overline{\text{TWS}} \text{ in cell } j=1^{(i)} \\ \vdots \\ \overline{\text{TWS}} \text{ in cell } J^{(i)} \end{pmatrix}_{J\times 1} = \mathbf{B}_{J\times n}\mathbf{H}_{n\times(10\cdot n+p)}(\mathbf{x}_k^{(i)-})_{(10\cdot n+p)\times 1} \tag{5.2}$$

$$= \begin{pmatrix} \frac{a_1}{A_1} & 0 & \frac{a_3}{A_1} & \ldots & 0 \\ 0 & \frac{a_2}{A_2} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \frac{a_n}{A_J} \end{pmatrix} \begin{pmatrix} \text{TWS in cell } 1^{(i)} \\ \vdots \\ \text{TWS in cell } n^{(i)} \end{pmatrix}.$$

Each row of $\mathbf{B}$ has elements $\frac{a_i}{A_j}$ with $i = 1\ldots n_j$ containing the area fraction for each of the n_j gridded values on the 0.5°×0.5° grid belonging to the larger cell j. The total

area of cell j is given by $A_j = \sum a_i$. As a limit case, for $J = 1$ the GRACE TWSA basin average is used. The vector on the left side of Eq. (5.2) therefore holds the model predicted TWS, which are combined with the TWSA observations from GRACE. Thus, in all approaches, i.e. EnKF, SQRA, and SEIK, the design matrix is replaced by the product of the aggregation **H** and the mapping operator **B**, as it was done in Eicker et al. (2014) and Schumacher et al. (2016b).

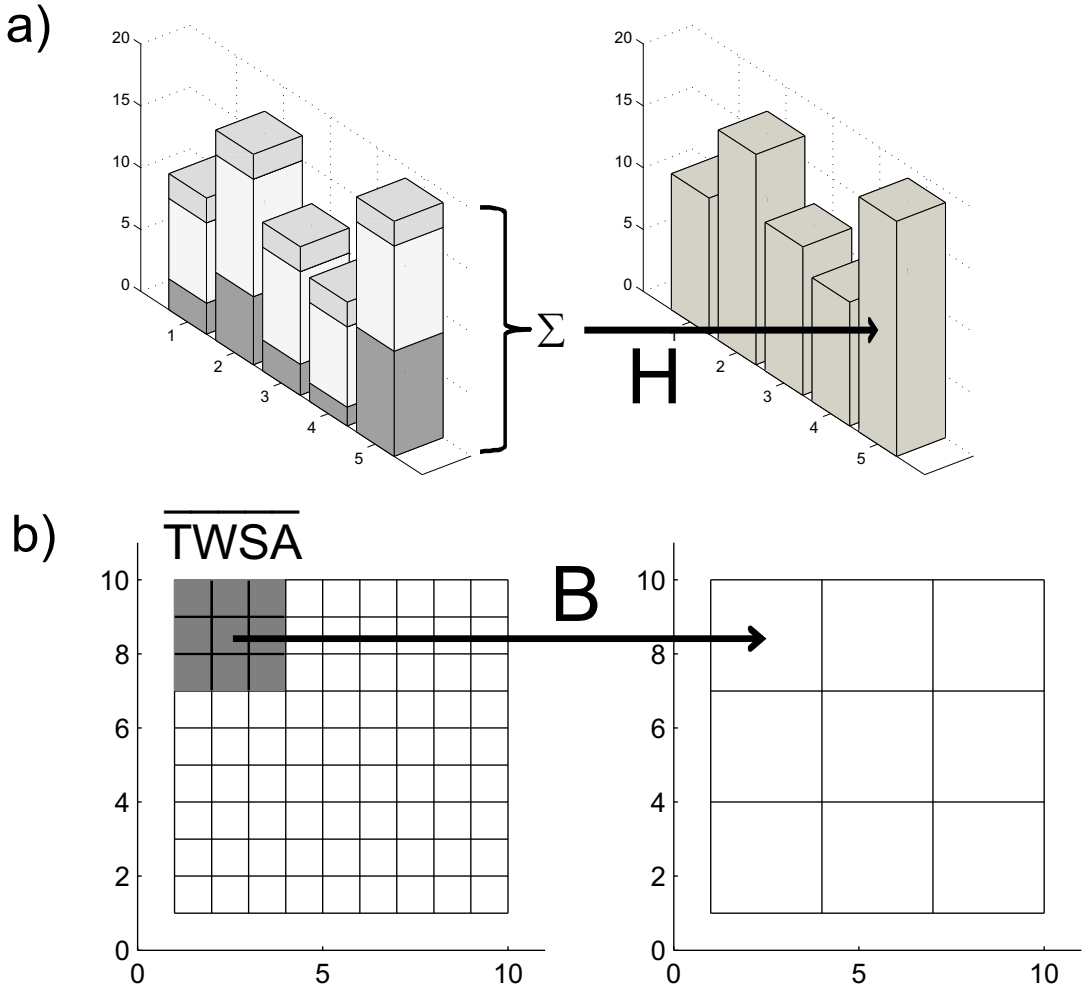

Figure 5.3: Schematic visualization of measurement and mapping operators, where **H** in a) vertically sums up the model storage compartments (left) to be comparable with GRACE TWSA (right), and in b) the operator **B** provides the spatial average of the model grids (left) to be equivalent with GRACE resolution (right). This figure is taken from Schumacher et al. (2016b).

5.2.3 Errors of Model Forward Integration

It was mentioned in section 1.2 that uncertainties in model parameters, forcing data, initial water states, and errors in the model structure have a significant impact on the accuracy of model simulations. Since the WGHM equations are non-linear, an ensemble of model runs is used to represent the errors of parameters, forcing and water states and to propagate them to the model output.

For this, WGHM is initialized over a minimum of five years to ensure a realistic fill level of the individual water compartments. Subsequently, an ensemble of N_e model runs is configured. An ensemble of model parameters (Tab. 2.1) according to predefined distributions is generated using the Latin-Hypercube method (Mckay et al., 2000). Most of the parameters are assumed to be triangular distributed with the parameter value that is used in the standard WGHM runs as mode (which is the highest value of a triangular distribution, see Tab. 2.1). The lower and upper limits are defined according to Döll et al. (2003), Kaspar (2004), Güntner et al. (2007) and Hunger and Döll (2008). The multipliers that are marked with (*) in Tab. 2.1 are not integrated in the standard WaterGAP 2.2 version, but are considered as calibration parameters within the C/DA framework.

Additionally, precipitation and temperature forcing fields are perturbed using random Monte Carlo sampling from triangular distributions. An additive error is assumed for temperature, centered at 0°C with the maximum limits of ± 2°C. A multiplicative error model is introduced for precipitation, centered at 1 with minimum and maximum limits of 0.7 and 1.3. These values were chosen based on personal communication with Prof. Dr. rer. nat. Petra Döll and Dipl.-Geogr. Hannes Müller Schmied (University of Frankfurt, Germany). Realizations of errors are drawn from triangular distributions using random Monte Carlo sampling. Additionally, a precipitation multiplier is calibrated which accounts for a systematic under- or overestimation of observed precipitation (systematic sensor error) and is identical for all grid cells within one basin. In contrast, the perturbations of the gridded precipitation and temperature input fields are different for each grid cell within a basin and account for random errors. It is worth mentioning that the ensembles of climate fields are only used to represent uncertainties in forcing input data to ensure a realistic representation of water state uncertainties. However, the climate fields are not included in the C/DA update.

Afterwards, a model spin-up phase over a few years (e.g., two years) with the parameter and forcing ensembles is performed to generate an ensemble of initial water states. As an example, in Fig. 5.4, the spread of the TWS ensemble over the Mississippi River Basin (used as test region in chapter 7 and section 8.1) after a two years spin-up phase (2002-2003) is shown for a three years model ensemble run (2004-2006). The model prediction error covariance matrix can be empirically determined by Eq. (4.4) or alternatively by Eq.(4.5) using the model output of monthly means of compartmental water states (canopy, snow, soil, local lake, global lake, local wetland, global wetland, reservoir, river and groundwater) and the calibration parameters.

In absence of reliable information about the errors in the model structure, a constant inflation factor (section 4.3.2) is used in the EnKF and in the SQRA approach to enlarge the ensemble spread prior to the filter update. Similarly, the inverse inflation factor, which is denoted as forgetting factor, is introduced in the SEIK filter (section 4.3.2). This factor also mitigates ensemble convergence. The range of possible values for the inflation factor are selected based on previous data assimilation experiments reported in Janjić et al. (2011) and Nerger et al. (2012). Subsequently, several C/DA runs are performed and a factor of 10% is selected as small as possible to avoid a strong influence on the model ensemble and large enough to ensure a contribution of the observations. Hendricks Franssen and Kinzelbach (2008) discussed the fact that fast ensemble convergence is problematic for the joint estimation of model states and parameters. Since previous GRACE

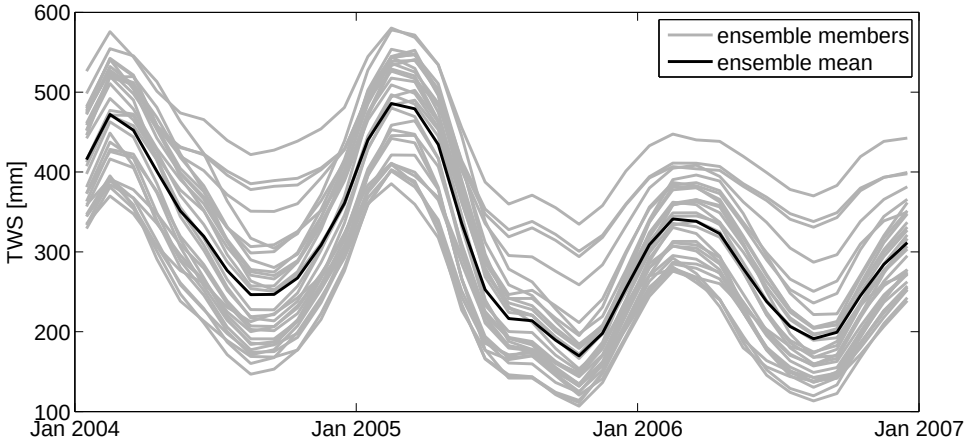

Figure 5.4: Ensemble mean and 30 ensemble members of TWS for a three years model ensemble run. The ensemble is generated by sampling of 22 calibration parameters and perturbed forcing data (precipitation and temperature). The run is performed after a two years spin-up phase to generate an ensemble of initial water states.

data assimilation studies focused on data assimilation and no attempt was undertaken to jointly update model parameters, fast ensemble convergence was not identified as a problem and no inflation factors were considered. Thus, the model uncertainty estimations did not include a description of errors in the model structure, for which the inflation factor is applied in this thesis.

In the ensemble filter updates, only the first and second order statistical moments are used to construct the Kalman gain matrix (Eq. (4.7)). As an example, the model error covariance matrix for 01/2004 is estimated (after the two years spin-up phase). In order to check whether the total and compartmental water states from WGHM are Gaussian distributed, their histograms are plotted. Alternatively, a Kolmogoroff-Smirnow test could have been applied to test the Gaussianity of the water states (Koch, 1999, p. 271). In case of Gaussian distributions, the histogram is bell-shaped, which is the case for the simulated TWSA shown in Fig. 5.5 a. Therefore, assuming a Gaussian distribution might be adequate for TWSA. The histogram of groundwater might be considered as a skewed

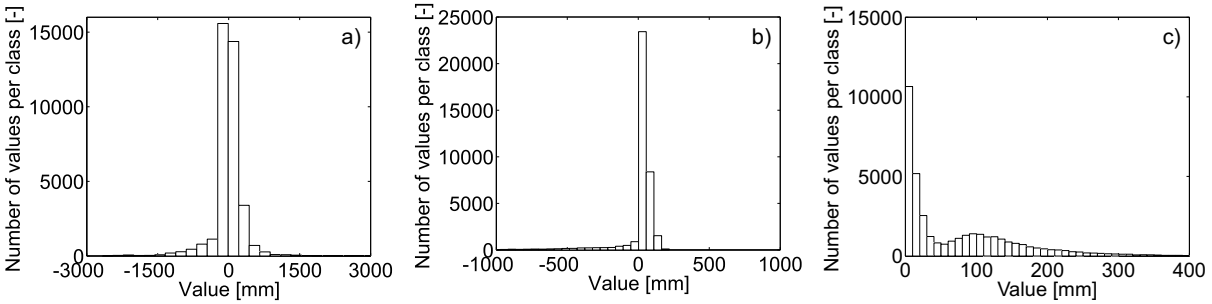

Figure 5.5: Histograms of simulated a) TWSA, b) groundwater, and c) soil water for 01/2004 after a two years spin-up phase. In this example, 30 ensemble members are generated and 1262 grids cells within the Mississippi River Basin are used for calculations. In each case 50 bins are selected for visualization of the histograms.

Gaussian distribution (Fig. 5.5 b), and that of the soil water compartment is found to

be clearly non-Gaussian distributed (Fig. 5.5 c). Therefore, the Gaussian assumption might not be adequate for the groundwater storage and in particular for the soil water compartment. Since the EnKF, SQRA, and SEIK update steps only take the first and second order statistical moments into account, the non-Gaussian parts will be neglected. In contrast, in the model forward integration non-Gaussian information is considered. Even by starting the model run with a Gaussian distributed ensemble of water states, the model will introduce non-Gaussianity as a result of its non-linear equations. Since the non-Gaussian parts are neglected in the update steps, the EnKF, SQRA and SEIK are "suboptimal" filters.

5.2.4 Errors of GRACE TWSA Observations

As described in Schumacher et al. (2016b), various levels of approximation can be applied to account for the uncertainty of GRACE TWSA observations in the C/DA procedure, three of which will result in white noise and the other two in spatially correlated errors (Fig. 5.6): The assumption of white noise can be made by either (1) using standard deviations based on literature, e.g., Wahr et al. (2006) (used, e.g., in Zaitchik et al., 2008, Su et al., 2010, Forman et al., 2012, Forman and Reichle, 2013), (2) propagating errors from standard deviations of GRACE level-2 potential coefficients or (3) propagating errors from the full covariance matrix of GRACE level-2 potential coefficients to standard deviations of TWSA. Correlated error samples can be generated from (4) error propagation of standard deviations of potential coefficients or from (5) propagation of the full error covariance matrix of potential coefficients to a full covariance matrix of TWSA (as shown in Eicker et al., 2014, Schumacher et al., 2016b,c). Alternatively, the three-cornered hat method (Tavella and Premoli, 1994) can be applied to quantify the errors in one GRACE product by comparing it to other products (used in van Dijk et al., 2014), in which the products are assumed to be independent from each other.

According to method (1), for instance, Zaitchik et al. (2008), Su et al. (2010), Forman et al. (2012), and Forman and Reichle (2013) assumed spatially uncorrelated errors for TWSA with variances depending on the size of river (sub-)basins, i.e. between 8^2 mm^2 for large areas of about 1.6 million km^2 and 20^2 mm^2 for smaller sub-basins of about 200000-300000 km^2. Tangdamrongsub et al. (2015) also used uncorrelated TWSA errors with a variance of 20^2 mm^2 but for 1 km $\times$ 1 km gridded observations. In this case, the assumption of white noise and the selected error variance are not realistic since GRACE does actually not resolve such small spatial scales. The signal and errors are both highly correlated in space, and the observation error is expected to be large, i.e. larger than the signal itself. Therefore, the gain matrix would exhibit negligible values for the observations and assimilation of GRACE would rarely influence the model simulations. In contrast, in this thesis, correlated GRACE TWSA errors according to method (5) are used, since it considers all available error information in the assimilation procedure in the best manner. The mathematical procedure of formal error propagation was described in section 2.1.4.

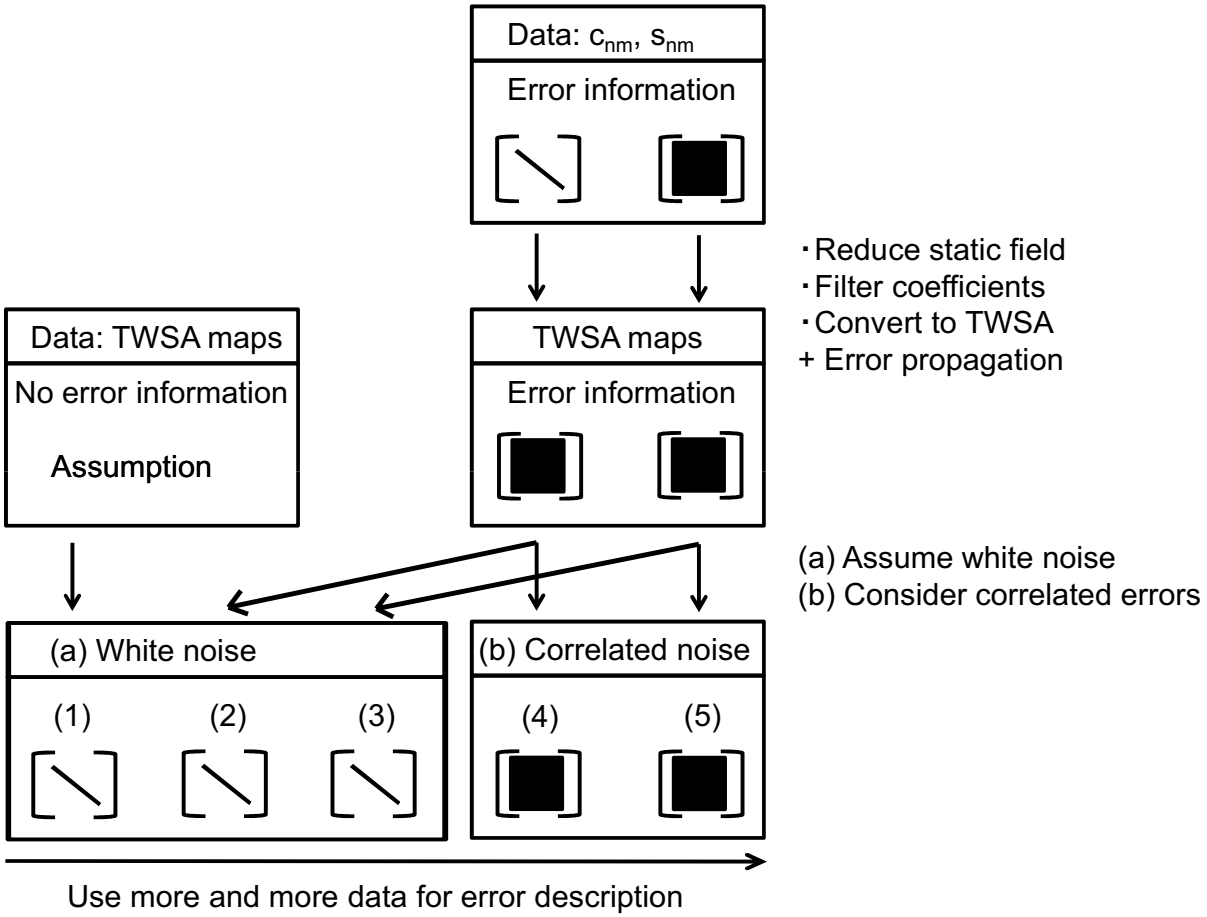

Figure 5.6: GRACE TWSA error description: (1) using standard deviations based on literature; propagating standard deviations of potential coefficients c_{nm} and s_{nm} (2) to standard deviations or (4) to correlated errors of TWSA, and propagating correlated errors of potential coefficients to (3) standard deviations or (5) correlated errors of TWSA. This figure is taken from Schumacher et al. (2016b).

5.3 Practical Implementation of the Filter Methods

In the following, pseudo codes are provided for the EnKF (section 4.2.1), SQRA (section 4.2.2), and SEIK (section 4.2.3) filter methods. These codes are not in a strict sense pseudo codes but should be better interpreted as "implementation guides". In addition, the Householder transformation is described, since it is an important part of the SEIK algorithm.

5.3.1 Implementing the EnKF

In this section, Algorithm 1 provides an implementation guide for determining the updated WGHM water states and parameters, as well as the corresponding error information using the classical EnKF. First, the ensemble mean $\overline{\mathbf{x}}^-$ of the model prediction vectors, which are stored in the matrix $\mathbf{X}^-$ with n rows (number of model states) and N_e columns (ensemble size), is calculated (line 1 of Algorithm 1). The operator $\mathbf{X}^-(:,i)$ refers to the

Algorithm 1: EnKF algorithm.

Data:
$\mathbf{X}^-$ ... Matrix of model prediction ensemble ($(10n+p)\times N_e$)
$\mathbf{Y}+\Delta\mathbf{Y}$ ... Matrix of observation ensemble ($J \times N_e$)
$\mathbf{\Sigma_{yy}}$ ... Error covariance matrix of observations ($J \times J$)
$\mathbf{H}$ ... Measurement operator ($n\times(10n+p)$)
$\mathbf{B}$ ... Mapping operator ($J \times n$)
f_c ... Inflation factor (1×1)
N_e ... Ensemble size (1×1)

Result:
$\mathbf{X}^+$ ... Matrix of model update ensemble ($(10n+p)\times N_e$)
$\mathbf{C}^e(\mathbf{x}^+)$ ... Error covariance matrix of model update ($(10n+p)\times(10n+p)$)

1: mean of model prediction ensemble
$\bar{\mathbf{x}}^- = N_e^{-1}\sum_{i=1}^{N_e}\mathbf{X}^-(:,i)$
2: ensemble inflation (not in classical EnKF)
for *i=1:*N_e **do**
$\mathbf{X}^-(:,i) = f_c(\mathbf{X}^-(:,i)-\bar{\mathbf{x}}^-)+\bar{\mathbf{x}}^-$
end
3: empirical covariance matrix of model prediction
a: reduce mean from each ensemble member
for *i=1:*N_e **do**
$\Delta\mathbf{X}^-(:,i) = \mathbf{X}^-(:,i)-\bar{\mathbf{x}}^-$
end
b: error covariance matrix of predicted water states and parameters
$\mathbf{C}^e(\mathbf{x}^-) = (N_e-1)^{-1}\Delta\mathbf{X}^-(\Delta\mathbf{X}^-)^T$
c: error covariance matrix of predicted TWS
$\mathbf{C}^e(\mathbf{x}^-_{\mathrm{TWS}}) = \mathbf{BHC}^e(\mathbf{x}^-)(\mathbf{BH})^T$
4: gain matrix $\mathbf{K} = \mathbf{K}'\mathbf{B}$
a: $\mathbf{W} = \mathbf{C}^e(\mathbf{x}^-_{\mathrm{TWS}})+\mathbf{\Sigma_{yy}}$
b: solve $\mathbf{L} = (\mathbf{BH})^T\mathbf{W}^{-1}$ for $\mathbf{L}$ (avoid explicit inversion of $\mathbf{W}$)
c: $\mathbf{K}' = \mathbf{C}^e(\mathbf{x}^-)\mathbf{L}$
5: EnKF update of model states and parameters
$\mathbf{X}^+ = \mathbf{X}^- + \mathbf{K}'((\mathbf{Y}+\Delta\mathbf{Y})-\mathbf{BHX}^-)$
6: mean of model update ensemble
$\bar{\mathbf{x}}^+ = N_e^{-1}\sum_{i=1}^{N_e}\mathbf{X}^+(:,i)$
7: covariance matrix of model update
$\mathbf{C}^e(\mathbf{x}^+) = (\mathbf{I}-\mathbf{K}'\mathbf{BBH})\mathbf{C}^e(\mathbf{x}^-)$

i-th column of the matrix. Optionally, an inflation of the ensemble perturbations may be applied by multiplication with the factor f_c >1 (line 2) in order to avoid fast ensemble convergence (see section 4.3.2). If the factor is equal to 1, no inflation is considered. The possibly inflated ensemble perturbations are collected in the matrix $\Delta\mathbf{X}^-$ (line 3a) and used to determine the empirical error covariance matrix of the WGHM predicted water states and model parameters (line 3b). In addition, by incorporating the measurement and mapping operators $\mathbf{H}$ and $\mathbf{B}$ (introduced in section 5.2.2.2) the covariance matrix of model predicted TWS is obtained (line 3c). The gain matrix $\mathbf{K}$ can be calculated using

Eq. (4.7), while avoiding the explicit inversion of the sum of model predicted and observed TWS covariances (line 4a,b). Since the mapping operator $\mathbf{B}$ is already incorporated with the perturbed measurements $\mathbf{Y} + \Delta\mathbf{Y}$, the gain matrix is assembled except for the multiplication with matrix $\mathbf{B}$ (line 4c). The individual EnKF update of each model prediction vector is estimated using Eq. (4.6) (line 5) and the final model update $\bar{\mathbf{x}}^+$ is obtained by evaluating the ensemble mean of the model update vectors (line 6). The error covariance matrix of the model update $\mathbf{C}^e(\mathbf{x}^+)$ is calculated using formal error propagation (line 7).

5.3.2 Implementing the SQRA Method

In section 4.2, the usage of alternative filter algorithms to update the WGHM water states and parameters by integrating GRACE TWSA was motivated. Therefore, in Algorithm 2, an implementation guide for the SQRA method is provided. Steps 1 to 3b, i.e. the determination of the model prediction ensemble mean and the corresponding empirical error covariance matrix (and possible inflation), are identical to those of the EnKF (Algorithm 1). However, to avoid an explicit calculation of the Kalman gain matrix in Eq. (4.9), first, the ensemble perturbations of the model predicted TWS are stored in the matrix $\mathbf{S}$ with J rows (number of observations) and N_e columns (ensemble size; line 4), which is then used to determine the sum of the covariances including model predicted and observed TWSA, i.e. $\mathbf{C}$ (line 5). This is equal to step 4a in Algorithm 1 except for the factor $(N_e - 1)$, i.e. $\mathbf{C} = (N_e - 1)\mathbf{W}$. An eigenvalue decomposition of $\mathbf{C}$ is performed (line 6) and used to determine several auxiliary variables ($\mathbf{y}_1$ to $\mathbf{y}_4$ in line 7). These are useful for an efficient implementation of the SQRA update of the ensemble mean of WGHM water states and model parameters $\bar{\mathbf{x}}^+$ (line 8). As described in section 4.2.2, the SQRA consists of two steps, and therefore after updating the model ensemble mean, an update of the ensemble perturbations is required (see Eq. (4.15)). From Eq. (4.8), it is obvious that the model update error covariance matrix $\mathbf{C}^e(\mathbf{x}^+)$ is not expressed by symmetric matrix operations. The following auxiliary matrices are therefore introduced to obtain such a symmetric expression for $\mathbf{C}^e(\mathbf{x}^+)$. An auxiliary matrix is constructed based on the ensemble perturbations of model predicted TWS and the decomposed sum of the error covariance matrices (line 9a in Algorithm 2). Then, a singular value decomposition of this matrix is performed (line 9b), which allows to represent $\mathbf{C}^e(\mathbf{x}^+)$ by symmetric matrix operations (see Eq. (4.14)). Additionally, a random orthonormal matrix $\boldsymbol{\Theta}$ is generated based on an eigenvalue decomposition of a random matrix $\mathbf{M}$ (line 9c-d). The matrix $\boldsymbol{\Theta}$ is subsequently used to update the ensemble perturbations according to the second part of Eq. (4.15). Finally, the update of each model ensemble member is estimated by adding the updated ensemble perturbations to the updated ensemble mean (line 10) and the corresponding error information is empirically determined from the updated ensemble perturbations (line 11).

5.3.3 Implementing the SEIK Filter

Algorithm 3 provides a guide for implementing the SEIK algorithm. As for the EnKF and SQRA, the algorithm starts with the determination of the ensemble mean of the

Algorithm 2: SQRA algorithm.

Data:
$\mathbf{X}^-$ ... Matrix of model prediction ensemble ($(10n+p)\times N_e$)
$\mathbf{y}$ ... Vector of observations ($J\times 1$)
$\mathbf{\Sigma_{yy}}$ ... Error covariance matrix of observations ($J \times J$)
$\mathbf{H}$ ... Measurement operator ($n\times(10n+p)$)
$\mathbf{B}$ ... Mapping operator ($J \times n$)
f_c ... Inflation factor (1×1)
N_e ... Ensemble size (1×1)

Result:
$\mathbf{X}^+$ ... Matrix of model update ensemble ($(10n+p)\times N_e$)
$\mathbf{C}^e(\mathbf{x}^+)$ ... Error covariance matrix of model update ($(10n+p)\times(10n+p)$)

1: mean of model prediction ensemble
$\bar{\mathbf{x}}^- = N_e^{-1}\sum_{i=1}^{N_e}\mathbf{X}^-(:,i)$
2: ensemble inflation (not in classical SQRA)
for *i=1:*N_e **do**
$\mathbf{X}^-(:,i) = f_c(\mathbf{X}^-(:,i) - \bar{\mathbf{x}}^-) + \bar{\mathbf{x}}^-$
end
3: empirical covariance matrix of model prediction
a: reduce mean from each ensemble member
for *i=1:*N_e **do**
$\Delta\mathbf{X}^-(:,i) = \mathbf{X}^-(:,i) - \bar{\mathbf{x}}^-$
end
b: error covariance matrix of predicted water states and parameters
$\mathbf{C}^e(\mathbf{x}^-) = (N_e-1)^{-1}\Delta\mathbf{X}^-(\Delta\mathbf{X}^-)^T$
4: matrix of ensemble perturbations of model predicted TWS
$\mathbf{S} = \mathbf{BH}\Delta\mathbf{X}^-$
5: sum of error covariance matrices of model predicted and observed TWS
$\mathbf{C} = \mathbf{SS}^T + (N_e-1)\mathbf{\Sigma_{yy}}$
6: eigenvalue decomposition of $\mathbf{C}$
$\mathbf{C} = \mathbf{ZLZ}^T$
7: auxiliary variables
$\mathbf{y}_1 = \mathbf{Z}^T(\mathbf{y} - \mathbf{BH}\bar{\mathbf{x}}^-)$
$\mathbf{y}_2 = \mathbf{L}^{-1}\mathbf{y}_1$
$\mathbf{y}_3 = \mathbf{Z}\mathbf{y}_2$
$\mathbf{y}_4 = \mathbf{S}^T\mathbf{y}_3$
8: SQRA update of ensemble mean of model states and parameters
$\bar{\mathbf{x}}^+ = \bar{\mathbf{x}}^- + \Delta\mathbf{X}^-\mathbf{y}_4$
9: SQRA update of ensemble perturbations
a: $\mathbf{X}_2 = \mathbf{L}^{\frac{1}{2}}\mathbf{Z}^T\mathbf{S}$
b: singular value decomposition of $\mathbf{X}_2$
$\mathbf{X}_2 = \mathbf{U}_2\mathbf{S}_2\mathbf{V}_2^T$
c: generate matrix $\mathbf{M}$ that holds uniformly distributed random numbers
d: generate random orthonormal matrix $\mathbf{\Theta}^T$ from eigenvalue decomposition of $\mathbf{M}$
$\mathbf{M} = \mathbf{\Theta\Lambda\Theta^T}$
e: $\Delta\mathbf{X}^+ = \Delta\mathbf{X}^-\mathbf{V}_2(\mathbf{I} - \mathbf{S}_2^T\mathbf{S}_2)^{\frac{1}{2}}\mathbf{\Theta}^T$
10: SQRA update of model state and parameter ensemble
$\mathbf{X}^+ = \bar{\mathbf{X}}^+ + \Delta\mathbf{X}^+$
11: covariance matrix of model update
$\mathbf{C}^e(\mathbf{x}^+) = \Delta\mathbf{X}^-(\mathbf{I} - \mathbf{X}_2^T\mathbf{X}_2)(\Delta\mathbf{X}^-)^T$

model states and parameters (line 1). However, the empirical error covariance matrix of the model prediction is not calculated explicitly but represented by the matrices $\mathbf{L}$ and $\mathbf{G}$ (defined in Eq.(4.16)) with the help of an auxiliary variable $\mathbf{T}$ (line 2a-c). In line 2a, $\mathbf{I}$ denotes the identity matrix and $\mathbf{1}$ is a matrix containing the number one for each element. Similar to the SQRA approach, the gain matrix is not determined explicitly but

Algorithm 3: SEIK algorithm.

Data:
$\mathbf{X}^-$ ... Matrix of model prediction ensemble ($(10n+p)\times N_e$)
$\mathbf{y}$ ... Vector of observations ($J\times 1$)
$\mathbf{\Sigma_{yy}}$ ... Error covariance matrix of observations ($J \times J$)
$\mathbf{H}$ ... Measurement operator ($n\times(10n+p)$)
$\mathbf{B}$ ... Mapping operator ($J \times n$)
f_c ... Inflation factor (1×1)
N_e ... Ensemble size (1×1)

Result:
$\mathbf{X}^+$ ... Matrix of model update ensemble ($(10n+p)\times N_e$)
$\mathbf{C}^e(\mathbf{x}^+)$ ... Error covariance matrix of model update ($(10n+p)\times(10n+p)$)

1: mean of model prediction ensemble
$\bar{\mathbf{x}}^- = N_e^{-1}\sum_{i=1}^{N_e}\mathbf{X}^-(:,i)$
2: empirical error covariance matrix of model prediction
a: $\mathbf{T} = \begin{pmatrix} \mathbf{I}_{[(N_e-1)\times(N_e-1)]} \\ \mathbf{0}_{[1\times(N_e-1)]} \end{pmatrix} - \frac{1}{N_e}\mathbf{1}_{[N_e\times(N_e-1)]}$
b: $\mathbf{G} = N_e^{-1}(\mathbf{T}^T\mathbf{T})^{-1}$
c: $\mathbf{L} = \mathbf{X}^-\mathbf{T}$
3: SEIK update of ensemble mean of model states and parameters and ensemble inflation (not in classical SEIK algorithm)
a: $\mathbf{U} = [(f_c^{-1}\mathbf{G}^{-1} + (\mathbf{BHL})^T\mathbf{\Sigma_{yy}^{-1}}\mathbf{BHL}]^{-1}$
b: $\mathbf{a} = \mathbf{U}(\mathbf{BHL})^T\mathbf{\Sigma_{yy}^{-1}}(\mathbf{y} - \mathbf{BH}\bar{\mathbf{x}}^-)$
c: $\bar{\mathbf{x}}^+ = \bar{\mathbf{x}}^- + \mathbf{La}$
4: covariance matrix of model update
$\mathbf{C}^e(\mathbf{x}^+) = \mathbf{LUL}^T$
5: SEIK update of ensemble perturbations
a: Cholesky decomposition of (N_e-1)-rank matrix $\mathbf{U}$
$\mathbf{U} = \mathbf{C}^T\mathbf{C}$
b: generate orthonormal matrix using Householder transformation
$\mathbf{\Omega}$ = generateOmega(N_e) (see Algorithm 4)
c: update of model state and parameter ensemble
$\mathbf{X}^+ = \bar{\mathbf{X}}^+ + \sqrt{N_e}\mathbf{LC}^T\mathbf{\Omega}^T$

the auxiliary matrix $\mathbf{U}$ and the vector $\mathbf{a}$ (line 3a-b) are used for an efficient estimation of the update of the ensemble mean of model states and parameters $\bar{\mathbf{x}}^+$ (line 3c). The error covariance matrix of the model update is obtained using formal error propagation (line 4). As for the SQRA method, the SEIK filter also requires an update of the ensemble perturbations. For this, Cholesky decomposition of the matrix $\mathbf{U}$ is performed (line 5a), which enables a representation of the model update error covariance matrix by symmetric

matrix operations (see Eq. (4.23)) as it was also required in the SQRA method. Finally, by means of the matrix $\mathbf{\Omega}$, whose columns are orthonormal and orthogonal to a vector that contains only ones (line 5b, see also Algorithm 4), the update of each individual ensemble member is calculated (line 5c).

Algorithm 4: Householder transformation.

Data:
N_e ... Ensemble size (1×1)

Result:
$\mathbf{\Omega}$ ... Orthonormal matrix whose columns are orthogonal to a vector containing only the number one ($N_e \times (N_e - 1)$)

1: Initialization
$\mathbf{S}$ =1 or -1
2: Generate orthonormal matrix
for *k=2:(N_e−1)* **do**
a: uniformly distributed random vector with length k and norm one
$\mathbf{z} = \text{rand}(k, 1)$
$\mathbf{z} = \mathbf{z}/\text{norm}(\mathbf{z})$
b: Householder matrix associated to vector $\mathbf{z}$
$\mathbf{J} = \mathbf{I} - \frac{1}{1-\mathbf{z}(1)} \begin{pmatrix} \mathbf{z}(1) - 1 \\ \mathbf{z}(2:\text{end}) \end{pmatrix} \begin{pmatrix} \mathbf{z}(1) - 1 & \mathbf{z}(2:\text{end})^T \end{pmatrix}$
c: matrix defined by $k - 1$ columns
$\mathbf{M} = \mathbf{J}(:, 2:\text{end})$
d: Householder transformation
$\mathbf{T} = \begin{pmatrix} \mathbf{z} & \mathbf{MS} \end{pmatrix}$
e: set matrix of current step to matrix of previous step
$\mathbf{S} = \mathbf{T}$
end
3: make columns of $\mathbf{S}$ orthogonal to a vector containing only the number one
a: vector that contains only the number one divided by N_e with length N_e and norm one
$\mathbf{e} = \text{ones}(N_e, 1)$
$\mathbf{e} = \mathbf{e}/\text{norm}(\mathbf{e})$
b: Householder matrix associated to vector $\mathbf{e}$
$\mathbf{J} = \mathbf{I} - \frac{1}{1-\mathbf{e}(1)} \begin{pmatrix} \mathbf{e}(1) - 1 \\ \mathbf{e}(2:\text{end}) \end{pmatrix} \begin{pmatrix} \mathbf{e}(1) - 1 & \mathbf{e}(2:\text{end})^T \end{pmatrix}$
c: matrix defined by $k - 1$ columns
$\mathbf{M} = \mathbf{J}(:, 2:\text{end})$
4: Householder transformation to obtain orthonormal matrix whose columns are orthogonal to a vector containing only the number one
$\mathbf{\Omega} = \mathbf{MS}$

To compute the matrix $\mathbf{\Omega}$ (line 5b of Algorithm 3), an auxiliary matrix $\mathbf{S}$ is initialized by randomly choosing 1 or -1 (line 1 of Algorithm 4). Then, a uniformly distributed normalized random vector is sampled (line 2a). The Householder matrix $\mathbf{J}$ is determined (line 2b), from which only the last $k - 1$ columns are used (line 2c) to calculate the orthonormal matrix $\mathbf{T}$ (line 2d), which is extended in the next step (line 2e). Finally, $\mathbf{S}$ is transformed in a way that its columns are orthogonal to a vector containing only ones. For this, a vector containing ones for each element is generated and normalized (line 3a).

