Güntner, A. (2008): Improvement of global hydrological models using GRACE data. -
Surveys in Geophysics, 29, 4-5, 375-397

DOI: 10.1007/s10712-008-9038-y.

Surveys in Geophysics

Improvement of global hydrological models using GRACE data

Andreas Güntner

GeoForschungsZentrum Potsdam (GFZ)

Telegrafenberg, 14473 Potsdam

Germany

guentner@gfz-potsdam.de

Abstract

After about six years of GRACE (Gravity Recovery and Climate Experiment) satellite mission operation, an unprecedented global data set on the spatio-temporal variations of the Earth's water storage is available. The data allow for a better understanding of the water cycle at the global scale and for large river basins. This review summarizes the experiences that have been made when comparing GRACE data with simulation results of global hydrological models and it points out the prerequisites and perspectives for model improvements by combination with GRACE data. When evaluated qualitatively at the global scale, water storage variations on the continents from GRACE agreed reasonably well with model predictions in terms of their general seasonal dynamics and continental-scale spatial patterns. Differences in amplitudes and phases of water storage dynamics revealed in more detailed analyses were mainly attributed to deficiencies in the meteorological model forcing data, to missing water storage compartments in the model, but also to limitations and errors of the GRACE data. Studies that transformed previously identified model deficiencies into adequate modifications of the model structure or parameters are still rare. Prerequisites for a comprehensive improvement of large-scale hydrological models are in particular the consistency of GRACE observation and model variables in terms of filtering, reliable error estimates, and a full assessment of the water balance. Using improvements in GRACE processing techniques, complementary observation data, multi-model evaluations and advanced methods of multi-objective calibration and data assimilation, considerable progress in large-scale hydrological modelling by integration of GRACE data can be expected.

Keywords

Satellite mission, time-variable gravity, continental water storage, global water cycle, water balance, hydrological model, land surface model, data assimilation, multi-objective validation

1. Introduction

By the GRACE satellite mission (Tapley et al. 2004a), time variations of the Earth's gravity field can be monitored for the first time with a resolution that integrates in space and time over a few hundred kilometres and over monthly down to 10-day time periods. Given the unprecedented accuracy of the GRACE data, mass variations on and below the Earth surface due to geophysical processes can be deduced from the gravity fields. Nearly six years of GRACE data have shown that water storage changes in continental hydrology are among the dominant mass variations that can be detected in the GRACE signal (e.g., Tapley et al. 2004b, Wahr et al. 2004; Schmidt et al. 2006b). Thus, a unique data source is available to quantify spatio-temporal variations of the Earth's water storage and to improve the understanding of the water cycle at regional up to global scales. Accurately closing the water budget at large scales can be considered an open problem and is subject of considerable current research (Swenson and Wahr 2006). Besides, it is obvious trying to use GRACE water storage changes in combination with hydrological simulation models for model evaluation and ultimately for model improvement. With enhanced models, benefits in terms of climate predictions, assessing the impact of environmental change on the water cycle and for water resources management for large areas can be expected.

Two main categories of hydrological models applied at the scale at continental to global scales are land surface models (LSMs), on the hand, and water balance models (WBMs) on the other hand. The purpose of LSMs basically is to represent the land surface in climate models and numerical weather prediction simulations (see Dirmeyer et al. 2006 for a LSM overview and comparison). LSMs usually represent the energy and water fluxes at the interface of atmosphere and land surface based on fully coupled heat and mass balance equations. Most LSMs, however, are confined to a limited depth of soil below the terrain surface and usually exclude model components for groundwater and water transport and storage in surface water bodies. Water balance models (see Widen-Nilsson et al. 2007 for a recent overview), in contrast, have mainly been developed for simulating streamflow at the outlet of river basins for purposes such as water resources assessment or flood forecasting. Usually, WBMs represent the terrestrial hydrological cycle in a more holistic way including all its components in order to close the water balance for the area of interest. In general terms, the continental water balance accounts for four major components, i.e., precipitation on the land surface is balanced by evapotranspiration, discharge and the temporal change in water storage. WBMs mainly use simplifying conceptual approaches to represent the processes of water fluxes and storage. With the overall goal to simulate vegetation dynamics and the land carbon cycle, also terrestrial biosphere models represent components of the global hydrological cycle with different degree of complexity (see Jung et al. 2007 for an overview and comparison). If not explicitly stated otherwise, the denotation *hydrological model* is used throughout this paper as a summary term for the different types of models introduced before.

Depending on their purpose, the hydrological models differ in terms of spatial and temporal resolution, detail in process representation, number of parameters and data demand. Common to all models at large scales, however, is that they suffer from considerable uncertainties in terms of model structure and process description, parameter values, and atmospheric forcing data used as model input. Consequently, simulation results for hydrological state variables and water fluxes on the continents vary considerably between the models (Dirmeyer et al. 2006).

A main reason for the model uncertainties is the lack of adequate data at the large spatial scales considered here. These deficiencies apply both to the data that are required to drive the models as well as to data needed to evaluate simulation results. Given that precipitation is the most important climate input variable to force hydrological models, unrealistic precipitation data are considered as one of the main factors causing deficient simulation results (Nijssen et al. 2001; Döll et al. 2003). Among the other components of the continental water balance that potentially could serve for model validation, evapotranspiration is not measured directly at large scales but may at the best be estimated with considerable uncertainties using indirect energy balance methods that rely on land surface temperatures determined from satellite data and on surface meteorological data (Roerink et al. 2000). Storage change could not be determined from measurements at large scales prior to the GRACE mission because existing monitoring systems provide either point observations that are hardly to transfer to larger areas, or cover only individual components of TWS on the continents. TWS is mainly composed of storage in the form of snow, ice, soil water, groundwater and surface water in rivers, lakes, wetlands and inundation areas. The term TWS is used in this sense throughout this paper (Ice mass variations of Greenland and Antarctica are not covered in this review).

The only observed water balance variable that has been available with acceptable accuracy and coverage for major parts of the global land surface was river discharge. As a consequence, hydrological models have mainly been validated at the global scale by comparing simulated to observed discharge for large river basins (Nijssen et al. 2001; Milly and Shmakin 2002; Döll et al. 2003; Widen-Nilsson et al. 2007; Hanasaki et al. 2007). However, showing that a hydrological model gives reasonable results in terms of discharge at the outlet of a river basin does not guarantee that the hydrological processes are correctly represented in the model and that water fluxes and storage are adequately partitioned between the different components of the water cycle. This can only be evaluated and improved by additional observation data on other hydrological state variables and fluxes. From this perspective it is evident that data on total continental water storage change as one of the four fundamental components of the continental water balance are extremely important to evaluate and enhance hydrological models. From this originates the unique potential that is seen with the availability of GRACE data for the field of hydrological modelling at large spatial scales.

The aim of this paper is to review the results that have been achieved during the first about six years of the GRACE mission in analyzing simulation results of global hydrological models relative to the GRACE data on continental water storage change (Chapter 4). A focus is on summarizing model deficits that have been revealed and model improvements that could be achieved by this comparison. In addition, the current expertise on the major conditions to be met for a consistent comparison and combination of GRACE data and hydrological models is summarized and perspectives towards the improvement of global hydrological models by the integration of GRACE data are outlined (Chapter 5). These chapters are preceded by short reviews on the state of the art in tuning global hydrological models with observation data (Chapter 2) and on basics about the processing and characteristics of water storage data from GRACE (Chapter 3). While the focus of this review is on global-scale modelling, it should be noted that many of the considerations addressed here are similarly applicable to hydrological models at regional scales or for individual large river basins.

