Werth, S., Güntner, A., Petrovic, S., Schmidt, R. (2009): Integration of GRACE mass variations into a global hydrological model. - Earth and Planetary Science Letters, 277, 1-2, 166-173

DOI: 10.1016/j.epsl.2008.10.021.

Integration of GRACE mass variations into a global hydrological model

S. Werth [*,a], A. Güntner [a], S. Petrovic [a], R. Schmidt [a]

[a] *Helmholtz Centre Potsdam - GFZ German Research Centre for Geosciences, Telegrafenberg, Potsdam, 14473, Germany.*

Abstract

Time-variable gravity data of the GRACE (Gravity Recovery And Climate Experiment) satellite mission provide global information on temporal variations of continental water storage. In this study, we incorporate GRACE data for the first time directly into the tuning process of a global hydrological model to improve simulations of the continental water cycle. For the WaterGAP Global Hydrology Model (WGHM), we adopt a multi-objective calibration framework to constrain model predictions by both measured river discharge and water storage variations from GRACE and illustrate it on the example of three large river basins: Amazon, Mississippi and Congo. The approach leads to improved simulation results with regard to both objectives. In case of monthly total water storage variations we obtained a RMSE reduction of about 25 mm for the Amazon, 6 mm for the Mississippi and 1 mm for the Congo river basin. The results highlight the valuable nature of GRACE data when merged into large-scale hydrological modeling. Furthermore, they reveal the utility of the multi-objective calibration framework for the integration of remote sensing data into hydrological models.

Key words: continental water cycle, total water storage change, GRACE, satellite gravity, time variable gravity, hydrological modeling, model calibration,

1 Introduction

By mapping time variations of the Earth's gravity field with the Gravity Recovery and Climate Experiment satellite mission (GRACE) since its launch in 2002, an unprecedented global data set of mass variations close to the Earth surface became available (Tapley et al., 2004). After removal of mass variations due to tides and non-tidal atmospheric and oceanic transport processes, the time-variable gravity data mainly represent water mass variations in continental hydrology, i.e., total water storage change (TWSC) on the continents (see a recent review by Schmidt et al. (2008a)). In specific regions, also mass variation from post glacial rebound (Tamisiea et al., 2007) and seismic activities (Chen et al., 2007) could be revealed from the GRACE data.

For the field of hydrology, the past six years of GRACE operation contributed to a significantly improved understanding of the spatio-temporal patterns of water storage variations on the continents because no comprehensive TWSC data were available before at large spatial scales due to the absence of adequate monitoring systems (Lettenmaier and Famiglietti, 2006). Thus, the GRACE TWSC data give new insights into the Earth's water cycle including the contribution of TWSC to sea level variations (Ramillien et al., 2008), the impact of climate variability or extremes on water storage (e.g. Andersen et al., 2005; Seitz et al., 2008), or melting of glaciers and ice caps (e.g. Chen et al., 2006; Luthcke et al., 2006). Numerous regional or river basin studies

* Corresponding author. Address: Tel.: +49 331 288 1563; Fax: +49 331 288 1570
Email address: swerth@gfz-potsdam.de (S. Werth).

analyzed GRACE TWSC from seasonal to inter-annual time scales (see a recent review by Schmidt et al. (2008a)). Others solved the water balance using TSWC from GRACE for other hydrological components such as evapotranspiration (Rodell et al., 2004; Ramillien et al., 2006) or runoff (Syed et al., 2007), or separated individual storage compartments such as groundwater (Rodell et al., 2007; Strassberg et al., 2007) or snow (Frappart et al., 2006; Niu et al., 2007).

Besides observation data, hydrological simulation models are an indispensable tool to assess the impact of environmental change on the continental water cycle and the particular processes mentioned above. Thus, in turn, they are a prerequisite for implementing measures of sustainable management of water-related issues in future. At continental to global scales, hydrological models are an integral part of atmospheric circulation models where they represent the land surface processes for climate and weather prediction simulations, see Dirmeyer et al. (2006) for an overview on land surface models and their comparison. In addition, water balance models are used to represent the full water cycle in river basins for purposes such as stream flow forecasting and water resources assessment (for a recent overview on global water balance models see Widen-Nilsson et al. (2007)). However, these large-scale hydrological models are known to suffer from uncertainties in terms of model structure, parameter values and climate forcing data. As a consequence, simulation results for hydrological state variables and water fluxes on the continents vary considerably between models (e.g. Dirmeyer et al., 2006). While river discharge has for a long time been the only observable to validate and calibrate global water balance models (Hunger and Döll, 2008), considerable model uncertainties remain for other components of the water cycle, e.g., water storage, evapotranspiration or groundwater recharge due to the lack of adequate observation data.