The Householder matrix associated with this vector is determined (line 3b) and its last $N_e - 1$ columns (line 3c) are used to transform the orthonormal matrix (line 4), and $\mathbf{\Omega}$ is then used in the SEIK filter (Algorithm 3).

6. Covariance Analysis and Sensitivity Study

Before developing the calibration and data assimilation (C/DA) framework (objective 1 of this thesis), a covariance analysis and sensitivity study is performed. In a sensitivity study, a large number of model parameter values is generated, the so-called parameter ensemble, and it is used to perform an ensemble of model runs. The ensemble of model outputs is then analyzed to determine those model parameters that exhibit the highest influence on the model outputs (Hamby, 1994). In the context of model calibration, a sensitivity study helps to decide which of the model parameters are crucial to be calibrated. In this thesis, the sensitivity analysis is also useful to assess the empirical error covariance matrix of WGHM simulations, which is used to estimate correlation coefficients between model parameters and water state outputs.

First, the sensitivity index (SI, Hoffman and Gardner, 1983), the Spearman's rank correlation coefficient (SRCC, Iman and Conover, 1979) and the empirical model covariance matrix are introduced in section 6.1 as measures of sensitivity. Second, an ensemble of WGHM runs is used to investigate whether it adequately represents the standard WGHM run (section 6.2.1). Then, the sensitivity of the water storage in individual compartments with respect to changes in model parameter values is investigated in section 6.2.2 and the results are presented for the Mississippi River Basin. A global sensitivity analysis was performed for the 33 largest river basins world-wide to assess which parameters dominantly influence the simulation of total water storage anomalies (TWSA). The results with respect to the Mississippi and Murray-Darling River Basins are presented in section 6.2.3. Both river basins will serve as test regions in chapters 7 and 8, where the C/DA framework is implemented. In section 6.3, the main findings of the sensitivity investigations are summarized. Parts of the regional and global sensitivity analysis have also been described in Schumacher et al. (2016a).

6.1 Sensitivity Study Set-Up

In this thesis, the sensitivity of storage changes in WGHM compartments is assessed 1) with respect to changes in one single model parameter, as well as 2) with respect to changes in all parameters reported in Tab. 2.1. First, N_e realizations of a single model parameter are generated, given the a priori probability density functions (PDF) in Tab. 2.1, while other model parameters are assumed to be constant (top left column of Fig. 6.1). An ensemble of free model runs, i.e. without considering any GRACE information, is generated for different parameter values and considering identical climate forcing and initial water states for 2002-2009 (center of Fig. 6.1). To evaluate the sensitivity of the modeled water storage compartments to the changes in one particular parameter, the so-called sensitivity index (SI, Hoffman and Gardner, 1983) is calculated. The SI describes the relative difference between the minimum ($S_{\min}$) and maximum ($S_{\max}$) water storage output of the

ensemble according to

$$I_{\mathrm{SI}}(i,j,k) = \frac{S_{\max}(j,k) - S_{\min}(j,k)}{S_{\max}(j,k)}, \tag{6.1}$$

with respect to parameter i, compartment j, and time step k. This is successively performed for all possible calibration parameters, i.e. for the 22 WGHM parameters listed in Tab. 2.1. SI is a simple measure, which is widely used due to its straight-forward interpretation. However, it strongly depends on a priori selections of the PDFs of model parameters, and it does not consider correlations between parameters.

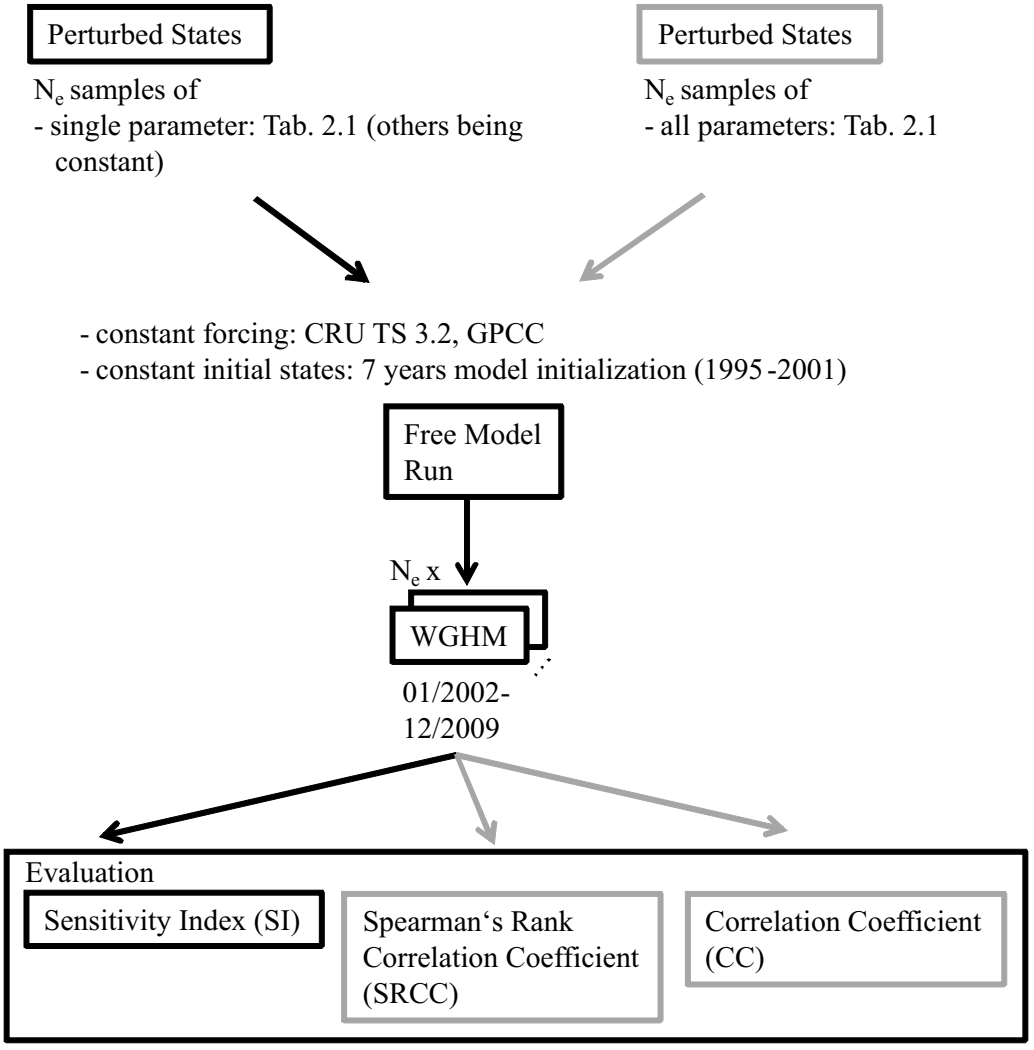

Figure 6.1: Overview of the set-up for the covariance and sensitivity analysis. The left column (black boxes and arrows) represents the first sensitivity set-up, in which only one parameter is modified for the N_e model runs and the others are considered as constant. The right column (gray boxes and arrows) represents the second set-up, in which ensembles of all parameters are generated simultaneously to perform the N_e model runs.

Secondly, realizations of all calibration parameters are therefore generated simultaneously (top right column of Fig. 6.1). Then, in total N_e model runs are performed with different values for each of the 22 parameters (center of Fig. 6.1). To evaluate the ensemble of

WGHM water state outputs, the Spearman's rank correlation coefficient (SRCC, Iman and Conover, 1979) is determined (middle entry of bottom row in Fig. 6.1). The SRCC allows to account for non-linear model equations by carrying out a rank transformation of the parameters and states, i.e. the input parameters and output states are sorted in ascending order according to their magnitudes, which defines their ranks. The Pearson's correlation coefficient between the ranks R_{P_i} of the i-th parameter and the ranks R_{S_j} of the water states in compartment j for each time step k is determined (Hamby, 1994) as follows

$$I_{\mathrm{SRCC}}(i,j,k) = \frac{\sum_{n=1}^{N_e}(R_{\mathrm{P}_{i,n}}(k) - \overline{R}_{\mathrm{P}_i}(k))(R_{\mathrm{S}_{j,n}}(k) - \overline{R}_{\mathrm{S}_j}(k))}{\sqrt{\sum_{n=1}^{N_e}(R_{\mathrm{P}_{i,n}}(k) - \overline{R}_{\mathrm{P}_i}(k))^2 \sum_{n=1}^{N_e}(R_{\mathrm{S}_{j,n}}(k) - \overline{R}_{\mathrm{S}_j}(k))^2}}. \tag{6.2}$$

Herein, $\overline{R}_{\mathrm{P}_i}$ and $\overline{R}_{\mathrm{S}_j}$ are the ensemble means (defined as unweighted means) of the ranks of parameter i and water states in compartment j, respectively.

In addition, the empirical covariance matrix of the calibration parameters and water state outputs, i.e.

$$\mathbf{C}^e(\mathbf{x}_k^-) = \begin{bmatrix} \mathbf{C}^e(\mathbf{w}_k^-, \mathbf{w}_k^-) & \mathbf{C}^e(\mathbf{w}_k^-, \mathbf{p}_k^-) \\ \mathbf{C}^e(\mathbf{p}_k^-, \mathbf{w}_k^-) & \mathbf{C}^e(\mathbf{p}_k^-, \mathbf{p}_k^-) \end{bmatrix},$$

is determined using Eqs. (4.4) and (3.30). The block $\mathbf{C}^e(\mathbf{w}_k^-, \mathbf{p}_k^-)$ contains the covariances between predicted model states $\mathbf{w}_k^-$ and parameters $\mathbf{p}_k^-$. The correlation coefficients (CC, hereafter named correlations unless stated otherwise) between predicted model states (averaged over the Mississippi River Basin) and parameters are calculated according to

$$I_{\mathrm{CC}}(i,j,k) = \frac{\sigma_{p_i w_j}(k)}{\sigma_{p_i}(k)\sigma_{w_j}(k)}, \tag{6.3}$$

where $\sigma_{p_i}(k)$ and $\sigma_{w_j}(k)$ denote the standard deviations of the parameter i and water states in compartment j at time step k, and $\sigma_{p_i w_j}(k)$ denotes their covariance (right entry of bottom row in Fig. 6.1). Since GRACE does not observe the parameters directly, the correlation values justify whether the observations will contribute in calibrating the model parameters.

6.2 Covariance and Sensitivity Analysis

6.2.1 Analysis of Parameter Distributions

It is obvious that the results of the covariance and sensitivity analysis directly depend on the chosen PDFs for the model parameters (see Tab. 2.1). Therefore, the ensemble mean of $N_e = 60$ free model runs is compared to the standard WGHM simulation to investigate whether the WGHM ensemble adequately represents the standard run. For 21 WGHM parameters, the simulated TWS time series differ less than 1 cm from each other. However, changes in the groundwater baseflow coefficient (IN=19 in Tab. 2.1) lead to a negative bias

in the ensemble mean of about 2 cm in groundwater and TWS compared to the standard model run (Fig. 6.2 a, b). To handle this, the PDF of the groundwater baseflow coefficient with 0.01 as mode and 0.006 and 0.1 as limits is modified to keep the difference between the ensemble mean and the standard model run less than 1 cm for total and individual water storage changes (Fig. 6.2 c, d). The mode of the new triangular distribution is empirically determined as 0.006 with 0.006 and 0.018 being the lower and upper limit values. In this case, the mode and the lower limit of the distribution coincide. This modified PDF results in lower baseflow and more groundwater storage, which might be a good choice given that WGHM tends to underestimate seasonality. However, a selection of other PDFs would have been possible.

In addition, the a priori limits of the triangular distribution of wetland depth are modified from 1 m and 20 m to 0.5 m and 5 m. Only in exceptional cases like the Amazon River Basin, a wetland depth of 20 m might occur. Therefore, the modified PDF fits better in most regions of the world, but this would not work for the Amazon River Basin.

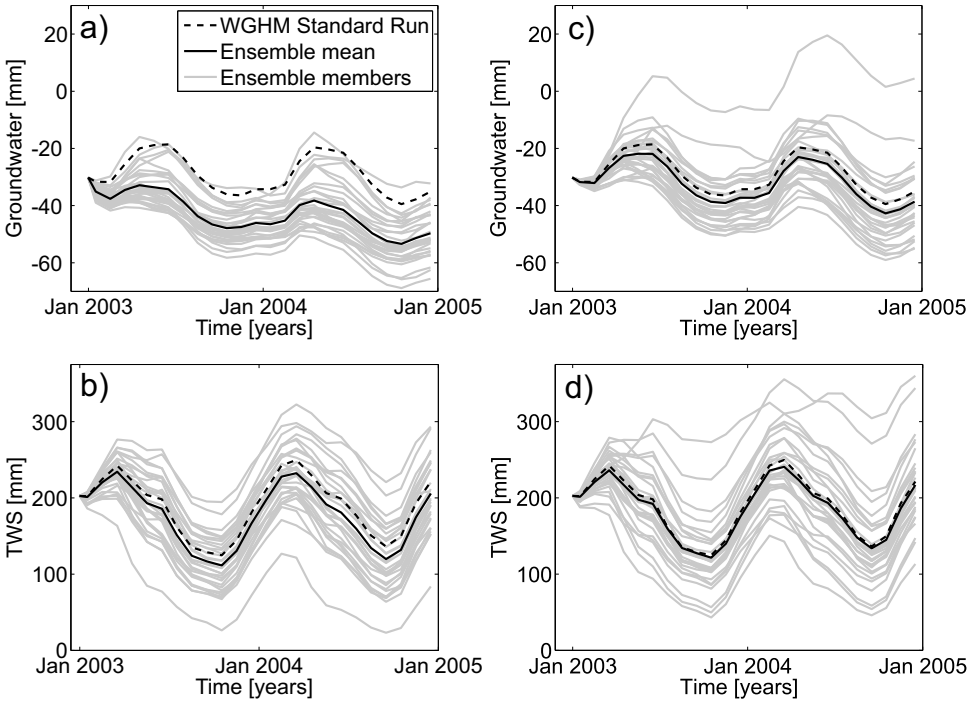

Figure 6.2: Time series of a) groundwater and b) TWS averaged over the entire Mississippi River Basin are shown, while using the a priori PDF in Tab. 2.1 to generate realizations of the groundwater outflow coefficients. Time series of c) groundwater, as well as d) TWS are shown, while using the modified PDF in section 6.2.1.

6.2.2 Sensitivity of Individual Water Storage Changes

To identify sensitive WGHM parameters, which predominantly influence the water storage compartments, the SI, SRCC, and CC between the water states averaged over the Mississippi River Basin and each model parameter for each month in 2002-2009 are determined (see Fig. 6.1). As an example in Fig. 6.3, the time evolution of the SI and CC

between the averaged snow and soil water storage and all model parameters are shown for 2008 (see also Schumacher et al., 2016a). Figure 6.3 a shows that considering SI as a measure of sensitivity, the snow storage is identified to be highly sensitive with respect to changes in the snow melt temperature (IN=13), the precipitation multiplier (IN=22) and the net radiation multiplier (IN=6) parameters. This can be understood by considering the physical meaning of these parameters: the precipitation multiplier represents a multiplication factor applied to the input forcing fields after interpolation to daily rainfall values; snow melts when the actual temperature exceeds the snow melt temperature; and the net radiation multiplier is used to scale solar radiation, which controls the potential evapotranspiration from land and water bodies. Note that between May and November no sensitivity is observed for the snow compartment, since there is usually no snow in the basin at this time.

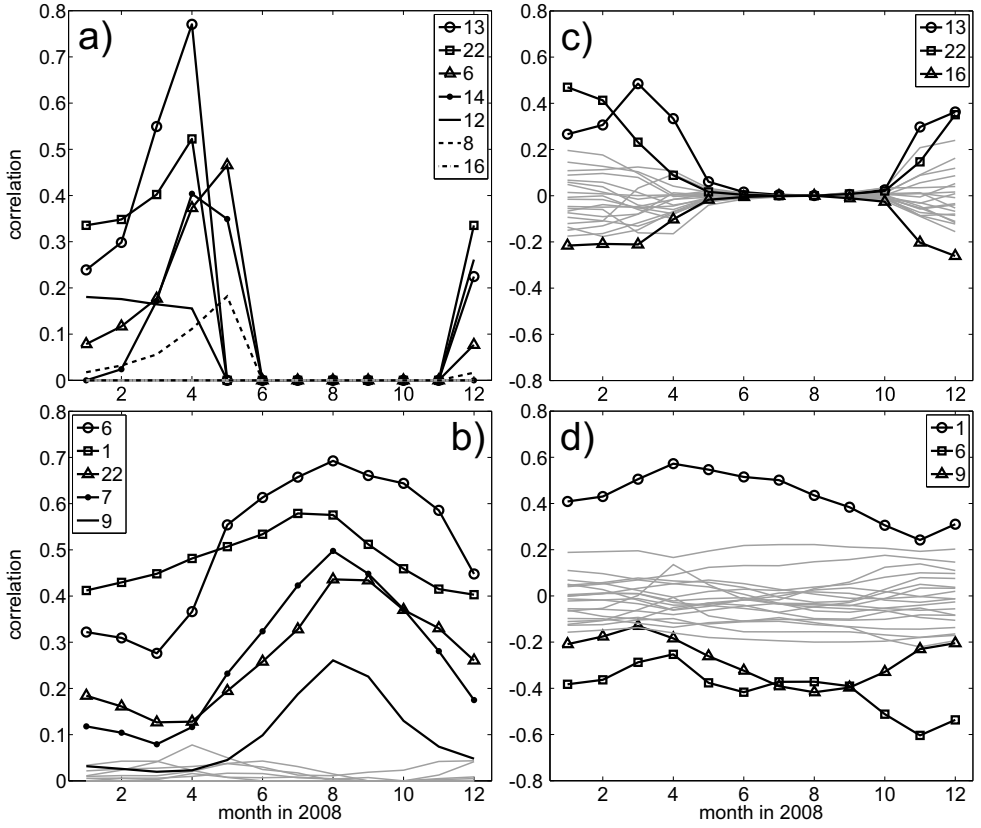

Figure 6.3: Time evolution of the SI between the 22 model parameters and the basin mean of the a) snow and b) soil compartment, and of the CC for the c) snow and d) soil compartment. The parameters with the highest correlations to the averaged compartment states are listed in the legend. The gray lines belong to the other parameters. See Tab. 2.1 for parameter names. This figure is taken from Schumacher et al. (2016a).

The CC confirms the high sensitivity of the snow storage change to the snow melt temperature and the precipitation multiplier (Fig. 6.3 c). However, high correlations are also found between the snow storage and the groundwater factor multiplier (IN=16). This parameter does not exhibit a direct physical relationship to the snow compartment. It appears this correlation is introduced through joint dependence of other perturbed parameters, and thus is only captured by CC and not by SI. The magnitude of the correlations is found to be different, e.g., the maximum correlation value for the snow melt temperature varies

between 0.5 in case of SI and 0.8 in case of CC and SRCC. However, the overall interpretation of all metrics leads to the same result: the snow melt temperature (IN=13) is the most important parameter with respect to the snow storage (see also Tab. 6.1).

For the soil compartment, the SI approach identifies high dependencies between soil water changes and the net radiation multiplier (IN=6), the root depth multiplier (IN=1), and the precipitation multiplier (IN=22) for the year 2008 (Fig. 6.3 b). The sensitive parameters for the soil water storage are confirmed by evaluating the CC and the SRCC (Fig. 6.3 d, here shown for CC).

Storage changes in the other individual compartments show lower variability of the SI, SRCC, and CC during the year 2008. Therefore, their time series are not shown here. The most sensitive parameters based on the analysis of SI, SRCC, and CC (temporal averaged during 2002-2009) are summarized in Tab. 6.1 as $\overline{I}_{\mathrm{SI}}$, $\overline{I}_{\mathrm{SRCC}}$, and $\overline{I}_{\mathrm{CC}}$.

Table 6.1: Most sensitive parameters are indicated for the Mississippi River Basin corresponding to the ten individual water storage compartments of WGHM. The identification numbers of the parameters (i=IN) can be found in Tab. 2.1. The overline denotes the temporal average. In case that only one parameter is provided, the index is zero for all other parameters.

	Three most sensitive parameters i		
Compartment j	$\overline{I}_{\mathrm{SI}}(i,j)$	$\overline{I}_{\mathrm{SRCC}}(i,j)$	$\overline{I}_{\mathrm{CC}}(i,j)$
canopy	10	10, 6, 7	10, 6, 7
snow	13, 22, 6	13, 22, 16	13, 22, 16
soil	6, 1, 22	1, 6, 3	1, 6, 9
local lake	3, 5, 6	3, 7, 6	3, 7, 6
local wetland	4, 5, 22	4, 5, 22	4, 22, 5
global lake	3	3, 7, 2	3, 7, 5
global wetland	4	4, 22, 9	4, 22, 5
reservoir	22, 6, 7	22, 6, 1	22, 6, 7
river	2, 6, 22	2, 6, 22	2, 6, 22
groundwater	6, 21, 17	21, 6, 7	19, 6, 21

Changes in the maximum canopy water height (IN=10), as well as the net radiation multiplier (IN=6), and the Priestley-Taylor coefficient for humid areas (IN=7) considerably influence the canopy water storage. All metrics suggest that the lake and wetland compartments are mostly sensitive to changes in the lake and wetland depth (IN=3 and IN=4), respectively. The precipitation multiplier (IN=22) and the surface water outflow coefficient (IN=5) also exhibit a high impact on the water storage in surface water bodies. Furthermore, the river storage is most sensitive to the river roughness coefficient (IN=2), the net radiation (IN=6) and the precipitation multiplier (IN=22), whereas the multipliers also show a high influence on reservoirs. For the groundwater compartment, all metrics indicate high influences of the net radiation (IN=6) and net groundwater abstraction multiplier (IN=21).

6.2.3 Sensitivity of TWSA

In Schumacher et al. (2016a), a sensitivity analysis for the 33 largest river basins world-wide was performed. For this, the SRCC between the calibration parameters and monthly GRACE TWSA averaged over the 33 basins were calculated. Different parameters were identified for these basins that exhibit the most dominant influence on simulated TWSA (Fig. 3, 4 in Schumacher et al., 2016a). Numerous river basins (e.g., the Mississippi River Basin) were found to react very sensitive to changes in the net radiation multiplier, the river roughness coefficient and the precipitation multiplier. Thus, it is concluded that these calibration parameters have an overall strong influence.

In Güntner et al. (2007) and Werth and Güntner (2010), a sensitivity study of WGHM parameters was reported. Their study used the WGHM version 2.1 f that is calibrated for 724 river basins. WGHM was forced by time series of the Climate Research Unit (CRU TS 2.0). In contrast, CRU TS 3.2 and precipitation data from GPCC are used in this thesis (see section 2.2.2.2). In Güntner et al. (2007) and Werth and Güntner (2010), $N_e = 2000$ ensemble members were generated from uniform, triangular and normal distributions for 36 calibration parameters (Kaspar, 2004), as well as for climate input fields comprising precipitation, the number of rain days, temperature and sunshine duration. The SRCC was used as a measure of sensitivity and was determined between the calibration parameters and the mean annual amplitude of TWS as a measure for sensitivity for the 22 largest river basins world-wide. To make the presented investigations comparable to their results, the global sensitivity analysis is repeated for the mean annual amplitude of TWS. However, the 33 largest river basins world-wide are considered and the climate input fields are not perturbed.

The parameters, that are found to be sensitive in this study, confirm some of the sensitive parameters that were listed in Güntner et al. (2007) as well as Werth and Güntner (2010) but not all of them. The reason for this is not clear but it might be explained due to differences in the study set-ups. As an example, for the Mississippi River Basin, the snow melt temperature (IN=13), the precipitation multiplier (IN=22), the groundwater baseflow coefficient (IN=19), the root depth multiplier (IN=1), the critical precipitation for groundwater recharge in semi-arid and arid regions (IN=18), and the maximum daily potential evapotranspiration (IN=9) are identified as most sensitive. Only the root depth multiplier and the snow melt temperature were also found to be sensitive in Güntner et al. (2007) as well as Werth and Güntner (2010). For the Murray-Darling River Basin, it was found that the net radiation multiplier (IN=6), the precipitation multiplier (IN=22), the root depth multiplier (IN=1), the Priestley-Taylor coefficient for arid areas (IN=8), the net abstraction surface water multiplier (IN=20), and the temperature gradient (IN=15) have a high impact on the simulation of TWSA. However, from these parameters only the Priestley-Taylor coefficient for arid areas was found to be sensitive in the study of Güntner et al. (2007) as well as Werth and Güntner (2010). In summary, the results of the global comparison only confirm some of those parameters with large model sensitivity in the world's largest river basins that were found in Güntner et al. (2007) as well as Werth and Güntner (2010). In their studies, parameters that govern radiation were identified as most sensitive, while in this study precipitation and net radiation multipliers are identified. Thus, a strong overall dependence of TWS on climate input is found in all investigated studies.

6.3 Discussion and Conclusions of the Sensitivity Investigations

The analysis of the SI, the SRCC, and the CC indicates that correlations between the model states and parameters exist and are mostly larger than 0.2, and thus enabling the parameter calibration by GRACE measurements. This supports hypothesis I of this thesis: The simulated water storage changes must be sensitive to changes in model parameters, and thus, enabling the calibration of parameters using GRACE TWSA via model state-parameter correlations. By performing the analysis for the Mississippi River Basin, it was demonstrated that the water compartments are sensitive to different model parameters.

However, it was noticed that some of the identified parameters do not have any direct physical relationship to the storage compartment (e.g., the groundwater factor multiplier for snow or the net radiation multiplier for groundwater). It is possible that spurious correlations exist when generating realizations of all calibration parameters simultaneously. This might be due to an over-parametrization, i.e. too many parameters are considered in the sensitivity analysis that influence each other. Some of these parameters might have a contrary effect on the model outputs, and therefore, their effects are compensated. This suggests investigations on the optimal selection and number of WGHM parameters to be calibrated within the EnKF update.

In addition, the time evolution of the parameter-state correlations indicates that the impact of GRACE TWSA in the calibration procedure changes over time. Therefore, calibrating the model parameters at least for one full year is required, since some parameters such as those related to the snow compartment cannot be updated during summer. Probably, this is the most obvious case, others may also exist.

The sensitivity analysis with respect to the SRCC could confirm some parameters, which were identified to be most sensitive in the previous studies of Güntner et al. (2007) as well as Werth and Güntner (2010). Based on the global sensitivity analysis of the 33 largest river basins world-wide (Schumacher et al., 2016a) and regarding the results in this thesis for the Mississippi and Murray-Darling River Basins, a basin-specific parameter calibration seems appropriate, since different sensitive parameters are found in different river basins.

Finally, it is worth mentioning that in traditional hydrological calibration 1000-2000 ensemble members are generated to perform a sensitivity analysis. Here, $N_e = 60$ samples were chosen to use more or less the same ensemble size that will be used in the application of the developed C/DA framework (chapters 7 and 8). It was shown that the calibration of WGHM parameters against GRACE TWSA seems possible. However, it should be stressed that a larger ensemble size will likely enable a more statistically stable estimation of the sensitive parameters.

7. Synthetic Experiment

A synthetic experiment was designed to test the developed calibration and data assimilation (C/DA) framework. The advantage of a simulation is that the true total and compartmental water states and fluxes are known by definition. This enables to compare the updated model states with true time series and to assess the performance of the C/DA framework under different configurations.

In section 7.1, the set-up of a twin experiment to assimilate synthetic GRACE total water storage anomalies (TWSA) into the WaterGAP Global Hydrology Model (WGHM) for the Mississippi River Basin is introduced. This basin was chosen, since it will also serve as a test region for the assimilation of observed GRACE TWSA in section 8.1. The impact of different configurations on the C/DA results is discussed in section 7.2. This includes considering a diagonal or full observation error covariance matrix in the ensemble Kalman filter (EnKF) update step to identify how the implementation of the correlated errors of GRACE TWSA affects the C/DA results (part 1 of objective 2, see section 7.2.1). Afterwards, the observations are aggregated to different spatial scales to investigate the effect of the spatial resolution of the observations on the C/DA results (part 2 of objective 2, see section 7.2.2). Furthermore, the square root analysis (SQRA) scheme and the singular evolutive interpolated Kalman (SEIK) filter technique are applied to update WGHM water states and parameters and these are compared to those of the EnKF. This is applied to understand the impact of alternative filtering methods on the C/DA results (objective 3, see section 7.2.3). A discussion on the calibration parameters is provided in section 7.2.4.

Parts of these assessments were presented in Schumacher et al. (2016b). As it is not trivial to choose an appropriate ensemble size for our study, here in addition, an analysis of the impact of the ensemble size on C/DA results is performed (section 7.2.5). Finally, the findings are summarized in section 7.3.

7.1 Twin Experiment Set-Up

An overview of the twin experiment set-up is given in Fig. 7.1. The experiment starts with the definition of "true" hydrological water states (left column in Fig. 7.1). For this, WGHM is driven by daily time series from the WFDEI meteorological data set (see section 2.2.2.2). The applied "true" model parameters are the calibrated values derived from the C/DA study by Eicker et al. (2014, see also section 8.1). WGHM is initialized over a period of nine years (1995-2003), followed by a three years model forward integration phase (2004-2006). Eventually, the monthly outputs of true total and compartmental water states serve as the basis to assess the C/DA results.

An imperfect representation of the truth in the analysis is realized by replacing the daily forcing with monthly time series of temperature, cloudiness and the number of wet days in months from CRU TS 3.2, as well as precipitation from GPCC (see section 2.2.2.2). In addition, the model parameters reported in Tab. 2.1 are used (second column in Fig. 7.1).

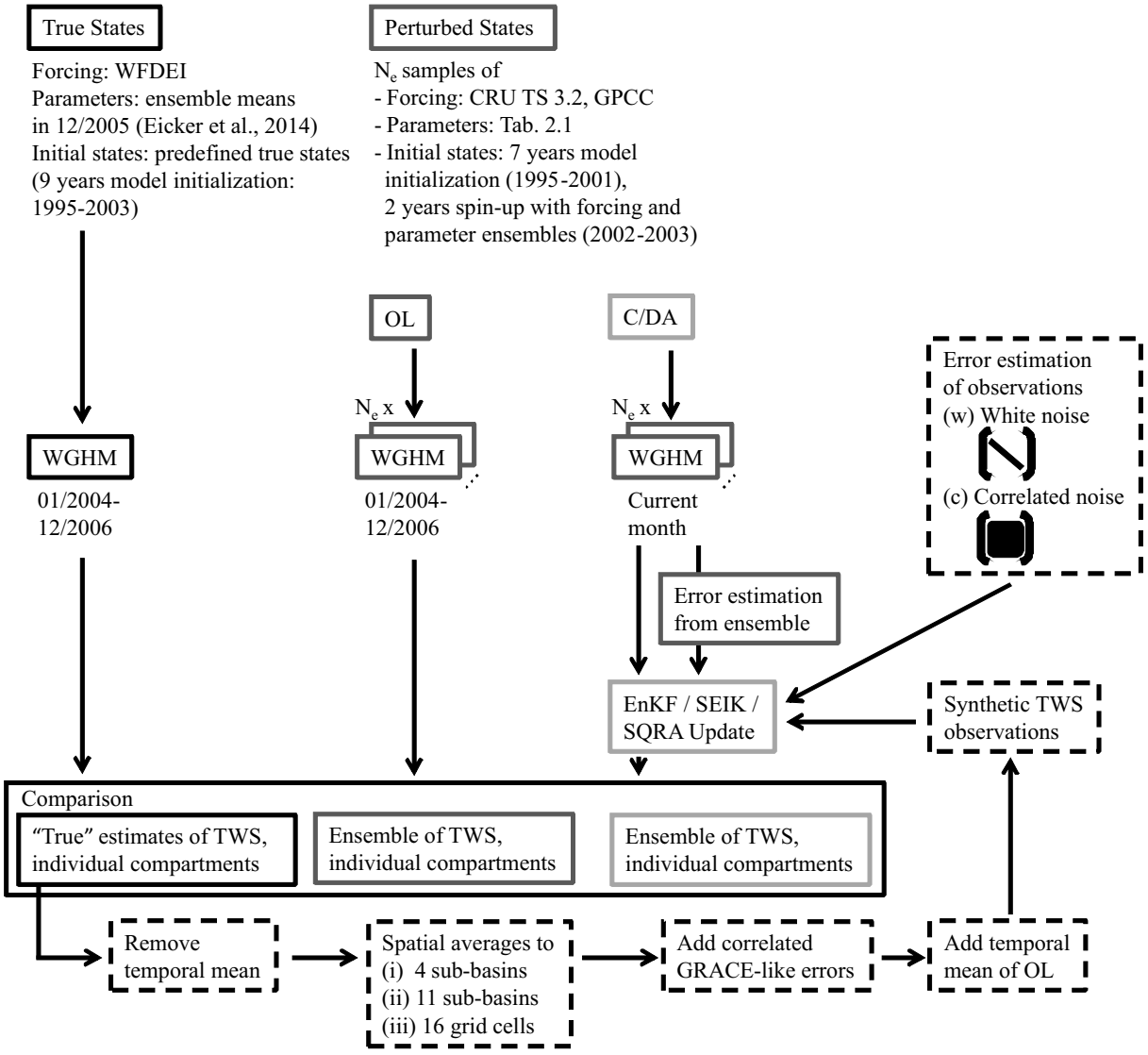

Figure 7.1: Overview of the twin experiment set-up, in which the true and perturbed model states are explained in the first row. Model prediction in open loop (OL) mode, i.e. without integrating GRACE data, and in C/DA mode for the N_e generated model ensemble members are shown in the second and third column. The generation of synthetic GRACE-like observations is described in the bottom row. OL and C/DA variants are compared to the true states. The performance of the C/DA variants is analyzed compared to the OL performance and compared to each other. This scheme is taken from Schumacher et al. (2016b).

An ensemble size of $N_e = 30$ samples is defined as a trade-off between computational costs, storage capacity and representative error statistics, and it is in accordance with previous GRACE data assimilation studies in hydrology, i.e. from five ensemble members in van Dijk et al. (2014) to 25 in Su et al. (2010), and 30 in Eicker et al. (2014). Consequently, errors of the model simulation are represented by an ensemble of 30 randomly perturbed precipitation and temperature input fields (see section 5.2.3 for details), and by an ensemble of calibration parameters according to the distributions provided in Tab. 2.1. For generating an ensemble of initial water states, the model initialization phase is shortened to seven years, and a spin-up phase of two years (2002-2003) is performed using the

parameter and climate input ensembles. The model outputs for the individual water compartments are introduced as initial values at the beginning of the C/DA phase. Open loop (OL) simulations are performed without integrating GRACE TWSA observations during 2004-2006 for each of the initial model ensemble members. The model prediction vector, Eqs. (4.6), (4.9) and (4.20), in this study is composed of monthly means of model derived water states in the ten individual water compartments for each of the 1262 grid cells in the Mississippi River Basin and the 20 WGHM calibration parameters. Please note that the lake depth and wetland depth (IN=3 and IN=4) are not included in this study.