2. Tuning of global hydrological models

For physically based land-surface models as introduced before, the parameter values of the model equations that represent water and energy fluxes on the continents can in principle be determined through direct observations. Global maps of, e.g., vegetation cover and soils are used for this purpose. However, a conceptual restraint of this approach is due to sub-grid variability of the environmental conditions. Therefore, setting a unique parameter value for a certain vegetation or soil property within an entire model grid cell may not allow for representing adequately its hydrological response. The sub-grid variability is either ignored in the parameterization of LSMs or represented by a distribution function that has to be adjusted to the estimated or observed sub-grid variability by a shape parameter (e.g., Liang et al. 1994). The more simple conceptual equations of water balance models (WBMs) often include parameters that cannot be directly inferred from observations and thus have to be specified independently. In addition, taking into consideration other sources of uncertainty such as of climate input data or model structure, it is obvious that one may need to tune or calibrate hydrological models in order to get reasonable simulation results that are close to observations. During the process of model calibration, repeated simulation runs with successively modified parameter values are performed until the simulation results correspond to the observations within in a predefined error criterion. Traditionally, river discharge is used for the calibration of hydrological models at the scale of river basins. As discussed above, discharge has been the only observed variable available in particular at large spatial scales. However, while most LSMs are not calibrated at all because, for instance, they do not include a routing approach that transforms runoff to river discharge by transport along the land surface and the river network, only few studies report on a tuning process of global-scale hydrological models against river discharge. Arnell 1999 (Macro-PDM) and Milly and Shmakin 2002 (LaD) adjusted model parameters in a globally uniform way while Nijssen et al. 2001 (VIC), Döll et al. 2003 and Hunger and Döll 2007 (WGHM), Hanasaki et al. 2007 and Widen-Nilsson et al. 2007 (WASMOD-M) performed a basin-specific calibration of one or more model parameters followed by a regionalisation approach to transfer parameter values to basins for which no discharge data were available. Mean annual river discharge was mainly used as a calibration value. Seasonal or inter-annual variations of river discharge were not directly taken into account during the tuning process except for Hanasaki et al. 2007 who adjusted their model to monthly streamflow time series. All studies agree in showing the value of discharge data for an overall improvement of the simulation results when considering a large number of observation points. Albeit at smaller spatial scale, also Wood et al. 1998 concluded from a comparison of several LSMs that model performance improved significantly for those models that were able to use river discharge for model calibration. Nevertheless, for the above studies at the global scale, model performance in terms of river discharge varied strongly between river basins. It was argued that the main reasons which precluded better model results were errors in the precipitation data used as model input, the lack of adequate data on runoff regulation by man-made reservoirs, and limitations of adequate process representation in particular in arid and semi-arid environments.

It should be noted that during the process of model calibration it is usually impossible to find a unique parameter set that gives the best simulation results, but instead several parameter sets give similarly good model results when evaluated by the observations in terms of a pre-defined performance criterion. This equifinality (Beven and Binley 1992) is usually due to the large number of parameters in a hydrological model that cannot be completely identified in an unambiguous way, given one observation variable or performance criterion only. This leaves the simulation results with considerable uncertainty and, as a drawback, it cannot be assured if all hydrological processes state variables are reasonably represented in the model. From this perspective, the goal of the calibration process is to "efficiently extract the information contained in the calibration data, rendering a reduction in model uncertainty" (Bastidas et al. 2002) by reducing the extent of the behavioural parameter space. Any additional independent information from observations may help to better constrain the model in form of a multi-objective model calibration or validation. In this context, water storage variations from GRACE are a promising observation variable besides river discharge.

Another approach to merge information from observations with hydrological models is by direct assimilation of the observed data into the model. In this approach, model state variables are continuously steered towards the observations. The degree of the corrections in the model is determined under consideration of the errors estimated for both the observations and the model state. Several observation variables have been assimilated into land surface models in recent years, such as land surface temperature, soil moisture, or snow cover extent (see a recent summary in Reichle et al. 2008 and the references therein). Well-defined error estimates for both model and observations are required to achieve superior model-based estimates of hydrological fields with than without data assimilation (Reichle et al. 2008). A prerequisite for data assimilation is the presence of a state variable in the model that corresponds to the observed quantity. This may be a limitation in assimilating water storage variations from GRACE because these data basically represent total continental water storage whereas in the models water storage usually is represented by different modules and state variables. Updating state variables in a data assimilation system may not preserve the water balance over the simulation period. While this is adequate for specific applications, it may be of limited value for an improved understanding of the hydrological cycle and of the closure of water budget at large scales.

3. GRACE data of continental water storage change

Fundamentals and the current status of the GRACE mission and of GRACE data processing is explained by Schmidt et al. 2008b (this issue) and the reader is referred to their review for details beyond the general outline given here. GRACE data are available from April 2002 until today. The most recent GRACE data generation based on the latest processing standards (release four "RL04") is provided by the three processing centres of the GRACE mission (GFZ, CSR and JPL) as an operational product in the form of global gravity fields with nominally monthly time steps. These fields are already reduced for several know gravity effects such as tides of the solid Earth, of oceans and the atmosphere, and non-tidal oceanic and atmospheric mass variations. Thus, formally, mainly the hydrological signal due to water mass variations on the continental area including mass variations of ice caps and glaciers can be expected to be left in the time-variable gravity fields. Other mass variation components still included in these reduced gravity fields are effects from post-glacial rebound and from seismic deformations of the Earth crust due to earthquakes. As required for hydrological applications, the global gravity fields represented by sets of spherical harmonic coefficients can be transferred into mass anomalies at the Earth surface with units of mm water equivalent, for instance.

Besides these operational global products, it should be pointed out that GRACE solutions based on alternative processing strategies from the raw GRACE mission data to the final product are subject of ongoing work (Han et al. 2005a; Han et al. 2005b; Rowlands et al.

2005; Luthcke et al. 2006; Schmidt et al. 2006a). Such alternative methods often focus on regional rather than global solutions. An alternative solution for global time-variable gravity fields was provided by an inversion approach of Ramillien et al. 2004 and Ramillien et al. 2005. While a higher accuracy of the alternative solutions has partly been stressed, there exists no final recommendation so far on the superiority of a certain GRACE product.

Errors in the GRACE data originate from (1) measurement errors in satellite instruments and orbit determination, (2) aliasing errors from imperfectly reducing non-hydrological mass contributions and (3) leakage errors from mass signals outside the area of interest (see Seo et al. 2006 for a comprehensive overview on the GRACE error budget in water storage estimates). The latter error term is due to the fact that the gravity acceleration felt by the GRACE satellites is an integral effect of all masses around the satellites, with the contribution of individual mass elements decreasing as a function of the squared distance to the satellites. Thus, in contrast to several other satellite products hydrologist may be familiar with, GRACE data cannot be attributed to a sharply delimited footprint area on the Earth surface. In fact, to extract mass variations for a region of interest such as a river basin from the continuous gravity fields, filter functions have to be applied (Swenson and Wahr 2002). These filters have to be optimized in a way such that they minimize at the same time (1) the contamination of the basin-average signal by signals from outside the region (leakage error) and (2) GRACE measurement errors which increase markedly with the spatial detail that is to be extracted from the GRACE fields. As a consequence, filters cannot follow exactly the boundary of the region of interest as this would involve unacceptably high GRACE errors, but they are gradual in space around the region and thereby incorporate some signal from the surroundings into the final basin-average estimates of water mass anomalies. A variety of filter techniques has been developed during the last years (see Kusche 2007 for a recent overview). In summary, so-called non-isotropic filter methods that also take into account correlations between the spherical harmonic coefficients of the gravity fields tend to be more effective in reducing errors in the GRACE data, such as North-South oriented striping. However, it is important to consider that the selection of an adequate filter tool may also strongly depend on the final field of application of the filtered GRACE data. This was demonstrated by Steffen et al. 2008 for an analysis of mass variations related to global isostatic adjustment where non-isotropic filters were considered less suitable than isotropic filtering. In a recent study tailored towards the requirements of hydrological applications, Werth et al. 2008 concluded that optimum filter types and parameters vary from river basin to river basin, depending on its geographical location, its shape, and the characteristics of the hydrological signals in and around the basin of interest.

Estimating errors for GRACE data of continental water storage change is a major challenge for ongoing GRACE processing activities (Schmidt et al. 2008b) and could not fully be achieved so far because no ground truth data sets on water storage with sufficiently large spatial extent, homogeneity and support (e.g. station density and monitoring coverage of all water storage compartments) exist. Instead, error estimates are constructed from the GRACE fields alone, including measurement errors but also simulation studies based on the underlying physics of data acquisition and transformation to include aliasing and leakage errors. In general terms, GRACE errors depend on latitude, i.e., tend to increase from poles to low latitudes (Wahr et al. 2006), they increase with decreasing size of the region of interest (Seo et al. 2006), increase with decreasing integration time of the solutions, and increase with decreasing similarity of the hydrological signal in the surroundings of the region of interest relative to the signal within that region (Swenson et al. 2003). Horwath and Dietrich 2006 point out that error correlations in the GRACE monthly gravity fields strongly affect error estimates for regional mass variations and cause higher errors in particular at high latitudes. As an order of magnitude, Wahr et al. 2006, using a filter function with 750 km smoothing radius, showed a global area-weighted average error for monthly gravity fields of 21 mm when expressed in water column equivalents. Schmidt et al. 2007 gave for similar filtering a global error estimate of 16 mm. Note, however, that errors may be higher for particular regions and months, and due to leakage errors after filtering to basin-average values.