In this context, GRACE provides a unique data set to evaluate and improve the simulation of TWSC on large scales and therewith to uncover shortcomings in model designs and parameters. Numerous studies compared GRACE-derived TWSC data with simulation results of hydrological models and concluded with a recommendation to use GRACE data as a model constraint (see a recent overview by Güntner (2008)). First attempts have been made to modify large-scale hydrological models and to evaluate the modifications with GRACE observations (Niu and Yang, 2006; Ngo-Duc et al., 2007) and very recently, Zaitchik et al. (2008) assimilated GRACE TWSC into a land surface model for the Mississippi river basin. A global integration of GRACE data with hydrological models to improve model performance by calibration has not been reported so far.

This motivated the present study to incorporate for the first time GRACE data into the tuning process of a global hydrological model (section 2.1). For this purpose, a multi-objective calibration scheme has been developed (see section 2.2). Calibration denotes the selection of model parameter values by evaluating the simulation performance via a model output objective against observations. In contrary to data assimilation, the system is tuned by determining model parameter values during a pre-defined time interval, and the resulting parameter set may be used for subsequent independent model runs. Multi-objective calibration denotes that more than one model output objectives are taken into consideration. In this study, two different types of measured data are used to constrain parameter sets (section 2.3). Improvements for the simulation of TWSC are analyzed (in section 3) and the value of calibration procedure using GRACE data towards enhanced predictions of the continental water cycle is outlined (section 4).

2 Methods and Data

2.1 Global Hydrological Model

The WaterGAP Global Hydrology Model (WGHM) is a conceptual water balance model which simulates the continental water cycle including the most important water storage components, i.e., interception, soil water, snow, groundwater and surface water. The major hydrological processes are simplified by conceptual formulations. WGHM has a 0.5°x0.5° spatial resolution and a daily computation time step. Information on land surface characteristics such as the spatial distribution of vegetation, soil types, land use, groundwater and surface water bodies is given in the model from global data sets. For details on model equations and their parameters see Döll et al. (2003). The model has widely been used to analyse continental water storage change (Güntner et al., 2007). In comparisons with GRACE TWSC, a general agreement of seasonal and other periodic characteristics of TWSC was found at the global scale, but amplitudes and phases in the model showed significant differences (larger than GRACE errors) in particular river basins (Ramillien et al., 2005; Schmidt et al., 2006, 2008b).

In this study, WGHM is driven by climate data (temperature, cloudiness and number of rain days per month) of the European Centre for Medium-Range Weather Forecast (ECMWF) and monthly precipitation data of the Global Precipitation Climatology Centre (GPCC). Precipitation is disaggregated to a daily resolution with the given number of rain days per month. The climate input data are available from 01/1992 until 12/2007 for this study. Antarctica and Greenland were excluded from the simulations.

We used the most recent WGHM version as described by Hunger and Döll (2008), who calibrated (i.e. tuned) the model against observed mean annual river runoff at 1235 discharge stations worldwide, by varying one runoff generation parameter. This model version is called the original version in the following. Overall, the model includes 26 process parameters. Their values in the original model as well as parameter ranges for the calibration are based on literature and qualitative reasoning (Kaspar, 2004), see Table 2 for the parameters calibrated in this study. Thereof, the parameter root depth is based on the global land cover distribution and can be calibrated by a multiplicative factor. The Priestley-Taylor coefficient is used in the corresponding approach to quantify potential evapotranspiration. The radiative fraction of the extraterrestrial radiation that reaches the Earth's surface is determined by cloud cover data and the radiation proportion parameter. The variability of snow melt temperature is due to different elevation and vegetation cover of different regions. A more detailed description of the model parameters is provided by Döll et al. (2003).

2.2 Calibration Technique

Combining both the present station-based accuracy of WGHM in terms of river discharge and the integrative nature of the GRACE data with global coverage, improved simulation results were expected from a multi-objective calibration approach. Calibration in the sense used here denotes an iterative method of testing different parameter values and selecting the best parameter sets based on performance criteria that evaluate simulation results against observation data. Calibration methods differ in their strategies to select parameter sets for each iteration from the given parameter space. Furthermore, multi-objective calibration denotes the selection of parameter values through evaluating model performance against more than one objective. In this study, these objectives are based on two observation data sets: river discharge and periodic TWSC (see section 2.3); hence, it is a two dimensional problem. Instead of a single optimum parameter set, such an approach will lead to a Pareto set of optimal solutions (Gupta et al., 1998). Each Pareto optimum of this set is an optimal solution from a multi-objective point of view in the sense that no other solution exists that provides a better simulation performance for both model output objectives. Hence, when moving from one Pareto solution to another, simulation performance increases for one objective while it decreases for the other objective. Without additional information it is not possible to undertake a ranking among the Pareto solutions. The trade-off (i.e. the spread) between the Pareto solutions reflects the minimum parameter uncertainty (Vrugt et al., 2003) caused by errors in the input and the measured data as well as by model structure.