The synthetic GRACE-like TWSA observations are generated in four steps (bottom row in Fig. 7.1): (1) 0.5°× 0.5° gridded monthly means of TWS outputs of the true model are reduced by their temporal mean over the C/DA period during 2004-2006. (2) These values are then spatially averaged to four (similar to Zaitchik et al., 2008) and eleven sub-basin means, as well as sixteen 5°× 5° grid cells (similar to Eicker et al., 2014), where the boundaries are taken from Fig. 7.2. (3) Spatially correlated errors of TWSA

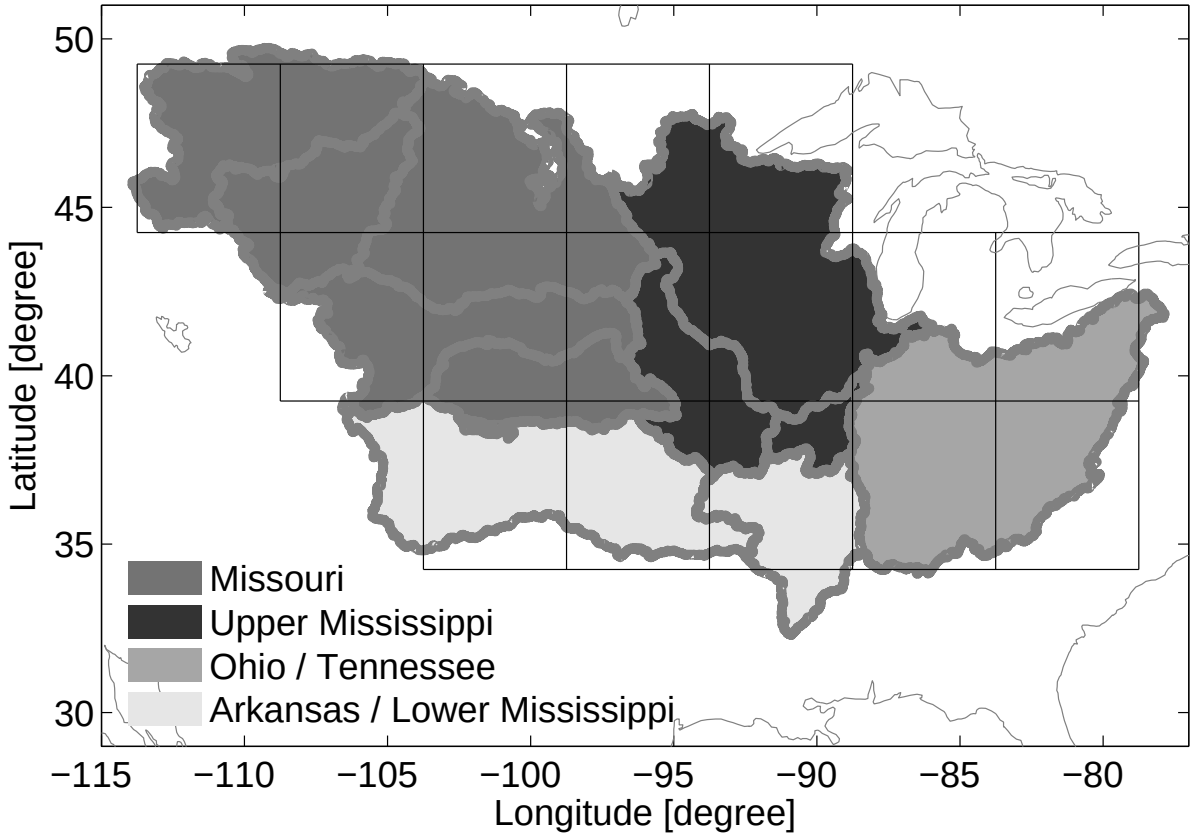

Figure 7.2: Sub-basins within the Mississippi River Basin. The four sub-basin definition is chosen similar to Zaitchik et al. (2008) and is shown with shaded areas. Eleven sub-basins are shown with the thick gray polygons. The grid definition is chosen similar to Eicker et al. (2014) and is shown using the thin black lines.

are generated by error propagation of the full ITG-Grace2010 error covariance matrix in 08/2003 (http://www.igg.uni-bonn.de/apmg/index.php?id=itg-grace2010). In this study, a time-constant observation error covariance matrix is assumed. The sub-basin/grid cell size influences the number of grid cells with error correlations, as well as the magnitude

of correlations, which increases when selecting smaller spatial averages. Correlations between GRACE TWSA errors of up to -0.5 exist when assimilating four sub-basin-averaged observations, up to 0.9 can be seen in case of eleven sub-basin averages, and the correlations exceed 0.9 in case of gridded observations being assimilated (for details see Fig. 6 in Schumacher et al., 2016b). The generated correlated errors are added to the TWSA time series derived in step 2. (4) Merging TWSA from the perturbed model states and the synthetic observations requires that the same temporal mean has been removed. Therefore, the temporal means of the OL simulations are added to the synthetic TWSA. As a result, corresponding to the number of sub-basins, the observation vector in Eqs. (4.6), (4.9), and (4.20) includes four, eleven or 16 sub-basin/grid cell averaged TWSA values.

To integrate the synthetic GRACE TWSA observations into WGHM, a range of configurations is defined (third and last column in Fig. 7.1, and Tab. 7.1): The EnKF, SQRA or SEIK is used (first column in Tab. 7.1), and the GRACE TWSA observations are averaged over four sub-basins, eleven sub-basins or 16 grid cells (second column in Tab. 7.1), while either white noise (indicated by "w") is assumed or correlated TWSA errors (indicated by "c") are considered in the update step (third and fourth column in Tab. 7.1). In addition, the ensemble size N_e is modified between 10, 30, 60 and 100 samples (last column in Tab. 7.1). In the ensemble filter updates, the principle of mass conservation, which is a basic assumption in hydrological models, is violated. Water mass might be added to or removed from the model. Additionally, an inflation factor of 10 % is introduced in all variants for representing errors in the model structure as well as in order to mitigate fast ensemble convergence. This factor is experimentally chosen as small as possible to avoid a strong influence on the model ensemble, and large enough to ensure that a contribution of the GRACE observations to the model update is guaranteed over the entire study period. An overview of the variants used in this study is given in Tab. 7.1. Here, in addition to the investigations reported in Schumacher et al. (2016b), the influence of alternative methods on the individual water storage compartments (not only on TWSA) is analyzed, and the influence of the ensemble size N_e on the performance of the C/DA framework is investigated.

Table 7.1: Calibration and data assimilation (C/DA) variants used in this study. For each case, an inflation factor of 10 % is used.

		Variant Name		
Filter Method	Discretization	White Noise	Correlated Noise	Ensemble Size
EnKF	4 sub-basins	4 w	4 c	30
EnKF	11 sub-basins	11 w	11 c	30
EnKF	16 grid cells	16 w	16 c	30
SQRA	11 sub-basins	Sq w	Sq c	30
SEIK	11 sub-basins	Se w	Se c	30
EnKF	11 sub-basins	-	10	10
EnKF	11 sub-basins	-	60	60
EnKF	11 sub-basins	-	100	100

7.2 Influence of C/DA Configurations on the Results

7.2.1 Choice of the Observation Error Model

In Fig. 7.3, the time series of total and compartmental water storage changes averaged over the Ohio/Tennessee Basin are shown: The C/DA results when introducing eleven sub-basin averaged TWSA observations and considering correlated TWSA errors in the EnKF update are shown as an example (variant 11 c in Tab. 7.1); the synthetic true and perturbed model states are included as references. The perturbed model states strongly overestimate the seasonal variations of TWSA compared to the synthetic truth (7.3 a). Assimilating synthetic GRACE observations into WGHM reduces the variability and

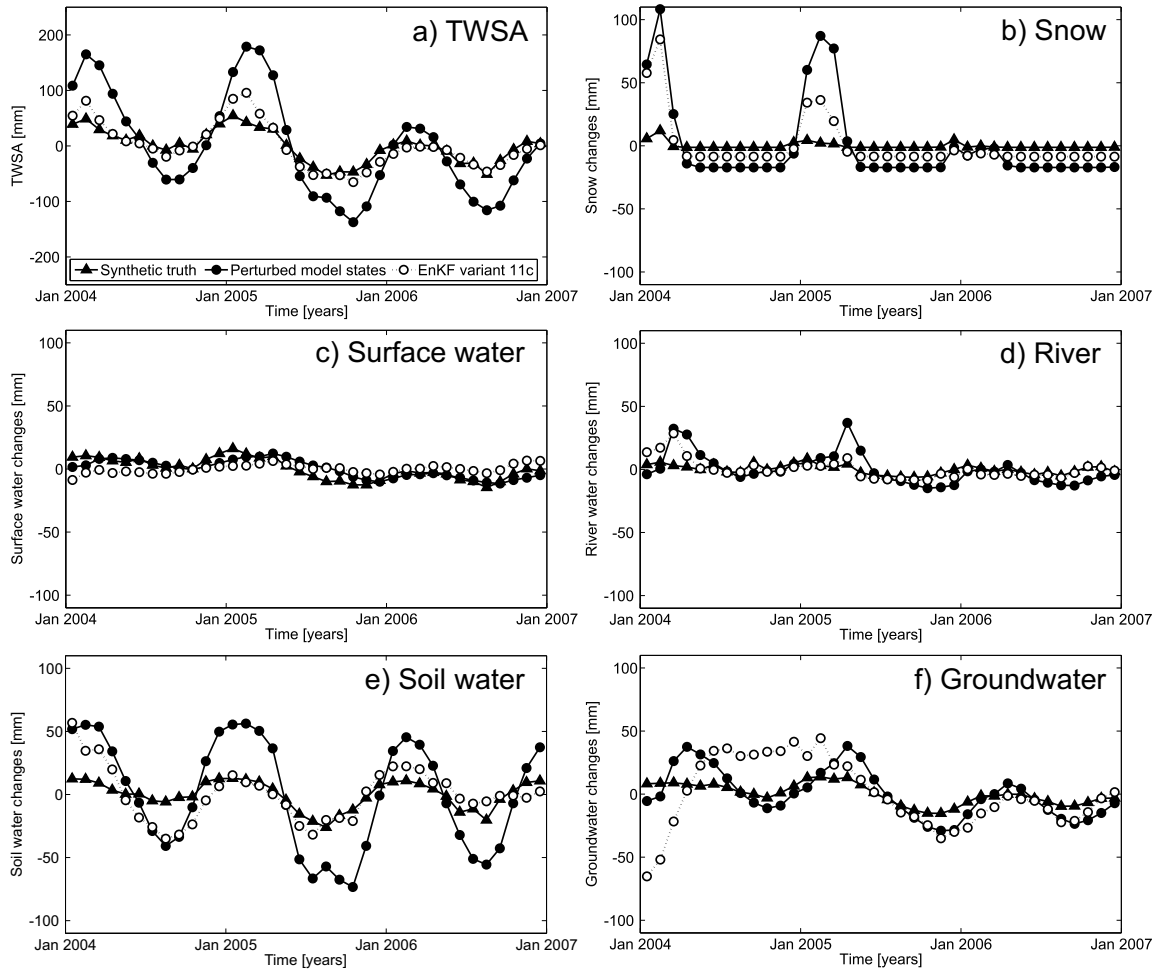

Figure 7.3: Time series of a) total and b)-f) compartmental water storage anomalies averaged over the Ohio/Tennessee Basin are shown for the synthetic truth, the perturbed model run, and the EnKF variant 11 c (see Tab. 7.1).

therefore results in a much better agreement with the synthetic truth. Water is stored in form of snow only during January to March, whereas the perturbed model run suggests a higher amount of snow water than the synthetic truth (7.3 b). Assimilation of GRACE TWSA also reduces the amount of snow. Thus, the EnKF variant 11 c (see Tab. 7.1) represents the true snow water states better than the OL variant. While changes in surface

water bodies and rivers only show small contributions to TWSA (7.3 c, d), soil water and groundwater variations influence TWSA throughout the whole year. During the first update steps in 2004, the soil water storage only obtains small update increments, such that the updated values are close to the perturbed model water states (7.3 e). In contrast, large update increments are associated with the groundwater storage (7.3 f). However, the updated values in 2004 do neither fit well to the true nor to the perturbed model water states. It seems that the vertical disaggregation of TWSA into soil water and groundwater does not fully succeed. The water mass that is reduced from the groundwater storage would actually be needed to be reduced from the soil water. In 2005 and 2006, the soil water changes follow the true values very closely after C/DA, while groundwater obtains smaller update increments resulting in values close to the perturbed model water states.

To evaluate the time series averaged over the Ohio/Tennessee Basin, several metrics are calculated with respect to the synthetic truth and are shown for the different EnKF variants in Fig. 7.4: (i) the root mean square error (RMSE), (ii) the correlation between residual curves after subtracting a linear trend, annual and semi-annual cycles, and (iii) the relative annual amplitudes, i.e. zero represents equal amplitudes. (i) Reductions in RMSE for the EnKF results compared to the OL results indicate that the temporal variability of total and individual water states is closer to the synthetic truth, while (ii) the residual correlation provides information on the phase (or timing) of the time series. A higher correlation indicates a better agreement with the true time series. In addition, (iii) the relative annual amplitude shows that the seasonal amplitude is closer to the truth. Metrics associated with TWSA are shown along the top row of Fig. 7.4, while the following rows correspond to the individual water compartments (snow, soil, surface water, river and groundwater). Each individual subplot contains the results from OL (shown in gray) and C/DA indicating the discretization level of assimilated TWSA observations. White bars correspond to the white observation noise introduced in the EnKF update step (additionally indicated by "w"), while black bars denote results from considering correlated observation errors (indicated by "c"). In all cases, TWSA are improved after C/DA compared to OL; yet assimilating GRACE data into WGHM does not guarantee an improved simulation of individual compartments. This is observed here for the surface water and groundwater in terms of RMSE, as well as for soil and groundwater in terms of residual correlations. Some metrics reveal that it is helpful to consider the full GRACE error covariance matrix, e.g., RMSE of surface water and river, as well as residual correlations of soil and groundwater in case of 16 c compared to 16 w (see Tab. 7.1), while others show a negative impact (e.g., RMSE of TWSA, soil and groundwater, as well as residual correlation of TWSA). These results indicate that the chosen observation error model has a considerable impact on the C/DA results in terms of TWSA as well as for several individual water storage compartments.

For all EnKF variants, water mass is removed from the model (see Fig. 8 in Schumacher et al., 2016b), which results in a more realistic representation of TWSA (see Fig. 7.3 a). Water is mainly removed from the soil water and from the river compartments, as well as from snow in winter months, while water is generally added to the groundwater compartment. The absolute mass change of TWSA, i.e. the sum of the absolute values of the update increments (difference between model update and prediction) over the entire study period, is about 300-400 mm depending on the EnKF variant. The largest parts of the mismatch between observed and simulated TWSA are associated with the soil water

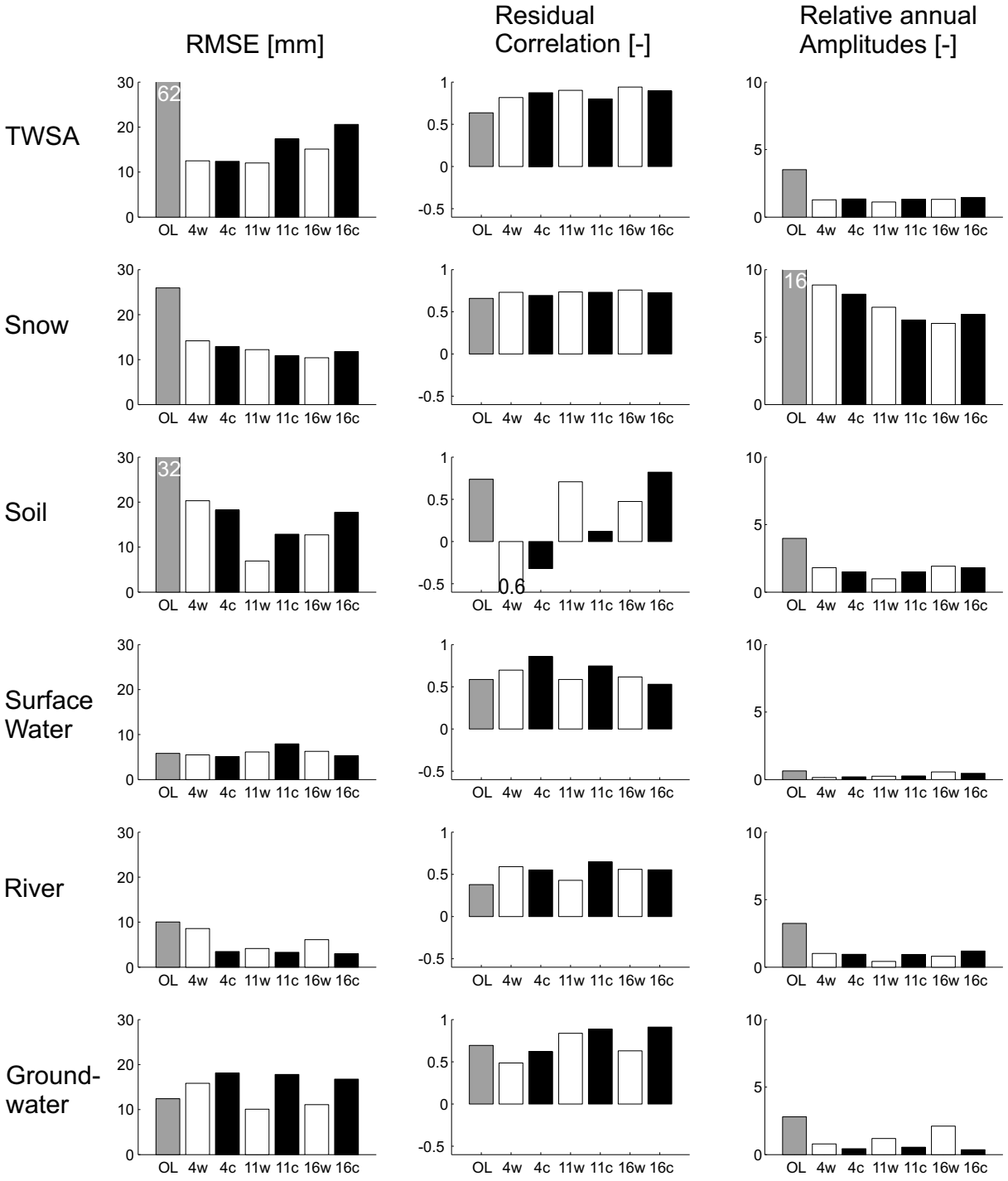

Figure 7.4: Root mean square error (RMSE), residual correlations, and relative annual amplitudes of total and individual water storage anomalies averaged over the Ohio/Tennessee Basin from open loop (OL) runs and C/DA results with respect to the truth. The time series and overall amplitudes are shown in Fig. 7.3 for EnKF variant 11 c (see Tab. 7.1). The definition of the names used in the x-axis can be found in Tab. 7.1. Some bars are truncated in order to fit the shown range. For these, the values are displayed at the top (or bottom) of the bar. This scheme is taken from Schumacher et al. (2016b).

and groundwater compartments with about 200-300 mm each. Since the EnKF update does not conserve water mass, the absolute mass changes indicate the degree of violating the principle of mass conservation.

The analysis is also performed for the spatial averages over the entire Mississippi River Basin and the remaining sub-basins (Fig. 9 in Schumacher et al., 2016b). The interpretation of the other regions confirms the findings for the Ohio/Tennessee Basin. In addition, higher impacts of the observation error model were found in those sub-basins with high EnKF update increments, i.e. in sub-basins with large discrepancies between modeled and observed TWSA or in sub-basins with accurate TWSA observations. Furthermore, it was demonstrated that sub-basins with north-south expansion are predominantly affected by the chosen observation error model in the EnKF.

7.2.2 Choice of the Spatial Resolution of Observations

Similar results were found for both EnKF variants 11 w and 16 w, which means that the change of the spatial resolution of the observations from eleven sub-basin averages to sixteen grid cell averages has only a small influence on the C/DA results (see Fig. 7.4). A similar conclusion holds for the cases 11 c and 16 c. Only the residual correlation between true soil changes and simulated soil changes after C/DA is considerably reduced to 0.1 in case 11 c. In contrast, switching from a diagonal to a full observation error covariance matrix in the EnKF update step has a clear impact on the C/DA results (compare e.g., 11 w and c, as well as 16 w and 16 c in Fig. 7.4). However, when assimilating synthetic TWSA aggregated to four sub-basins (case 4 w and 4 c), the effect of changing the observation error model in the EnKF update is found to be small.

7.2.3 Choice of Alternative Filtering Methods

The SQRA and SEIK algorithms are also applied to assimilate the eleven sub-basin observations into WGHM. As for the application of the classical EnKF, some metrics indicate that it is helpful to consider correlated errors of GRACE, while it seems to have an adverse impact on others (Fig. 7.5). Implementing the SQRA and SEIK has only a small influence on the RMSE of TWSA with respect to the classical EnKF when considering white noise in the update step (less than 2 mm). In case of SQRA, the residual correlation is even degraded by 0.1. After considering correlated errors in the SEIK filter, the RMSE is improved by 6 mm and the residual correlation by 0.1. The application of SQRA and SEIK only marginally influences the snow compartment, as well as the RMSE and relative annual amplitude of surface water changes compared to the classical EnKF. In contrast, both methods help to considerably increase the residual correlation of soil water from 0.12 to 0.75 and 0.70 in case of using correlated observation errors in the update step (compare 11 c with Sq c and Se c in Fig. 7.5). In case of SQRA, even an improvement compared to the OL was achieved. The residual correlation of groundwater is however reduced. This confirms the assumption that dependencies between soil water and groundwater time series exist, which complicate the vertical disaggregation of TWSA into these two compartments.

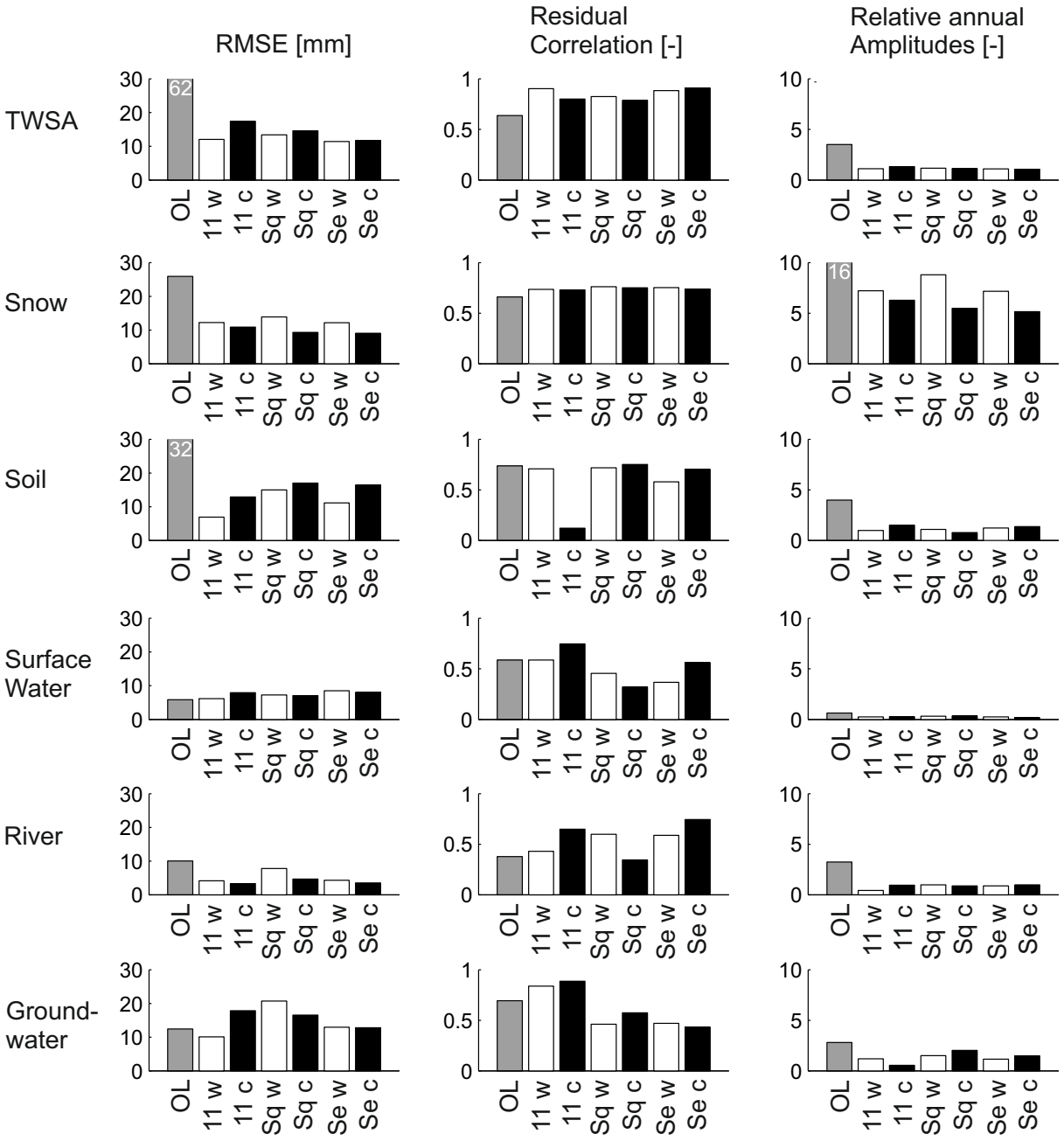

Figure 7.5: Root mean square error (RMSE), residual correlations, and relative annual amplitudes of total and individual water storage changes averaged over the Ohio/Tennessee Basin from open loop (OL) runs and C/DA results based on EnKF (11 w and 11 c), SQRA (Sq w and Sq c) or SEIK (Se w and Se c) with respect to the truth. The time series and overall amplitudes are shown in Fig. 7.3 for EnKF variant 11 c. The definition of the names used in the x-axis can be found in Tab. 7.1. Some bars are truncated in order to fit the shown range. For these, the values are displayed at the top of the bar.

7.2.4 Calibration Parameters

The impact of GRACE TWSA on model calibration parameters is influenced by the number of observations that are assimilated into WGHM. Those parameters are defined as sensitive to TWSA assimilation, whose ensemble spread is reduced to less than 25 % of its initial spread after 18 months (see Tab. 6 in Schumacher et al., 2016b). Using synthetic observations averaged over four sub-basins does not affect the parameter estimation. Enhancing the spatial resolution to smaller sub-basins indicates that 15 % of the calibration parameters are sensitive to GRACE C/DA (case 11 w). Using gridded GRACE TWSA increases the impact from 15 % to 55 % in case of 16 c. The number of sensitive parameters is found to be higher when introducing the correlated observation error model in the update step. The number of sensitive parameters also grows up to 40 % when applying the SQRA and SEIK algorithms. However, it cannot be stated with certainty that individual parameter values improve after C/DA. The reason for this is equifinality, i.e. different parameter sets may result in a similar optimal simulation of WGHM water states.

7.2.5 Choice of the Ensemble Size

To investigate how the ensemble size affects the C/DA performance within the classical EnKF, eleven sub-basin average observations are also assimilated into WGHM using 10, 60 and 100 model ensemble members, while considering the full observation error covariance matrix. The RMSE, the residual correlation, and the relative annual amplitudes with respect to the synthetic truth are shown in Fig. 7.6. Generating only ten ensemble members does not improve the simulation of TWSA in terms of RMSE and in terms of residual correlation with respect to the run with 30 samples. This was expected, since enlarging the ensemble size likely leads to a more stable computation of the model error covariance matrix, and thus, yields a more realistic weighting of model and observations in the EnKF update. It is reasonable that an ensemble of ten samples is too small to well represent the model error statistics. When performing C/DA with 60 or 100 samples, the representation of TWSA in terms of RMSE is improved by 3 mm and 4 mm compared to the run with 30 samples and in terms of residual correlation by 0.1 in both cases. The difference between the metrics when using 60 or 100 samples is marginal, i.e. less than 1 mm in terms of RMSE and there is no difference in residual correlation. The representation of snow has not become better with respect to OL when considering only 10 samples in the assimilation. However, improvements are achieved for the other compartments, except for the residual correlation of soil water and for the relative annual amplitude of the river. The usage of 60 or 100 samples improves the RMSE of all individual compartments compared to the C/DA run with 30 samples. The residual correlation is better for snow, soil water and river but not for surface water and groundwater. A comparison of the metrics with respect to TWSA for 100 samples and for the application of the SEIK filter (case Se c in Fig. 7.5) reveals less than 1 mm difference in terms of RMSE, and less than 0.02 in terms of residual correlation. The performance metrics of the individual compartments exhibit also similar values.

In the following, the computational issues related to the choice of the ensemble size are discussed. The computations have been performed on the cluster of the APMG/TG groups

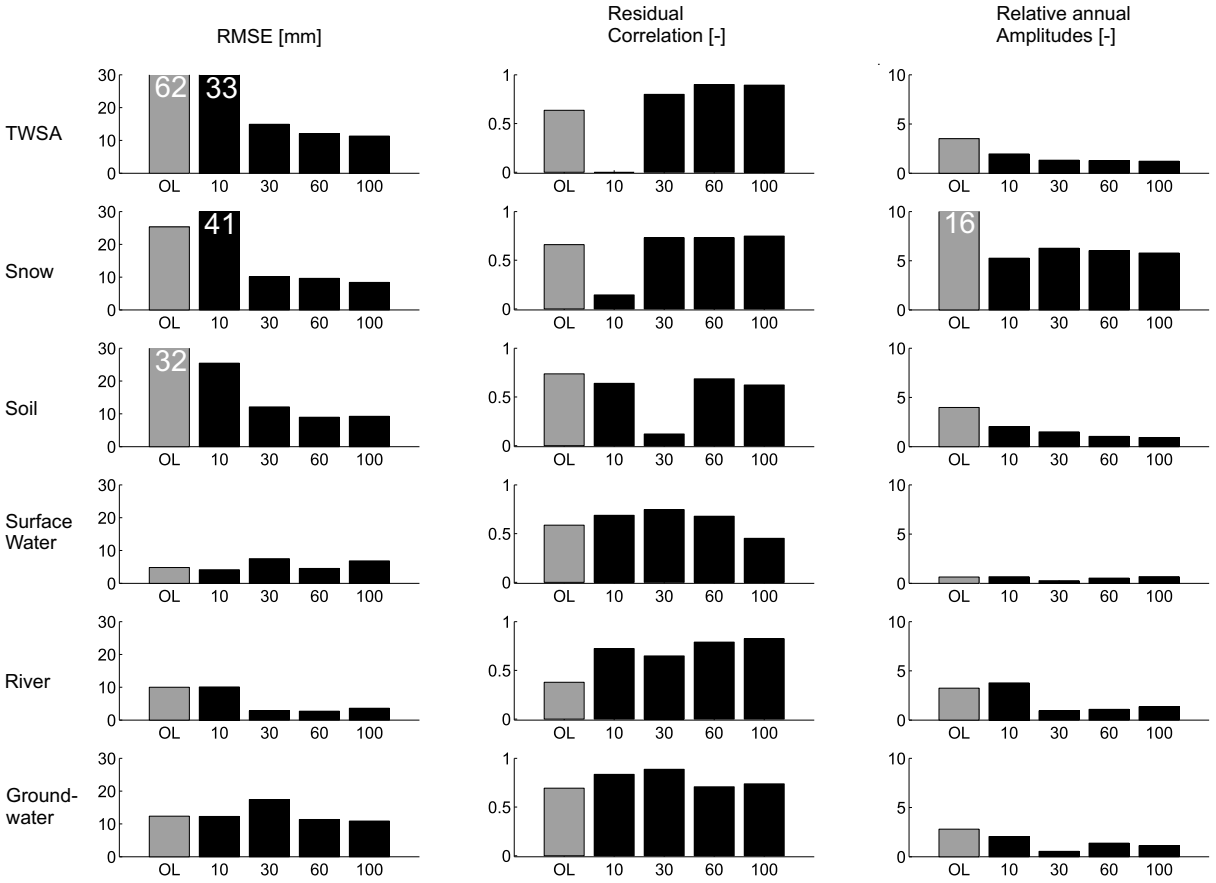

Figure 7.6: Root mean square error (RMSE), residual correlations, and relative annual amplitudes of total and individual water storage changes averaged over the Ohio/Tennessee Basin from open loop (OL) runs and C/DA after applying the classical EnKF, and varying the ensemble size (numbers in the x-axis; case 30 denotes the EnKF variant 11 c in Tab. 7.1). The time series and overall amplitudes are shown in Fig. 7.3. Some bars are truncated in order to fit the shown range. For these, the values are displayed at the top of the bar.

at the Institute of Geodesy and Geoinformation, University of Bonn. This cluster has 32 cores, from which 30 cores were used to run the 30 model samples in parallel. For merging GRACE and WGHM over three years for the Mississippi River Basin, the computation time was about 7 hours (Fig. 7.7). Decreasing the number of observations to 10 shortened the computation time to three hours, which corresponds to approximately 43 %. Increasing the ensemble size to 60, lengthened the computation time by 71 % to 11 hours and by 143 % in case of 100 ensemble members. The required storage capacity also grows with the number of ensemble members, e.g. from 53 GB in case of 30 samples to 171 GB in case of 100 samples. For this purpose, the source code needs to be optimized for future applications.

7.3 Discussion and Conclusions of the Synthetic Study

The results of the synthetic experiment show that C/DA of GRACE TWSA into WGHM improves the representation of simulated TWSA (hypothesis II a of this thesis). How-

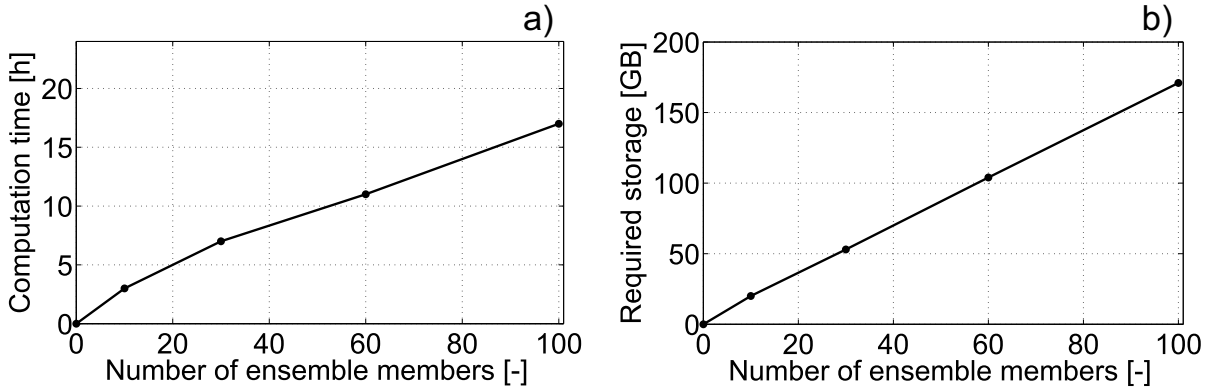

Figure 7.7: a) The computational time (in hours) and b) the required storage (in GB) depending on the ensemble size are shown for the assimilation of synthetic GRACE data into WGHM for the Mississippi River Basin over three years.

ever, it does not guarantee improvements for individual water storage compartments (hypothesis II b of this thesis). Insufficiently resolved or numerically introduced correlations between the individual states, as reflected in the error covariance matrix of the model (having a rank-defect and showing large condition numbers), might result in a deterioration of individual water compartment estimates. Investigations on estimating data-adaptive empirical and analytical covariance functions might help to tackle this challenge.

The findings suggest that error correlations of GRACE should not be neglected, since they affect the C/DA results considerably (hypothesis IV of this thesis). However, no general rule can be formulated that states whether applying correlated GRACE errors in the update step improves the agreement with the synthetic truth. Yet, the description of the GRACE errors is more realistic. For GRACE assimilation, the vertical and horizontal disaggregation is challenging, because the spatial scales of GRACE TWSA error correlations of a few hundred km are similar to the scales of physical correlations of land surface and groundwater variables in hydrological models. In addition, the assumption of spatially uncorrelated observation errors results in higher weights for the GRACE data in the update step. Therefore, the model update is usually pulled closer towards GRACE TWSA than with the correlated noise model; yet this does not always mean that the metrics of total and individual water states improve. It is thus difficult to directly compare experiments with and without (or with partly) implementing error correlations. This is made worse by the fact that in limited-size ensemble approaches the model covariance matrix has a rank-defect. A full-rank observation error covariance matrix enables a numerically stable solution of the EnKF update equation. Therefore, assuming spatially uncorrelated errors, i.e. a diagonal covariance matrix, has a regularization effect.

In summary, it has been demonstrated that investigations on the observation error model (which has not been done so far) are at least as important as studies on the choice of observation discretization (which has been tackled in previous GRACE data assimilation studies). The investigations also indicate that the observation error model in the filter update step affects the C/DA results on the three selected spatial scales, i.e. when assimilating GRACE TWSA as four or eleven sub-basin averages and 16 grid cells. The effect of changing the observation error model is found to be large for TWSA and several individual storage compartments, when assimilating TWSA with a fine spatial discretization.

Regarding e.g., groundwater, the effect of changing the observation error model is large for all selected spatial scales.

Promising results were found after applying alternative ensemble filter methods. Particularly, applying the SEIK filter and considering correlated GRACE errors improved the RMSE and the residual correlation of TWSA by 6 mm and 0.10 with respect to the classical EnKF. A likely explanation for this is that the EnKF relies on an observation ensemble to update each model sample individually, which introduces sampling errors. The SEIK and SQRA filter avoid this by updating the model ensemble mean. In addition, the SEIK filter uses the minimum second order exact sampling method to update the model ensemble perturbations, and thus ensures that no additional sampling errors are introduced.

The number of parameters that are sensitive to assimilation of GRACE TWSA into WGHM increases particularly when considering TWSA averaged over small sub-basins. However, the uncertainties of several parameters do not decrease during the ensemble updates. The choice of a smaller sub-set of the 22 parameters might enlarge the impact of GRACE TWSA assimilation on model parameters.

Two strategies were identified that almost always improve the performance of the C/DA procedure. First, increasing the number of ensemble members from 10 to 30, and then to 60 and 100 improves the C/DA results of TWSA and the majority of the individual compartments. However, using 100 samples instead of 60 has marginal influence on the metrics. Second, implementing the SEIK filter with 30 ensemble members shows a similar performance compared to the application of the classical EnKF with 100 ensemble members. Therefore, alternative methods can help to keep the ensemble size small which is necessary to minimize computational costs and required storage.

8. Case Studies

To test the calibration and data assimilation (C/DA) framework with observed data, GRACE total water storage anomalies (TWSA) are assimilated into the WaterGAP Global Hydrology Model (WGHM) for the Mississippi River Basin (USA) for one year (section 8.1). Afterwards, the model is run for the following three years with the parameter values calibrated from the previous one year assimilation. Subsequently, the simulated TWSA states are compared to GRACE measurements. Furthermore, independent in-situ and remotely-sensed measurements are considered for validating individual water compartments and fluxes. In line with objective 4 of this thesis, the assimilation is then performed for the Murray-Darling River Basin (Australia) to investigate whether the framework is transferable to other regions that face different climatic and anthropogenic conditions (section 8.2). Seven years of observed GRACE data (2003-2009) are assimilated into WGHM while keeping the ensemble Kalman filter (EnKF) settings that were found optimal for the Mississippi River Basin. Here, independent observations of groundwater wells are used to validate the C/DA results.