Another important consequence of the filter applications to GRACE gravity fields is that filtering alters the variability in the resulting signal. In particular for most hydrological applications, filtering has a smoothing effect that reduces the seasonal amplitude of the final TWS change signal and may also shift the phase of annual variations (Chen et al. 2007; Klees et al. 2007). As an example, Chen et al. 2007 found for the application of a Gaussian filter with 800km filter radius (which is a standard procedure in GRACE processing) that the seasonal amplitudes of basin-average water storage variations were reduced by 25% in the Amazon and Mississippi basins, and 35% in the Ganges and Zambezi basins. The magnitude of amplitude attenuation depends on the water storage variations inside and outside of the area of interest and on the filter type and its parameters. In addition, as the spatial patterns of water storage variability change with time (Winsemius et al. 2006) so will do the filter-induced bias. With increasing filter radius, the amplitudes tend to be biased towards lower values. Without correcting for this bias, “it is hardly possible to enhance hydrological models using GRACE” (Klees et al. 2007).

As a consequence, several studies proposed correction factors to re-scale biased water storage amplitudes for selected areas (Chen et al. 2007; Swenson and Wahr 2007; Klees et al. 2007). A limitation is that filter biases are strongly region-dependent, and, thus, general functions for the bias-correction cannot be given. Klees et al. 2007 suggested using reliable regional information such as from a well calibrated regional model to estimate the bias correction factor. In many cases, however, this information is not available. Instead, they demonstrate that taking the uncertain information of a global hydrological model as a-priori information for bias correction is still preferred to a case where no correction is performed at all. Alternatively, to make water storage data of hydrological models comparable to GRACE, the model data can be treated in the same way as it was done with the GRACE data, i.e., transforming the model data into a spherical harmonic representation as common for global gravity fields and subsequently extracting the region-average water storage variations with the same filter that was used for the GRACE data. This approach was followed by several of the comparison studies described in the next chapter.

4. Comparison of global hydrological model results with GRACE data

4.1 General overview

In a first phase after the start of the GRACE mission, the focus of studies that analyzed GRACE data in conjunction with hydrological model data was to explore the extraordinary value of observation data that became available with GRACE, and to illustrate that GRACE is sensitive to mass variations in continental hydrology (see Table 1 for an overview on the studies). In this context, water storage change as simulated by the models was used as a reference, but no attempts were made in reverse towards an in-depth analysis of the hydrological models themselves in terms of modelling approaches or uncertainties, for instance. This similarly applies for a second group of GRACE publications where the focus was on specific topics of GRACE data processing, such as regional solution strategies or filter techniques. Nevertheless, the fundamental benefit of these studies was that continental water storage was presented and discussed as a state and output variable of global hydrological models and, thus, a general idea about how the models simulate spatio-temporal fields of water storage was given. This has rarely been done before because the interest of earlier studies usually was in other variables such as river discharge or fluxes between the land surface and the atmosphere.

In some more recent studies, the focus was moved from geodetic aspects of GRACE data processing or pure detection and analysis of hydrological signals towards an explicit evaluation of hydrological models by means of GRACE data. However, the number of these studies with a hydrological modelling perspective is still small (Table 1). Water storage changes simulated with the land surface schemes of five climate models and with Global Land data Assimilation System (GLDAS) were evaluated by means of GRACE data in the studies of Swenson and Milly 2006 and Syed et al. 2008, respectively. Niu and Yang 2006a included an additional step towards model improvement by comparing both a standard and a modified version of the NCAR Community Land Model (CLM) to GRACE data. In the modified model version, more realistic process descriptions concerning interception, runoff generation and frozen soil dynamics were incorporated. The value of these modifications could in general be justified by finding a better correspondence with GRACE data for the updated as compared to the original model version. Similary, Ngo-Duc et al. 2007 concluded that an improved version of the ORCHIDEE LSM could be achieved after incorporation of a transfer scheme to route runoff components through reservoirs with different residence times within river basins. This result was based on the observation that the modified model led to predicted water storage variations that were comparable to GRACE, contrary to the original model version.

Many studies performed a simple qualitative comparison of GRACE versus model data, e.g., by visually comparing global maps or basin-average time series of water storage anomalies. In line with a more detailed comparison of different modelling approaches or different GRACE processing strategies, however, quantitative measures were applied to describe the correspondence of model and GRACE data. Among these were root-mean-square differences between GRACE and model estimates of monthly water storage (Chen et al. 2005a; Schmidt et al. 2006b; Swenson and Milly 2006; Klees et al. 2007), bias and RMS differences between simulated and GRACE-based evapotranspriation, correlation maps between monthly patterns of GRACE and model data (Ngo-Duc et al. 2007), summarized also by spatial correlation coefficients (Andersen and Hinderer 2005), or Taylor diagrams that evaluate jointly standard deviations and correlations of observed and modelled basin-average water storage time series

(Ngo-Duc et al. 2007). After a principal component analysis, Rangelova et al. 2007 used summed differences in loading patterns and correlation coefficients of principal component time series for a quantitative evaluation of GRACE and model consistency.

The comparisons of water storage variations from GRACE and from global hydrological models in general showed that GRACE-based data agreed reasonably well with model predictions in terms of the general seasonal dynamics and their continental-scale spatial patterns. The most pronounced seasonal signals were found for northern South America, South-East Asia and the Ganges/Brahmaputra river basin, tropical Africa north and south of the Equator and over larger parts of northern Eurasia and North America. These signals were similarly represented in the models when compared in a qualitative way. On the other hand, also some obvious differences were found, such as differences in the amplitude and phase of seasonal water storage variations. Concerning differences in the spatial patterns, north-south oriented stripe-like features in the GRACE data not present in hydrological data were most frequently mentioned in the literature. These could be explained as error artefacts in the GRACE data and can be attenuated by adequate filter techniques (see Chapter 3).

In general, correspondence between model and GRACE data tends to degrade if going to smaller spatial scales (Seo et al. 2006). This mainly points out the limits of GRACE applicability due to the increasing signal degradation by GRACE errors with increasing resolution. For example, Klees et al. 2007 stated that the bias due to filtering increases for smaller target areas. Rodell et al. 2007b concluded for the Mississippi basin that extraction of groundwater storage variations from GRACE worked reasonably well for areas larger than 900000 km² but was poor for smaller areas at the sub-basin scale.

While most studies compared GRACE data to hydrological model output in terms of continental water storage variations, some studies evaluated other water balance components (evapotranspiration) or storage variations in individual compartments such as snow and groundwater (Table 1). A critical limitation when extracting a particular storage compartment is the need to subtract other water mass variations from the integral storage signal that is provided by GRACE (Güntner et al. 2007a). Except for rare cases (Yeh et al. 2006), adequate in-site observation data usually are not available at the required spatial scale and thus uncertain model data have to be used for the reduction. Consequently, the final data are not strictly observation-based anymore and, ultimately, using this information again towards model improvement may be of limited value unless an adequate error assessment is done in parallel. In this context, Rodell et al. 2007b showed for the Mississippi basin that reasonable seasonal groundwater storage variations can be inferred from GRACE when taking the model uncertainty ranges into consideration. In their study, uncertainties were estimated from the differences between three LSMs that were used for subtracting snow and soil water storage from the GRACE TWS signal.

Alternatively, an iterative inverse approach for the decomposition of global GRACE gravity fields into individual components (Ramillien et al. 2005) has been used by Frappart et al. 2006 to isolate snow storage changes from the integral GRACE signal. Initial fields of the components of interest are required as a first guess for the inversion approach. As no adequate observation data are available, these initial fields usually have to be based on hydrological model results. Nevertheless, the fact that according to Frappart et al. 2006 the final GRACE-based snow mass variations agreed better to other models (GLDAS, LaD) than to the model that was used for the first guess (WGHM) proofs the potential of this method for signal separation.

With TWS change from GRACE and observed precipitation (P) and discharge data, the continental water balance equation at the scale of river basins can be solved for evapotranspiration (ET). Rodell et al. 2004 were the first to show the possible skill of GRACE-based ET estimates to evaluate simulated ET from hydrological models. While several precipitation products are available at the global scale albeit subject to high uncertainties (e.g., Sheffield et al. 2006), data availability of recent discharge data in the

GRACE period is limited in many parts of the world and restricts the applicability of this method. Therefore, Ramillien et al. 2006 used simulated river discharge which in turn incorporates considerable uncertainties into the results. As a further alternative, Swenson and Wahr 2006 pointed out the possibility to use P-ET from water balance calculations (with TWS change from GRACE and observed discharge) to validate model output. While P-ET simulations of a climate model corresponded well to the observation-based data, they found a systematic overestimation of P-ET in the results of a LSM (GLDAS-Noah) which they attributed to errors in the precipitation data used to force the model.