The calibration of a number of model parameters against more than one objective depicts a highly non-linear optimization problem and requires a global optimization method. Furthermore, only stochastic methods like a multi-start simulated annealing or an evolutionary algorithm assure a feasible computing time for the calibration of the global hydrological model WGHM. Therefore, to handle the complexity of a multi-objective and multi-parameter calibration problem as well as the computational demands we select the ϵ-Non-dominated-Sorting-Genetic-Algorithm-II (ϵ-NSGAII) (Kollat and Reed, 2006), which ranks among the most effective and efficient multi-objective optimization methods (Tang et al., 2006). This global optimization algorithm solves multi-objective problems using the concept of evolutionary parameter variation (mutation, crossover and selection). It is an elitist algorithm with a Pareto ranking routine. Furthermore, as an extension of NSGAII (Deb et al., 2000) by the concept of ϵ-dominance, it allows to specify the accuracy to be fulfilled by each objective. For this study, we parameterize its operators as proposed by Kollat and Reed (2006). Furthermore, we use a population size of $N = 8$ and an ϵ-resolution of 0.05 for both objectives and stop the optimization after 400 iterations.

The calibration of WGHM is exemplarily done for the Amazon, the Mississippi and the Congo river basins in this study. These basins were selected because of their large size of over three million km^2. The period 01/2003-12/2006 was used for WGHM calibration.

Güntner et al. (2007) showed that WGHM parameter sensitivity for TWSC simulations varies considerably between the river basins. This inter-basin variability of parameter sensitivity can be explained by differences of the climatic conditions (represented in the model by the climate input data and parameters steering evaporation or snow melt processes, for instance) and of the land surface properties (represented by, e.g., vegetation or soil parameters) between the river basins. This results in different water flow and storage characteristics in the basins. In particular, different storage components dominate the individual river basin response, e.g., snow storage in higher latitude areas or surface water storage in some tropical areas with large inundation zones. Thus, also the sensitivity of model parameters used to govern these individual dominant storage processes varies between the river basins. Consequently, ahead of the calibration work, a sensitivity study was undertaken by a Latin Hypercube sampling for 2000 parameter sets and by an analysis scheme going back to Hornberger and Spear (1981), who selected sensitive parameters based on their ability to provide behavioural model simulations. For each river basin, we selected the six most sensitive parameters for calibration against TWSC and river discharge (see row (e) and row (f) of Table 1). Parameter values and ranges are documented in Table 2.

For the Amazon basin, three of these parameters concern the process of surface water transport, because of the high water volume during an important flood season. In contrast, evaporation is most important in the tropical Congo river basin with a distinct dry season. A diverse set of important processes (e.g. snow, evaporation and surface water) provides the most sensitive parameter of the Mississippi river basin, due to its location in three different climate regions (cold in the north, subtropical in the southeast and dry in the southwest).

The evaluation of model performance for each iteration is effected by the following four steps: 1) Model simulation of monthly global TWSC fields and river discharge with the current parameter set. 2) Application of a GRACE-equivalent filter procedure, which comprises the conversion of WGHM TWSC fields into the frequency domain, i.e. spherical harmonic coefficients, followed by Gaussian smoothing (Jekeli, 1981) and the computation of basin averages of TWSC according to Wahr et al. (1998). 3) Fitting amplitudes and phases of significant periods which were determined from GRACE data (see section 2.3.2) to the simulated basin averages of TWSC and reconstruction of a basin-average time series of TWSC from these periods. 4) Evaluation of each calibration objective (discharge and TWSC) by computation of the Nash-Sutcliffe-efficiency coefficient (NSC) (Nash and Sutcliffe, 1970) as a criterion of agreement between modeled and measured time-series.

NSC is a simulation performance measure that normalizes the squared difference of a predicted (P) to an observed (O) time series by the variance of the observed values with n time steps:

$$NSC = 1 - \frac{\sum_{i=0}^{n}(O_i - P_i)^2}{\sum_{i=0}^{n}(O_i - \bar{O})^2}, \quad (1)$$

where $\bar{O}$ is the mean of the observations over the examined period. NSC evaluates both phase and amplitude agreement between two time series. It ranges from $-\infty$ to 1 (optimal fit), with a value of 0 indicating a simulated time series that performs as well as a model being equal to the mean of the observable. Therefore Pareto solutions are restricted to NSC values greater than 0.