8.1 Test Region: Mississippi River Basin

The C/DA framework was established in chapter 5 and is tested here to assess whether the integration of GRACE TWSA into WGHM enables a more realistic representation of the water cycle, including total and individual water compartments and fluxes, related to hypotheses II a and II b of this thesis. The C/DA of 5°×5° gridded GRACE TWSA is examined and compared to the C/DA of basin averages to analyze the potential of highly resolved observations for an improvement of the fit with GRACE TWSA (hypothesis III of this thesis). Therefore, in the first part of the study, basin averaged and gridded monthly GRACE TWSA are assimilated into WGHM for one year, while simultaneously calibrating 22 model parameters, i.e. following the strategy reported in Eicker et al. (2014). Afterwards, the model is run for the following three years with the standard model parameters and the calibrated values. The set-up of this experiment is described in section 8.1.1. Updated TWSA states of both C/DA versions are then compared to TWSA measured by GRACE (section 8.1.2.1). In the second part of the study, the performance of the C/DA process is assessed for the individual water compartments and fluxes including soil moisture, groundwater, river discharge, and surface water extent (section 8.1.2.2).

The Mississippi River Basin is located in the eastern part of the USA with a mostly humid climate. It is the third largest basin world-wide with an area of about 3 million km^2, draining 41 % of the USA and providing a source of water to more than 18 million people. Water is intensively used for agriculture, especially in the High Plains Aquifer, covering parts of the Arkansas and Missouri Basins in the south-west of the basin, where groundwater is abstracted for irrigation purposes resulting in groundwater depletion (e.g., Rodell et al., 2007, Strassberg et al., 2009, Döll et al., 2012, 2014). Several data sets on water storage variability over the entire basin exist, including in-situ observations of, e.g., snow, soil moisture and groundwater, as well as remotely-sensed observations of, e.g., soil

moisture, surface water storage changes and inundation that are used for validation. In addition, several studies on the water cycle within the basin have been published, which can also be used for comparison. The main findings of this study are summarized in section 8.1.3.

8.1.1 Set-up of Assimilating Observed TWSA

An overview of the C/DA set-up for assimilating observed GRACE data for the Mississippi River Basin can be found in Fig. 8.1. First, a WGHM ensemble of $N_e = 30$ members is defined as described in section 5.2.3 by generating sets of forcing input data, model parameters, and initial water states (top left box in Fig. 8.1). It should be stressed that modified PDFs for the parameters wetland depth (IN=4) and groundwater baseflow coefficient (IN=19) are implemented as described in section 6.2.1. Then, the model is run in open loop (OL) mode for each of the ensemble members for 2005-2008 (first column), i.e. without integrating GRACE TWSA. This run will be denoted by WGHM_MRB-1 in the following. MRB serves as an abbreviation for Mississippi River Basin and the last digit is a consecutive number (Tab. 8.1). The OL is followed by model runs in C/DA mode (second column), i.e. GRACE TWSA are used to correct model states and parameters for each month in 2005. The C/DA variant that assimilates basin averaged observations into WGHM is called EnKF_MRB-1, and the variant derived by assimilating gridded TWSA is referred to as EnKF_MRB-2 (Tab. 8.1). Consistent with Eicker et al. (2014), perturbation vectors are not generated for the observations (see Eq. 4.6). This has to be kept in mind when interpreting the uncertainty estimations of the C/DA results.

Table 8.1: Overview of model simulations and assimilation runs that are analyzed in this study.

Run	Description	Reference
WGHM_MRB-1	WaterGAP 2.2 using monthly climate input (CRU TS 3.2, GPCC v6 precipitation), calibration against mean annual river discharge, consideration of human water use	Müller Schmied et al. (2014)
EnKF_MRB-1	Similar to WGHM_MRB-1 but assimilated by monthly GRACE TWSA averaged over the MRB from 01/2005 to 12/2005 (C/DA) and free (i.e. without assimilation) model run in 2006-2008 (validation)	Eicker et al. (2014)
EnKF_MRB-2	Similar to WGHM_MRB-1 but with assimilation of monthly 5°×5° gridded GRACE TWSA over 01/2005 to 12/2005 (C/DA) and free model run in 2006-2008 (validation)	Eicker et al. (2014)

In this study, monthly GRACE TWSA estimated from ITG-Grace2010 spherical harmonic coefficients (http://www.igg.uni-bonn.de/apmg/index.php?id=itg-grace2010) are used. These coefficients are provided with full variance-covariance information for the time span 08/2002-08/2009. For the ITG-Grace2010 solutions, a 500 km Gaussian filter is applied to smooth its spherical harmonic coefficients of up to degree/order 60, and con-

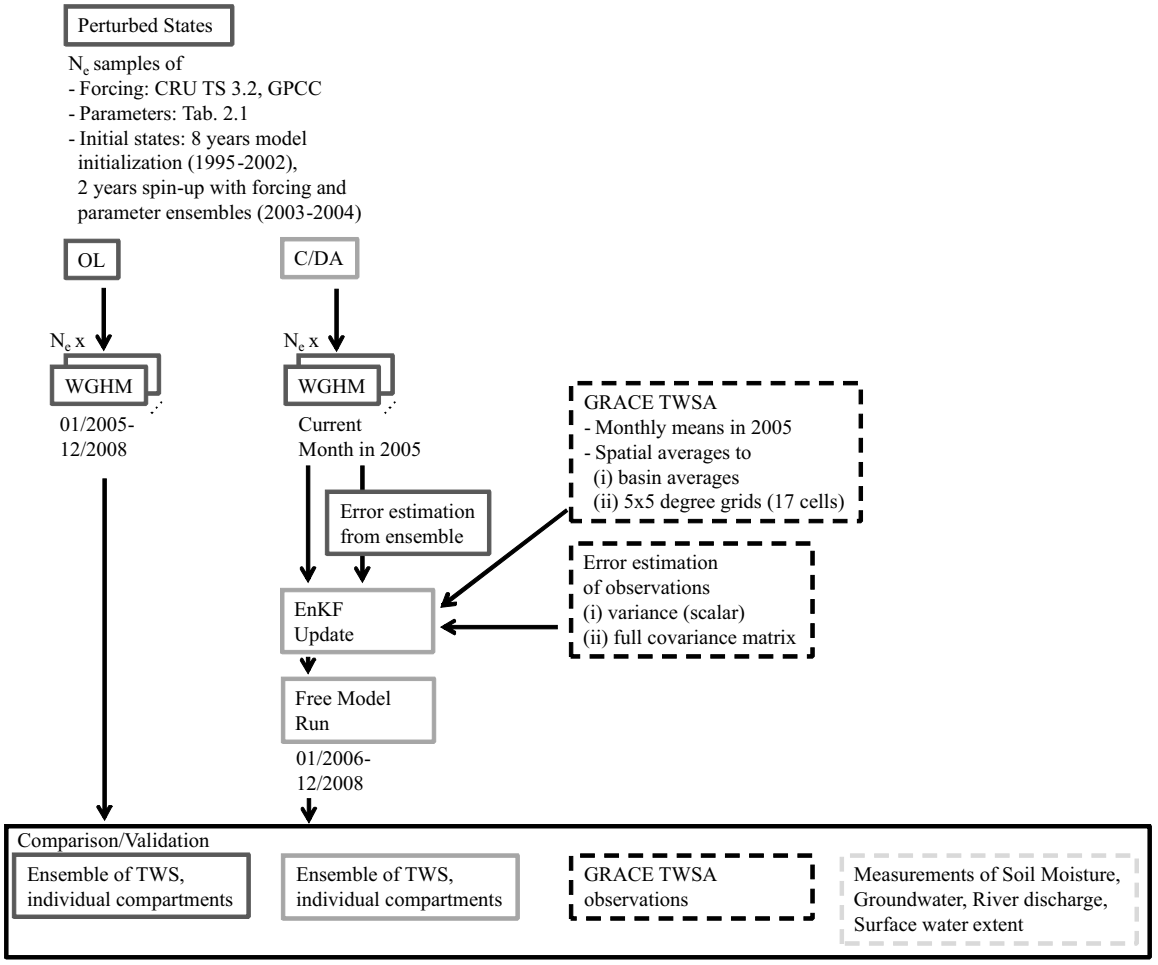

Figure 8.1: An overview of the set-up for assimilating observed TWSA into WGHM for the Mississippi River Basin. After running the model in open loop (OL) mode for each of the $N_e = 30$ ensemble members (left column), GRACE TWSA are assimilated into WGHM in 2005 as either basin averages or gridded TWSA (second column). Then, free model runs, i.e. without assimilating GRACE TWSA, are performed for 2006-2008. The C/DA updated total and compartmental water states are validated against independent measurements (last row).

verted to gridded values of TWSA (on the $0.5°\times0.5°$ grid used by WGHM). To account for the signal attenuation effect of the filtering process, an average re-scaling factor of 1.1 is determined for the Mississippi River Basin from analyzing basin-averaged filtered vs. unfiltered monthly WGHM model outputs in 2005. This factor is applied to the gridded TWSA values, and the GRACE observations are then aggregated to spatial averages over the entire Mississippi River Basin and over $5°\times5°$ grid cells (see black crosses in Fig. 8.2). Formal variance-covariance error propagation is carried out to obtain the full observation error covariance matrices for the gridded TWSA values from the full error covariance matrices of the potential coefficients (last column in Fig. 8.1). The methodology is described in section 2.1 for filtering and scaling (section 2.1.2) as well as for the error propagation (section 2.1.4).

The GRACE TWSA are assimilated into WGHM and model parameters are simultaneously calibrated in 2005. The principle of mass conservation is not maintained, since water is introduced to or removed from WGHM. Then, a three year model forward integration

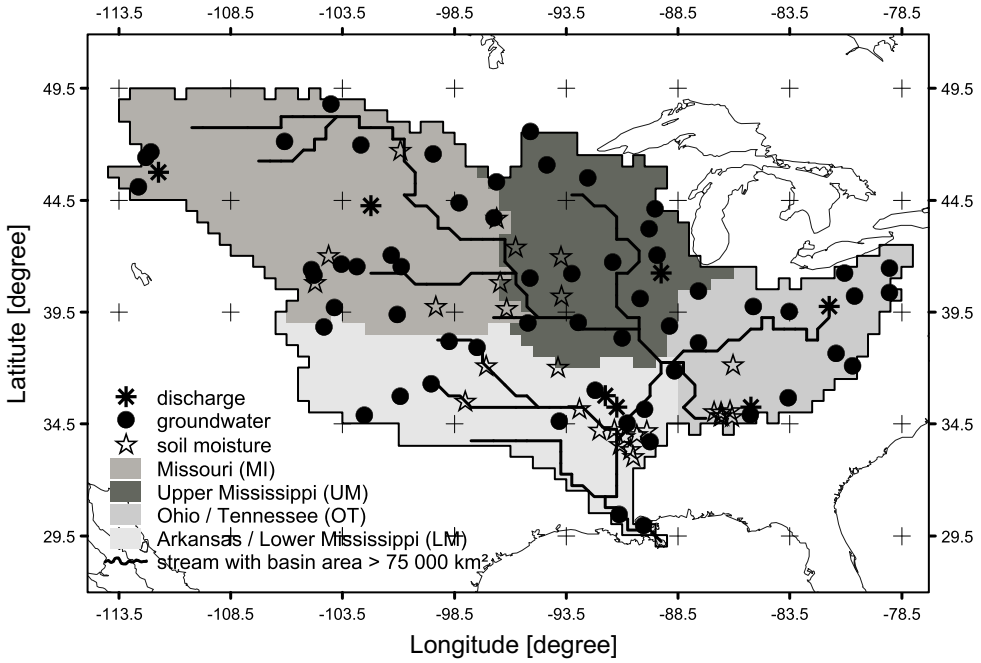

Figure 8.2: Overview of the Mississippi River Basin (black polygon), its four major sub-basins (gray shaded areas) and the 5°×5° grid (black crosses), which are used for averaging GRACE observations. The distribution of the in-situ measurement stations is also shown.

with updated initial states and calibrated parameters is performed for 2006-2008, i.e. the free model run in the middle column in Fig. 8.1. Finally, OL and C/DA variants are compared to GRACE TWSA and validated against measurements of individual water compartments and fluxes (last row; see Tab. 8.2). For the first time, four independent data sets are used for validating individual water states and fluxes after GRACE data assimilation. These are USGS in-situ groundwater table observations (as in Zaitchik et al., 2008), SCAN in-situ soil moisture data (as in Houborg et al., 2012), and river discharge from the Global Runoff Data Centre (GRDC) (as in Li et al., 2012). New is the evaluation of assimilation results with the Global Inundation Extent from Multi-Satellite (GIEMS) data that was created for comparisons with GRACE and WGHM simulations in Papa et al. (2008).

Table 8.2: Overview of independent observations that were used for validation of C/DA for the Mississippi River Basin.

Name	Description	Reference	Details in
SCAN	In-situ soil moisture	Schaefer et al. (2007)	section 2.3
USGS	In-situ groundwater table in observation wells	Rodell et al. (2007)	section 2.3
GRDC	In-situ river discharge	http://grdc.bafg.de/	section 2.3
GIEMS	Multi-satellite product of surface water extent	Papa et al. (2008)	section 2.3

8.1.2 First C/DA Results

8.1.2.1 Comparison to GRACE TWSA

Monthly time series of TWSA, averaged over the entire Mississippi River Basin, are shown in Fig. 8.3 for the standard and C/DA variants of WGHM, as well as for GRACE. The observed annual variations are found to be smaller than the simulated variations, especially in the C/DA phase in 2005. According to Llovel et al. (2010), the Mississippi River Basin lost an unusual large amount of water in the dryer period of 2005. Assimilating either GRACE TWSA averaged over the entire Mississippi River Basin (EnKF_MRB-1) or averaged over 5°×5° grid cells (EnKF_MRB-2) into WGHM nudged the updated water states towards the GRACE observations, resulting in smaller annual variations. During the C/DA phase, the temporal root mean square error (RMSE) of the basin averaged time series with respect to GRACE is reduced from 28 mm to 8 mm and 9 mm, respectively. Even during the validation phase in 2006-2008, i.e. without any further assimilation of GRACE data, the prediction of TWSA has significantly improved, resulting in a temporal RMSE of 20 mm and 19 mm instead of 27 mm. The positive impact of assimilating spatially high resolved GRACE observations is visible in the spatial domain: the spatio-temporal RMSE is reduced by 47 % (66 mm instead of 126 mm) for EnKF_MRB-2, while EnKF_MRB-1 improves by 25 %. In the validation phase, the effect is smaller with reductions in the RMSE values of about 40 % for both C/DA variants. This can also be seen when analyzing the spatial pattern of the C/DA results, which was described in more detail in Eicker et al. (2014).

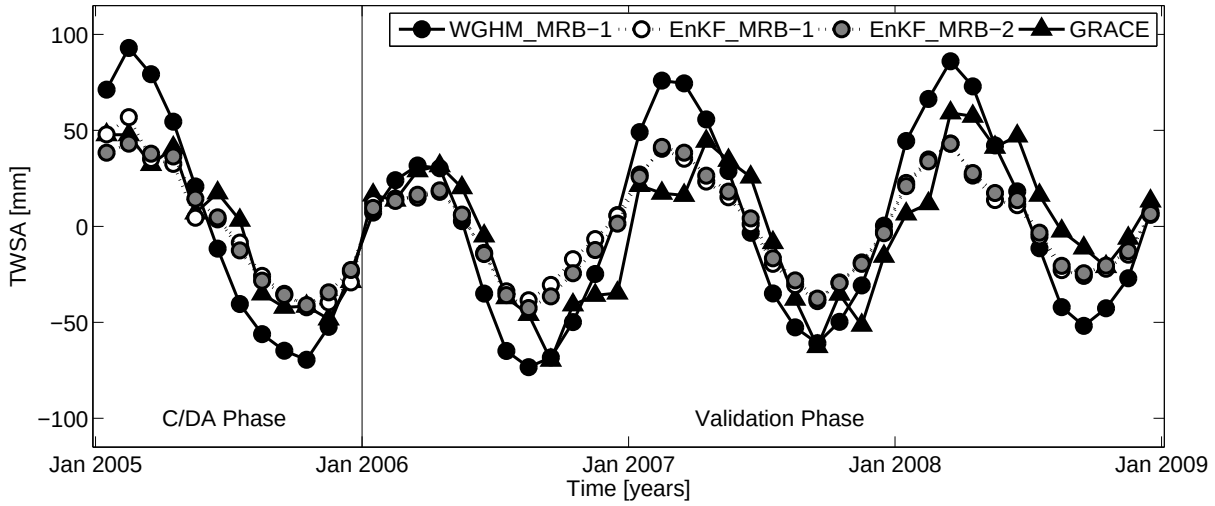

Figure 8.3: Monthly time series of basin-averaged TWSA (in mm) from the standard and two C/DA variants of WGHM, and from GRACE observations during the C/DA phase in 2005 and the validation phase in 2006-2008.

In Fig. 8.4, time series of the update increments for the total and individual water storage changes, averaged over the entire Mississippi River Basin, are shown; i.e. the difference between the model prediction and the model update. The update increment varies among the individual water compartments and in time. An analysis of the spatial pattern of the

update increments also showed variations in the spatial domain (see Eicker et al., 2014). Snow storage content is updated during the winter months, mainly in the mountainous regions of the Mississippi River Basin, while soil water content is updated throughout the year over the entire basin. The water level in surface water bodies (lakes, wetlands, reservoirs and rivers) is updated only in some parts of the basin, e.g., along the Mississippi River. The groundwater storage is predominantly updated in the High Plains Aquifer, where the present simulation of human water consumption introduced large uncertainties to the model simulation. In general, the spatial and the temporal patterns of the increments are found to be similar to those of the water storage uncertainties. The model uncertainties decrease gradually, since the ensemble spread of water states and parameters becomes less. Therefore, the contribution of GRACE on updating the WGHM water states also decreases with time. The amount of water that is assimilated into the model run depends also on the difference between the observed and the model predicted TWSA values. However, hydrological models rely on the concept of mass conservation. Therefore, the amount of total water mass that is assimilated into WGHM also represents the degree to which mass conservation is violated.

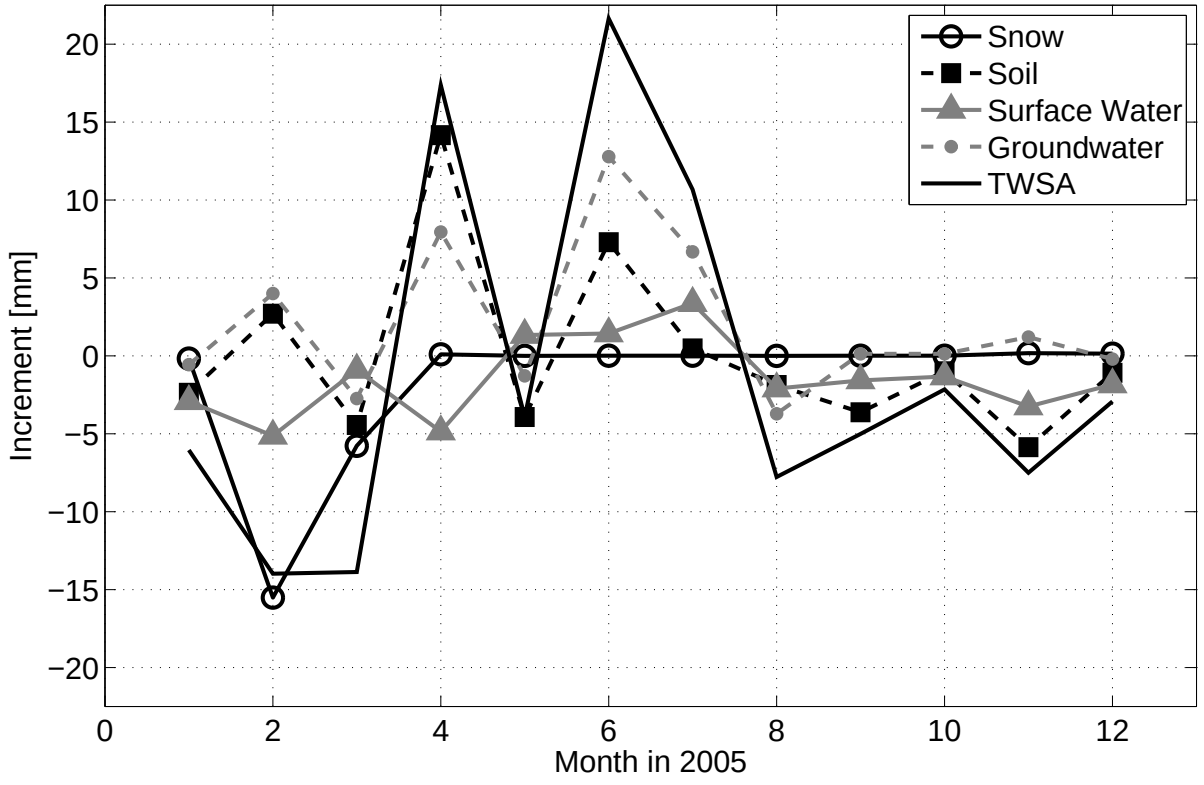

Figure 8.4: Time series of the basin-averaged monthly update increments (in mm) for total water storage and disaggregation to individual water compartments during the C/DA phase in 2005.

In Fig. 8.5 a, the percentages of TWSA from the individual water compartments are shown for the ensemble mean of the open loop (OL) run. The surface water storage significantly contributes in the Mississippi River Basin, followed by water stored in soil and ground. The changes in the percentages of TWSA after C/DA (Fig. 8.5 b) are caused by the

disaggregation of assimilated TWSA into the individual compartments (see Fig. 8.4) and by updating the model parameter values. After C/DA, the variability in surface water and groundwater is reduced, while the snow and soil water reflect more or less the same contribution to TWSA compared to the OL.

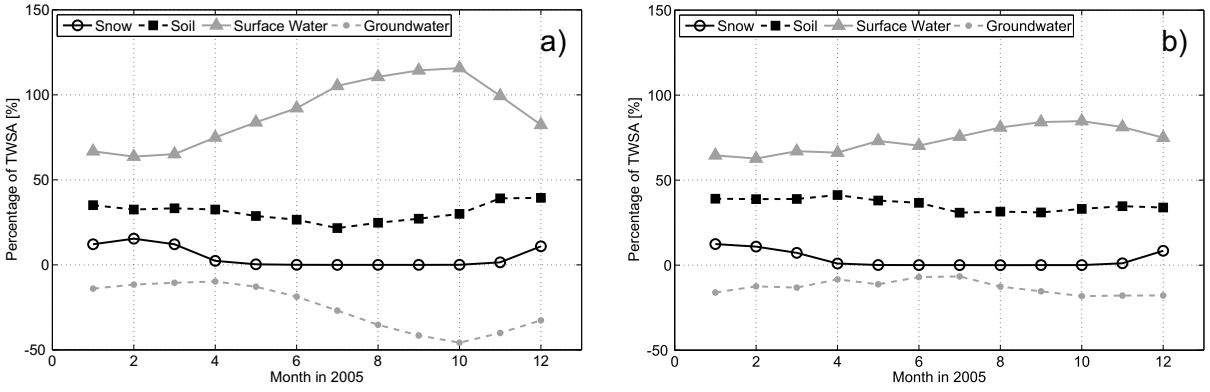

Figure 8.5: Time series of the percentages of basin-averaged TWSA for the individual water compartments of a) the OL run and b) the C/DA run for variant EnKF_MRB-2 during the C/DA phase in 2005.

In section 8.1.2.2, the individual water storage compartments of the standard and C/DA variants of WGHM are validated with independent measurements of soil moisture, groundwater, river discharge and surface water inundation to investigate whether the vertical disaggregation of the total water mass as illustrated in Fig. 8.4 allows for a more realistic representation of compartmental water storage changes.

In Fig. 8.6, time series for the estimates of the six most sensitive parameters are shown for the run EnKF_MRB-2. During the winter months, the calibration suggests a snow melt temperature below 0° (Fig. 8.6 a), while during April to December the parameter value changes marginally. This is in line with the findings in chapter 6, saying that snow parameter values are only updated in the winter months. Also for the other parameters directly related to snow only marginal updates are obtained (not shown here). The small changes in the parameter values between April and December are likely caused by numerically introduced correlations due to the finite ensemble size and due to cross-correlations of the 22 calibration parameters.

Four of the six sensitive parameters converge rather fast against their upper or lower limits (Fig. 8.6 b, c, e, f). A possible reason for this is that the ensemble spread of the parameters is often already strongly reduced to a high extent during the first two update steps. An adjustment of the model parameters in the following months is not obtained, since their uncertainties are too small to be updated in the EnKF. Including an inflation factor will help to avoid the fast ensemble convergence.

8.1.2.2 Validation Using Independent Data Sets

Figure 8.7 shows time series of soil water, groundwater, and surface water averaged over the entire Mississippi River Basin, as well as river discharge for one example station

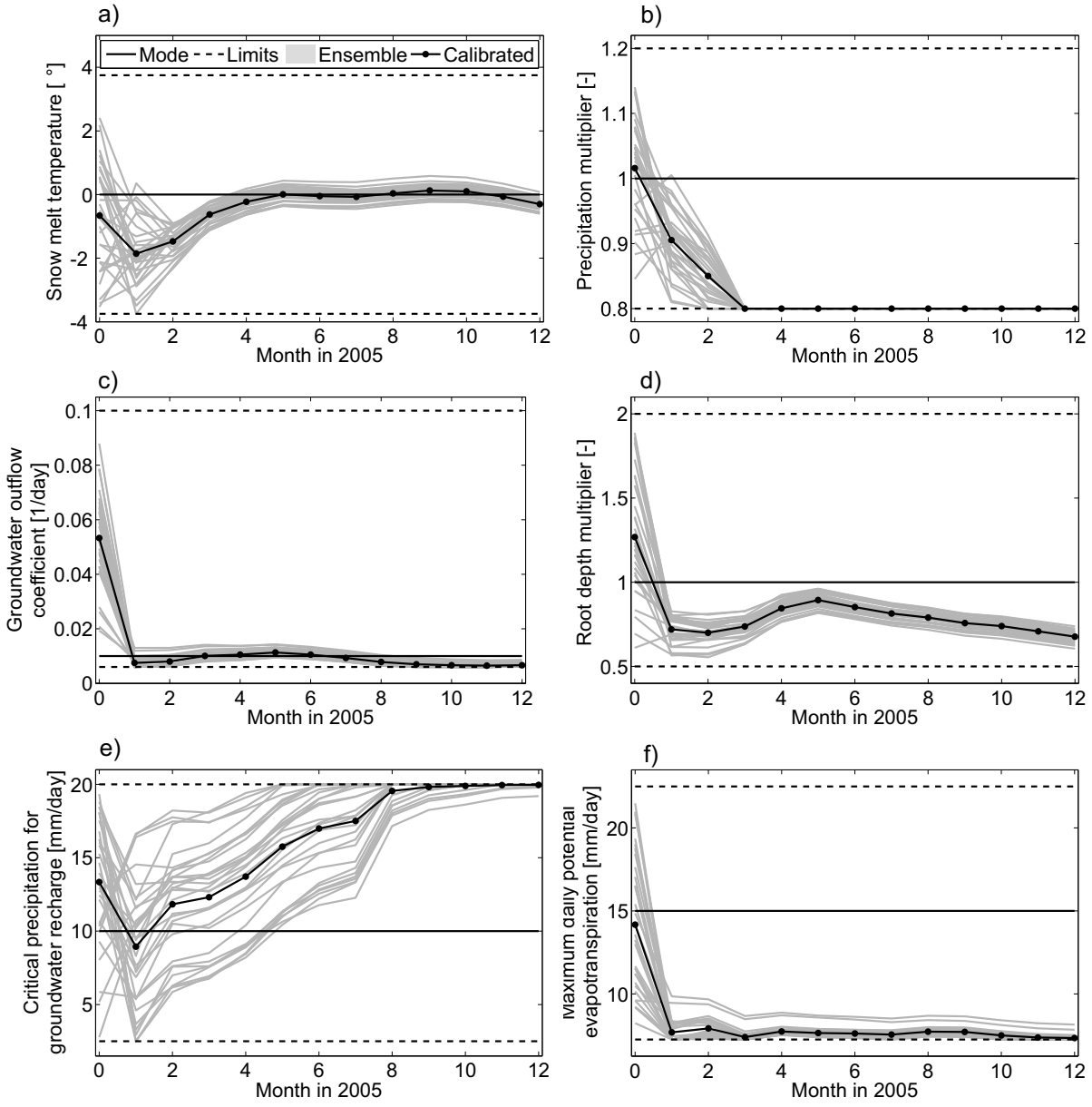

Figure 8.6: Time series of calibrated parameters (ensemble mean) and ensemble members that are found to be most sensitive in chapter 6. The initial parameter ensemble is shown for month "0".

(GRDC station number 4123400). The simulated water changes from the standard WGHM variant (WGHM_MRB-1) and the C/DA variants (EnKF_MRB-1 and EnKF_MRB-2) are compared to independent observations.

In-situ observations and model simulations of soil water changes are spatially averaged over the entire Mississippi River Basin (Fig. 8.7 a). Estimates of soil water changes become better during the C/DA phase. The correlation of model outputs and in-situ soil moisture measurements from SCAN is improved from 0.81 for WGHM_MRB-1 to 0.97 and 0.93 in case of EnKF_MRB-1 and EnKF_MRB-2, respectively. This indicates a better representation of the seasonal phase of the observations. However, in the validation phase, the correlations drop from 0.83 to 0.75 and 0.78. In terms of RMSE, which is used to test the

reproducibility of the seasonal amplitude, only during the C/DA phase of EnKF_MRB-2 the simulation of soil water changes are improved (36 mm instead of 40 mm). This can be explained by the very dry conditions in 2005 that might not reflect the hydrological conditions for the three following years. This leads to an underestimation of the seasonal amplitudes of the soil water storage. Thus, only adapting the model parameters in 2005 does not guarantee improvements between 2006-2008. Uncertainty information for the soil moisture measurements are not available. Instead, the time series of the in-situ observations at all 29 stations in the Mississippi River Basin are shown in Fig. 8.7 a, which illustrates the high spatial variability of soil moisture in the basin. Soil water changes are also averaged over the four major sub-basins of the Mississippi River Basin to investigate the influence of GRACE data assimilation on a smaller spatial scale. During the C/DA phase, the correlation of EnKF_MRB-1 and EnKF_MRB-2 with in-situ soil moisture was increased up to 0.41 in three of four sub-basins but was not improved in the validation phase.

The observed groundwater time series from USGS, prepared for investigations in Rodell et al. (2007), and the study period overlap only for the year 2005. A comparison of the observed and simulated basin averaged time series (Fig. 8.7 b) shows that WGHM_MRB-1 underestimates the annual amplitude by a factor of 2.2 (17 mm instead of 38 mm). In case of EnKF_MRB-1 the representation of the annual amplitude has not considerably changed, while EnKF_MRB-2 leads to a larger amplitude. This seems to disagree with the reduction of the annual amplitude that is found for TWSA. However, while integration of GRACE into WGHM results in larger amplitudes for groundwater, the amplitudes of soil water changes are considerably reduced (see Fig. 8.7 a). Nonetheless, the observed groundwater amplitude is still underestimated by EnKF_MRB-2 by a factor of 1.9. WGHM_MRB-1 shows a high correlation of 0.95 for the Mississippi River Basin. When basin-averaged or gridded TWSA are assimilated, the correlation decreased to 0.36 and 0.80. The sub-basin averages are also analyzed. These time series are shown in Fig. 8.7 b as well, since their values were spatially averaged to determine the basin averaged time series. The investigations showed higher agreements with the measurements in two of four sub-basins in case of EnKF_MRB-2. Compared to previous GRACE data assimilation experiments into NASA's catchment land surface model for the Mississippi River Basin improvements in terms of correlations also for two or three of the four sub-basins were reported (Houborg et al., 2012, Zaitchik et al., 2008).

Simulated river discharge at seven stations was validated against measurements provided by GRDC. In Fig. 8.7 c, for example, the time series at one particular river discharge station (GRDC station number 4123400) located in the eastern part of the Ohio/Tennessee Basin (Fig. 8.2) is shown. Uncertainty information is not available for the river discharge observations. WGHM_MRB-1 overestimates the high flows occurring in spring dramatically, while the low flows are represented almost perfectly. In case of EnKF_MRB-1, the simulation of high flows is nudged to the observed river discharge. Even further improvements are achieved in case of EnKF_MRB-2 in terms of the Nash-Sutcliffe coefficient (NSC). In the C/DA phase, NSC is improved from -1.48 to -0.22, which still represents a poor fit with the observations. Then, during the validation phase, the NSC increases from 0.43 to 0.63 (Fig. 8.7 c). Since NSC is very sensitive to high flows, better values are achieved for EnKF_MRB-1 and EnKF_MRB-2, although the low flows are better represented by WGHM_MRB-1. A perfect fit between simulated and observed river discharge

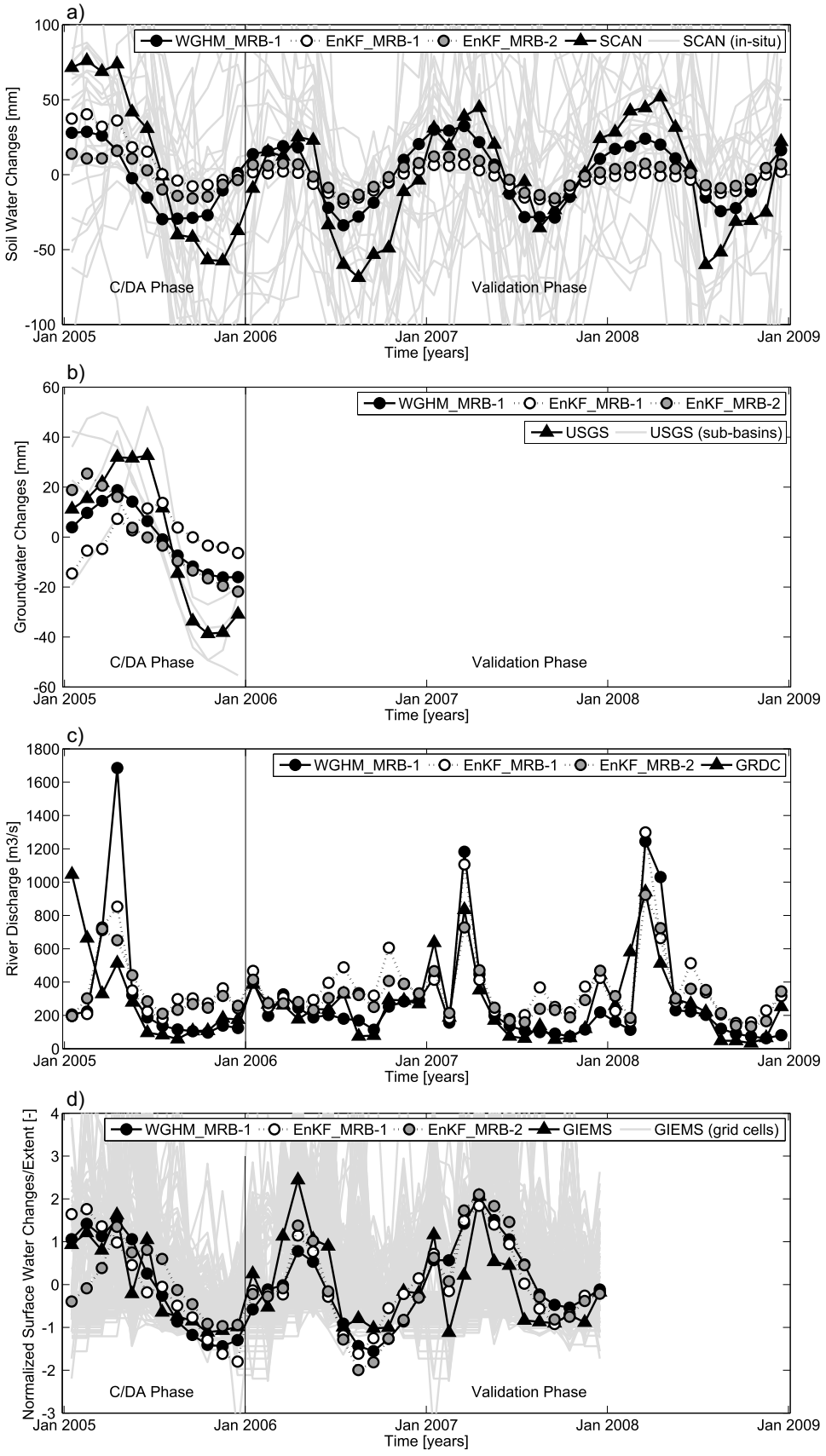

Figure 8.7: Time series of a) soil water and b) groundwater changes, c) river discharge at GRDC station 4123400, and d) normalized surface water changes and extent are shown from the standard and C/DA variants of WGHM averaged over the entire Mississippi River Basin (except for river discharge). Corresponding independent measurements are also shown, which are used for validation.

would result in a NSC of 1. The investigations indicate that the simulation of river discharge during the C/DA phase is improved for up to five stations for both C/DA variants and for four stations in case of EnKF_MRB-2, even during the validation phase. C/DA of TWSA into WGHM shows potential to also exhibit a positive impact on water fluxes.