An alternative method to assess TWS change is by the combined atmospheric-terrestrial water balance approach (Seneviratne et al. 2004; Hirschi et al. 2006). It circumvents the need to explicitly estimate precipitation and evapotranspiration by resolving for TWS change from atmospheric data on the change of water content and water vapour flux divergence in the atmosphere, and from observed river discharge. The accuracy is limited by errors of the atmospheric data obtained from atmospheric circulation models. Nevertheless, the method is appealing because it is an independent way of estimating TWS changes, and some studies included these estimates in addition to GRACE and hydrological model TWS data into the comparison (Table 1). Finding very good correspondence of TWS time series from this water balance approach with those simulated with a regional hydrological for the upper Zambezi basin in Africa, Winsemius et al. 2006 were confirmed to conclude that inconsistencies between model and GRACE data could be attributed to GRACE errors. An alternative way to solve the combined atmospheric-terrestrial water balance is for discharge with water storage change from GRACE. Syed et al. 2007 used that approach to estimate large-scale runoff into the Arctic Ocean. In the absence of comprehensive observation data of river discharge at the continental scale, these may in turn be used to evaluate runoff simulations of hydrological models.

Besides TWS changes from GRACE or the combined water balance approach described in the last paragraph, only few studies used additional observation data to assess model performance in an even broader sense. Niu and Yang 2006a and Niu et al. 2007b were the first to evaluate model modifications against water storage from GRACE and additionally against a hybrid model-observation-based runoff climatology (Fekete et al. 2002). They showed closer agreement of the modified model with respect to both variables.

4.2 Lessons learned from differences between GRACE and hydrological model data

As a valuable starting point towards an improvement of hydrological models, the differences found in the comparison of simulation results and GRACE-based observations can give clues about existing model deficits. To this end, this chapter summarizes in more detail the inconsistencies found in previous studies and the explanations given for them in combined investigations of GRACE and hydrological models. Besides model deficiencies, differences are also due to limitations of the GRACE monitoring technique and to data errors. While a general overview is given in Chapter 3, some more specific points will be mentioned in the following to further illustrate the limitations that modellers may be confronted with when using GRACE data.

When interpreting discrepancies between model and GRACE data, signal leakage is one of major factors that have to be taken into account in terms of GRACE data uncertainty. Niu et al. 2007b stressed the strong corruption of the TWS change signal of the Orinoco basin by leakage from the dominant Amazon. Similarly, noticeable disagreement found in arid regions between model predictions and GRACE (Ngo-Duc et al. 2007) can be attributed to leakage from strong storage signals from surrounding areas, but also to GRACE data errors that are high in relative terms in such areas where the real signal is small. Winsemius et al. 2006 argued that in addition to leakage effects from surrounding areas due to filtering, regional

TWS changes extracted from the global gravity fields may be biased because of aliasing effects caused by irregularly spaced passes of the GRACE satellites over the target area or its vicinity. For a regional study in the Zambezi river basin in Africa they showed by means of GRACE ground tracks that in some months mainly areas outside their target area with different water storage conditions than in the river basin of interest (e.g., induced by intense rainfall events) were directly covered by the satellite tracks. Building also on a high confidence in their regional hydrological model that was fed with regional observation data, they consequently attributed differences between modelled TWS change and GRACE data to these limitations of GRACE data coverage for their regional hydrological application. As another factor for poor correspondence between GRACE and model data on the side of the satellite mission, the lower accuracy of GRACE data in the first period of the mission until mid of 2003 was quoted (Rangelova et al. 2007).

In very general terms, many studies attributed differences in amplitude or phase of seasonal storage variations to the fact that one or more water storage compartments were missing in the model they compared to GRACE data (e.g., Tapley et al. 2004b). According to Schmidt et al. 2006b, for instance, lower seasonal amplitudes can be due to the absence of surface water storage in the model which may be an important component for TWS in several river basins (Güntner et al. 2007b). For the GLDAS-Noah LSM, Syed et al. 2008 suppose that a smaller simulated magnitude of water storage changes and less variability at the monthly scale can at least partly be attributed to neglecting soil moisture below 2 meter depth, groundwater and surface water in the model, apart from GRACE errors. Similarly, Zeng et al. 2008 argued that underestimated seasonal storage amplitudes are due to the limitation of their LSMs to the uppermost soil zone. An earlier phase of seasonal storage dynamics in the model may indicate that storage components with a delayed dynamics are missing in the model, such as groundwater (Chen et al. 2005a, Yamamoto et al. 2007). Rangelova et al. 2007 found for

GLDAS which did not include surface water and groundwater storage that its annual cycle preceded that of GRACE by one month in North America. Depending on the model and the region, the phases from the hydrological models were found to be about 1 to 6 weeks ahead of the GRACE estimate in the study of Schmidt et al. 2008a which was attributed to similar deficiencies in the hydrological models such as the lack of a surface water storage module.

In their comparison of the simulation results of five climate models, Swenson and Milly 2006 showed that all models reasonably reproduced the seasonal patterns of water storage changes as seen by GRACE when analyzed qualitatively at the global scale. A more detailed analysis, however, exhibited significant regional model failures that were averaged out when looking at global means. For example, the different magnitudes of storage amplitudes for the outstanding tropical areas of South America, Africa and South-East Asia were not correctly reproduced by the models, with each model outperforming other models in different areas. These deficiencies could partly be explained by errors in the precipitation data. In addition, systematic model biases at low latitudes were found in terms of too early seasonal storage maxima, while the high-latitude storage variations tend to be better reproduced. The absence of a river runoff routing component and of a representation of inundation storage during high flows that delays water export from a river basin was highlighted by Swenson and Milly 2006 as a possible reason for early storage maxima in some of the models relative to GRACE. Also, the huge water masses stored seasonally in inundation areas along the Amazon river that were disregarded by all models may cause the underestimation of storage amplitudes in particular for that region.

Nevertheless, underestimated seasonal storage amplitudes cannot in all cases simply be explained by missing storage compartments in the hydrological models. For example, a net compensation for a missing component may already be implicitly included in the parameterization of a model by increasing water residence times in one of the existing storage components (Swenson and Milly 2006), and interferences of different compartments in terms of storage volume and phasing can even cause a decrease of the overall seasonal amplitude of TWS, such as by delayed groundwater variations (Güntner et al. 2007a, Güntner et al. 2007b).

Among the first studies that were directly dedicated towards the modification of a hydrological model including its validation with GRACE data, Niu and Yang 2006b developed a new version of CLM with a modified approach to model frozen-soil processes. They evaluated the new relative to the original model version by mean of GRACE-based TWS variations for several large river basins in high latitudes. Niu and Yang 2006b concluded from the comparison that model performance improved. A main reason was that more snow melt water was allowed to infiltrate into the soil, leading to a shift of the seasonal maximum of water storage by about one month which corresponded better to the observed TWS dynamics from GRACE. However, the full validation of the model modifications was restricted by the fact that the model did not include a surface water storage component. This component may cause a similar shift in water storage dynamics at the basin scale than the one produced by the snow model modifications. Only if the better model agreement to GRACE data can be unambiguously devoted to the modifications of the frozen-soil algorithms, it can be stated from a process-based perspective that a real model improvement has been achieved. Additionally, similar to several other comparison studies (see Table 1), model data in the study of Niu and Yang 2006b were not filtered in the same way as it was done for GRACE. This makes the data sets hardly comparable, at least not to an extent that is required for detailed model validation. As described in Chapter 3, filtering can significantly bias the amplitude and phase of GRACE water storage variations. In this context it should also be pointed out that the good agreement of seasonal storage amplitudes between filtered GRACE data and unfiltered model data that was found in some studies may be caused by two bias contributors cancelling out to some extent in the final filtered GRACE signal in certain regions (Klees et al. 2007).

Differences between model results and GRACE data for TWS variations are often attributed to errors in precipitation forcing data (e.g., Chen et al. 2006, Swenson and Milly 2006). This is an obvious and plausible rationale because the existence of considerable uncertainties and biases in large-scale precipitation data is known (e.g., Milly and Dunne 2002), and precipitation variability usually has a much larger impact on hydrological processes and the hydrological budget than other meteorological variables (Ngo-Duc et al. 2005, Crowley et al. 2008). It is also in line with the reasoning behind unsatisfactory simulation results of global hydrological models when validated by river discharge data (see Chapter 2). Towards an improvement of future model applications, a more explicit analysis that clearly traces errors in simulated TWS or other water balance components back to the precipitation input is required. For GRACE-related studies, an example is given by Swenson and Wahr 2006 which by comparing four precipitation products for the Mississippi basin could show that a model bias revealed by GRACE-based P-ET estimates was due to overestimated precipitation in the model forcing data they used.

For high-latitude areas with a dominance of snow accumulation and melt processes, Ngo-Duc et al. 2007 partly explained deficiencies of simulated storage variations relative to GRACE also by errors in air temperature as atmospheric forcing data.