Figure 8.7 d presents the time series for surface water bodies (summarizing lakes, wetlands, reservoirs and rivers) for 2005-2007 averaged over the entire Mississippi River Basin. The time series of the individual grid cells that contribute to the basin average are also shown. Since WGHM simulates water storage changes, represented as equivalent water heights (in mm), and GIEMS specifies the area of a grid cell that is inundated (in km^2), the time series are normalized and their correlation coefficients are computed following Papa et al. (2008). A high correlation of 0.90 exists in the C/DA phase and a moderate correlation of 0.64 in the validation phase between WGHM_MRB-1 and GIEMS. These values are found to be higher than the correlation of 0.51 that was reported in Papa et al. (2008), who compared surface water extent from GIEMS to model simulations from WGHM for the Mississippi River Basin for 2003-2004. Papa et al. (2008) explained the low to moderate correlations by snow melt in the northern part of the basin (Missouri, Upper Mississippi, and northern part of Ohio/Tennessee) which causes spring floods that complicate the hydrological processes within the basin. Assimilation of GRACE TWSA results in decreased correlation coefficients of 0.85 and 0.66 during the C/DA phase, while improvements in the validation period are achieved (0.74 and 0.71). Analyzing the four sub-basins of the Mississippi River Basin shows that in each case of the C/DA variants two sub-basins were improved during the assimilation phase. Then, the surface water simulation of two sub-basins for EnKF_MRB-1 and three for EnKF_MRB-2 were enhanced in the validation phase.

8.1.3 Discussion and Conclusions for the Mississippi Case Study

Assimilation of GRACE TWSA into WGHM improves the fit of its TWSA simulations with GRACE in the C/DA phase in 2005, as well as in the validation phase during 2006-2008. However, this does not necessarily hold for the individual water states and river discharge. For some compartments, assimilation improves the fit to observations during the C/DA period, while for a few the fit becomes worse. Over the validation period, the calibrated parameters often lead to a worse fit as compared to the (already calibrated) standard model output. In summary, hypothesis II a of this thesis is supported. It was shown that assimilation of GRACE TWSA enables a more realistic representation of TWSA even after the C/DA phase. Hypothesis II b, stating that assimilation of GRACE TWSA also positively impacts individual water compartments and fluxes, cannot completely be proven in this study. In case of in-situ soil moisture and groundwater data, the locations and the rather sparse station networks complicate the assessment of the C/DA results. Even so the groundwater observation wells were chosen based on their location in unconfined or semi-confined aquifers (Rodell et al., 2007), local effects such as pumping might influence the measurements, which means they are likely to be more seasonally dynamic. This might result in overestimated observed seasonal amplitudes (Strassberg et al., 2009). Additionally, in case of spatial averaging, i.e. interpolation to the areas between the soil moisture or groundwater observation stations, seasonal amplitudes of

(sub-)basin averages may be over- or underestimated (Döll et al., 2012). Nevertheless, our investigations show improvements of groundwater and soil water changes at least during the C/DA phase in half of the sub-basins of the Mississippi River Basin or more, which suggests that GRACE data contain valuable information that helps constraining these individual model compartments. Integrated GRACE information on a 5°×5° grid (EnKF_MRB-2) performs better for the groundwater compartment, while integration of basin-averaged GRACE TWSA data (EnKF_MRB-1) shows higher agreement with soil moisture observations. Therefore, hypothesis III of the thesis could not completely be proven in this C/DA experiment. However, it is noticed that the informative value of a one year validation is questionable, as it is presented for the groundwater compartment.

A large amount of river discharge during spring in the modeled time series at stations in the north-eastern part of the Mississippi River Basin could not be confirmed by the observations. It is likely that this is caused by overestimated snow melt in spring for this region, which is reduced by TWSA assimilation, especially in case of EnKF_MRB-2. Therefore, it was shown that GRACE data assimilation enables detecting limitations in the representation of water changes in the model and might even reduce their effect on the model outputs.

EnKF_MRB-1 and EnKF_MRB-2 partly show better fits to GIEMS in terms of correlation coefficients, but partly worse fits compared to WGHM_MRB-1. However, the mean annual cycle was analyzed which revealed a non-linear relation between water storage extent and changes (not shown here). This suggests the application of a rank correlation coefficient, which takes into account the non-linearity and therefore might be better suited for interpretation than the application of the linear correlation coefficient.

From these findings, one can conclude that assimilation of GRACE observations into WGHM over only one year already enables an improved prediction of TWSA in the following years, but this is apparently not the case for individual water compartments. The parameters have been calibrated when the hydrology in the Mississippi River Basin exhibited an extreme state, i.e. the dry conditions in 2005, and therefore are not valid for the following years. A longer period for C/DA would smooth these extreme events, and calibrated parameter values might likely be more representative for the following period. Therefore, it is preferable to assimilate GRACE TWSA data whenever they are available. However, this would require the introduction of an inflation factor to avoid fast ensemble convergence (see section 4.3). It is noticed that the C/DA phase was probably chosen too short to allow the model parameters to transfer the seasonal behavior of the observations to WGHM. Another difficulty might be the fact that 22 model parameters are calibrated simultaneously. Choosing any sub-set of these parameters for calibration and freezing the others might improve the results by avoiding dependencies between the calibration parameters and thereby leading to a better representation of individual water compartment changes as well as water fluxes.

8.2 Transfer C/DA to Murray-Darling River Basin

The C/DA framework is tested for the Murray-Darling River Basin in the south-east of Australia, that faces different climatic and anthropogenic conditions compared to the Mississippi River Basin. The motivation is to prove the general application of the proposed C/DA framework, which is formulated in objective 4 of this thesis. The Murray-Darling River Basin is one of the driest basins over the world. A long drought period occurred over the basin during 2001-2009, the so-called Millennium Drought (Leblanc et al., 2012). As a result, the water storage in the region steadily decreased, attributed to a combination of dry meteorological conditions and extensive irrigation for agriculture. Simulations of WGHM are used to estimate linear trends in soil water, surface water and groundwater compartments, as well as in TWSA during 2003-2009. However, the model is not able to capture the effect of the Millennium Drought on the storage in compartments, likely due to missing processes in dry regions or uncertainties in climate forcing data. Therefore, in the following it is investigated whether assimilating GRACE TWSA into WGHM enables a more realistic representation of the water decline.

In section 8.2.1, an overview of the study set-up is given. Firstly, the linear trend of TWSA from different GRACE solutions and with different post-processing (filtering) is analyzed (section 8.2.2.1). Secondly, the EnKF and singular evolutive interpolated Kalman filter (SEIK) are then applied to integrate GRACE TWSA along with its full error covariance information into WGHM during 2003-2009 (section 8.2.2.2). Finally, independent observations of groundwater are used to validate the model outputs after C/DA (section 8.2.2.4). The findings of the study are summarized in section 8.2.3.

Parts of these assessments were presented in Schumacher et al. (2016c). Results were shown for 2003-2010, in which both GRACE data and forcing data for WGHM are available. In this thesis, the focus is on the Millennium Drought period between 2003-2009. Additionally, the SEIK filter is applied and a detailed discussion of the update increments is provided, which describe the violation of the mass conservation principle.

8.2.1 Set-up of C/DA for the Murray-Darling River Basin

Similar to the synthetic experiment in chapter 7 and to the Mississippi case study in section 8.1, the WaterGAP version 2.2 is used here. The model has already been calibrated against mean annual river discharge at 1319 gauge stations, of which 11 stations are located in the Murray-Darling River Basin (Müller Schmied et al., 2014). Monthly forcing fields of temperature, cloud cover, and the number of wet days are used from the Climate Research Unit's Time Series (CRU TS 3.2) and precipitation is taken from the Global Precipitation Climatology Center (GPCC v6; see section 2.2.2.2). An overview of the set-up for transferring the developed C/DA approach to the Murray-Darling River Basin is given in Fig. 8.8.

The initial ensemble of WGHM water states as well as the ensemble of model parameters and forcing data are generated using the same procedure as for the Mississippi case study (section 8.1). Here, the model is initialized during 1995-2000 and the model spin-up phase

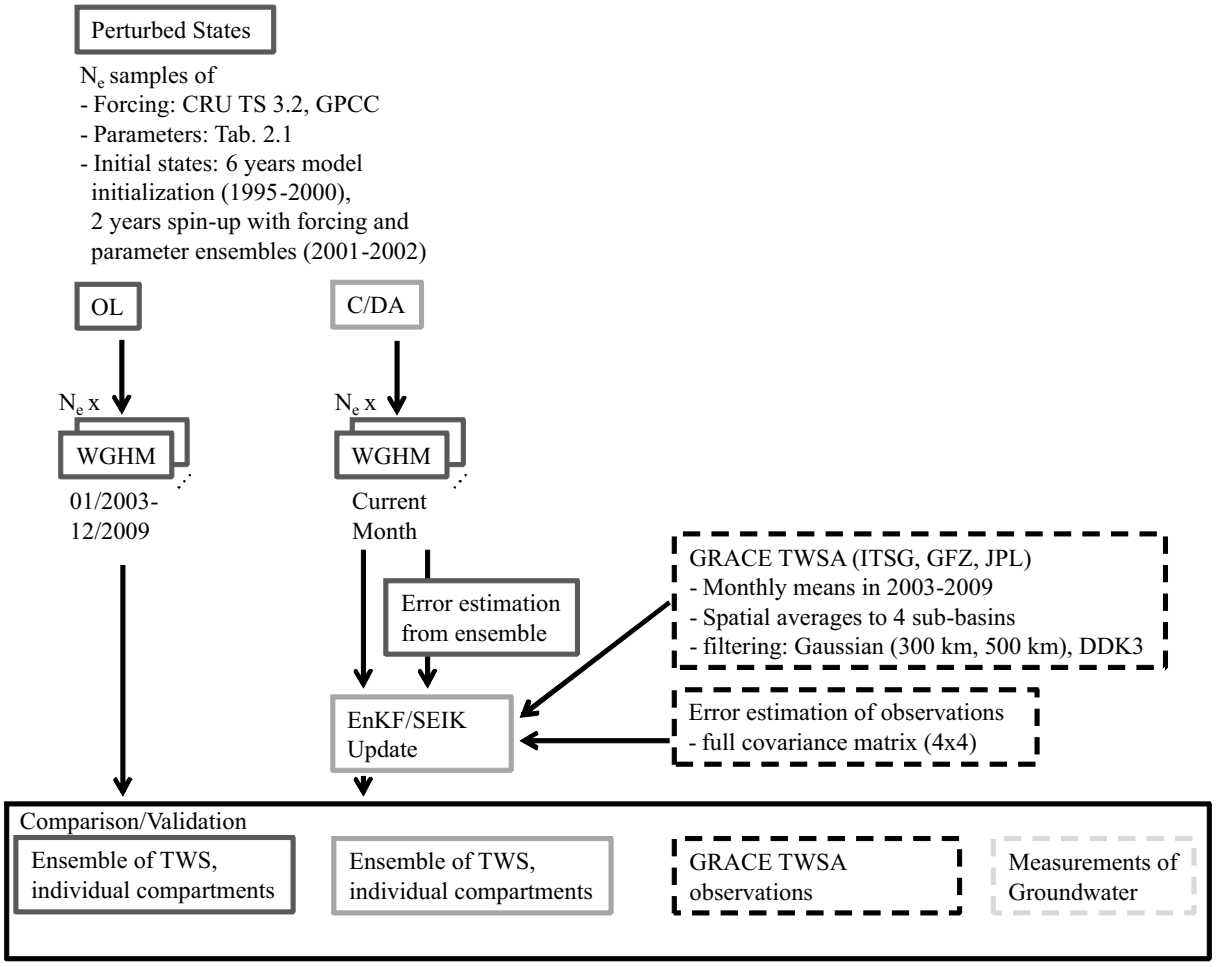

Figure 8.8: An overview of the set-up of the transfer study for the Murray-Darling River Basin. First, $N_e = 30$ open loop (OL) model runs are performed for 2003-2009 (left column). Then, GRACE TWSA are assimilated into WGHM considering different configurations (center and right column). To assess the C/DA results, simulated TWSA and groundwater changes are compared to GRACE and to independent groundwater well measurements.

with the ensembles for the parameters, temperature and precipitation is performed for 2001-2002 to generate the ensemble of initial water states. This is followed by an open loop (OL) run of WGHM during 2003-2009 (first column in Fig. 8.8 and Tab. 8.3). Monthly GRACE observations along with their full error covariance information are assimilated into WGHM using the EnKF, as well as the SEIK filter during 2003-2009 (second column in Fig. 8.8 and Tab. 8.3).

In this study, monthly GRACE spherical harmonic coefficients of the ITSG-Grace2014 time series are used (https://www.tugraz.at/institute/ifg/downloads/gravity-field-models/itsg-grace2014/), which are provided with full variance-covariance information for the time span 02/2003-06/2014. The GRACE solutions are evaluated up to spherical harmonic degree/order 90, smoothed using a 300 km Gaussian filter, and converted to gridded values of TWSA on the $0.5°\times0.5°$ grid used by WGHM. A formal variance-covariance error propagation is applied to obtain the observation error covariance matrices for the gridded values. The full error covariance matrix of the po-

tential coefficients provided for 05/2003 is used and it is assumed that the error information of TWSA is temporally constant. To study the effect of different filter approaches on the C/DA results, the GRACE solution is also prepared with a 500 km Gaussian filter, as well as the anisotropic DDK3 filter (Kusche et al., 2009). In addition, the influence of different GRACE products on the results is investigated. Therefore, GFZ RL05 (ftp://podaac.jpl.nasa.gov/allData/grace/L2/GFZ/RL05/) and JPL RL05 (ftp://podaac.jpl.nasa.gov/allData/grace/L2/JPL/RL05/) official products are considered in this case study, filtered using a DDK3 filter. In order to isolate the influence of changing the TWSA observation vector, in all cases the covariance matrix of the ITSG-Grace2014 for 300 km Gaussian filtering is considered. No correction for signal damping and spatial leakage due to filtering is applied, since the re-scaling factors at the scale of the four sub-basins are estimated to be close to 1 (see Schumacher et al., 2016c).

With respect to the findings in the Mississippi River Basin and considering the spatial resolution of GRACE data, the GRACE TWSA observations are aggregated to spatial averages over the four major sub-basins of the Murray-Darling River Basin (Fig. 8.9). The Darling Basin in the North is divided into a western and an eastern part with areas of about 240000 km^2 and 380000 km^2 and the Murray Basin in the South into two areas of about 215000 km^2 each.

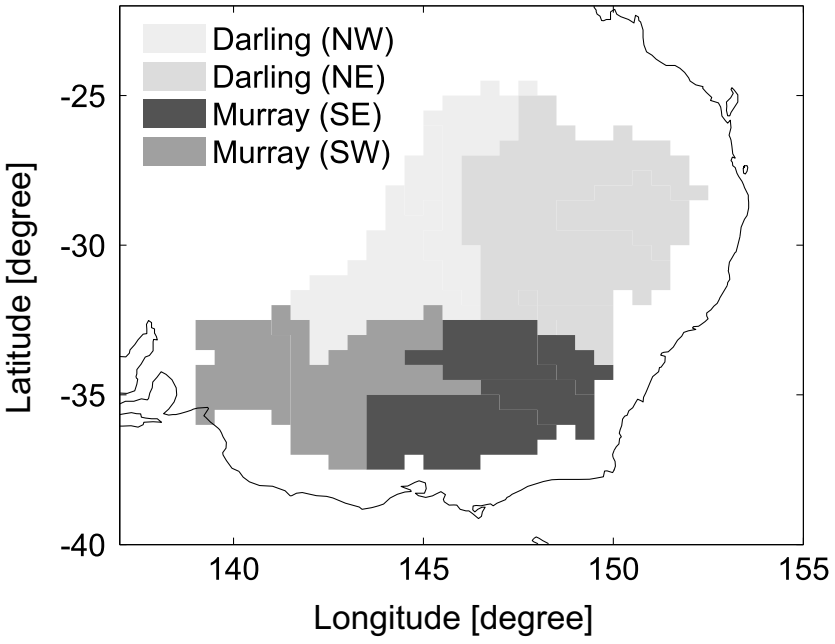

Figure 8.9: The north-western (NW) and north-eastern (NE) Darling Basins as well as the south-eastern (SE) and south-western (SW) Murray Basins are defined as the four major sub-basins of the Murray-Darling River Basin. GRACE TWSA are spatially averaged over these regions and introduced as observations in the C/DA.

The C/DA is performed during 2003-2009. In contrast to the Mississippi case study, an inflation factor of 10 % is used for representing errors in the model structure and to ensure a contribution of GRACE TWSA to the updated water states as well as parameters over the entire study period (see section 7.1 of the synthetic experiments for details on the chosen inflation factor). Here, again water mass is not conserved in the EnKF/SEIK filter

update steps. Seven experiments are carried out, for which a range of configurations is defined: (i) different GRACE products are used (ITSG, GFZ, JPL), (ii) various filters are applied to the ITSG-Grace2014 data product (300 km and 500 km Gaussian filter, as well as DDK3), and (iii) the classical EnKF algorithm as well as the SEIK filter are applied to evaluate their effects on the C/DA results (see section 4.2.1 and 4.2.3). The abbreviations and applied options are provided in Tab. 8.3.

Table 8.3: Overview of model simulations and assimilation runs that are analyzed in this study.

Run	Method	GRACE solution	Spatial Filtering
OL	Open Loop	-	-
ITSG-300km	EnKF	ITSG-Grace2014	300 km Gaussian
ITSG-500km	EnKF	ITSG-Grace2014	500 km Gaussian
ITSG-DDK3	EnKF	ITSG-Grace2014	DDK3
GFZ-DDK3	EnKF	GFZ RL05	DDK3
JPL-DDK3	EnKF	JPL RL05	DDK3
ITSG-SEIK	SEIK	ITSG-Grace2014	DDK3

To assess the performance of the C/DA results, the simulated TWSA and groundwater changes are compared to GRACE and to independent groundwater well measurements (see section 2.3), respectively.

8.2.2 Hydrological Characteristics of the River Basin

8.2.2.1 Meteorological and Hydrological Drought

Precipitation data from the Bureau of Meteorology (BoM, Australia; http://www.bom.gov.au/climate/data/) are analyzed to identify dry and wet periods in the Murray-Darling Basin. Figure 8.10 a shows the annual rainfall during 1981-2013 averaged over the entire region compared to the temporal mean of 477 mm over the time period. During the first two decades, there have been regularly drier and wetter years, while between 2001-2009 the annual rainfall remained always below the long term temporal mean, including two years that belong to the driest years over the whole period (2002, 2007). The annual rainfall sum over the basin was up to 41 % smaller between 2001-2009 than the three decade average rainfall. In contrast, the years 2010-2012 have been wetter than the long term average. Especially, in 2010 the annual rainfall was 66 % larger than the average. These investigations show that a meteorological drought over the Murray-Darling River Basin occurred between 2001-2009 (see also Forootan et al., 2016). In Fig. 8.10 b, the monthly TWSA from GRACE averaged over the entire Murray-Darling River Basin are shown between 2003-2013. During 2003-2007, TWSA decreased and was below the mean up to the end of 2009. The strong rainfall events in the two following years resulted in an increase of the total water mass in 2010-2011. Afterwards, the water mass was still above the temporal mean during

2003-2013. In addition, the time series of WGHM simulated TWSA is shown during 1995-2010 in Fig. 8.10 b. The strong water decline in 2002 and the strong increase of TWSA in 2010 are clearly visible in WGHM showing the model reacting to these extreme rainfall events, although, the increase in model simulations of TWSA is smaller compared to GRACE TWSA. In 2006-2007, WGHM simulates no further decline in TWSA, which is clearly suggested by GRACE observations. It seems that the WGHM water storage compartments are already quite empty. Therefore, an improved representation of the TWSA decline between 2003-2009 is expected by merging GRACE and WGHM in the C/DA framework.

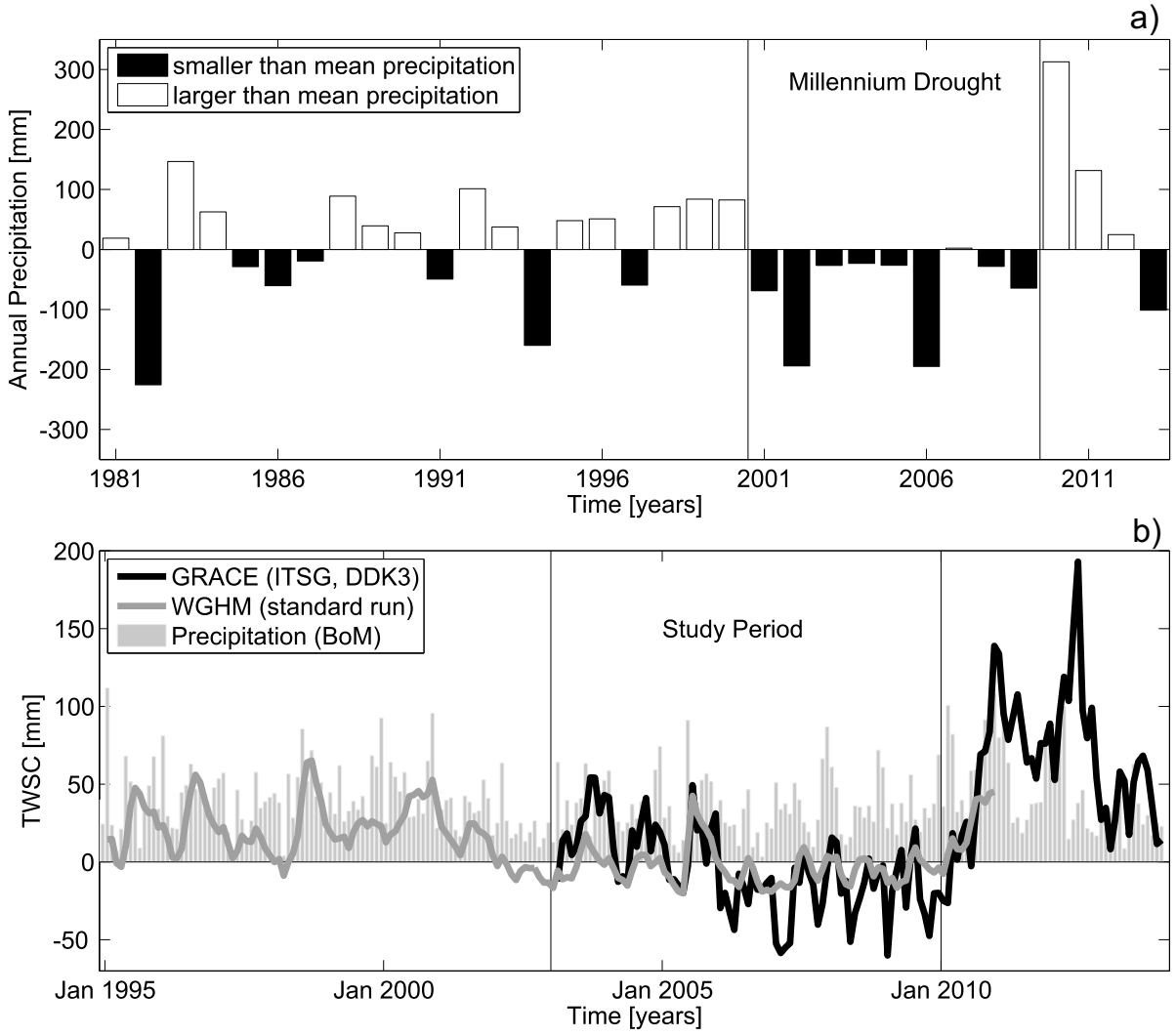

Figure 8.10: a) Annual precipitation from BoM is shown averaged over the entire Murray-Darling River Basin with respect to the temporal mean of 477 mm during 1981-2013. b) TWSA time series from ITSG-Grace2014 during 2003-2013 and WGHM standard run during 1995-2010 are plotted, as well as monthly precipitation from BoM during 1995-2013 averaged over the entire Murray-Darling River Basin.

The estimation of linear trends over 2003-2009 from ITSG-DDK3 shows a decrease of -7.6 mm/year over the entire Murray-Darling River Basin, ranging from -2.9 mm/year in the north-eastern Darling Basin to -14.0 mm/year in the south-eastern Murray Basin

(Tab. 8.4). The application of different filters to smooth GRACE TWSA has a small impact on the linear trend estimation in the Darling sub-basins (differences of around 0.3 mm/year, see Tab. 8.4) and a higher influence in the Murray sub-basins (differences of up to 3.0 mm/year, see Tab. 8.4). Using different GRACE products for the trend estimation has a similar effect on the results. However, all analyzed GRACE data sets show negative trends in TWSA for the entire Murray-Darling River Basin and its four major sub-basins.

Table 8.4: Linear trend in TWSA during 2003-2009 in mm/year from different GRACE products and after applying different filter methods. The trends are provided for the spatial averages over the entire Murray-Darling River Basin (MDB), and its four major sub-basins (columns 2-6). Averaged trends and their uncertainties are estimated after applying different filtering techniques (column 7), as well as from different GRACE products (column 8).

Basin	ITSG 300 km	ITSG 500 km	ITSG DDK3	GFZ DDK3	JPL DDK3	Filtering	Product
MDB	-7.0	-5.7	-7.6	-5.6	-4.6	-6.8 ± 1.0	-5.9 ± 1.5
Darling (NW)	-4.5	-4.3	-3.8	-2.2	-2.0	-4.2 ± 0.3	-2.7 ± 1.0
Darling (NE)	-3.3	-3.5	-2.9	-0.6	1.2	-3.2 ± 0.3	-0.8 ± 2.1
Murray (SE)	-11.2	-8.0	-14.0	-11.3	-9.8	-11.1 ± 3.0	-11.7 ± 2.1
Murray (SW)	-12.0	-8.7	-13.5	-12.6	-12.4	-11.4 ± 2.4	-12.8 ± 0.6

8.2.2.2 Update of TWSA

The TWSA time series from OL, GRACE and the C/DA variant ITSG-DDK3 are shown in Fig. 8.11, averaged over the entire Murray-Darling River Basin and its four major sub-basins. The C/DA results show a much better agreement with the GRACE TWSA than the OL variant of WGHM. In terms of RMSE, the fit for the entire basin is improved by 50 % (from 21.4 mm to 10.7 mm), ranging from 45 % in the north-western Darling Basin to 53 % in both Murray sub-basins (Tab. 8.5). The choice of different filters or different GRACE products leads to improvements for the entire basin up to 60 %. The low up to moderate correlations between the OL and GRACE observations, e.g., 0.21 in the north-western Darling to 0.68 in the south-eastern Murray Basin, are considerably increased after GRACE assimilation (Tab. 8.6). In case of assimilating ITSG-DDK3, high correlations for all basins are achieved, ranging from 0.75 in the north-western Darling to 0.95 in the south-eastern Murray Basin. In summary, all EnKF variants of WGHM outperform the OL in terms of RMSE and correlation for the entire Murray-Darling River Basin as well as for all sub-basins. The seasonal variability which is assessed in terms of RMSE, and the seasonal cycle of TWSA, which is assessed in terms of correlation coefficients, are better represented after C/DA.

The estimation of linear trends in TWSA from the OL and C/DA variants of WGHM are summarized in Tab. 8.7 for the spatial averages over the entire Murray-Darling River Basin and its four major sub-basins. A comparison of the trends after C/DA to the trends

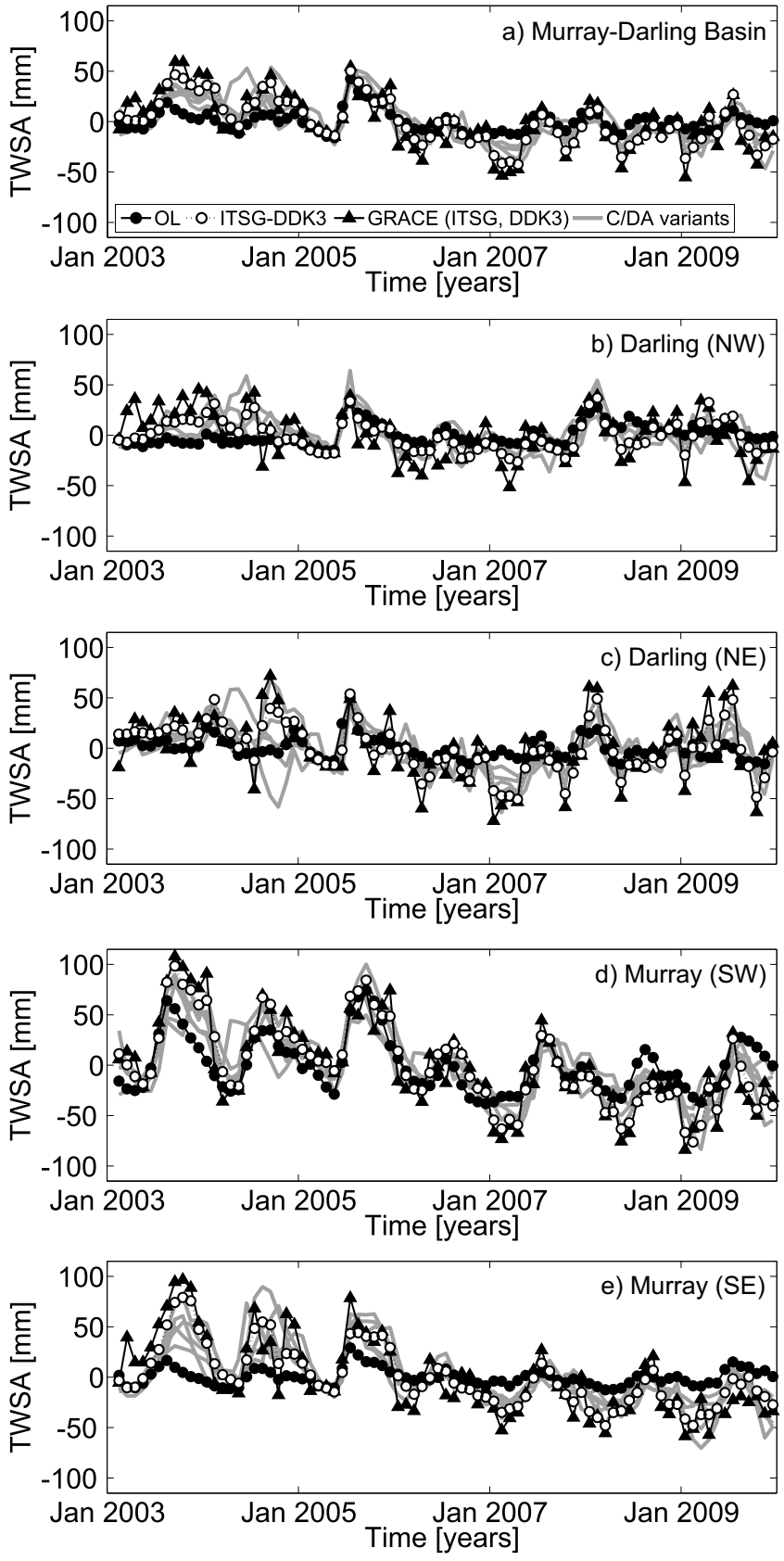

Figure 8.11: Time series of TWSA (in mm) from open loop (OL) simulations from WGHM and from ensemble filter updates averaged over a) the entire Murray-Darling River Basin and the four major sub-basins b)-e).

Table 8.5: Agreement between model predicted and observed TWSA in terms of RMSE in mm averaged over the entire Murray-Darling River Basin (MDB) and over its four major sub-basins. The values in brackets indicate the improvement of RMSE in mm compared to the OL run. The OL run is compared to each assimilated GRACE data set (see Tab. 8.3 for names of C/DA variants): The average of the RMSE values is provided in column 2 and its range in brackets.

Basin	OL	ITSG-300km	ITSG-500km	ITSG-DDK3
MDB	21.4 (±3.9)	15.5 (-4.5)	8.2 (-7.4)	10.7 (-11.0)
Darling (NW)	28.5 (±7.3)	34.8 (-3.1)	11.5 (-7.9)	15.7 (-7.6)
Darling (NE)	28.2 (±6.6)	20.3 (-12.9)	9.1 (-8.2)	14.7 (-13.1)
Murray (SE)	28.9 (±5.6)	22.3 (-10.4)	9.3 (-9.8)	13.7 (-16.5)
Murray (SW)	34.6 (±8.1)	23.3 (-12.0)	10.1 (-11.6)	16.1 (-17.7)
		GFZ-DDK3	JPL-DDK3	ITSG-SEIK
MDB		17.0 (-9.0)	16.9 (-6.6)	17.3 (-4.4)
Darling (NW)		23.4 (-6.9)	23.5 (-7.9)	20.8 (-2.4)
Darling (NE)		20.1 (-13.3)	17.4 (-11.8)	26.0 (-1.8)
Murray (SE)		19.6 (-12.9)	18.6 (-11.5)	20.6 (-9.6)
Murray (SW)		20.2 (-18.2)	28.2 (-15.4)	24.6 (-9.2)

Table 8.6: Agreement between model predicted and observed TWSA in terms of correlation averaged over the entire Murray-Darling River Basin (MDB) and over its four major sub-basins. The values in brackets indicate the improvement of correlation compared to the OL run. The OL run is compared to each assimilated GRACE data set (see Tab. 8.3 for names of C/DA variants): The average of the correlation values is provided in column 2 and its range in brackets.

Basin	OL	ITSG-300km	ITSG-500km	ITSG-DDK3
MDB	0.58 (±0.05)	0.78 (+0.18)	0.92 (+0.29)	0.92 (+0.31)
Darling (NW)	0.21 (±0.07)	0.39 (+0.26)	0.79 (+0.62)	0.75 (+0.52)
Darling (NE)	0.44 (±0.05)	0.84 (+0.44)	0.90 (+0.39)	0.89 (+0.44)
Murray (SE)	0.68 (±0.06)	0.84 (+0.24)	0.94 (+0.19)	0.95 (+0.22)
Murray (SW)	0.50 (±0.04)	0.80 (+0.32)	0.92 (+0.38)	0.91 (+0.39)
		GFZ-DDK3	JPL-DDK3	ITSG-SEIK
MDB		0.83 (+0.30)	0.80 (+0.27)	0.76 (+0.15)
Darling (NW)		0.69 (+0.39)	0.69 (+0.49)	0.46 (+0.23)
Darling (NE)		0.85 (+0.47)	0.85 (+0.40)	0.54 (+0.09)
Murray (SE)		0.89 (+0.24)	0.90 (+0.22)	0.88 (+0.15)
Murray (SW)		0.89 (+0.35)	0.81 (+0.37)	0.77 (+0.25)

estimated from OL and from the different GRACE data sets shows that C/DA introduces or intensifies negative trends in TWSA. The mean difference from the C/DA variants compared to GRACE is 1.5 mm/year, while the mean difference to OL is 5 mm/year. Only in case of applying the SEIK filter, the decline is close to the equally weighted average of OL and GRACE decline in TWSA.

The differences in TWSA trends, estimated from the different C/DA variants (Tab. 8.7), are found to have a similar magnitude as the differences in TWSA trends estimated from the different GRACE products and after applying different filtering (Tab. 8.4). This indicates that the choice of the GRACE solution and the filter to smooth GRACE TWSA affect the C/DA results.

Table 8.7: Linear trend (in mm/year) in TWSA during 2003-2009 from OL and C/DA variants. The trends are provided for the spatial averages over the entire Murray-Darling River Basin (MDB) and over its four major sub-basins in the Darling (D) and Murray (M) basins (columns 2-8). See Tab. 8.3 for names of C/DA variants. Averaged trends and their uncertainties are estimated after applying different filtering techniques (column 9), as well as from different GRACE products (column 10).

Basin	OL	ITSG-300km	ITSG-500km	ITSG-DDK3	GFZ-DDK3	JPL-DDK3	ITSG-SEIK	Filtering	Product
MDB	-0.9	-6.2	-4.4	-6.5	-5.8	-3.5	-5.2	-5.7 ± 1.1	-5.3 ± 1.6
D (NW)	2.1	-3.0	-2.0	-1.0	-1.7	0.4	0.2	-2.0 ± 1.0	-0.8 ± 1.0
D (NE)	-1.6	-4.0	-3.4	-4.2	-2.8	0.0	-4.8	-3.9 ± 0.4	-2.3 ± 2.1
M (SE)	-3.7	-9.8	-6.3	-13.0	-11.6	-8.2	-11.1	-9.7 ± 3.4	-10.9 ± 2.5
M (SW)	-0.4	-10.1	-6.9	-10.0	-10.1	-9.1	-6.0	-9.0 ± 1.8	-9.7 ± 0.6

8.2.2.3 Update of Individual Water Storage Compartments

In Fig. 8.12, the update increments, i.e. the difference between model prediction and model update, are shown for the total and the individual water compartments averaged over the entire Murray-Darling River Basin. The amount of water mass that is introduced into WGHM varies over time. The TWSA update increments have similar magnitudes compared to the update increments in the Mississippi River Basin (see Fig. 8.4). However, after a few update steps the introduced water mass decreased considerably in the Mississippi Basin. In contrast, even after six full years of C/DA, water mass is still introduced into the model in 2009 for the Murray-Darling River Basin. This is ensured through the application of an inflation factor that keeps the model ensemble spread large enough, and thereby guarantees a contribution of GRACE observations over the entire study period.

In 2003, the TWSA update increment is predominantly associated with the soil water compartment (Fig. 8.12 b). This was also the case for the one year C/DA phase in the Mississippi case study (section 8.1). Then, in 2004 a large portion of the TWSA update increment is introduced to the surface water storage compartments (Fig. 8.12 a), while

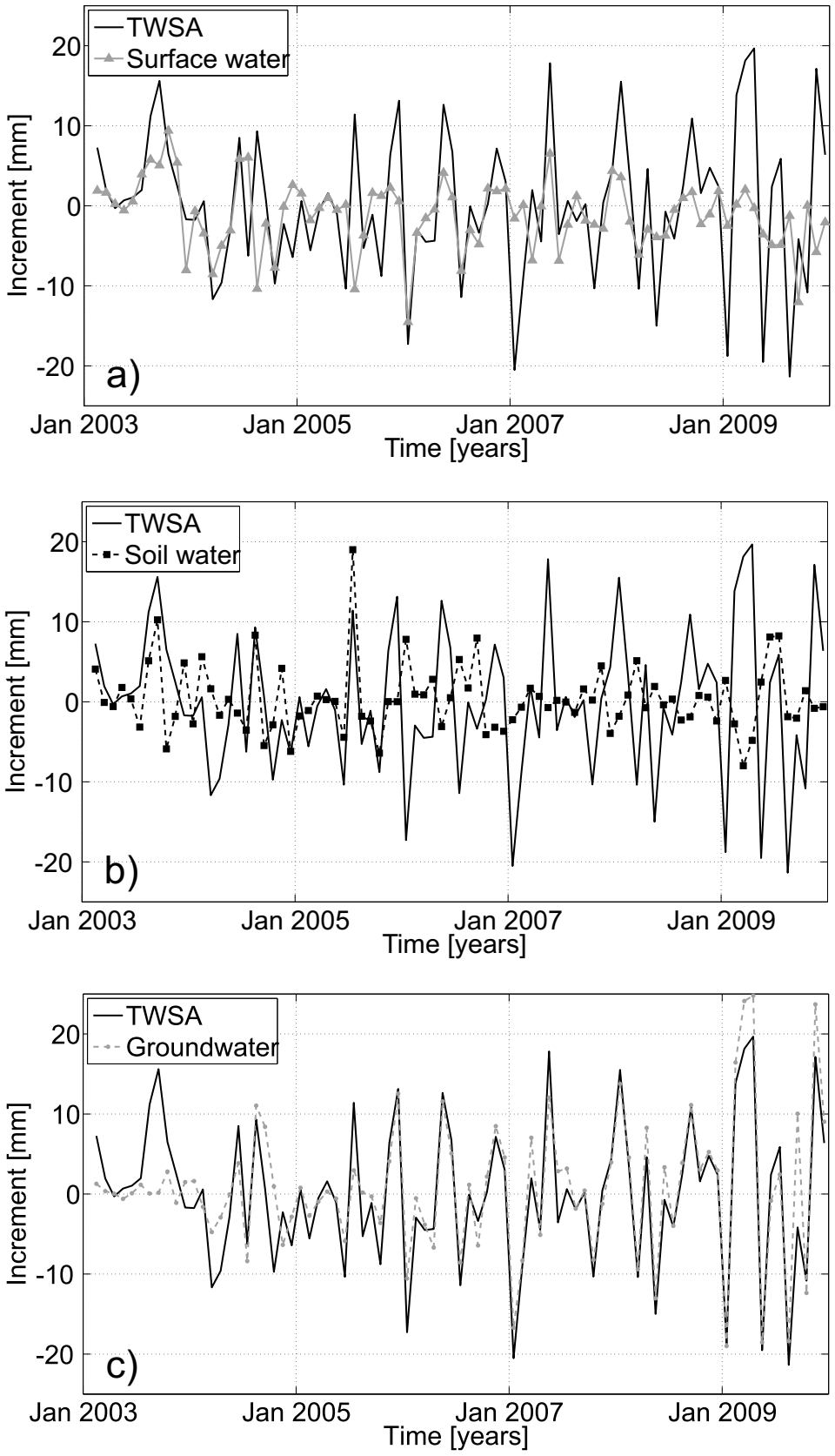

Figure 8.12: Time series of monthly update increments (in mm) for TWSA and a) surface water storage, b) soil water storage, and c) groundwater storage averaged over the entire Murray-Darling River Basin. Here, ITSG-Grace2014 (DDK3) is assimilated into WGHM.

between 2005-2009 the TWSA update increment is predominantly associated with the groundwater storage (Fig. 8.12 c).

The spatial distribution of the update increments of the total and individual water storage compartments were presented in Schumacher et al. (2016c). The updates of soil water are higher in the east and south-east of the Murray-Darling River Basin and decrease in western direction. For groundwater, the same spatial pattern is visible but the amount of water mass associated with the groundwater compartment is considerably larger. Only small update increments were found for lakes, which seems to be reasonable, since only a few small surface water bodies are located in the Murray-Darling River Basin. However, large increments are found for the wetland and the river compartment alongside the Murray and Darling Rivers, which was not expected. Wetlands in the Murray-Darling River Basin are usually large in area, at least for Australian standards, but small in depth (up to about 0.5 m), and therefore water evaporates quickly.

The sum of the update increments over all time steps shows how much water is introduced to or subtracted from WGHM in total. This is a measure how strong the concept of mass conservation is violated by C/DA. In Fig. 8.13, the sum of increments averaged over the entire Murray-Darling River Basin and its four major sub-basins is shown for TWSA, as well as soil water, surface water and groundwater.

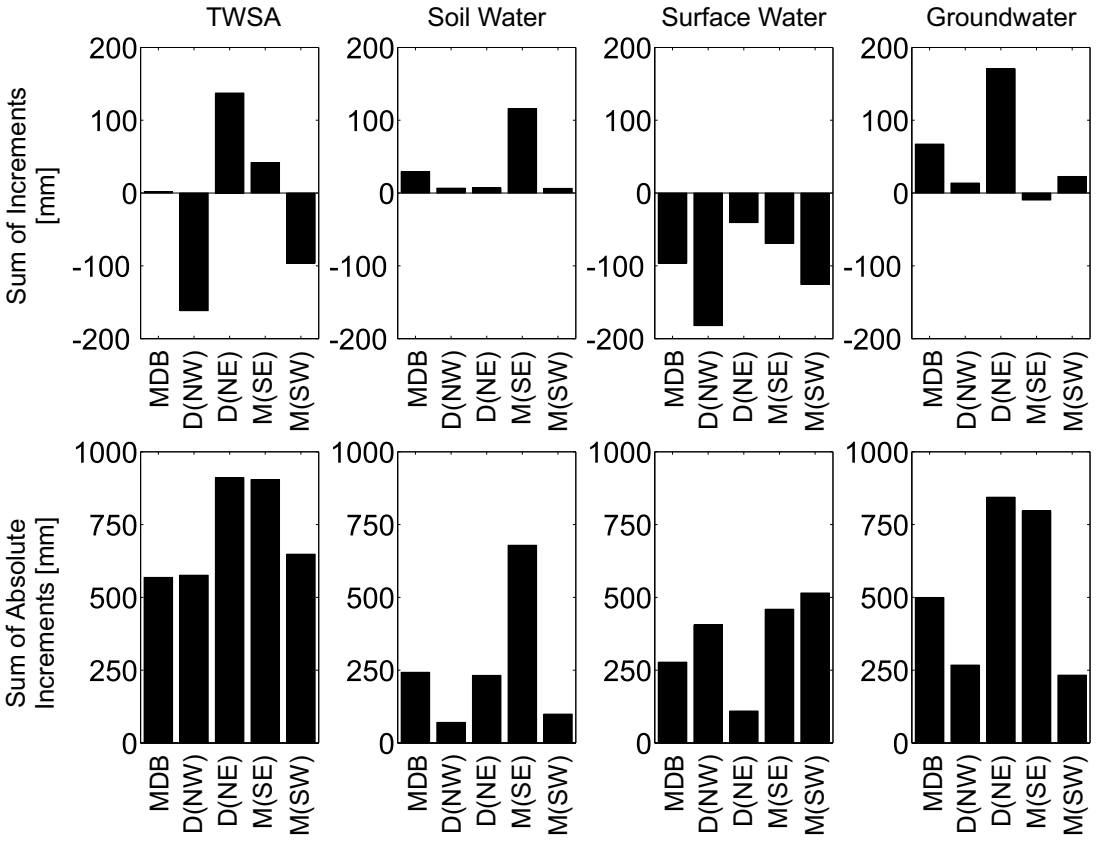

Figure 8.13: Sum of update increments (in mm) during 2003-2009 for total and for individual water storage compartments averaged over the entire Murray-Darling River Basin and over its major sub-basins in the Darling (D) and the Murray (M) basins (first row). Absolute values are used to calculate the sum (second row). Here, ITSG-Grace2014 (DDK3) is assimilated into WGHM.

During 2003-2009, the TWSA mass is more or less conserved over the entire Murray-Darling River Basin: Water is reduced in the north-western Darling and south-western Murray Basin, while water mass is added to the eastern parts of the basin, which are closer to the coast lines. In general, water is added to the soil water storage and reduced from the surface water storage all over the basin. Also, water is mostly added to the groundwater storage. The sum of the absolute water increment values shows that the sub-basins, which are located closer to the coast line, receive larger update increments for TWSA, as well as soil water and groundwater. The surface water storage is predominantly affected by C/DA in the western sub-basins. A spatial analysis of the update increment revealed that water is mainly added to or reduced from the Darling River and Murray River as well as the surrounding wetlands (see Fig. 13 in Schumacher et al., 2016c).

The update increments of the individual water compartments change their percentage of TWSA considerably. In case of the ensemble mean of OL, soil water has the highest contribution to TWSA during 2003-2009, which varies seasonally between 30 % and 75 % (Fig. 8.14 a, c). Surface water contributes with less than 5 %, while groundwater has a negative contribution to TWSA of up to -20 % between 2006-2009. The negative contribution of the groundwater changes to TWSA is caused by the model realization of the groundwater compartment that allows for negative storage values to simulate groundwater depletion (see section 2.2.2.3). After C/DA, the soil water storage has still a high contribution to TWSA in 2003 and 2004. Then, starting in 2006 the decline of TWSA is predominantly generated by the loss of groundwater. However, the interpretation of the compartment percentage of TWSA is not straightforward for the C/DA results, since large negative values exist for the groundwater storage and partially small positive values occur for TWS causing high percentage values (> 3000 %). Therefore, the values are not shown here.

In Fig. 8.14 b, it is clearly visible that the negative trend in GRACE TWSA is strongly associated with the groundwater storage compartment. Linear trends are calculated for the changes in soil water, surface water and groundwater averaged over the entire Murray-Darling River Basin and its four major sub-basins. The results show that the water decline in the updated TWSA, introduced into WGHM by assimilation of GRACE TWSA, mainly affects the groundwater compartment and is much less associated with the surface water and soil water storage compartments. This agrees with the findings in Leblanc et al. (2009), who did not discover any trend in surface water and soil moisture since 2003. In contrast, van Dijk et al. (2013) reported a decrease in public reservoirs between 2006-2007. In the following, the decline of groundwater storage changes is further investigated.

8.2.2.4 Groundwater Depletion

In this section, groundwater observations from wells are used to investigate the C/DA results for the groundwater storage compartment (see section 2.3 for details on the observations). The observations are given in terms of groundwater levels, which have to be converted into equivalent water heights by considering a factor known as the specific yield (Tregoning et al., 2012). However, the specific yield is usually unknown and cannot be measured. In the following, it is demonstrated that the choice of the specific yield values affects the validation of the C/DA results. First, an assumption of 0.1 as a typical value for water aquifers for the specific yield in each $1° \times 1°$ grid cell is used, which was proposed

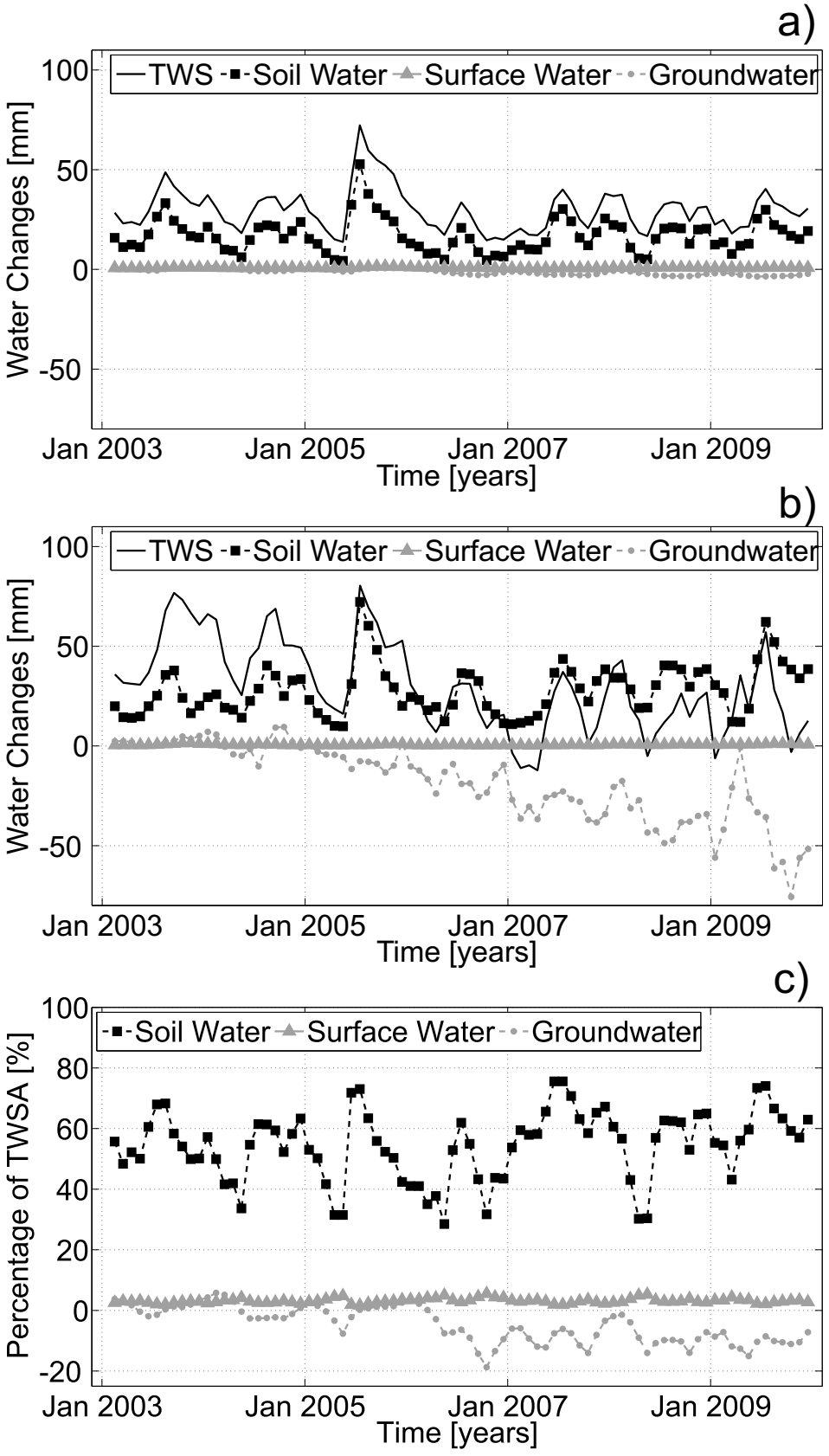

Figure 8.14: Time series of total and individual water storage changes a) of the OL and b) after C/DA. Here, ITSG-Grace2014 (DDK3) is assimilated into WGHM. In c), the percentages of TWSA for the individual water storage compartments are shown for the OL run.

by Tregoning et al. (2012). In addition, specific yield values ranging from 0.06 to 0.30 are applied based on geology maps (personal communication with Dr. Russell Crosbie, CSIRO Land and Water, Australia). Finally, those groundwater time series were identified that exhibit the highest RMSE compared to the (sub-)basin averaged time series. These are neglected for calculation of the (sub-)basin averages to investigate the influence of spatial averaging on the validation results.

The time series of the obtained three data sets are shown in Fig. 8.15. These indicate that the annual minimum and maximum values strongly depend on the chosen specific yield values and on the considered grid cells to calculate the basin averages. The seasonal variability might be overestimated due to local effects such as pumping and recharge from irrigation close to the considered observation wells (Tregoning et al., 2012).

The correlation coefficients of the OL and C/DA time series with respect to the groundwater observation time series, while considering the specific yield map, are shown in Tab. 8.8. The C/DA variants outperform the OL simulation of groundwater in all (sub-)basins, ex-

Table 8.8: Agreement between model predicted and observed groundwater in terms of correlation during 2003-2009 averaged over the entire Murray-Darling River Basin (MDB) and over its four major sub-basins. The values in brackets indicate the improvement of correlation compared to the OL run.

Basin	OL	ITSG-300km	ITSG-500km	ITSG-DDK3
MDB	0.53	0.71 (+0.17)	0.63 (+0.10)	0.66 (+0.13)
Darling (NW)	-0.01	0.63 (+0.64)	0.72 (+0.73)	0.74 (+0.75)
Darling (NE)	0.32	0.44 (+0.12)	0.31 (-0.01)	0.16 (-0.17)
Murray (SE)	0.01	0.37 (+0.36)	0.30 (+0.28)	0.36 (+0.34)
Murray (SW)	-0.05	0.60 (+0.66)	0.66 (+0.72)	0.77 (+0.82)
		GFZ-DDK3	JPL-DDK3	ITSG-SEIK
MDB		0.69 (+0.16)	0.42 (-0.12)	0.52 (-0.01)
Darling (NW)		0.77 (+0.78)	0.56 (+0.56)	0.36 (+0.37)
Darling (NE)		0.15 (-0.17)	-0.12 (-0.44)	0.23 (-0.09)
Murray (SE)		0.38 (+0.37)	0.17 (+0.16)	-0.12 (-0.13)
Murray (SW)		0.74 (+0.79)	0.60 (+0.66)	0.60 (+0.65)

cept for the entire Murray-Darling River Basin and two sub-basins in case of ITSG-SEIK. The improvements range from 0.10 to 0.17 for the average over the entire region and from 0.66 to 0.82 in the south-western Murray Basin. This clearly confirms the positive effect of GRACE data assimilation on the seasonal representation of the groundwater compartment in the Murray-Darling River Basin. In Fig. 8.16, the normalized time series of groundwater are shown, i.e. the temporal mean is reduced and the simulated and observed values are divided by their root mean square (RMS) of the time series. This allows a better interpretation of the seasonal cycles of groundwater changes.The OL simulation exhibits a phase shift compared to the observed groundwater changes, especially over the Murray

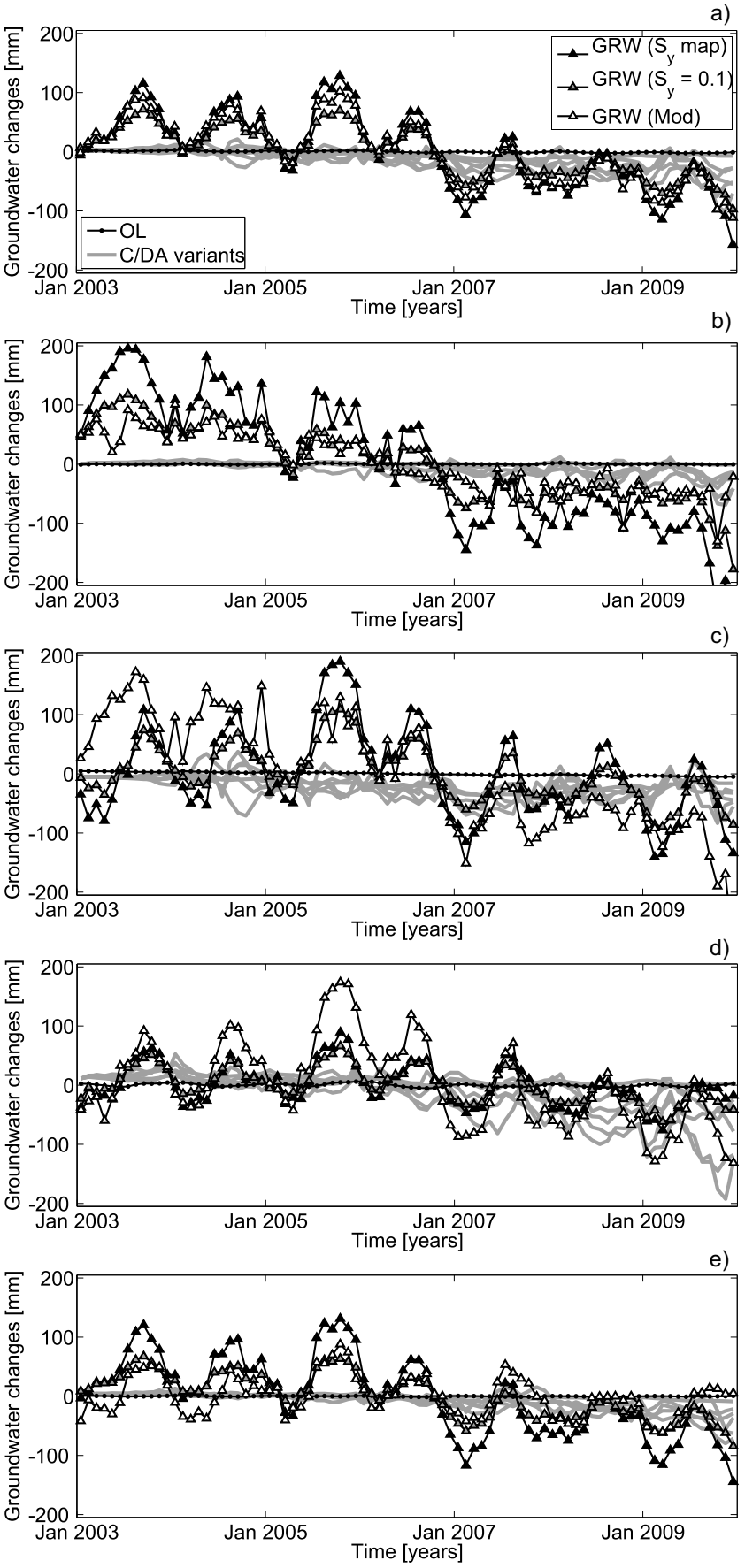

Figure 8.15: Time series of groundwater changes (in mm) from open loop (OL) simulations from WGHM and ensemble filter updates averaged over a) the entire Murray-Darling River Basin, b) the Darling (NW), c) the Darling (NE), d) the Murray (SE), and e) the Murray (SW).

sub-basins. This leads to their small correlation coefficients. After C/DA, the phase shift is reduced over all regions except for the north-eastern Darling Basin. The improvements occur mainly between 2006-2009, which are reflected in the higher correlation coefficients. However, the inter-annual variability between 2003-2005 seems to be underestimated in all regions (see also Fig. 8.15).

The estimation of linear trends in groundwater from the OL variant of WGHM indicates no or only a small decline with less than 1.5 mm/year in the four sub-basins (Tab. 8.9).

Table 8.9: Linear trends (in mm/year) in groundwater changes during 2003-2009 computed for the entire Murray-Darling River Basin (MDB) and the four sub-basins in the Darling (D) and Murray (M) basins. The linear trends estimated from groundwater (GRW) well measurements are provided assuming a specific yield map (column 2), a specific yield value of 0.1 (column 3), and modifying the sub-basin averages (column 4). The results of WGHM OL and after C/DA are shown in columns 5-11. The names of the C/DA variants are provided in Tab. 8.3.

Basin	GRW S_y map	GRW S_y=0.1	GRW Mod	OL	ITSG-300km	ITSG-500km	ITSG-DDK3	GFZ-DDK3	JPL-DDK3	ITSG-SEIK
MDB	-23.8	-16.1	-21.6	-0.6	-4.1	-5.4	-8.3	-9.9	-5.1	-1.7
D (NW)	-49.5	-28.7	-21.7	0.1	-2.1	-3.0	-3.6	-5.6	-4.3	-0.1
D (NE)	-14.4	-12.6	-40.4	-1.4	-3.6	-4.9	-6.4	-6.9	-2.3	-3.1
M (SE)	-7.4	-8.4	-19.1	-0.4	-7.5	-9.5	-19.2	-20.7	-9.0	-1.3
M (SW)	-24.1	-14.9	-5.1	-0.1	-3.9	-4.7	-5.8	-8.9	-6.9	-1.4

Depending on the preparation of the observations, i.e. the choice of specific yield values and the choice of $1^\circ \times 1^\circ$ grid cells contributing to the (sub-)basin averages, the water storage decreased about -11.6 ± 6.5 mm/year in the south-eastern Murray Basin up to -33.3 ± 14.5 mm/year in the north-western Darling Basin (Tab. 8.10). In contrast, GRACE provides

Table 8.10: The average of linear trends in groundwater during 2003-2009 in mm/year and standard deviations are shown for the entire Murray-Darling River Basin (MDB) and its four sub-basins: from different post-processing strategies (i.e. specific yield and spatial averaging) for the groundwater well measurements (column 2), after applying different filtering techniques for GRACE (column 3), and for different GRACE products (column 4).

Basin	Groundwater Observations	GRACE Filtering	GRACE Product
MDB	-20.5 ± 4.0	-5.9 ± 2.2	-7.7 ± 2.4
Darling (NW)	-33.3 ± 14.5	-2.9 ± 0.8	-4.5 ± 1.0
Darling (NE)	-22.5 ± 15.6	-5.0 ± 1.4	-5.2 ± 2.5
Murray (SE)	-11.6 ± 6.5	-12.1 ± 6.3	-16.3 ± 6.3
Murray (SW)	-14.7 ± 9.5	-4.8 ± 1.0	-7.2 ± 1.6

the largest TWSA decline in the south-eastern and south-western Murray Basins, while the north-eastern and north-western Darling Basins show a smaller decline in TWSA. The large differences in the linear trend of the measured groundwater imply large uncertainties

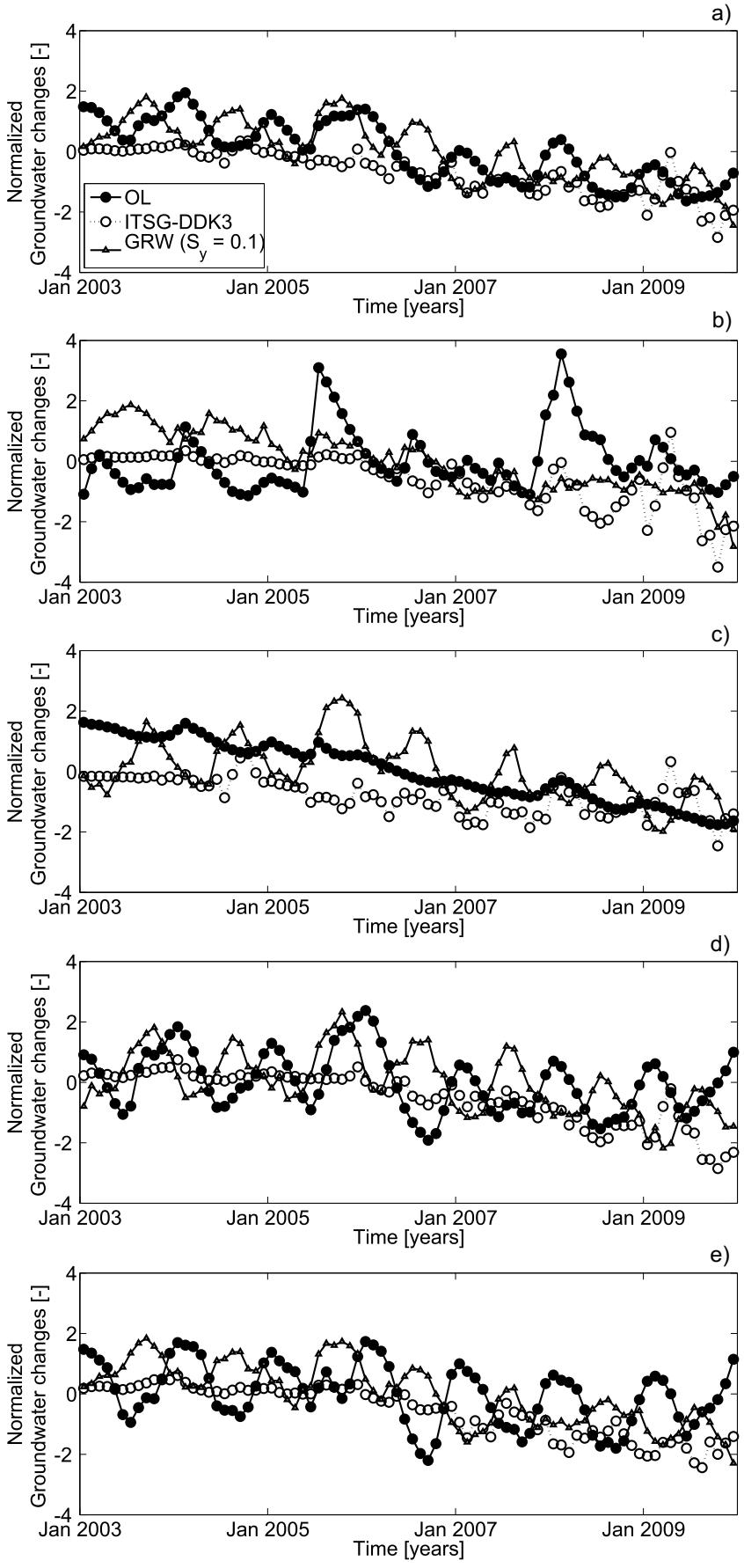

Figure 8.16: Normalized time series of groundwater changes (in mm) from open loop (OL) simulations from WGHM and EnKF update ITSG-DDK3 averaged over a) the entire Murray-Darling River Basin, b) the Darling (NW), c) the Darling (NE), d) the Murray (SE), and e) the Murray (SW).

due to the application of the specific yield and due to the interpolation of point measurements to create (sub-)basin averages. The C/DA variants of WGHM suggest that the strongest negative trend occurs in the south-eastern Murray Basin. This is caused by the stronger decline in the south observed by GRACE. The ranking of the basins with respect to the decline of groundwater is still different after C/DA compared to the groundwater measurements. Depending on the GRACE post-processing variant (i.e. filtering), the trend is estimated to be about -12.1 ± 6.3 mm/year, and depending on the GRACE product the trend is about -16.3 ± 6.3 mm/year. The effect of the chosen GRACE product and the choice of the filtering are therefore considerably smaller than the a priori choices of the specific yield and the spatial averaging for the groundwater observations.

8.2.2.5 Calibration of Model Parameters

In total 22 parameters are calibrated during the C/DA procedure. In Fig. 8.17, the time series of six parameters after C/DA using the ITSG-DDK3 product are shown: Three parameters were found to be most sensitive in Schumacher et al. (2016a) and are related to the soil water compartment and to the calculation of evapotranspiration (Fig. 8.17 a, b, c); The other three parameters are related to the surface water and groundwater compartments (Fig. 8.17 d, e, f).

The root depth multiplier is a factor to regulate the depth of the soil compartment, and it is equal to one in the standard WGHM run. After one year, the ensemble mean of the calibrated values is larger than one (Fig. 8.17 a). Therefore, the model allows a higher amount of water to be stored in the soil compartment before it is saturated. By this, the calibrated root depth multiplier likely contributes to a larger variability in soil water changes after C/DA, and thus leads to a better agreement with GRACE TWSA compared to the OL. However, the calibrated parameter neither converges to a constant value nor is its uncertainty reduced.

The net radiation multiplier and the Priestley-Taylor (PT) coefficient for arid regions are used for the computation of potential evapotranspiration. The net radiation multiplier converges to a value close to the lower limit of 0.5, and its uncertainty clearly becomes smaller except for one sample (Fig. 8.17 b). In contrast, the PT coefficient does not show a clear pattern over the C/DA period and its uncertainty remains large (Fig. 8.17 c).

The net abstraction groundwater multiplier controls the amount of water that is consumed and is set to one in the standard WGHM run. Until the end of 2007, the calibrated value is mostly larger than one suggesting a higher anthropogenic water use (Fig. 8.17 d). It seems likely that during the drought period more water is consumed than originally simulated by the WaterGAP sub-models. However, this would also be expected for the years 2008-2009, in which the value is smaller than one. The uncertainty of the calibrated parameter also does not decrease over time.

The wetland depth decreases with increasing the number of EnKF updates (Fig. 8.17 e). This seems to be reasonable, since the wetlands are not very deep within the Murray-Darling River Basin. Finally, for the river roughness coefficient multiplier, which is closely related to the river velocity, no clear pattern is visible over the study period (Fig. 8.17

f). The estimated parameter values in the EnKF updates generally do not converge to a constant value and their uncertainties do not become smaller with time.

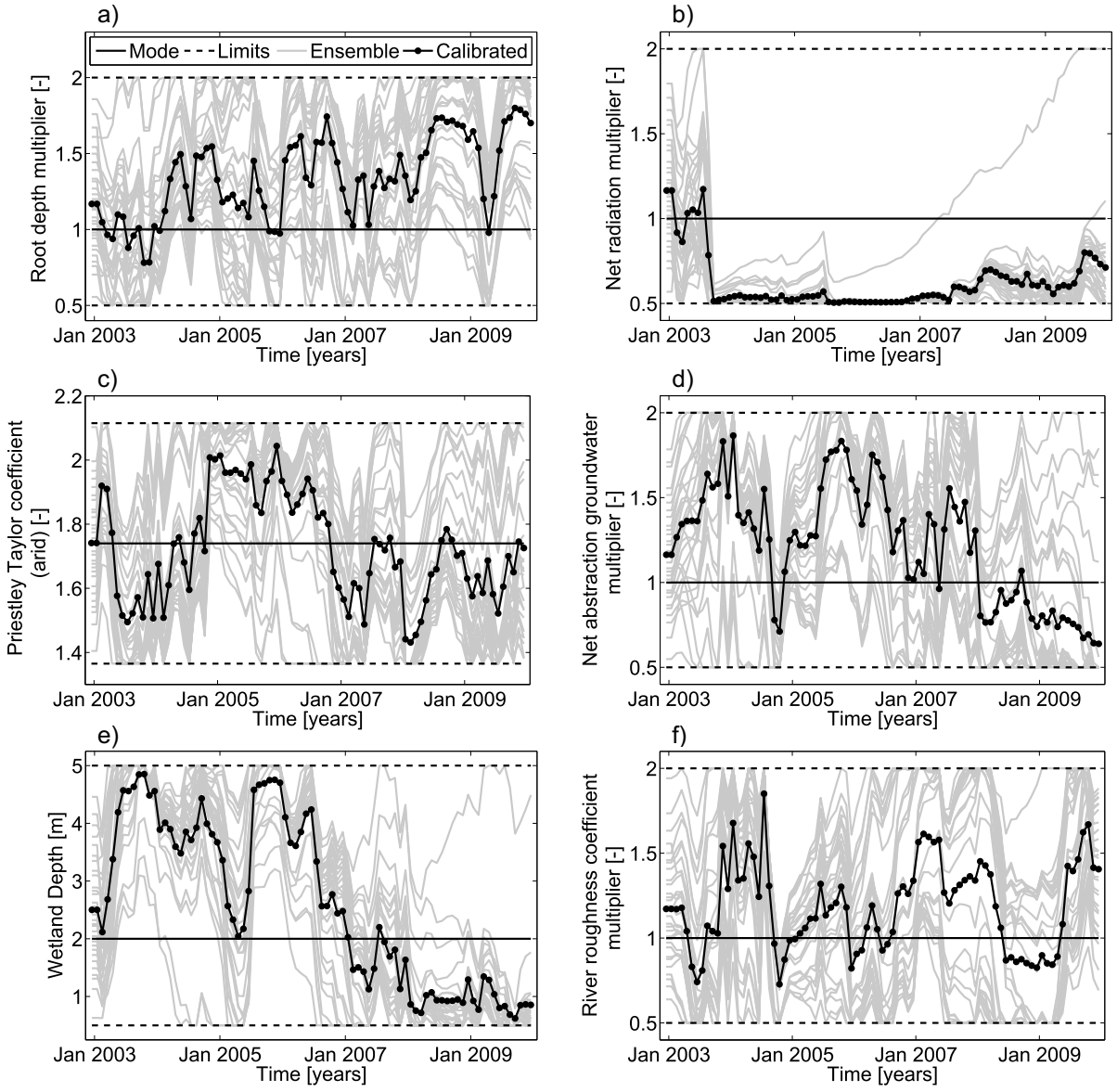

Figure 8.17: Time series of six selected calibration parameters (ensemble mean) and ensemble members. The initial parameter ensemble is shown for month "Dec 2002".

8.2.3 Discussion and Conclusions for the Murray-Darling Case Study

In this section, the developed C/DA framework of this PhD is successfully transferred to the Murray-Darling River Basin. It is shown that the assimilation of GRACE TWSA products into WGHM over 2003-2009 considerably improves the representation of TWSA. Decline in TWSA, resulting from the Millennium Drought over the region, is successfully introduced or intensified by C/DA using GRACE data. The trends are found to be predominantly associated with the changes in groundwater storage, which is confirmed by

validation with in-situ groundwater observations. However, the ranking of the trends in the four sub-basins coincides with the distribution of linear trends in GRACE TWSA but it is somewhat different from the ranking suggested by groundwater observations. In addition, C/DA improves the representation of the seasonality of the groundwater compartment in terms of phase shift, which is assessed via correlation coefficients. The inter-annual variability of groundwater changes has a better quality after 2006 but is still underestimated during 2003-2005. In summary, hypothesis VI of this thesis is supported. However, conclusions on the representation of surface and soil water cannot be made here, since the validation with independent measurements has to be done in future work. It is worth mentioning that the representation of river discharge deteriorated after C/DA (Schumacher et al., 2016c). Therefore, a multi-criteria C/DA is desirable, e.g. including GRACE TWSA and river discharge measurements to simultaneously constrain water storage changes and water fluxes.