Apart from seasonal dynamics that were the focus of most comparison studies, inter-annual variations, secular trends and non-periodic mass changes from GRACE have been investigated in conjunction with hydrological models in some first studies (Table 1). Clearly, this type of analysis is still limited due to the short life time of the GRACE mission so far. Andersen and Hinderer 2005 analyzed the inter-annual gravity changes between the first two years of the GRACE mission (2002-2003) and found that among four hydrological models GLDAS water storage variations had the highest correlation with the GRACE inter-annual signal. Possible reasons for the performance differences between the models were however not further explored. For a strong drought situation in Central Europe in 2003 with low water storage, Andersen et al. 2005 found a good agreement between GLDAS estimates and GRACE data. In contrary, for the year 2002 where extreme floods hit Central Europe in summer, poorer correspondence was found. According to the authors this may be an indication that extreme floods are more difficult to represent in hydrological models than drought situations. Although not based on GRACE data, Zeng et al. 2008 point out that land surface models may considerably underestimate inter-annual variations of water storage because of insufficient water storage capacity in the soil and groundwater zone and deficits for extracting deep soil water for evapotranspiration. Given longer GRACE observation periods in future, related sub-surface and vegetation parameters may be adjusted.

In a Principal Component Analysis for North America, Rangelova et al. 2007 revealed long-term mass changes in hydrological models that found good correspondence in the GRACE data, although the latter may partly be contaminated by remaining signals from imperfectly removing other mass contributions, in particular due to post-glacial rebound. Interestingly, also specific non-periodic mass variations appeared in both GRACE and model principal component time series. A relation to precipitation anomalies and flooding was suggested. The observed time shift between GRACE and hydrological models for these anomalies may in future analyses give further insight into the appropriateness of representing such events in the hydrological models. For parts of Eastern Europe, Steffen et al. 2008 revealed secular trends in the GRACE data that were of a size that could not be attributed to GRACE errors or to other insufficiently reduced geophysical signals such as from glacial isostatic adjustment. As these trends were not present in the hydrological models WGHM and LaD either, Steffen et al. 2008 argued that this gives and indication that hydrological models may have to be improved for representing secular components. Based on a principal component analysis, Schmidt et al. 2008a figured out besides semi-annual and annual variations significant longer-term periods in both GRACE and hydrological model data for several large river basins worldwide.

The joint analysis of simulation results of several hydrological models relative to GRACE data can be in particular efficient to identify on basis of the different model performances any advantages or deficits of a certain model and the particular reasons for it (see, for instance, the comparison of five climate models by Swenson and Milly 2006 discussed above). In another example applying a principal component analysis of water storage variations in North America on three hydrological models besides GRACE, Rangelova et al. 2007 could clearly relate the different EOF pattern of one model (CPC) in the first mode to its missing snow component. For higher modes, however, this model showed a higher correlation to the GRACE patterns than GLDAS. Such differences may be a starting point to identify in future studies those model components that may improve long-term simulation capabilities. Comparing three hydrological models in their ability to represent northern hemisphere snow storage variations, Frappart et al. 2006 supposed that worse results of WGHM compared to LaD and GLDAS may be due to the simple snow algorithm used in WGHM that does not explicitly solve for the energy balance of the snow cover, contrary to the other models. For target areas in southern Africa, Klees et al. 2007 found a superior quality of a regional hydrological model (LEW) compared to a global model (CPC) when evaluated against GRACE storage data. This can similarly be expected for other study areas where a regional model has been adapted or calibrated to the specific hydrological conditions of that area, such as with regard to a higher spatial resolution and the inclusion of surface water transport, storage and human water use in the study of Klees et al. 2007.

5. Conclusions and perspectives

Based on the information content for continental water storage variations that has been revealed from GRACE data during the first about six year of the satellite mission, and based on the experiences so far on comparing these data with outputs from hydrological models, it can be concluded similar to earlier studies (e.g, Lettenmaier and Famiglietti 2006, Niu and Yang 2006a, Swenson and Milly 2006, Ngo-Duc et al. 2007, Syed et al. 2008) that GRACE-based hydrological data have the potential for improving global hydrology models. As seen from the review in this paper, partial steps have been made into that direction by (1) demonstrating deficiencies of the model output relative to GRACE data, (2) identifying sources of these model deficiencies in terms of model input data, model structure or model parameters, and by (3) modifying hydrological models and evaluating the modifications against GRACE observations. In particular for the last aspect, only first attempts in very few studies have been made so far.

To progress further towards a comprehensive and consistent improvement of large-scale hydrological models, the studies so far have highlighted several prerequisites that have to be fulfilled if GRACE data should be adequately used in conjunction with hydrological models, as summarized in the following:

- *Consistency of water storage components*

  A sound comparison or integration of hydrological model data and GRACE data can only be achieved if the variables described in the two data sets are congruent. As GRACE data basically give the integral value of water storage variations in all storage components on the continents while models often represent only a selection of storage components, several of the comparisons listed in the previous chapters are flawed and do not allow for a full evaluation of model results. A consistent comparison in future studies has to be achieved by (1) using a model that incorporates the full set of storage components to simulate total water storage change as in the GRACE data, (2) separating the storage component to be improved in the hydrological model from the integral GRACE signal, or by (3) excluding components from the model if reliable information exists that supports the assumption that these components do not significantly contribute to water storage variations in the area of interest. The choice of an approach will depend on the study area, data availability and the objectives of the study.

- *Consistency of data filtering*

  For a fair comparison of water storage variations, hydrological model data have to be treated in the same way as the GRACE data. While filtering is necessary to reduce noise in the GRACE data and to extract area-average storage change data for a region of interest, several studies have shown that filtering may considerably modify the signal. A comparison of filtered GRACE data with unfiltered model data is therefore of limited value for model improvements. Data consistency can be achieved by (1) applying correction factors to re-scale for the filter-induced bias or (2) by filtering hydrological model data in the same way as GRACE data. When working at the scale of river basins as for most hydrological applications, it has to be noted for both approaches that a river basin may not be seen separately from others because signal modification due to filtering also depends on the variability in the surroundings of the target area. While these spatial dependencies can approximately be captured when using a global-scale model, limitations exist for regional studies if the hydrological model covers only the river basin of interest.

- *Error estimates for GRACE data*

  The uncertainty ranges of GRACE data must be smaller than differences between hydrological models to ensure that GRACE data are useful for validation and improvement of modelling schemes. Furthermore, advanced techniques of data assimilation and model calibration require reasonable error estimates in addition to the observation data themselves. Error budgets for hydrological data from GRACE should encompass various components, i.e., satellite measurement errors, uncertainties due to different solution techniques, errors due to imperfect signal separation, leakage and filtering.

- *Consistency of the water balance – multi-criterial evaluation*

  While for certain purposes it may be sufficient to enhance the model quality in terms of water storage simulations, hydrological model improvement ideally seeks for a broader objective, i.e., an improved representation of the water cycle as a whole. Most studies summarized above, however, compared only a single model variable to GRACE data, mainly water storage change, in few studies basin-average evapotranspiration. While adjusting the model to the GRACE-based storage variations, model performance for other variables such as runoff may degrade, and/or, by chance or any compensation effect, the ostensible model improvement was due to the wrong reason in terms of the parameters or process descriptions that have been adjusted in the model. Thus, a multi-criterial evaluation of simulation results is required in future studies to check for a consistent simulation of all water balance components. River discharge may be a primary choice in this respect because of its widespread availability (although still limited at the moment for the recent years of the GRACE mission) and its integrative nature at the outlet of a river basin. Other variables to be added in the multi-criterial framework at large spatial scales are remote sensing based data sets on evapotranspiration, snow cover, surface water storage or surface soil moisture, for instance. Being related to individual storage components or water fluxes, this type of information is particularly valuable when compared to the respective simulation results in order to cross-check if real model improvements from a process-based perspective have been achieved.

At the time of writing, two preliminary studies were available that point out the direction of future work going beyond the mere comparison of GRACE and hydrological model data towards a full integration of GRACE data into the modelling process: For the global water balance model WGHM, Güntner et al. 2007c developed a multi-objective calibration approach to adjust model parameters in a way that model performance improves at the scale of large river basins both for water storage variations from GRACE and for observed river discharge. In another study, Rodell et al. 2007a assimilate water storage data from GRACE into the Catchment Land Surface Model by updating a catchment water deficit variable in the model by Ensemble Kalman Filtering. The value of both combination strategies for model improvement also in view of some critical issues raised about the methods in Chapter 2 will have to be discussed in future studies. Some other main topics that will be a focus of future work towards hydrological model improvement with GRACE data can be summarized from the review as follows:

- *Signal separation:* Separating the integral mass variation signal observed by GRACE into individual signal components is fundamental for hydrological applications. Future work should be devoted to (1) a more accurate separation of unwanted non-hydrological mass contributions from atmosphere, ocean and dynamics of the Earth's interior and (2) the separation of water mass variations within the field of continental hydrology by auxiliary ground-based and remote sensing observations. Information on individual storage components will be extremely valuable towards improving hydrological models efficiently for individual process descriptions. Progress in signal separation requires combined analyses using new observations, modelling, inversion techniques (e.g., Ramillien et al. 2004) and advanced statistical methods for identifying dominant and significant signal components (e.g., Schrama et al. 2007; Schmidt et al. 2008a).
- *Multi-model comparisons:* Previous studies have shown that in particular comparisons that took into account a larger number of hydrological models were helpful to identify uncertainties and deficiencies of individual data sets or model structures and to get clues about preferences of certain approaches for specific conditions. As proposed by Syed et

al. 2008, future work should for example be directed to compare GRACE observations with multiple hydrological models with and without explicit representation of certain storage components to fully characterize their role in TWS variations. Model intercomparison studies such as GSWP (Dirmeyer et al. 2006) may explicitly include continental water storage variations in their investigations of variability among the models and try to identify how this relates to model approaches and parameters.

- *Longer time scales:* Given the longer GRACE observation period that will be available with ongoing mission operation and the gradual improvement of gravity models, inter-annual variations, non-periodic events and secular trends on continental water storage will be within reach of being significantly detected by GRACE. As noted by Zeng et al. 2008, even if a model simulates a good seasonal cycle (as it was the focus of most studies so far), this does not guarantee a good simulation of inter-annual variability. GRACE observations allow in future using characteristic longer-term dynamics of water storage as an additional means of improving hydrological models.

Recognizing on the hand that hydrological models of different types are far from being comprehensive and reliable enough to represent reliably the continental water cycle and, on the other hand, that extracting hydrologically relevant data from the GRACE gravity mission in a best possible way is not a finished process, there is huge amount of potential for progress to be explored in future. Best progress may be expected from a combined and iterative approach of hydrology, geodesy and other geophysical disciplines together, where application-specific requirements and feedbacks are combined towards optimized GRACE processing in terms of regional solutions, filtering, or error assessment, for instance, signal separation, and ultimately an improved understanding and model-based representation of the global water cycle.

Tables

Table 1. Studies comparing simulation results of hydrological models with GRACE data

Study	Focus of study	Hydrological model	GRACE data	GRACE filter	Model filter	Study area	Temporal variability	Other variables for comparison?
Andersen et al. 2005	TWS	GLDAS	CSR	Gauss	+	Central Europe	Inter-annual	CATWB Gravimeter
Andersen and Hinderer 2005	TWS	Au&Chao CPC GLDAS LaD	CSR	Gauss	+	Global	Inter-annual	
Boronina and Ramillien 2008	ET	ORCHIDEE WGHM GLDAS LaD	Ramillien et al. 2005	Truncation	-	Chad aquifer	Monthly	ET from energy balance
Chen et al. 2005a	Processing	GLDAS-Noah	CSR	Gauss	-	Amazon, Mississippi, Zambezi, Ganges, Ob, Victoria	Monthly	
Chen et al. 2005b	Filter	GLDAS	CSR	Gauss	+	Global	Monthly	
Chen et al. 2006	Filter	GLDAS-Noah	CSR-RL01	Own	+	Global, Alaska, South America	Monthly	
Chen et al. 2007	Filter	GLDAS-Noah	CSR-RL01	Gauss	+	Amazon, Mississippi, Ganges, Zambezi	Seasonal	
Fengler et al. 2007	Filter	CPC LaD	CSR	Wavelets	+	Global	Monthly	
Frappart et al. 2006	Snow	WGHM GLDAS-Noah LaD	Ramillien et al. 2005, CSR	Truncation	+	High northern latitude	Monthly	Snow from SSM/I
Güntner et al. 2007a	GW	WGHM*	GFZ	Gauss	+	global	Monthly	
Han et al. 2005c	Filter	CPC GLDAS	CSR	Own	+	Global	Monthly	
Hu et al. 2006	TWS	GLDAS CPC	n.s.	Gauss	+	Global and Yangtze	Monthly	
Klees et al. 2007	TWS, filter	LEW CPC	GFZ-RL03	Gauss	+	Zambezi and Southern Africa	Monthly	
Lettenmaier and Famiglietti 2006	TWS	VIC	n.s.	n.s.	n.s.	Mississippi	Monthly	CATWB
Luthcke et al. 2006	Processing	GLDAS	own	-	n.s.	global	Monthly	

Ngo-Duc et al. 2007	Model	ORCHIDEE	Ramillien et al. 2005	Truncation	+	Global and 8 large basins	Monthly	
Niu and Yang 2006a	Model	GLDAS-CLM	CSR	different	-	Global and selected basins	Monthly	Q
Niu et al. 2007a	Snow	GLDAS-CLM*	CSR-RL04	Seo et al. 2006	+	High northern latitude, Lena, Ob, Mackenzie, Yenisei	Monthly	Snow from AMSR-E
Niu et al. 2007b	Model	GLDAS-CLM	CSR	different	n.s.	Global and selected basins	Monthly	Q
Niu and Yang 2006b	Model	GLDAS-CLM	CSR	different	n.s.	Six largest northern-latitude basins	Monthly	Q
Ramillien et al. 2005	Processing, TWS, ET	LaD WGHM	own	truncation	+	global	Monthly	
Ramillien et al. 2006	ET	WGHM LaD GLDAS ORCHIDEE	Ramillien et al. 2005	Truncation	-	16 basins worldwide	Monthly	
Rangelova et al. 2007	TWS	CPC GLDAS LaD	CSR-RL01	Own and Gauss	+	North America	Monthly	
Rodell et al. 2007b	GW	GLDAS/CLM* GLDAS/Noah* GLDAS/Mosaic*	CSR	Chen et al. 2006	n.s.	Mississippi	Monthly	Groundwater level
Rowlands et al. 2005	Processing	GLDAS	Own (Mascons)	-	n.s.	South America	10 days	
Schmidt et al. 2006a	Processing	LaD	Own	-	+	Amazon	Monthly	
Schmidt et al. 2006b	TWS	CPC LaD WGHM	GFZ	Gauss	+	global	Monthly	
Schmidt et al. 2008a	TWS	WGHM LaD GLDAS	GFZ-RL04	Gauss	+	global, Amazon, Mississippi, Ganges	Semi-annual - interannual	
Schrama et al. 2007	Processing	GLDAS	CSR-RL04	own	-	global	Semi-annual	
Seo et al. 2006	Filter	GLDAS	CSR, GFZ	various	n.s.	Global, 12 basins	Monthly, interannual	
Steffen et al. 2008	GIA	WGHM LaD	CSR, JPL, GFZ RL04	Gauss	+	Fennoscandinavia	Secular trends	
Swenson and Milly 2006	Model	GFDL CCSR-NIES MRI	CSR	Gauss	+	global	Monthly, seasonal	
Swenson and Wahr 2006	P-ET, ET	NCEP GLDAS-Noah	CSR	Swenson and Wahr 2002	-	Mississippi	Monthly	
Syed et al. 2008	TWS	GLDAS-Noah	CSR-RL01	Gauss	+	Global and large basins	Monthly,	CATWB

							seasonal	
Tapley et al. 2004b	TWS	GLDAS	CSR	Gauss	+	Global, South America	Monthly, seasonal	
Wahr et al. 2004	TWS	CPC	CSR	Gauss, Swenson and Wahr 2002	+	Global, Mississippi, Amazon, Ganges	seasonal	
Werth et al. 2008	Filter	WGHM GLDAS LaD	GFZ-RL04	several	+	22 large basins worldwide	monhtly	
Winsemius et al. 2006	TWS	LEW	GFZ-RL03	Gauss	+	Zambezi	monthly	CATWB
Yamamoto et al. 2007	TWS	SiB+TRIP	CSR-RL02 GFZ-RL03 JPL-RL02	Swenson and Wahr 2002	n.s.	4 basins in Indochina	monthly	
Zeng et al. 2008	TWS	SLand CPC CLM	n.s.	Gauss	-	Amazon, Mississippi	monthly	CATWB, Q

Explanations:

n.s.: not specified

Column 2: Focus of study:
TWS: Estimation of total continental water storage change
ET: Estimation of basin-average evapotranspiration
GW: Estimation of groundwater storage change
Snow: Estimation of groundwater storage change
Processing: GRACE data processing
Filter: Development or analysis of filter methods