The analysis of the model parameters shows that large dependencies exist even after seven years of C/DA. Obviously, several of the 22 calibrated parameters seem to compensate each other due to their dependencies, e.g., the net radiation multiplier and the PT coefficient that both influence the calculation of evapotranspiration. However, the calibrated set of parameters contributes to a good representation of TWSA and groundwater as shown in sections 8.2.2.2 and 8.2.2.4. Therefore, the presented interpretations of the parameters in Fig. 8.17 should be considered with caution. Based on these investigations, a final answer cannot be given whether individual parameter values have improved. Calibrating a parameter sub-set might result in a stronger constraint of GRACE TWSA on these parameters, and therefore might lead to improvements of individual values. Furthermore, the results suggest to carefully investigate the assumptions about the parameter distributions, which are used to generate a parameter ensemble. For example, to create an ensemble of WGHM runs, samples of wetland depth are realized using a triangular distribution with the most probable value of 2 m and minimum and maximum limits of 0.5 and 5 m. In the Murray-Darling River Basin, wetlands are usually of small depth. Therefore, several WGHM runs might overestimate the amount of water that is stored in the wetland compartment. In the C/DA procedure, this leads to rather high uncertainties for the wetland depth parameter as well as for the wetland water storage compartment. As a result, the update increments are likely overestimated. An improved representation of surface water might thereby not be achieved after C/DA.

A comparison of the C/DA results, while assimilating different GRACE products and applying different filtering strategies, revealed that the differences in the estimated trends are of the same magnitude as the differences in the estimated trends from the GRACE products. However, a comparison of TWSA from the C/DA results to GRACE or the validation with independent groundwater measurements could not identify a particular GRACE product or a particular filtering approach that always produced superior results. Therefore, a recommendation on the "best" product or filtering approach cannot be given based on the presented investigations.

The application of the SEIK filter indicates a smaller effect of GRACE on the WGHM water states. The filter update appears to trust the model simulation more than it is the case for the application of the EnKF. This is related to smaller uncertainties for the model prediction in case of the SEIK filter, although a forgetting factor was introduced

to enlarge the model ensemble spread for each update step similar to the inflation factor in the EnKF. Inflation and forgetting factor do not seem to have identical impacts on the model covariance matrices.

The validation with groundwater observations suggested that the inconsistencies between C/DA variants of WGHM and measurements are not only due to model uncertainties, uncertainties in GRACE products or due to the disaggregation of TWSA into individual water storage compartments during the C/DA update step, but also associated with uncertainties in the groundwater measurements and their post-processing. However, the validation confirmed that the improvement of groundwater decline and its seasonality is possible when assimilating GRACE TWSA into WGHM. This outcome provides a first step towards transferring the developed C/DA framework also to assimilate TWSA into hydrological models over data-sparse regions, where a validation of the C/DA results is limited.

A comparison of the Murray-Darling River Basin case study to the case study for the Mississippi River Basin, in which one year of GRACE TWSA were assimilated into WGHM, shows that the developed C/DA framework can be applied in different river basins. In contrast to the Mississippi study, a constant inflation factor was introduced to avoid fast ensemble convergence here. The application of C/DA over a period of seven years, revealed considerable improvements with respect to the standard WGHM version. The outcomes indicate that a long period of C/DA helps clearly to improve water state simulations. Therefore, in further studies, a longer period of C/DA for the Mississippi River Basin might be considered.

9. Conclusions and Outlook

This PhD study aimed at developing a flexible calibration and data assimilation (C/DA) framework to merge model predicted water states with GRACE total water storage anomalies (TWSA), while simultaneously calibrating model parameters.

The main results of this thesis can briefly be summarized as:

- GRACE TWSA are beneficial for improving hydrological model simulations (objective 1). This is always the case for the representation of TWSA. Yet, a general statement for the individual water compartments cannot be formulated. A better quality of the simulation of water states in individual storage compartments depends on the C/DA analysis strategy, quality of observations, and the characteristics of the basin.

- Several configurations of the C/DA strategy were applied and the C/DA results indicated a significant influence of the observation error model and the spatial resolution of GRACE TWSA (objective 2).

- Application of alternative ensemble filter approaches helped to improve the C/DA performance (objective 3).

- The transferability of the developed C/DA framework was demonstrated by applying the approach to two different case studies, the Mississippi River Basin, USA, and the Murray-Darling River Basin, Australia, that are distinguishable in terms of hydrological and climatic characteristics (objective 4).

A strategy to investigate the C/DA framework and its benefit for hydrological model simulations was proposed in chapter 1. After introducing GRACE TWSA products and providing a detailed description of the WaterGAP Global Hydrology Model (WGHM) in chapter 2, the principle of variational and sequential data assimilation using linear or non-linear models was discussed in chapter 3. Ensemble filter algorithms were introduced in chapter 4. The similarities and differences between the classical ensemble Kalman filter (EnKF) and the alternative filter formulations, i.e. the square root analysis scheme (SQRA) and the singular evolutive interpolated Kalman (SEIK) filter, were addressed. Filter tuning techniques were also presented. In chapter 5, the implementation of GRACE data assimilation into WGHM and the model parameter calibration was described, including strategies to handle the temporal and spatial resolution mismatch between observations and model, as well as the description of measurement and model uncertainties. A sensitivity and covariance analysis were presented in chapter 6 to investigate whether a combined C/DA approach is appropriate to improve WGHM simulations. In chapters 7 and 8, the application of the established C/DA framework was investigated via a simulation experiment and two real case studies. In this chapter, the main conclusions are drawn from this thesis and an outlook on future research is given.

9.1 Conclusions

Traditionally, hydrological models have been calibrated against river discharge measurements (Gupta et al., 1998). A parameter that regulates the amount of land runoff has already been calibrated in the case of WGHM against mean annual river discharge at 1319 gauge stations (Müller Schmied et al., 2014). In Werth (2010), a multi-criteria calibration against river discharge and GRACE TWSA has been performed. In addition, a number of studies have suggested to assimilate GRACE TWSA into hydrological models (Zaitchik et al., 2008, Su et al., 2010, Forman et al., 2012, Houborg et al., 2012, Li et al., 2012, Forman and Reichle, 2013, van Dijk et al., 2014, Tangdamrongsub et al., 2015, Reager et al., 2015, Girotto et al., 2016, Kumar et al., 2016). In this thesis, for the first time, a combined C/DA approach was proposed to assimilate remotely-sensed TWSA of GRACE into a hydrological model, while simultaneously calibrating model parameters (see also Schumacher, 2012, Eicker et al., 2014, Schumacher et al., 2016a,b,c). The mathematical aspects of C/DA and their implementations were described in detail using a simple model and the more complex global WGHM.

Benefit of GRACE TWSA for Model Parameter Calibration

A fundamental question regarding the assimilation of GRACE TWSA into hydrological models was "whether it has a benefit for simultaneously calibrating model parameters". Therefore, a sensitivity and covariance analysis was performed to test the sensitivity of WGHM water storage variability with respect to changes in the model parameter values and to assess the correlations between water states and parameters (see chapter 6 and also Schumacher et al., 2016a). The investigations indicated that high correlations between water states and several model parameters exist, which enable a calibration of parameters against GRACE TWSA. This supports hypothesis I of this thesis. Numerical simulations confirmed that indeed the large correlations between water states and model parameters are responsible for the observed big contribution of GRACE TWSA in the ensemble filer update. Therefore, a combined C/DA approach seems to be beneficial for improving hydrological model simulations. In addition, basin-specific calibration of model parameters appears to be more appropriate than a global calibration, since each river basin exhibits different climatic and anthropogenic characteristics, and therefore different parameters were found to be sensitive to TWSA. Moreover, it was found that calibrating a large number of parameters, e.g., 20-22 parameters, might not always be beneficial and the improvement of some parameters might be compensated by others. This might lead to changes in the parameter values without any improvements of the water compartment simulation. While the calibrated parameters as a set contributed to a good representation of TWSA as well as to several individual water storage compartments in the presented simulation and real case studies, the results indicated that individual parameters have not always been improved after C/DA. Therefore, it is concluded that if dependencies between model parameters exist, it seems more appropriate to calibrate only one of these parameters.

Benefit of GRACE TWSA for Improving Hydrological Water States and Fluxes

To answer the research question "whether the assimilation of GRACE TWSA into hydrological models and simultaneous parameter calibration improves the representation of total and individual water storage changes", the established C/DA framework was tested within a simulation experiment designed for the Mississippi River Basin. Using this, one can evaluate the C/DA results by comparison to the realistically defined synthetic "truth" (see chapter 7 and also Schumacher et al., 2016b). Evaluation of the metrics RMSE and of the correlation coefficients of residual curves (i.e. after subtracting trend, annual and semi-annual cycles) showed that model simulations of TWSA were always improved after C/DA as expected, which confirms hypothesis II a of this thesis. However, this did not necessarily hold for the individual water compartments and water fluxes. Especially, the vertical disaggregation into soil water and groundwater was not always performed successfully. A possible reason might be the similarity of the inter-annual cycles of soil water and groundwater that complicates the separation of GRACE TWSA signal. In addition, residual issues in the C/DA update might exist, such as wrong correlations introduced by the finite approximation of error statistics with a limited model ensemble size. The investigations show the potential of GRACE TWSA for improving the simulation of individual water states and fluxes but do not absolutely confirm hypothesis II b of this thesis.

Observed GRACE data were assimilated into WGHM for one year for the Mississippi River Basin (section 8.1). Then, the model was run forward for the next three years, which indicated that an improved prediction of TWSA is possible after the assimilation (see also Eicker et al., 2014). During the data assimilation period, integrating GRACE TWSA was found to have a positive impact on several individual water compartments and on the river discharge but not on all water states and fluxes. During the three following years without additional GRACE data assimilation, the performance of individual water compartment simulations rather degraded compared to the standard model simulation. The length of the period of C/DA influences the performance for individual water compartments and fluxes and therefore should be chosen according to the availability of observation data, as it was followed in this thesis when transferring the C/DA framework to the Murray-Darling River Basin (see section 8.2).

Impact of TWSA Discretization and GRACE Error Model Settings on C/DA Results

The question "whether and how accounting for GRACE TWSA error correlations and exploiting the spatial resolution of GRACE observations influence the C/DA results" was investigated in this thesis. In the theoretical part of this work, it was described how GRACE TWSA can be spatially averaged. The impact of spatial discretization was subsequently examined when assimilating observed GRACE TWSA into WGHM for the Mississippi River Basin. It was shown that the assimilation of $5° \times 5°$ gridded data or subbasin averages was found to be superior compared to a basin average assimilation (see section 8.1). This supports hypothesis III of this thesis, which states that exploring the spatial resolution of GRACE TWSA leads to better C/DA results.

It the theoretical part of this thesis it was described how correlated errors of GRACE TWSA can be handled (section 2.1.4). The formal error propagation of the full error

covariance matrix of potential coefficients to a full error covariance matrix of TWSA was demonstrated. In the application part of this study, a simulation experiment was designed to investigate the effect of introducing either white or correlated observation noise in the ensemble filter update step (see chapter 7 and also Schumacher et al., 2016b). The results indicated that the choice of the observation error model has a considerable influence on the C/DA results with respect to the three selected spatial scales (four and eleven sub-basin averages, and sixteen grid cells within the Mississippi River Basin). This supports hypothesis IV of this PhD study, which states that the error model of observations matters significantly. However, applying GRACE TWSA error correlations did not always lead to a better representation of total and compartmental water states. Assuming white noise for GRACE TWSA results in a higher Kalman gain in the update step, and therefore the model states are usually pulled closer towards the observations. Nevertheless, this did not always result in superior metrics of total and compartmental water storage (see sections 7.2.1-7.2.3). Since an accurate representation of observation errors is desirable to estimate a realistic Kalman gain, investigations regarding the effect on the C/DA results are important and it is recommended to implement the GRACE TWSA error correlations. Yet, a final answer cannot be provided whether accounting for GRACE TWSA error correlations helps improving the performance of the C/DA framework.

Impact of Alternative Filter Algorithms on C/DA Results

The C/DA framework based on the classical EnKF was extended by variants of ensemble filter algorithms to answer the research question "whether alternative algorithms help to improve the performance of the ensemble filter update". The SQRA and SEIK methods were introduced as alternative filter algorithms in the theory part of this work. Their similarities and differences with respect to the classical EnKF were discussed (chapter 4). While alternative algorithms have already been used in meteorological and oceanographic data assimilation studies (e.g., Nerger, 2003), in this thesis, these filters were introduced for the first time to GRACE data assimilation. The application of SQRA and SEIK was tested in a simulation experiment and their C/DA performance for total and individual water storage was compared to the classical EnKF (see chapter 7 and also Schumacher et al., 2016b). Especially after applying the SEIK filter, improvements were found compared to the classical EnKF. These are likely attributed to smaller sampling errors in the filter update step: The SEIK filter does not generate an ensemble of observation perturbations, which is required in the EnKF. The SEIK filter, when using 30 ensemble members, was found to show a similar performance as the EnKF with 100 ensemble members. These results indicate that the application of alternative filters can also help to keep the ensemble size small. Eventually, it was shown that alternative filter algorithms improve the C/DA performance, and thus, hypothesis V is supported.

Transferability of the C/DA Framework

Finally, the research question "whether the C/DA framework is transferable to river basins with different climatic and hydrological conditions" was investigated in this thesis. After testing the established C/DA framework using a synthetic experiment and

observed GRACE TWSA assimilation for the Mississippi River Basin, the framework was successfully applied to the Murray-Darling River Basin, which exhibits different climatic and anthropogenic characteristics compared to the Mississippi River Basin (see section 8.2 and also Schumacher et al., 2016c). Identical initial settings were used to test whether the framework is in principle transferable to any region world-wide without tuning. For the Murray-Darling River Basin study, the focus was on the drought period during 2003-2009, in which WGHM did not show significant TWSA trends. The C/DA results revealed considerable improvements of TWSA and groundwater after assimilating GRACE products supporting hypothesis VI of this thesis. Evaluating linear trends in TWSA indicated that GRACE data assimilation indeed introduced or intensified negative trends, which were missing before. The validation of the groundwater simulations with independent groundwater measurements showed that the trend was correctly associated with changes in the groundwater storage. The ranking of the linear trends in the four sub-basins coincided with the spatial distribution of linear trends in GRACE TWSA but differed somewhat from the distribution suggested by groundwater well measurements. Seasonal variations did not agree with in-situ observations even though C/DA achieved to modify the seasonal variations. In addition, C/DA led to an overestimation of river discharge during high-flow periods. As also found in the simulation and by the real case studies for the Mississippi River Basin, the potential of C/DA in vertically disaggregating and downscaling TWSA, and in adapting model parameters has to be better exploited in further work. Generally, the applications presented in chapter 8 provide a first step towards transferring the established C/DA framework to data-sparse regions, where validation using independent measurements is limited.

Independent Validation of C/DA Results

To assess the C/DA results for individual water states and fluxes, a combined use of different observations and model simulations is important. This is, however, difficult due to limited ground-based observations, including sparse networks and data gaps, as well as spatial and temporal coverage and resolution mismatch between ground- or satellite-based measurements and model outputs. Furthermore, uncertainties exist for both model simulations and observations. The cause of the mismatch between model outputs and observations might therefore not always be identified.

9.2 Outlook

Improving the Model Parameter Calibration

The investigations on the calibrated model parameters in sections 7.2.4, 8.1.2.1, and 8.2.2.5 showed that some parameters likely compensate each other, and thereby a simultaneous calibration is not possible. Therefore, studies on the optimal number and selection of calibration parameters will be conducted in future research. The intention is to avoid dependencies of calibration parameters and therefore to improve the calibration of individual model parameter values in the ensemble filter update. It will also be investigated to what extent a desirable parameter convergence and the application of

inflation factors to ensure a contribution of observations during the entire C/DA phase contradict each other. In addition, it will be analyzed how a more efficient sampling of the parameter space might contribute in avoiding fast ensemble convergence (Pham, 2001, Evensen, 2004, Hendricks Franssen and Kinzelbach, 2008). Tests on an optimal length of C/DA will also be performed. For the future basin-specific calibration of smaller parameter sets, adapting the initial model parameter distributions will be tested to address the individual climatic and anthropogenic conditions in the different river basins. Iterative C/DA will also be investigated. The calibrated parameters after a first C/DA phase will be used to start the C/DA procedure once again and to iteratively improve the parameter values and water state estimates. For this, the calibration and the data assimilation might be separated and performed sequentially or a combined C/DA might be used.

Constraining Individual Water States and Fluxes

Since the real case studies in section 8.1 and 8.2 showed that assimilation of GRACE TWSA does not automatically improve individual water compartments, a multi-criteria data assimilation approach that considers further data sets in the C/DA update is desirable. These extra observations may include lake and river levels from altimetry observations, in-situ river discharge, in-situ groundwater measurements or remotely-sensed soil moisture (e.g., from the Soil Moisture Ocean Salinity (SMOS) satellite mission).

Improving the Disaggregation of GRACE TWSA

To improve the performance of the ensemble filter updates with respect to the vertical and horizontal disaggregation of GRACE TWSA into individual water storage compartments, a range of tuning techniques will be analyzed. For a more stable computation of the model error covariance matrix, the model ensemble size will be reviewed, the application of covariance localization will be implemented to suppress unphysical correlations, e.g. between large-distant points, and the sampling of model uncertainties will be revised. The latter will help to realize a more realistic, i.e. physics-based description of hydrological model uncertainties. For this, adapting distributions for model parameters, using spatially and temporally correlated error models for climate forcing data, in particular for precipitation, and applying the minimum second order exact sampling technique (see section 4.2.3) for initial water states will be analyzed in future work. On the other hand, the adequate handling of GRACE TWSA observations should be further developed. This comprises a more sophisticated method to distribute the monthly updated water states to daily values, or alternatively, assimilating weekly or daily GRACE data products into the model. Alternative to formal error propagation, triple collocation error estimation will be examined. Moreover, model improvements may also be obstructed by considering errors of the background models for other Earth system components that have already been reduced from GRACE products (see e.g., Forootan et al., 2014a). Such errors are worth to be investigated, since they represent additional uncertainty sources of TWSA observations.

Improving the Data Assimilation Methodology

The benefit of extending the C/DA framework based on the classical EnKF by alternative filtering methods such as the SQRA and SEIK approaches was demonstrated in this

thesis. For future work, the EnKS will be applied and methods that perform a full Bayesian update, e.g., the Particle Filter (Del Moral, 1997) or a combination of particle and ensemble Kalman filter (Hoteit et al., 2008a,b, Stordal et al., 2011) will be tested for hydrological data assimilation. This provides a chance to include statistical moments of higher order in the ensemble update as well as probability distributions that are not Gaussian. One might also consider the introduction of a bias estimation, e.g., via random walk Markov processes, or constraint ensemble methods that ensure mass conservation, such as the Constraint EnKF (CEnKF) using the terrestrial water balance as additional constraint in the EnKF update step (Pan and Wood, 2006). A systematic investigation on the optimal inflation factor also has to be done.

Future Applications

A comparison of the C/DA results to other data assimilation strategies (e.g., van Dijk et al., 2014) or inversion techniques (e.g., Forootan et al., 2014b) might help to better interpret the updated model water states. Applications of the framework to analyze the impact of climatological conditions on the basin hydrology and to assess natural hazards such as droughts and floods will be considered in further work. Studies on the prediction skills of WGHM after C/DA are planned. A merged TWSA product might bridge the possible gap between the still operating GRACE satellite mission and its follow-on mission, whose launch is planned for the end of 2017. Moreover, establishing an (operational) daily GRACE TWSA data assimilation approach would contribute in monitoring and early warning systems for natural hazards, especially droughts and floods. Since the WaterGAP model has already been used, e.g., in the IPCC Report in 2013 for estimating hydrological impacts of climate change, eventually, the developed C/DA framework allows to provide a globally improved WGHM version delivering more reliable hydrology and climate predictions in future.

A. Acronyms

ANU	**A**ustralian **N**ational **U**niversity
APMG	**A**stronomical **P**hysical and **M**athematical **G**eodesy Group, University of Bonn
AWI	**A**lfred **W**egener **I**nstitute
BfG	**B**undesanstalt **f**ür **G**ewässerkunde
BLUE	**B**est **L**inear **U**nbiased **E**stimator
BoM	**B**ureau **o**f **M**eteorology
C/DA	**C**alibration / **D**ata **A**ssimilation
CLSM	**C**atchment **L**and **S**urface **M**odel
COFFEE	**C**omplementary **O**rthogonal subspace **F**ilter **F**or **E**fficient **E**nsembles
CRU TS 3.2	**T**ime **S**eries of the **C**limate **R**esearch **U**nit
CSIRO	**C**ommonwealth **S**cience and **I**ndustrial **R**esearch **O**rganisation
CSR	**C**entre for **S**pace **R**esearch
DAAD	**D**eutscher **A**kademischer **A**ustausch **D**ienst
DDM	**D**rainage **D**irection **M**ap
DFG	**D**eutsche **F**orschungs**g**emeinschaft
DLR	**D**eutsches Zentrum für **L**uft- und **R**aumfahrt
EAKF	**E**nsemble **A**djustment **K**alman **F**ilter
ECMWF	**E**uropean **C**entre for **M**edium-Range **W**eather **F**orecasts
EKF	**E**xtended **K**alman **F**ilter
EnKF	**En**semble **K**alman **F**ilter
EnKS	**E**nsemble **K**alman **S**moother
ESSE	**E**rror **S**ubspace **S**tatistical **E**stimation
ETKF	**E**nsemble **T**ransform **K**alman **F**ilter
GFZ	**G**eo**F**orschungs**Z**entrum
GIA	**G**lacial **I**sostatic **A**djustment
GIEMS	**G**lobal **I**nundation **E**xtent from **M**ulti-**S**atellite
GLDAS	**G**lobal **L**and **D**ata **A**ssimilation **S**ystem
GMAO	**G**lobal **M**odeling and **A**ssimilation **O**ffice
GPCC	**G**lobal **P**recipitation **C**limatology **C**entre
GPS	**G**lobal **P**ositioning **S**ystem
GRACE	**G**ravity **R**ecovery **A**nd **C**limate **E**xperiment
GRDC	**G**lobal **R**unoff **D**ata **C**entre
GRGS	**G**roupe de **R**echerche de **G**éodésie **S**patiale (**S**pace **G**eodesy **R**esearch **G**roup)
GSFC	**G**oddard **S**pace **F**light **C**enter
GWCR	**G**lobal **W**ater **C**ycle **R**eanalysis
GWSWUSE	**G**round**W**ater and **S**urface **W**ater **USE** model
HBV-96	**H**ydrologiska **B**yråns **V**attenbalansavdelning model
ICA	**I**ndependent **C**omponent **A**nalysis
IPCC	**I**ntergovernmental **P**anel on **C**limate **C**hange
IRD	**I**nstitut de **R**echerche pour le **D**éveloppement
Irstea	**I**nstitut national de **r**echerche en **s**ciences et **t**echnologies pour l'**e**nvironnement et l'**a**griculture
JPL	**J**et **P**ropulsion **L**aboratory

KBR	**K**-**B**and microwave **R**anging
KF	**K**alman **F**ilter
LEGOS	**L**aboratoire d'**E**tude en **G**éophysique et **O**céanographie **S**patiales
MCMC	**M**arkov **C**hain **M**onte **C**arlo
NASA	**N**ational **A**eronautics and **S**pace **A**dministration
NSC	**N**ash-**S**utcliffe **C**oefficient
NSSC	**N**ational **S**oil **S**urvey **C**enter
PCA	**P**rincipal **C**omponent **A**nalysis
PD	**P**rivat **D**ozent (Assistant Professor)
PDF	**P**robability **D**istribution **F**unction
PF	**P**article **F**ilter
PoEnKF	**P**artially **o**rthogonal **En**semble **K**alman **F**ilter
RMS	**R**oot **M**ean **S**quare
RMSE	**R**oot **M**ean **S**quare **E**rror
SCAN	**S**oil **C**limate **A**nalysis **N**etwork
SEIK	**S**ingular **E**volutive **I**nterpolated **K**alman filter
SI	**S**ensitivity **I**ndex
SLR	**S**atellite **L**aser **R**anging
SMOS	**S**oil **M**oisture **O**cean **S**alinity
SQRA	**SQ**uare **R**oot **A**nalysis scheme
SRCC	**S**pearman's **R**ank **C**orrelation **C**oefficient
SVD	**S**ingular **V**alue **D**ecomposition
TG	**T**heoretical **G**eodesy
TWS	**T**otal **W**ater **S**torage
TWSA	**T**otal **W**ater **S**torage **A**nomaly
UK	**U**nited **K**ingdom
USA	**U**nited **S**tates of **A**merica
USGS	**U**nited **S**tates **G**eological **S**urvey
WaterGAP	**Water G**lobal **A**ssessment and **P**rognosis
WFDEI	**W**ATCH **F**orcing **D**ata methodology applied to **E**RA-**I**nterim data
WGHM	**W**aterGAP **G**lobal **H**ydrology **M**odel
W3RA	**W**orld-**W**ide **W**ater **R**esources **A**ssessment
3D/4D-Var	3-**D**imensional / 4-**D**imensional **Var**iational assimilation

List of Figures

List of Tables

References

Alcamo, J., P. Döll, T. Henrichs, F. Kaspar, B. Lehner, T. Rösch and S. Siebert (2003). *Development and testing of the WaterGAP 2 global model of water use and availability.* Hydrolog Sci J, 48:317–337. doi:10.1623/hysj.48.3.317.45290.

Anderson, J.L. (2001). *An Ensemble Adjustment Kalman Filter for Data Assimilation.* Mon Wea Rev, 129:2884–2903. doi:http://dx.doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

Anderson, J.L. (2007). *An adaptive covariance inflation error correction algorithm for ensemble filters.* Tellus A, 59(2):210–224. doi:10.1111/j.1600-0870.2006.00216.x.

Anderson, J.L. (2009). *Spatially and temporally varying adaptive covariance inflation for ensemble filters.* Tellus A, 61(1):72–83. doi:10.1111/j.1600-0870.2008.00361.x.

Baumgartner, A. and H. J. Liebscher (1990). *Allgemeine Hydrologie.* Gebrüder Borntraeger, Berlin, Stuttgart.

Bennett, A. (1992). *Inverse methods in physical oceanography.* Cambridge University Press.

Bennett, A. (2002). *Inverse modeling of the ocean and atmosphere.* Cambridge University Press.

Bennett, A. and B.S. Chua (1994). *Open-Ocean Modeling as an Inverse Problem: The Primitive Equations.* Mon Weather Rev, 122:1326–1336. doi:http://dx.doi.org/10.1175/1520-0493(1994)122<1326:OOMAAI>2.0.CO;2.

Bennett, A., B.S. Chua and L.M. Leslie (1996). *Generalized Inversion of a Global Numerical Weather Prediction Model.* Meteorol Atmos Phys, 60:165–178. doi:10.1007/BF01029793.

Bishop, C.H., B.J. Etherton and S.J. Majumdar (2001). *Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects.* Mon Wea Rev, 129:420–436. doi:http://dx.doi.org/10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.

Blewitt, G. (2013). *Self-consistency in reference frames, geocenter definition, and surface loading of the solid Earth.* J Geophys Res, 108(B2). doi:10.1029/2002JB002082.

Burgers, G., P.J. van Leeuwen and G. Evensen (1998). *Analysis scheme in the ensemble Kalman filter.* Mon Weather Rev, 126:1719–1724. doi:10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Chambers, D.P. (2006). *Observing seasonal steric sea level variations with GRACE and satellite altimetry.* J of Geophys Res-Oceans, 111:C03010. doi:10.1029/2005JC002914.

Chen, J., J.S. Famiglietti, B.R. Scanlon and M. Rodell (2016). *Groundwater Storage Changes: Present Status from GRACE Observations.* Surv Geophys, 37(2):397–417. doi:10.1007/s10712-015-9332-4.

Chen, Y. and D. Zhang (2006). *Data assimilation for transient flow in geologic formations via Ensemble Kalman Filter.* Adv Water Resour, 29(8):1107–1122. doi:10.1016/j.advwatres.2005.09.007.

Cheng, M. and J.C. Ries (2012). *Monthly estimates of C20 from 5 SLR satellites.* Technichal Report. `ftp://podaac.jpl.nasa.gov/allData/grace/docs/TN-05_C20_SLR.txt`.

Cohn, S.E., N.S. Sivakumaran and R. Todling (1994). *A Fixed-Lag Kalman Smoother for Retrospective Data Assimilation.* Mon Wea Rev, (122):2838–2867. doi:10.1175/1520-0493(1994)122<2838:AFLKSF>2.0.CO;2.

Collilieux, X., T. van Dam, J. Ray, D. Coulot, L. Metivier and Z. Altamimi (2011). *Strategies to mitigate aliasing of loading signals while estimating GPS frame parameters.* J Geod, 86:1–14. doi:10.1007/s00190-011-0487-6.

Courtier, P. and O. Talagrand (1987). *Variational Assimilation of Meteorological Observations with the Adjoint Vorticity Equation. II: Numerical Results.* Quart J Roy Meteorol Soc, 113(478):1329–1347. doi:10.1002/qj.49711347813.

Courtier, P., J.N. Thépaut and A. Hollingsworth (1994). *strategy for operational implementation of 4D-Var, using an incremental approach.* Q J Roy Meteor Soc, 120(519): 1367–1387. doi:10.1002/qj.49712051912.

Daley, R. (1991). *Atmospheric data analysis.* Cambridge University Press, Cambridge.

de Zeeuw, J.W. (1973). *Hydrograph analysis for areas with mainly groundwater runoff*, chapter 16, 321–358. Drainage Principle and Applications, Vol. II, Chapter 16, Theories of field drainage and watershed runoff. Publication 16, International Institute for Land Reclamation and Improvement (ILRI), Wageningen, The Netherlands.

Deardorff, J.W. (1978). *Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation.* J Geophys Res, 83(C4):1889–1903. doi:10.1029/JC083iC04p01889.

Del Moral, P. (1997). *Nonlinear filtering: Interacting particle resolution.* Comptes Rendus de l'Académie des Sciences - Series I - Mathematics, 325(6):653–658. doi:10.1016/S0764-4442(97)84778-7.

Döll, P. and K. Fiedler (2008). *Global-scale modeling of groundwater recharge.* Hydrol Earth Syst Sci, 12:863–885. doi:10.5194/hess-12-863-2008.

Döll, P. and B. Lehner (2002). *Validation of a new global 30-min drainage map.* J Hydrol, 258(1-4):214–231. doi:Validation of a new global 30-min drainage map.

Döll, P., F. Kaspar and B. Lehner (2003). *A global hydrological model for deriving water availability indicators: model tuning and validation.* J Hydrol, 207:105–134. doi:10.1016/S0022-1694(02)00283-4.

Döll, P., H. Hoffmann-Dobrev, F.T. Portmann, S. Siebert, A. Eicker, M. Rodell, G. Strassberg and B.R. Scanlon (2012). *Impact of water withdrawals from groundwater and surface water on continental water storage variations.* J Geodyn, 59-60:143–156. doi:10.1016/j.jog.2011.05.001.

Döll, P., H. Müller Schmied, C. Schuh, F.T. Portmann and A. Eicker (2014). *Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites.* Water Resour Res, 50(7):5698–5720. doi:10.1002/2014WR015595.

Döll, P., H. Douville, A. Güntner, H. Müller Schmied and Y. Wada (2016). *Modelling Freshwater Resources at the Global Scale: Challenges and Prospects.* Surv Geophys, 37: 195–221. doi:10.1007/s10712-015-9343-1.

Drécourt, J.P. (2004). *Data assimilation in hydrological modelling.* PhD Thesis, Technical University of Denmark.

Drécourt, J.P., H. Madsen and D. Rosbjerg (2006). *Calibration framework for a Kalman filter applied to a groundwater model.* Adv Water Resour, 29:719–734. doi:10.1016/j.advwatres.2005.07.007.

Duane, S., A.D. Kennedy, B.J. Pendleton and D. Roweth (1987). *Hybrid Monte Carlo.* Phys Lett B, 195:216–222. doi:doi:10.1016/0370-2693(87)91197-X.

Dziewonski, A. M. and D. L. Anderson (1981). *Preliminary reference Earth model (PREM).* Phys Earth Planet Int, 25:297–356. doi:10.1016/0031-9201(81)90046-7.

Eicker, A., M. Schumacher, J. Kusche, P. Döll and H. Müller Schmied (2014). *Calibration/Data Assimilation Approach for Integrating GRACE Data into the WaterGAP Global Hydrology Model (WGHM) Using an Ensemble Kalman Filter: First Results.* Surv Geophys, 35(6):1285–1309. doi:10.1007/s10712-014-9309-8.

Evensen, G. (1994). *Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics.* J Geophys Res, 99(C5):10143–10162. doi:10.1029/94JC00572.

Evensen, G. (2003). *The ensemble Kalman filter: Theoretical formulation and practical implementation.* Ocean dynam, 53:343–367. doi:10.1007/s10236-003-0036-9.

Evensen, G. (2004). *Sampling strategies and square root analysis schemes for the EnKF.* Ocean Dynam, (54):539–560. doi:10.1007/s10236-004-0099-2.

Evensen, G. (2007). *Data assimilation. The Ensemble Kalman Filter.* Springer, Berlin, Heidelberg.

Evensen, G. and P.J. van Leeuwen (2000). *An Ensemble Kalman Smoother for Nonlinear Dynamics.* Mon Wea Rev, 128:1852–1867. doi:10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2.

Famiglietti, J.S. and M. Rodell (2013). *Water in the balance.* Science, 340(6138):1300–1301. doi:10.1126/science.1236460.

Famiglietti, J.S., A. Cazenave, A. Eicker, J.T. Reager, M. Rodell and I. Velicogna (2015). *Satellites provide the big picture.* Science, 349(6249):684–685. doi:10.1126/science.aac9238.

Farrell, W.E. (1972). *Deformation of the Earth by surface loads.* Rev Geophys, 10 (3): 761–797. doi:10.1029/RG010i003p00761.

Fishman, G. (1996). *Monte Carlo. Concepts, Algorithms, and Applications.* Springer, Heidelberg, New York.

Flörke, M., E. Kynast, I. Bärlund, S. Eisner, F. Wimmer and J. Alcamo (2013). *Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development: A global simulation study.* Global Environ Change, 23:144–156. doi:10.1016/j.gloenvcha.2012.10.018.

Forman, B.A. and R.H. Reichle (2013). *The spatial scale of model errors and assimilated retrievals in a terrestrial water storage assimilation system.* Water Resour Res, 49: 7457–7468. doi:10.1002/2012WR012885.

Forman, B.A., R.H. Reichle and M. Rodell (2012). *Assimilation of terrestrial water storage from GRACE in a snow-dominated basin.* Water Resour Res, 48:W01507. doi:10.1029/2011WR011239.

Forootan, E. (2014). *Statistical Signal Decomposition Techniques for Analyzing Time-Variable Satellite Gravimetry Data.* PhD Thesis, University of Bonn, Germany.

Forootan, E., O. Didova, J. Kusche and A. Löcher (2013). *Comparisons of atmospheric data and reduction methods for the analysis of satellite gravimetry observations.* J Geophys Res-Sol Ea, 118 (5):2382–2396. doi:10.1002/jgrb.50160.

Forootan, E., O. Didova, M. Schumacher, J. Kusche and B. Elsaka (2014a). *Comparisons of atmospheric mass variations derived from ECMWF reanalysis and operational fields.* J Geod, 88:503–514. doi:10.1007/s00190-014-0696-x.

Forootan, E., R. Rietbroek, J. Kusche, M.A. Sharifi, J.L. Awange, M. Schmidt, P. Omondi and J. Famiglietti (2014b). *Separation of large scale water storage patterns over Iran using GRACE, altimetry and hydrological data.* Rem Sens Environ, 140:580–595. doi:10.1016/j.rse.2013.09.025.

Forootan, E., K. Khandu, J. Awange, M. Schumacher, R. Anyah, A.I.J.M. van Dijk and J. Kusche (2016). *Quantifying the impacts of ENSO and IOD on rain gauge and remotely sensed precipitation products over Australia.* Remote Sens Environ, 172:50–66. doi:10.1016/j.rse.2015.10.027.

Fritsche, M., P. Döll and R. Dietrich (2012). *Global-scale validation of model-based load deformations from water mass and atmospheric pressure variations using GPS.* J Geodyn, 59-60:133–142. doi:10.1016/j.jog.2011.04.001.

Gamerman, D. and H.F. Lopes (1997). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference.* Chapman & Hall.

Gaspari, G. and S.E. Cohn (1999). *Construction of correlation functions in two and three dimensions.* Q J R Meteorol Soc, 125(554):723–757. doi:10.1002/qj.49712555417.

Girotto, M., G.J.M. De Lannoy, R.H. Reichle and M. Rodell (2016). *Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model.* Water Resour Res, 52(5):4164–4183. doi:10.1002/2015WR018417.

Greybush, S.J., E. Kalnay, T. Miyoshi, K. Ide and B.R. Hunt (2011). *Balance and Ensemble Kalman Filter Localization Techniques.* Mon Wea Rev, 139:511–522. doi:10.1175/2010MWR3328.1.

Güntner, A., J. Stuck, S. Werth, P. Döll, K. Verzano and B. Merz (2007). *A global analysis of temporal and spatial variations in continental water storage.* Water Resour Res, 43: W05416. doi:10.1029/2006WR005247.

Gupta, H.V., S. Sorooshian and P.O. Yapo (1998). *Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information.* Water Resour Res, 34(4):751–763. doi:10.1029/97WR03495.

Hamby, D.M. (1994). *Review of Techniques for Parameter Sensitivity Analysis of Environmental Models.* Environ Monit Assess, 32(2):135–154. doi:10.1007/BF00547132.

Hamill, T.M. and C. Snyder (2002). *Using improved Background-Error Covariances from an Ensemble Kalman Filter for Adaptive Observations.* Mon Wea Rev, (130):1552–1572. doi:10.1175/1520-0493(2002)130<1552:UIBECF>2.0.CO;2.