Column 3: *: model used only for signal separation, not for comparison

Column 6: Model filter: +: Same filter for model data as for GRACE data used
-: Model data not filtered

Column 9: CATWB: Water storage change from combined atmospheric-terrestrial water balance
Q: River discharge

References

Andersen OB, Hinderer J (2005) Global inter-annual gravity changes from GRACE: Early results. Geophys. Res. Lett. 32, L01402, doi:10.1029/2004GL020948

Andersen OB, Seneviratne SI, Hinderer J, Viterbo P (2005) GRACE-derived terrestrial water storage depletion associated with the 2003 European heat wave. Geophys. Res. Lett. 32, L18405, doi:10.1029/2005GL023574

Arnell NW (1999) A simple water balance model for the simulation of streamflow over a large geographic domain. Journal of Hydrology 217:314-335

Bastidas LA, Gupta HV, Sorooshian S (2002) Emerging paradigms in the calibration of hydrogloical models. In: Singh VP, Frevert DK (eds) Mathematical model of large watershed hydrology. Water Resources Publications, LLC, Highlands Ranch, Colorado, USA:25-66

Beven KJ, Binley A (1992) The future of distributed models: model calibration and uncertainty predictions. Hydrological Processes 6:349-368

Boronina A, Ramillien G (2008) Application of AVHRR imagery and GRACE measurements for calculation of actual evapotranspiration over the Quaternary aquifer (Lake Chad basin) abd validation of groundwater models. Journal of Hydrology 348:98-109

Chen JL, Wilcox CR, Seo KW (2006) Optimized smoothing of Gravity Recovery and Climate Experiment (GRACE) time-variable gravity observations. Journal of Geophysical Research 111(B06408)

Chen JL, Rodell M, Wilson CR, Famiglietti JS (2005a) Low degree spherical harmonic influences on Gravity Recovery and Climate Experiment (GRACE) water storage estimates. Geophys. Res. Lett. 32(L14405):doi:10.1029/2005GL022964

Chen JL, Wilson CR, Famiglietti JS, Rodell M (2005b) Spatial sensitivity of Gravity Recovery and Climate Experiment (GRACE) time-variable gravity observations. Journal of Geophysical Research 110(B08408):doi:10.1029/2004JB003536

Chen JL, Wilson CR, Famiglietti JS, Rodell M (2007) Attenuation effect on seasonal basin-scale water storage changes from GRACE time-variable gravity. Journal of Geodesy 81(4):237-245

Crowley JW, Mitrovica JX, Bailey RC, Tamisiea ME, Davis JL (2008) Annual variations in water storage and precipiation in the Amazon basin. Journal of Geodesy 82:9-13

Dirmeyer PA, Gao X, Zhao M, Guo Z, Oki T, Hanasaki N (2006) GSWP-2: Multimodel analysis and implications for our perception of the land surface. Bull. Am. Meteorol. Soc. 87:1381-1397

Döll P, Kaspar F, Lehner B (2003) A global hydrological model for deriving water availability indicators: model tuning and validation. Journal of Hydrology 270:105-134

Fekete BM, Vorosmarty CJ, Grabs W (2002) High-resolution fields of global runoff combining observed river discharge and simulated water balance. Global Biochemical Cycles 16(3):1042, 10.1029/1999GB001254

Fengler MJ, Freeden W, Kohlhaas A, Michel V, Peters T (2007) Wavelet modeling of regional and temporal variations of the earth's gravitational potential observed by GRACE. Journal of Geodesy 81(1):5-15

Frappart F, Ramillien G, Biancamaria S, Mognard NM, Cazenave A (2006) Evolution of high-latitude snow mass derived from the GRACE gravimetry mission (2002-2004). Geophysical Research Letters 33:L02501, doi:10.1029/2005GL024778

Güntner A, Schmidt R, Döll P (2007a) Supporting large-scale hydrogeological monitoring and modelling by time-variable gravity data. Hydrogeology Journal 15(1):167-170

Güntner A, Stuck J, Werth S, Döll P, Verzano K, Merz B (2007b) A global analysis of temporal and spatial variations in continental water storage. Water Resources Research 43(5):W05416

Güntner A, Werth S, Schmidt R, Petrovic S, Wünsch J, Barthelmes F (2007c) Multi-objective calibration of a global hydrology model using GRACE water storage variations. In: Eos Trans. AGU, 88(52), Fall Meet. Suppl., Abstract H31A-0124

Han SC, Shum CK, Braun A (2005a) High-resolution continental water storage recovery from low-low satellite-to-satellite tracking. Journal of Geodynamics 39(1):11-28

Han SC, Shum CK, Jekeli C, Alsdorf D (2005b) Improved estimation of terrestrial water storage changes from GRACE. Geophysical Research Letters 32(7):1-5

Han SC, Shum CK, Jekeli C, Kuo CY, Wilson C, Seo KW (2005c) Non-isotropic filtering of GRACE temporal gravity for geophysical signal enhancement. Geophysical Journal International 163:18-25

Hanasaki N, Kanae S, Oki T, Masuda K, Motoya K, Shen Y, Tanaka K (2007) An integrated model for the assessment of global water resources - Part 1: Input meteorological forcing and natural hydrological cycle modules. Hydrology and Earth System Sciences Discussions 4:3535-3582

Hirschi M, Seneviratne SI, Schär C (2006) Seasonal variations in terrestrial water storage for major mid-latitude river basins. Journal of Hydrometeorology 7(1):39-60

Horwath M, Dietrich R (2006) Errors of regional mass variations inferred from GRACE monthly solutions. Geophys. Res. Lett. 33(L07502)

Hu XG, Chen JL, Zhou YH, Huang C, Liao XH (2006) Seasonal water storage change of the Yangtze River basin detected by GRACE. Science In China Series D-Earth Sciences 49(5):483-491

Hunger M, Döll P (2007) Value of river discharge data for global-sale hydrological modeling. Hydrol. Earth Syste. Sci. Discuss. 4:4125-4173

Jung M, Le Maire G, Zaehle S, Luyssaert S, Vetter M, Churkina G, Ciais P, Viovy N, Reichstein M (2007) Assessing the ability of three land ecosystem models to simulate gross carbon uptake of forests from boreal to Mediterranean climate in Europe. Biogeosciences 4(4):647-656

Klees R, Zapreeva EA, Winsemius HC, Savenije HHG (2007) The bias in GRACE estimates of continental water storage variations. Hydrology and Earth System Sciences 11(4):1227-1241

Kusche J (2007) Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models. Journal of Geodesy 81(11):733-749

Lettenmaier DP, Famiglietti JS (2006) Hydrology - Water from on high. Nature 444(7119):562-563

Liang X, Lettenmaier DP, Wood EF, Burges SJ (1994) A simple hydrologically based model of land surface and energy fluxes for general circulation models. Journal of Geophysical Research 99(D7):14415-14428

Luthcke SB, Rowlands DD, Lemoine FG, Klosko SM, Chinn D, McCarthy JJ (2006) Monthly spherical harmonic gravity field solutions determined from GRACE inter-satellite range-rate data alone. Geophysical Research Letters 33:L02402, doi:10.1029/2005GL024846

Milly PCD, Dunne KA (2002) Macroscale water fluxes: 1. Quantifying errors in the estimation of basin mean precipitation. Water Resources Research 38(10):231-2314

Milly PCD, Shmakin AB (2002) Global modeling of land water and energy balances. Part I: The Land Dynamics (LaD) Model. Journal of Hydrometeorology 3(3):283-299

Ngo-Duc T, Polcher J, Laval K (2005) A 53-year forcing data set for land surface models. Journal of Geophysical Research D: Atmospheres 110(6):1

Ngo-Duc T, Laval K, Ramillien G, Polcher J, Cazenave A (2007) Validation of the land water storage simulated by Organising Carbon and Hydrology in Dynamic Ecosystems (ORCHIDEE) with Gravity Recovery and Climate Experiment (GRACE) data. Water Resources Research 43(4)

Nijssen B, O'Donnell GM, Lettenmaier DP, Lohmann D, Wood EF (2001) Prediciting the discharge of global rivers. Journal of Climate:3307-3323

Niu G-Y, Yang Z-L (2006a) Assessing a land surface model's improvements with GRACE estimates. Geophysical Research Letters 33:L07401

Niu G-Y, Seo KW, Yang Z-L, Wilson C, Su H, Chen J, Rodell M (2007a) Retrieving snow mass from GRACE terrestrial water storage change with a land surface model. Geophysical Research Letters 34:L15704

Niu GY, Yang ZL (2006b) Effects of frozen soil on snowmelt runoff and soil water storage at a continental scale. Journal of Hydrometeorology 7(5):937-952

Niu GY, Yang ZL, Dickinson RE, Gulden LE, Su H (2007b) Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data. Journal Of Geophysical Research-Atmospheres 112(D7)

Ramillien G, Frappart F, Cazenave A, Güntner A (2005) Time variation of land water storage from an inversion of 2 years of GRACE geoids. Earth and Planetary Science Letters 235:283-301

Ramillien G, Cazenave A, Reigber C, Schmidt R, Schwintzer P (2004) Recovery of global time-variations of surface water mass by GRACE geoid inversion. In: GGSM2004, Porto:

Ramillien G, Frappart F, Güntner A, Ngo-Duc T, Cazenave A, Laval K (2006) Time-variations of the regional evapotranspiration rate from Gravity Recovery and Climate Experiment (GRACE) satellite gravimetry. Water Resources Research 42:W10403

Rangelova E, van der Wal W, Braun A, Sideris MG, Wu P (2007) Analysis of Gravity Recovery and Climate Experiment time-variable mass redistribution signals over North America by means of principal component analysis. Journal Of Geophysical Research-Earth Surface 112(F3)

Reichle RH, Crow WT, Keppenne CL (2008) An adaptive ensemble Kalman filter for soil moisture data assimilation. Water Resources Research 44:W03423

Rodell M, Zaitchik BF, Reichle RH (2007a) Assimilation of GRACE dervied terrestrial water storage data into a hydrological model. In: Eos Trans. AGU, 88(52), Fall Meet. Suppl., Abstract G31A-04

Rodell M, Chen JL, Kato H, Famiglietti JS, Nigro J, Wilson CR (2007b) Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. Hydrogeology Journal 15(1):159-166

Rodell M, Famiglietti JS, Chen J, Seneviratne SI, Viterbo P, Holl S, Wilson CR (2004) Basin scale estimates of evapotranspiration using GRACE and other observations. Geophysical Research Letters 31:L20504, doi:10.1029/2004GL020873)

Roerink GJ, Su Z, Menenti M (2000) S-SEBI: A simple remote sensing algorithm to estimate the surface energy balance. Physics and Chemistry of the Earth (B) 25(2):147-157

Rowlands DD, Luthcke SB, Klosko SM, Lemoine FG, Chinn DS, McCarthy JJ, Cox CM, Anderson OB (2005) Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements. Geophysical Research Letters 32:L04310

Schmidt M, Han SC, Kusche J, Sanchez L, Shum CK (2006a) Regional high-resolution spatiotemporal gravity modeling from GRACE data using spherical wavelets. Geophysical Research Letters 33(8)

Schmidt R, Petrovic S, Güntner A, Barthelmes F, Wünsch J (2008a) Periodic components of water storage changes from GRACE and global hydrology models. Journal of Geophysical Reaseach – Solid Earth, submitted

Schmidt R, Flechtner F, Meyer U, Neumayer K-H, Dahle C, König R, Kusche J (2008b) Hydrological signals observed by the GRACE satellite mission. Survey in Geophysics, submitted

Schmidt R, Flechtner F, Güntner A, König R, Meyer U, Neumayer K-H, Reigber C, Rothacher M, Petrovic S, Zhu S-Y (2007) GRACE time-variable gravity accuracy assessment. In: Rizos C, Tregoning P (eds) Dynamic planet - Monitoring and understanding a dynamic planet with geodetic and oceanographic tools. Springer Berlin Heidelberg, Cairns/Australien, August 22 - 27, 2005:237-243

Schmidt R, Schwintzer P, Flechtner F, Reigber C, Güntner A, Döll P, Ramillien G, Cazenave A, Petrovic S, Jochmann H, Wünsch J (2006b) GRACE observations of changes in continental water storage. Global and Planetary Change 50(1-2):112-126

Schrama EJO, Wouters B, Lavallée DA (2007) Signal and noise in Gravity Recovery and Climate Experiment (GRACE) observed surface mass variations. Journal of Geophysical Research B: Solid Earth 112:B08407

Seneviratne SI, Viterbo P, Lüthi D, Schär C (2004) Inferring changes in terrestrial water storage using ERA-40 reanalysis data: The Mississippi River basin. Journal of Climate 17:2039-2057

Seo KW, Wilson CR, Famiglietti JS, Chen JL, Rodell M (2006) Terrestrial water mass load changes from gravity recovery and climate experiment (GRACE). Water Resources Research 42(5)

Sheffield J, Goteti G, Wood EF (2006) Development of a 50-year high resolution global dataset of meterological forcings for land surface modeling. Journal of Climate 19:3088-3111

Steffen H, Denker H, Müller J (2008) Glacial isostatic adjustment in Fennoscandia from GRACE data and comparison with geodynmical models. Journal of Geodynamics, accepted

Swenson S, Wahr J (2002) Methods for inferring regional surface-mass anomalies from Gravity Recovery and Climate Experiment (GRACE) measurements of time-variable gravity. Journal of Geophysical Research 107(B9):2193

Swenson S, Wahr J (2006) Estimating large-scale precipitation minus evapotranspiration from GRACE satellite gravity measurements. Journal of Hydrometeorology 7(2):252-270

Swenson S, Wahr J (2007) Multi-sensor analysis of water storage variations of the Caspian Sea. Geophysical Research Letters 34(16)

Swenson S, Wahr J, Milly PCD (2003) Estimated accuracies of regional water storage variations inferred from the Gravity Recovery and Climate Experiment (GRACE). Water Resources Research 39(8):1223, doi:10.1029/2002WR001808

Swenson SC, Milly PCD (2006) Climate model biases in seasonality of continental water storage revealed by satellite gravimetry. Water Resources Research 42:W03201, doi:10.1029/2005WR004628

Syed TH, Famiglietti JS, Zlotnicki V, Rodell M (2007) Contemporary estimates of Pan-Arctic freshwater discharge from GRACE and reanalysis. Geophysical Research Letters 34(19)

Syed TH, Famiglietti JS, Rodell M, Chen J, Wilson CR (2008) Analysis of terrestrial water storage changes from GRACE and GLDAS. Water Resources Research 44(W02433)

Tapley BD, Bettadpur SV, Watkins M, Reigber C (2004a) The gravity recovery and climate experiment: Mission overview and early results. Geophys. Res. Lett. 31:L09607, doi:10.1029/2004GL019920

Tapley BD, Bettadpur SV, Ries JC, Thompson PF, Watkins MM (2004b) GRACE measurements of mass variability in the Earth system. Science 305:503-505

Wahr J, Swenson S, Velicogna I (2006) Accuracy of GRACE mass estimates. Geophys. Res. Lett. 33(L06401)

Wahr J, Swenson S, Zlotnicki V, Velicogna I (2004) Time-variable gravity from GRACE: First results. Geophysical Research Letters 31:L11501, doi:10.1029/2004GL019779

Werth S, Güntner A, Schmidt R, Kusche J (2008) Evaluation of GRACE filter tools from a hydrological perspective. Geophysical Journal International, submitted

Widen-Nilsson E, Halldin S, Xu CY (2007) Global water-balance modelling with WASMOD-M: Parameter estimation and regionalisation. Journal of Hydrology 340(1-2):105-118

Winsemius HC, Savenije HHG, Van de Giesen N, Van den Hurk BJJM, Zapreeva EA, Klees R (2006) Assessment of Gravity Recovery and Climate Experiment (GRACE) temporal signature over the upper Zambezi. Water Resources Research 42(W12201)

Wood EF, Lettenmaier DP, Liang X, Lohmann D, Boone A, Chang S, Chen F, Dai YJ, Dickinson RE, Duan QY, Ek M, Gusev YM, Habets F, Irannejad P, Koster R, Mitchel KE, Nasonova ON, Noilhan J, Schaake J, Schlosser A, Shao YP, Shmakin AB, Verseghy D, Warrach K, Wetzel P, Xue YK, Yang ZL, Zeng QC (1998) The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) phase 2(c) Red-Arkansas River basin experiment: 1. Experiment description and summary intercomparisons. Global and Planetary Change 19(1-4):115-135

Yamamoto K, Fukuda Y, Nakaegawa T, Nishijima J (2007) Landwater variation in four major river basins of the Indochina peninsula as revealed by GRACE. Earth Planets And Space 59(4):193-200

Yeh PJF, Swenson SC, Famiglietti JS, Rodell M (2006) Remote sensing of groundwater storage changes in Illinois using the Gravity Recovery and Climate Experiment (GRACE). Water Resources Research 42(12)

Zeng N, Yoon JH, Mariotti A, Swenson S (2008) Variability of basin-scale terrestrial water storage from a PER water budget method: The Amazon and the Mississippi. Journal of Climate 21:248-265