Hamill, T.M., J.S. Whitaker and C. Snyder (2001). *Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter.* Mon Wea Rev, 129 (2001):2776–2790. doi:10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.

Harris, I., P. Jones, T. Osborn and D. Lister (2013). *Updated high-resolution grids of monthly climatic observations-the CRU TS3.10 Dataset.* Int J Climatol, 34(3):623–642. doi:10.1002/joc.3711.

Heemink, A.W., M. Verlaan and A.J. Segers (2001). *Variance Reduced Ensemble Kalman Filtering.* Mon Wea Rev, (129):1718–1728. doi:10.1175/1520-0493(2001)129<1718:VREKF>2.0.CO;2.

Heiskanen, W.A. and H. Moritz (1967). *Physical geodesy.* Freeman, San Francisco.

Hendricks Franssen, H.J. and W. Kinzelbach (2008). *Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem.* Water Resour Res, 44:W09408. doi:10.1029/2007WR006505.

Hoffman, F.O. and R.H. Gardner (1983). *Evaluation of Uncertainties in Environmental Radiological Assessment Models.* In: Till, J.E. and H.R. Meyer (Eds.), *Radiological Assessment: a Textbook on Environmental Dose Assessment*, Vol. Report No. NUREG/CR-3332. Washington, DC: U.S. Nuclear Regulatory Commission.

Hoteit, I., D.T. Pham and J. Blum (2002). *A simplified reduced order Kalman filtering and application to altimetric data assimilation in Tropical Pacific.* J Marine Syst, 36: 101–127. doi:10.1016/S0924-7963(02)00129-X.

Hoteit, I., D.T. Pham, G. Korres and G. Triantafyllou (2008a). *Particle Kalman filtering for data assimilation in meteorology and oceanography.* In: 3rd WCRP International Conference on Reanalysis. Tokyo, Japan, 6 pages.

Hoteit, I., D.T. Pham, G. Triantafyllou and G. Korres (2008b). *A New Approximate Solution of the Optimal Nonlinear Filter for Data Assimilation in Meteorology and Oceanography.* Mon Wea Rev, 136:317–334. doi:http://dx.doi.org/10.1175/2007MWR1927.1.

Houborg, R., M. Rodell, B. Li, R.H. Reichle and B.F. Zaitchik (2012). *Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations.* Water Resour Res, 48:W07525. doi:10.1029/2011WR011291.

Houtekamer, P.L. and H.L. Mitchell (2001). *A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation.* Mon Wea Rev, 129:123–137. doi:10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.

Houtekamer, P.L., H.L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek and B. Hansen (2005). *Atmospheric Data Assimilation with an Ensemble Kalman Filter: Results with Real Observations.* Mon Wea Rev, 133(3):604–620. doi:10.1175/MWR-2864.1.

Hunger, M. and P. Döll (2008). *Value of river discharge data for global-scale hydrological modeling.* Hydrol Earth Syst Sci, 12:841–861. doi:10.5194/hess-12-841-2008.

Hunt, B.R., E.J. Kostelich and I. Szunyogh (2007). *Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter.* Physica D, 230:112–126. doi:10.1016/j.physd.2006.11.008.

Iman, R.L. and W.J. Conover (1979). *The Use of the Rank Transform in Regression.* Technometrics, 21:499–509. doi:10.1080/00401706.1979.10489820.

Janjić, T., N. Nerger, A. Albertella, J. Schröter and S. Skachko (2011). *On Domain Localization in Ensemble-Based Kalman Filter Algorithms.* Mon Wea Rev, 139:2046–2060. doi:http://dx.doi.org/10.1175/2011MWR3552.1.

Jekeli, C. (1981). *Alternative methods to smooth the Earth's gravity field.* Reports of the Department of Geodetic Science, 327, Ohio State University, Columbus, Ohio.

Johnson, A.I. (1967). *Specific yield: compilation of specific yields for various materials.* U.S. Geological Survey Water Supply Paper, 1662-D:74 p.

Kalman, R. E. (1960). *A new approach to linear filtering and prediction problems.* Trans ASME J Basic Eng, 82(D):35–45.

Kalman, R. E. and S. Bucy (1961). *New Results in Linear Filtering and Prediction Theory.* J Basic Eng, 83(1):95–108. doi:10.1115/1.3658902.

Kalnay, E. (2003). *Atmospheric modeling, data assimilation and predictability.* Cambridge University Press, Cambridge.

Kaspar, F. (2004). *Entwicklung und Unsicherheitsanalyse eines globalen hydrologischen Modells (in German).* PhD Thesis, University of Kassel, Germany.

Kirchgessner, P., L. Nerger and A. Bunse-Gerstner (2014). *On the Choice of an Optimal Localization Radius in Ensemble Kalman Filter Method.* Mon Wea Rev, 142:2165–2175. doi:10.1175/MWR-D-13-00246.1.

Klees, R., E.A. Zapreeva, H.C. Winsemius and H.H.G. Savenije (2006). *The bias in GRACE estimates of continental water storage variations.* Hydrology and Earth System Sciences Discussions, 3 (6):3557–3594.

Klees, R., A. Revtova, B.C. Gunter, P. Ditmar, E. Oudman, H.C. Winsemius and H.H.G. Savenije (2008). *The design of an optimal filter for monthly GRACE gravity models.* Geophys J Int, 175:417–432. doi:0.1111/j.1365-246X.2008.03922.x.

Klinker, E., F. Rabier, G. Kelly and J.F. Mahfouf (2000). *The ECMWF operational implementation of four-dimensional variational assimilation. III: Experimental results and diagnostics with operational configuration.* Q J Roy Meteor Soc, 126(564):1191–1215. doi:10.1002/qj.49712656417.

Koch, K.R. (1999). *Parameter estimation and hypothesis testing in linear models.* Springer, New York.

Koch, K.R. (2007). *Introduction to Bayesian statistics.* Springer, New York.

Kumar, S.V., B.F. Zaitchik, C.D. Peters-Lidard, M. Rodell, R.H. Reichle, B. Li, M. Jasinski, D. Mocko, A. Getirana, G.J.M. De Lannoy, M.H. Cosh, C.R. Hain, M. Anderson, K.R. Arsenault, Y. Xia and M. Ek (2016). *Assimilation of gridded GRACE terrestrial water storage estimates in the North American Land Data Assimilation System.* J Hydrometeorol, online first. doi:10.1175/JHM-D-15-0157.1.

Kurtenbach, E. (2011). *Entwicklung eines Kalman-Filters zur Bestimmung kurzzeitiger Variationen des Erdschwerefeldes aus Daten der Satellitenmission GRACE (in German).* PhD Thesis, University of Bonn, Germany.

Kurtenbach, E., T. Mayer-Gürr and A. Eicker (2009). *Deriving daily snapshots of the Earth's gravity field from GRACE L1B data using Kalman filtering.* Geophys Res Lett, 36:L17102. doi:10.1029/2009GL039564.

Kusche, J. (2003). *A Monte-Carlo technique for weight estimation in satellite geodesy.* J Geod, 76(11-12):641–652. doi:10.1007/s00190-002-0302-5.

Kusche, J. (2007). *Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models.* J Geod, 81(11):733–749. doi:10.1007/s00190-007-0143-3.

Kusche, J., R. Schmidt, S. Petrovic and R. Rietbroek (2009). *Decorrelated GRACE time-variable gravity solutions by GFZ, and their validation using a hydrological model.* J Geod, 83 (10):903–913. doi:10.1007/s00190-009-0308-3.

Kusche, J., V. Klemann and W. Bosch (2012). *Mass distribution and mass transport in the Earth system.* J Geodyn, 59-60:1–8. doi:10.1016/j.jog.2012.03.003.

Landerer, F. W. and S.C. Swenson (2012). *Accuracy of scaled GRACE terrestrial water storage estimates.* Water Resour Res, 48(4):W04531. doi:10.1029/2011WR011453.

Landerer, F. W., J. O. Dickey and A. Güntner (2010). *Terrestrial water budget of the Eurasian pan-Arctic from GRACE satellite measurements during 2003-2009.* J Geophys Res-Atmos, 115:D23115. doi:10.1029/2010JD014584.

Leblanc, M., P. Tregoning, G. Ramillien, S.O. Tweed and A. Fakes (2009). *Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia.* Water Resour Res, 45 (4):W04408. doi:10.1029/2008WR007333.

Leblanc, M., S. Tweed, A.I.J.M. van Dijk and B. Timbal (2012). *A review of historic and future hydrological changes in the Murray-Darling Basin.* Global Planet Change, 80-81: 226–246. doi:10.1016/j.gloplacha.2011.10.012.

Lermusiaux, P.F.J. and A.R. Robinson (1999). *Data Assimilation via Error Subspace Statistical Estimation. Part I: Theory and Schemes.* Mon Wea Rev, 127:1385–1407. doi:10.1175/1520-0493(1999)127<1385:DAVESS>2.0.CO;2.

Li, B., M. Rodell, B. F. Zaitchik, R.H. Reichle, R. D. Koster and T. 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.* J Hydrol, 446-447:103–115. doi:10.1016/j.jhydrol.2012.04.035.

Li, H., E. Kalnay and T. Miyoshi (2009). *Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter.* Quart J Roy Meteor Soc, 135: 523–533. doi:10.1002/qj.371.

Liu, Y., A.H. Weerts, M. Clark, H.J. Hendricks Franssen, S. Kumar, H. Moradkhani, D.J. Seo, D. Schwanenberg, P. Smith, A.I.J.M. van Dijk, N. van Velzen, M. He, H. Lee, S.J. Noh, O. Rakovec and P. Restrepo (2012). *Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities.* Hydrol Earth Syst Sci, 16:3863–3887. doi:10.5194/hess-16-3863-2012.

Llovel, W., M. Becker, A. Cazenave, J.-F. Crétaux and G. Ramillien (2010). *Global land water storage change from GRACE over 2002–2009; Inference on sea level.* G R Geosci, 342(3):179–188. doi:10.1016/j.crte.2009.12.004.

Longuevergne, L., B.R. Scanlon and C.R. Wilson (2010). *GRACE hydrological estimates for small basins: evaluating processing approaches on the high plains aquifer, USA.* Water Resources Research, 46:W11517. doi:10.1029/2009WR008564.

Losa, S.N., S. Danilov, J. Schröter, L. Nerger, S. Maßmann and F. Janssen (2012). *Assimilating NOAA SST data into the BSH operational circulation model for the North and Baltic Sea: Inference about the data.* J Marine Syst, 105-108:152–162. doi:10.1016/j.jmarssys.2012.07.008.

Mayer-Gürr, T. (2008). *Gravitationsfeldbestimmung aus der Analyse kurzer Bahnbögen am Beispiel der Satellitenmissionen CHAMP und GRACE (in German)*. PhD Thesis, University of Bonn, Germany.

Mckay, M.D., R.J. Beckman and W.J. Conover (2000). *A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code*. Technometrics, 42(1):55–61. doi:10.1080/00401706.2000.10485979.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller (1953). *Equation of State Calculations by Fast Computing Machines*. J Chem Phys, 21:1087–1092. doi:http://dx.doi.org/10.1063/1.1699114.

Miyoshi, T. (2011). *The Gaussian Approach to Adaptive Covariance Inflation and Its Implementation with the Local Ensemble Transform Kalman Filter*. Mon Wea Rev, 139 (5):1519–1535. doi:http://dx.doi.org/10.1175/2010MWR3570.1.

Moradkhani, H., S. Sorooshian, H.V. Gupta and P.R. Houser (2005). *Dual state-parameter estimation of hydrological models using ensemble Kalman filter*. Adv Water Resour, 28: 135–147. doi:10.1016/j.advwatres.2004.09.002.

Müller Schmied, H., S. Eisner, D. Franz, M. Wattenbach, F.T. Portmann, M. Flörke and P. Döll (2014). *Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration*. Hydrol Earth Syst Sci, 18:3511–3538. doi:10.5194/hess-18-3511-2014.

Müller Schmied, H., L. Adam, S. Eisner, G. Fink, M. Flörke, H. Kim, T. Oki, F.T. Portmann, R. Reinecke, C. Riedel, Q. Song, J. Zhang and P. Döll (2016). *Variations of global and continental water balance components as impacted by climate forcing uncertainty and human water use*. Hydrol Earth Syst Sci Discussions. doi:10.5194/hess-2015-527.

Nash, J. and J. Sutcliffe (1970). *River flow forecasting through conceptual models part I: A discussion of priciples*. J Hydrol, 10:282–290. doi:10.1016/0022-1694(70)90255-6.

Nerger, L. (2003). *Parallel Filter Algorithms for Data Assimilation in Oceanography*. PhD Thesis, University of Bremen, Germany.

Nerger, L., S. Danilov, W. Hiller and J. Schröter (2006). *Using sea level data to constrain a finite-element primitive-equation ocean model with a local SEIK filter*. Ocean Dyn, 56:634–649. doi:10.1007/s10236-006-0083-0.

Nerger, L., S. Danilov, G. Kivman, W. Hiller and J. Schröter (2007). *Data assimilation with the Ensemble Kalman Filter and the SEIK filter applied to a finite element model of the North Atlantic*. J Marine Syst, 65:288–298. doi:10.1016/j.jmarsys.2005.06.009.

Nerger, L., T. Janjić, J. Schröter and W. Hiller (2012). *A regulated localization scheme for ensemble-based Kalman filters*. Q J Roy Meteor Soc, 138(664):802–812. doi:10.1002/qj.945.

Nichols, N.K. (2010). *Mathematical Concepts of Data Assimilation*, 13–39. Springer, Heidelberg, Dordrecht, London, New York.

Ott, E., B.R. Hunt, I. Szunyogh, A.V. Zimin, E.J. Kostelich, M. Corazza, E. Kalnay, D.J. Patil and J.A. Yorke (2004). *A local ensemble Kalman filter for atmospheric data assimilation.* Tellus A, 56(5):415–428. doi:10.1111/j.1600-0870.2004.00076.x.

Palmer, T.N., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G.J. Shutts, M. Steinheimer and A. Weisheimer (2009). *Stochastic parametrization and Model Uncertainty.* ECMWF Tech Mem, 598:42 pages.

Pan, M. and E.F. Wood (2006). *Data Assimilation for Estimating the Terrestrial Water Budget Using a Constrained Ensemble Kalman Filter.* J Hydrometeor, 7:534–547. doi:10.1175/JHM495.1.

Papa, F., C. Prigent, W.B. Rossow, B. Legresy and F. Remy (2006). *Inundated wetland dynamics over boreal regions from remote sensing: The use of Topex-Poseidon dual-frequency radar altimeter observations.* Int J Remote Sens, 27(21):4847–4866. doi:10.1080/01431160600675887.

Papa, F., A. Güntner, F. Frappart, C. Prigent and W.B. Rossow (2008). *Variations of surface water extent and water storage in large river basins: A comparison of different global data sources.* Geophys Res Lett, 35:L11401. doi:10.1029/2008GL033857.

Papa, F., C. Prigent, F. Aires, C. Jimenez, W.B. Rossow and E. Matthews (2010). *Interannual variability of surface water extent at the global scale, 1993-2004.* J Geophys Res, 115:D12111. doi:10.1029/2009JD012674.

Pham, D.T. (2001). *Stochastic Methods for Sequential Data Assimilation in Strongly Nonlinear Systems.* Mon Wea Rev, 129:1194–1207. doi:10.1175/1520-0493(2001)129<1194:SMFSDA>2.0.CO;2.

Pham, D.T., J. Verron and M.C. Roubaud (1998). *A singular evolutive extended Kalman filter for data assimilation in oceanography.* J Marine Syst, 16(3-4):323–340. doi:10.1016/S0924-7963(97)00109-7.

Pierce, R., J. Leitch, M. Stephens, P. Bender and R. Nerem (2008). *Intersatellite range monitoring using optical interferometry.* Appl Optics, 47(27):5007–5019. doi:10.1364/AO.47.005007.

Priestley, C.H.B. and R.J. Taylor (1972). *On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters.* Mon Wea Rev, 100:81–92. doi:10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2.

Prigent, C., F. Papa, F. Aires, W.B. Rossow and E. Matthews (2007). *Global inundation dynamics inferred from multiple satellite observations, 1993-2000.* J Geophys Res, 112: D12107. doi:10.1029/2006JD007847.

Rabier, F., H. Järvinen, E. Klinker, J.F. Mahfouf and A. Simmons (2000). *The ECMWF operational implementation of four-dimensional variational assimilation. 1: Experimental results with simplifed physics.* Quart J Roy Meteor Soc, 126:1143–1170. doi:10.1002/qj.49712656415.

Ramillien, G.L., L. Seoane, F. Frappart, R. Biancale, S. Gratton, X. Vasseur and S. Bourgogne (2012). *Constrained Regional Recovery of Continental Water Mass Time-variations from GRACE-based Geopotential Anomalies over South America.* Surv Geophys, 33:887–905. doi:10.1007/s10712-012-9177-z.

Ramillien, G.L., F. Frappart, S. Gratton and X. Vasseur (2015). *Sequential estimation of surface water mass changes from daily satellite gravimetry data.* J Geod, 89(3):259–282. doi:10.1007/s00190-014-0772-2.

Reager, J.T., A.C. Thomas, E.A. Sproles, M. Rodell, H.K. Beaudoing, B. Li and J.S. Famiglietti (2015). *Assimilation of GRACE Terrestrial Water Storage Observations into a Land Surface Model for the Assessment of Regional Flood Potential.* Remote Sens, (7):14663–14679. doi:10.3390/rs71114663.

Reichle, R.H. (2008). *Data assimilation methods in the Earth sciences.* Adv Water Resour, 31(11):1411–1418. doi:10.1016/j.advwatres.2008.01.001.

Reichle, R.H. and R.D. Koster (2003). *Assessing the impact of horizontal error correlations in background fields on soil moisture estimation.* J Hydrometeorol, 4(6):1229–1242. doi:10.1175/1525-7541(2003)004<1229:ATIOHE >2.0.CO;2.

Rietbroek, R. (2014). *Retrieval of sea level and surface loading variations from geodetic observations and model simulations: an integration approach.* PhD Thesis, University of Bonn, Germany.

Rietbroek, R., S.E. Brunnabend, C. Dahle, F. Flechtner, J. Kusche, J. Schröter and R. Timmermann (2009). *Changes in total ocean mass derived from GRACE, GPS, and ocean modeling with weekly resolution.* J Geophys Res-Oceans, 114:C11004. doi:10.1029/2009JC005449.

Rietbroek, R., M. Fritsche, S.-E. Brunnabend, I. Daras, J. Kusche, J. Schröter, F. Flechtner and R. Dietrich (2012). *Global surface mass from a new combination of GRACE, modelled OBP and reprocessed GPS data.* Journal of Geodynamics, 59-60:64–1.

Ripley, B.D. (1987). *Stochastic Simulation.* Wiley, New York.

Risken, H. and T. Frank (1996). *The Fokker-Planck Equation: Methods of Solutions and Applications.* Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-642-61544-3.

Rodell, M., J. Chen, H. Kato, J.S. Famiglietti, J. Nigro and C.R. Wilson (2007). *Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE.* Hydrogeology Journal, 15:159–166. doi:10.1007/s10040-006-0103-7.

Rodell, M., I. Velicogna and J.S. Famiglietti (2009). *Satellite-based estimates of groundwater depletion in India.* Nature, 460:999–1002. doi:10.1038/nature08238.

Sakumura, C., S. Bettadpur and S. Bruinsma (2014). *Ensemble prediction and intercomparison analysis of GRACE time-variable gravity field models.* Geophys Res Lett, 41: 1389–1397. doi:10.1002/2013GL058632.

Sasaki, Y. (1970a). *Some basic formalisms in numerical variational analysis.* Monthly Weather Review, 98(12):875–883. doi:10.1175/1520-0493(1970)098<0875:SBFINV>2.3.CO;2.

Sasaki, Y. (1970b). *Numerical variational analysis with weak constraint and application to surface analysis of severe storm gust.* Mon Weather Rev, 98:899–910.

Sasaki, Y. (1970c). *Numerical variational analysis formulated under the constraints as determined by longwave equations and a low-pass filter.* Mon Weather Rev, 98(12): 884–898.

Saynisch, J., I. Bergmann-Wolf and M. Thomas (2015). *Assimilation of GRACE-derived oceanic mass distributions with a global ocean circulation model.* J Geod, 89(2):121–139. doi:10.1007/s00190-014-0766-0.

Schaefer, G.L., M.H. Cosh and T.J. Jackson (2007). *The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN).* J Atmos Ocean Tech, 24(12): 2073–2077. doi:10.1175/2007JTECHA930.1.

Schmeer, M., M. Schmidt, W. Bosch and F. Seitz (2012). *Separation of mass signals within GRACE monthly gravity field models by means of empirical orthogonal functions.* J Geodyn, 59-60:124–132. doi:10.1016/j.jog.2012.03.001.

Schmidt, R., F. Flechtner, U. Meyer, K.H. Neumayer, C. Dahle, R. König and J. Kusche (2008). *Hydrological signals observed by the GRACE satellites.* Surv Geophys, 29 (4-5): 319–334. doi:10.1007/s10712-008-9033-3.

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese and B. Rudolf (2014). *GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle.* Theor Appl Climatol, 115:15–40. doi:10.1007/s00704-013-0860-x.

Schrama, E.J.O., B. Wouters and D.A. Lavallée (2007). *Signal and noise in Gravity Recovery and Climate Experiment (GRACE) observed surface mass variations.* J Geophys Res-Sol Ea, 112:B08407. doi:10.1029/2006JB004882.

Schumacher, M. (2012). *Assimilation of GRACE data into a global hydrological model using an ensemble Kalman filter.* Master Thesis, University of Bonn, Germany.

Schumacher, M., A. Eicker, J. Kusche, H. Müller Schmied and P. Döll (2016a). *Covariance analysis and sensitivity studies for GRACE assimilation into WGHM.* IAG Symp 2013, 143:241–248. doi:10.1007/1345_2015_119.

Schumacher, M., J. Kusche and P. Döll (2016b). *A Systematic Impact Assessment of GRACE Error Correlation on Data Assimilation in Hydrological Models.* J Geod, 90(6): 537–559. doi:10.1007/s00190-016-0892-y.

Schumacher, M., E. Forootan, A.I.J.M. van Dijk, H. Müller Schmied, R.S. Crosbie, J. Kusche and P. Döll (2016c). *GRACE assimilation into a global water resources and use model improves representation of drought-induced groundwater decline within the Murray-Darling Basin, Australia.* Remote Sens Environ (under review).

Shuttleworth, W.J. (1993). *Evaporation*, chapter 4.1–4.53. Handbook of hydrology. McGraw-Hill Inc.

Shutts, G., M. Leutbecher, A. Weisheimer, T. Stockdale, L. Isaksen and M. Bonavita (2007). *Representing model uncertainty: Stochastic parametrizations at ECMWF.* ECMWF Newsletter, 129:19–24.

Siebert, S., P. Döll, J. Hoogeveen, J.-M. Faures, K. Frenken and S. Feick (2005). *Development and validation of the global map of irrigation areas.* Hydrol Earth Syst Sci, 9: 535–547. doi:10.5194/hess-9-535-2005.

Sood, A. and V. Smakhtin (2015). *Global hydrological models: a review.* Hydrolog Sci J, 60(4):549–565. doi:10.1080/02626667.2014.950580.

Stocker, T.F., D. Qin, G.K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (Eds.). *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Changes.* Cambridge University Press, Cambridge (United Kingdom), NewYork (USA), 1535 pp.

Stordal, S., H.A. Karlsen, G. Nævdal, H.J. Skaug and B. Vallès (2011). *Bridging the ensemble Kalman filter and particle filters: the adaptive Gaussian mixture filter.* Comput Geosci, 15:293–305. doi:10.1007/s10596-010-9207-1.

Strassberg, G., B.R. Scanlon and D. Chambers (2009). *Evaluation of groundwater storage monitoring with the GRACE satellite: case study of the High Plains aquifer, central United States.* Water Resour Res, 45:W05410. doi:10.1029/2008WR006892.

Su, H., Z.L. Yang, R.E. Dickinson, C.R. Wilson and G.Y. Niu (2010). *Multisensor snow data assimilation at the continental scale: The value of gravity recovery and climate experiment terrestrial water storage information.* J Geophys Res, 115:D10104. doi:10.1029/2009JD013035.

Swenson, S. and J.M. Wahr (2002). *Methods for inferring regional surface-mass anomalies from Gravity Recovery and Climate Experiment (GRACE) measurements of time-variable gravity.* J Geophys Res-Sol Ea, 107(B9):2193. doi:10.1029/2001JB000576.

Swenson, S. and J.M. Wahr (2006). *Post-processing removal of correlated errors in GRACE data.* Geophysi Res Lett, 33(8). doi:10.1029/2005GL025285.

Talagrand, O. (2010). *Variational Assimilation.* In: *Data Assimilation. Making Sense of Observations*, 41–67. Springer, Berlin, Heidelberg.

Talagrand, O. and P. Courtier (1987). *Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory.* Quart J Roy Meteorol Soc, 113: 1311–1328. doi:10.1002/qj.49711347812.

Tangdamrongsub, N., S.C. Steele-Dunne, B.C. Gunter, P.G. Ditmar and A.H. Weerts (2015). *Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin.* Hydrol Earth Syst Sci, 19:2079–2100. doi:10.5194/hess-19-2079-2015.

Tapley, B. D., S. Bettadpur, M. Watkins and C. Reigber (2004). *The gravity recovery and climate experiment: mission overview and early results.* Geophysical Research Letters, 31:L09607.

Tavella, P. and A. Premoli (1994). *Estimating the instabilities of n clocks by measuring differences of their readings.* Premoli A, 30 (5):479. doi:10.1002/joc.4346.

Tippett, M.K., J.L. Anderson, C.H. Bishop, T.M. Hamill and J.S. Whitaker (2003). *Ensemble Square Root Filters.* Mon Wea Rev, (131):1485–1490. doi:10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2.

Tregoning, P., S. McClusky, A.I.J.M. van Dijk, R.S. Crosbie and J.L. Peña-Arancibia (2012). *Assessment of GRACE satellites for groundwater estimation in Australia.* National Water Commission, Waterlines report No 71, Canberra. doi:http://archive.nwc.gov.au/library/waterlines/71.

van Dijk, A.I.J.M., L.J. Renzullo and M. Rodell (2011). *Use of Gravity Recovery and Climate Experiment terrestrial water storage retrievals to evaluate model estimates by the Australian water resources assessment system.* Water Resour Res, 47:W11524. doi:10.1029/2011WR010714.

van Dijk, A.I.J.M., H.E. Beck, R.S. Crosbie, R.A.M. de Jeu, Y.Y. Liu, G.M. Podger, B. Timbal and N.R. Viney (2013). *The Millennium Drought in south-east Australia (2001-2009): Natural and human causes and implications for water resources, ecosystems, economy, and society.* Water Resour Res, 49(2):1040–1057. doi:10.1002/wrcr.20123.

van Dijk, A.I.J.M., L.J. Renzullo, Y. Wada and P. Tregoning (2014). *A global water cycle reanalysis (2003-2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble.* Hydrol Earth Syst Sci, (18):2955–2973. doi:10.5194/hess-18-2955-2014.

Vassolo, S. and P. Döll (2005). *Global-scale gridded estimates of thermoelectric power and manufacturing water use.* Water Resour Res, 41:W04010. doi:10.1029/2004WR003360.

Verlaan, M. and A.W. Heemink (2001). *Nonlinearity in Data Assimilation Applications: A Practical Method for Analysis.* Mon Wea Rev, 129:1578–1589. doi:10.1175/1520-0493(2001)129<1578:NIDAAA>2.0.CO;2.

Vrugt, J., C.G.H. Dirks, H.V. Gupta, W. Bouten and J.M. Verstraten (2005). *Improved treatment of uncertainty in hydrologic modeling: Combining strengths of global optimization and data assimilation.* Water Resour Res, 41:W01017. doi:10.1029/2004WR003059.

Vrugt, J.A. (2004). *Towards improved treatment of parameter uncertainty in hydrologic modeling.* PhD Thesis, University of Amsterdam, The Netherlands.

Wahr, J.M., M. Molenaar and F. Bryan (1998). *Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE.* J Geophys Res-Sol Ea, 108(B12):30205–30229. doi:10.1029/98JB02844.

Wahr, J.M., S. Swenson and I. Velicogna (2006). *Accuracy of GRACE mass estimates.* Geophys Res Lett, 33 (6):L06401. doi:10.1029/2005GL025305.

Watkins, M.M., D.N. Wiese, D.-N. Yuan, C. Boening and F. W. Landerer (2015). *Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons.* J Geophys Res-Sol Ea, 120(4):2648–2671. doi:10.1002/2014JB011547.

Weedon, G.P., G. Balsamo, N. Bellouin, S. Gomes, M.J. Best and P. Viterbo (2014). *The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data.* Water Resour Res, 50(9):7505–7514. doi:10.1002/2014WR015638.

Werth, S. (2010). *Calibration of the global hydrological model WGHM with water mass variations from GRACE gravity data.* PhD Thesis, University of Potsdam, Germany.

Werth, S. and A. Güntner (2010). *Calibration analysis for water storage variability of the global hydrological model WGHM.* Hydrol Earth Syst Science, 14:59–78. doi:10.5194/hess-14-59-2010.

Whitaker, J.S. and T.M. Hamill (2002). *Ensemble Data Assimilation without Perturbed Observations.* Mon Weather Rev, 130:1913–1924. doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

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.* Rep Prog Phys, 77:116801. doi:10.1088/0034-4885/77/11/116801.

Zaitchik, B. F., M. Rodell and R.H. Reichle (2008). *Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model: Results for the Mississippi River Basin.* J Hydrometeorol, 9(3):535–548. doi:10.1175/2007JHM951.1.

Zhang, H., H.-J. Hendricks-Franssen, X. Han, J. Vrugt and H. Vereecken (2016). *Joint State and Parameter Estimation of Two Land Surface Models Using the Ensemble Kalman Filter and Particle Filter.* Hydrol Earth Syst Sc Discussions. doi:10.5194/hess-2016-42.

Zuletzt sind in der Schriftenreihe des Instituts für Geodäsie und Geoinformation der Rheinischen Friedrich-Wilhelms-Universität Bonn erschienen:

Heft 52 2016	Maike Schumacher Methods for assimilating remotely-sensed water storage changes into hydrological models
Heft 51 2015	Christoph Holst Analyse der Konfiguration bei der Approximation ungleichmäßig abgetasteter Oberflächen auf Basis von Nivellements und terrestrischen Laserscans
Heft 50 2015	Lutz Rolf Roese-Koerner Convex Optimization for Inequality Constrained Adjustment Problems
Heft 49 2015	Jan Martin Brockmann On High Performance Computing in Geodesy Applications in Global Gravity Field Determination
Heft 48 2015	Judith Leek The application of impact factors to scheduling VLBI Intensive sessions with twin telescopes
Heft 47 2015	Thomas Artz Determination of Sub-daily Earth Rotation Parameters from VLBI Observations
Heft 46 2015	Roelof Rietbroek Retrieval of Sea Level and Surface Loading Variations from Geodetic Observations and Model Simulations: an Integrated Approach
Heft 45 2014	Ehsan Forootan Statistical Signal Decomposition Techniques for Analyzing Time-Variable Satellite Gravimetry Data
Heft 44 2014	Erich Weiß Lebensbilder der preußischen Verwaltung des 19. und 20. Jahrhunderts im Wandel Eine Sammlung biographischer Miniaturen
Heft 43 2014	Neysa Jacqueline Setiadi Assessing People´s Early Warning Response Capability to Inform Urban Planning Interventions to Reduce Vulnerability to Tsunamis Case Study of Padang City, Indonesia
Heft 42 2013	Nils Leber Entwicklungsperspektiven metropolitaner Peripherien im Rahmen Stadtregionaler Planungs- und Entwicklungsprozesse am Beispiel Nordrhein-Westfalen
Heft 41 2013	Sophie Schetke Socio-environmental impacts of settlement growth under conditions of fostered infill development: a methodological framework for a multicriteria assessment
Heft 40 2013	Ribana Roscher Sequential Learning Using Incremental Import Vector Machines for Semantic Segmentation
Heft 39 2013	Michael Ying Yang Hierarchical and Spatial Structures for Interpreting Images of Man-made Scenes Using Graphical Models
Heft 38 2013	Sabine Daniela Bauer Automatische Detektion von Krankheiten auf Blättern von Nutzpflanzen

Heft 37 2013	Martin Drauschke Ein hierarchischer Ansatz zur Interpretation von Gebäudeaufnahmen
Heft 36 2013	Timo Dickscheid Robust Wide-Baseline Stereo Matching for Sparsely Textured Scenes
Heft 35 2013	Alexander Barth Vehicle Tracking and Motion Estimation Based on Stereo Vision Sequences
Heft 34 2013	Richard Steffen Visual SLAM from image sequences acquired by unmanned aerial vehicles
Heft 33 2013	Till Rumpf Finding spectral features for the early identification of biotic stress in plants
Heft 32 2012	Christian Siemes Digital Filtering Algorithms for Decorrelation within Large Least Squares Problems
Heft 31 2012	Silvia Becker Konsistente Kombination von Schwerefeld, Altimetrie und hydrographischen Daten zur Modellierung der dynamischen Ozeantopographie
Heft 30 2013	Annette Eicker / Jürgen Kusche (eds.) Lecture Notes from the Summer School of DFG SPP1257 Global Water Cycle
Heft 29 2012	Matthias Siemes Ein Beitrag zur koordinatengesteuerten Aussaat von Rübenpflanzen mittels Multi-Sensor-System und Filteransatz
Heft 28 2012	Jörg Schmittwilken Attributierte Grammatiken zur Rekonstruktion und Interpretation von Fassaden
Heft 27 2012	Markus Rembold Die Anerkennung und Feststellung von Grundstücksgrenzen Ein Beitrag zur Entwicklung des Liegenschaftskatasters im Lande Nordrhein-Westfalen in Vergangenheit, Gegenwart und Zukunft
Heft 26 2012	Lihua Li Separability of deformations and measurement noises of GPS time series with modified Kalman filter for landslide monitoring in real-time
Heft 25 2012	Benedikt Frielinghaus Ökonomisches Entscheidungstool zur Wohnbaulandentwicklung Wirtschaftlichkeitsanalysen potenzieller Wohnbauflächen auf der Ebene des Flächennutzungsplanes
Heft 24 2011	Enrico Kurtenbach Entwicklung eines Kalman-Filters zur Bestimmung kurzzeitiger Variationen des Erdschwerefeldes aus Daten der Satellitenmission GRACE
Heft 23 2011	Sarah Böckmann Robust determination of station positions and Earth orientation parameters by VLBI intra-technique combination
Heft 22 2011	20th Meeting of the European VLBI Group for Geodesy and Astronomy Proceedings
Heft 21 2011	Philipp Zeimetz Zur Entwicklung und Bewertung der absoluten GNSS-Antennenkalibrierung im HF-Labor
Heft 20 2011	Alessandra Roy Effects on the Geodetic-VLBI Observables Due to

	Polarization Leakage in the Receivers
Heft 19 2011	Dietmar Weigt Auswirkungen von Flughäfen insbesondere von Fluglärm auf den Immobilienmarkt am Beispiel des Marktsegments „individuelles Wohnen“
Heft 18 2011	Anno Löcher Möglichkeiten der Nutzung kinematischer Satellitenbahnen zur Bestimmung des Gravitationsfeldes der Erde
Heft 17 2010	Basem Elsaka Simulated Satellite Formation Flights for Detecting the Temporal Variations of the Earth’s Gravity Field
Heft 16 2010	2nd International Conference on Machine Control & Guidance Proceedings
Heft 15 2009	Alexandra Weitkamp Brachflächenrevitalisierung im Rahmen der Flächenkreislaufwirtschaft
Heft 14 2008	Akbar Shabanloui A New Approach for a Kinematic-Dynamic Determination of Low Satellite Orbits Based on GNSS Observations
Heft 13 2008	Frank Friesecke Stadtumbau im Konsens!? Zur Leistungsfähigkeit und Fortentwicklung des städtebaulichen Instrumentariums unter Schrumpfungsbedingungen
Heft 12 2008	Heinz Rütz Zur Kostenanalyse der privaten Umlegung als Teil der konsensualen integrierten Baulandentwicklung
Heft 11 2008	Gaby Alexandra Boele-Keimer Kommunales Kennzahlenmanagement am Beispiel von Vermessungs- und Katasterämtern in Nordrhein-Westfalen
Heft 10 2008	Annette Eicker Gravity Field Refinement by Radial Basis Functions
Heft 9 2008	Torsten Mayer-Gürr Gravitationsfeldbestimmung aus der Analyse kurzer Bahnbögen
Heft 8 2008	Boris Kargoll On the Theory and Application of Model Misspecification Tests
Heft 7 2008	Hamza Alkhatib On Monte Carlo Methods
Heft 6 2008	Klaus Borchard Annäherungen an Städtebau und Raumentwicklung
Heft 5 2008	Jens Jähnke Zur Teilmarktbildung beim Landerwerb der öffentlichen Hand
Heft 4 2008	Atef Abd-Elhakee Makhloof The Use of Topographic Isostatic Mass Information
Heft 3 2008	Markus Vennebusch Singular Value Decomposition and Cluster Analysis
Heft 2 2007	Christian Beder Grouping Uncertain Oriented Projective Geometric Entities
Heft 1 2007	Klaus Börger Geodäsie und Quantenphysik

Vertrieb Rheinische Friedrich-Wilhelms-Universität Bonn
Institut für Geodäsie und Geoinformation
· Bibliothek ·
Nußallee 17
53115 Bonn

Tel.: #49 (0)228 73-3566
Fax: #49 (0)228 73-2988

Internet: http://www.igg.uni-bonn.de

ISSN 1864-1113