2.3 Calibration data

2.3.1 River basin discharge: Objective 1

River discharge data of Amazon, Mississippi and Congo from the most downstream gauging station were used (Table 1). We computed monthly mean values for the calibration period. For the Congo river where no up-to-date measurements were available, we assigned the monthly mean discharge of earlier observations to the calibration period.

2.3.2 GRACE TWSC: Objective 2

Reconstructed significant periodic parts of basin-averaged TWSC resulting from the investigation presented in Schmidt et al. (2008b) are used as calibration input for this study. These data are chosen, because errors in the GRACE original data and the difficulty to separate the errors from real signals mark the greatest challenge for application of satellite gravity solutions.

Schmidt et al. (2008b) developed a technique to extract significant water storage change information from GRACE data by three steps: 1) Identification of the dominant spatio-temporal patterns in mass variations derived from GRACE observations through a principal component analysis (applied at the scale of the river basins to grids previously filtered by a Gaussian smoothing with a 500 km averaging radius), 2) Identification of significant periods of TWSC contained in the principal components without fixing a priori the period lengths, and 3) Reconstruction of (error-reduced) basin-average time series of TWSC from the significant periods.

As a basis, monthly GRACE-only time series of global gravity fields generated as spherical harmonic expansions up to degree and order 120 at the GFZ German Research Center for Geosciences (GRACE Level-2 products, version GFZ-RL04, Schmidt et al., 2008a) for the time period from 02/2003 until 12/2006 (excluding unavailable months 06/2003 and 01/2004) were used. The noise contained in the spherical harmonics increases with the degree of the expansion terms, and the noise/signal ratio reaches unacceptably high values in higher-degree terms. In the space domain this noise becomes visible in the form of the typical meridional-oriented spurious gravity signals ("stripes") (e.g. Swenson and Wahr, 2006; Schmidt et al., 2008a). Hence, a spatial filtering is mandatory when computing water storage variations from GRACE gravity field models in order to reduce these errors. For the present study a widely used Gaussian smoothing (Jekeli, 1981) with an averaging radius of 500 km was applied. Mass variations (TWSC) were derived relative to a mean field (i.e. in the form of mass anomalies) for the considered data period applying the procedure presented by Swenson and Wahr (2002).

Since the effects of the atmospheric and the oceanic circulations were previously removed in the course of the gravity field recovery from the raw GRACE data by applying appropriate geophysical models (Flechtner, 2007), the major part of the signal contained in the derived grids of mass anomalies can be attributed to hydrological variations. Due to the rather short time period covered by the available GRACE data, the long-term trends determined both from the hydrology model WGHM and from the GRACE gravity fields should be regarded as less reliable than the periodic components resulting from the same data. Therefore, as the last preparatory step, the data used in this study have been de-trended.

Subsequently, the three-step strategy for the detection of significant periodic components, depicted at the beginning of this section, was realized, see (Schmidt et al., 2008b) for more details. It is important to note, that the period search was not a-priori constrained to seasonal or other postulated variations. For all three river basins, considered in this study, two periods resulted to be significant with respect to their signal proportion and an uncertainty study. Corresponding amplitudes and phases used for the calibration are given in Table 1, row (g). TWSC of all three basins exhibit a seasonal period. A second period of inter-annual scale (about 2.5 years) occurs for the Amazon as well as the Mississippi and of semi-annual scale for the Congo river basin. The cumulative variability of the reconstructed periodic components dominates the integral GRACE signal (see Table 1 row (h) for percentage proportion).

Error estimations of GRACE data differ between several studies. For example, using a Gaussian smoothing with an averaging radius of 750 km Wahr et al. (2006) derived latitude-dependant errors of GRACE mass estimates ranging from 8 mm near the poles up to 25-27 mm at low latitudes, when expressed in water column equivalents. This results in a global area-weighted mean of

21 mm. Schmidt et al. (2007) gave for a 500 km Gaussian filtering a global error estimate of 24-30 mm water column. According to Schmidt et al. (2008a) the accuracy of the GFZ-RL04 used in this study is approximately two times better than the accuracy of the earlier releases used in both cited studies. However, it should be taken into account that errors may be higher for particular regions and months, and are also influenced by leakage errors after forming basin-average values.

3 Results and Discussion

The multi-objective calibration of WGHM with GRACE TWSC and river discharge led to improved simulation results in all three river basins (Figure 1). Each Pareto solution (on the red line) is superior to the original model version (green dot) with regard to both objectives.

Best results were obtained for the Amazon basin. NSC performances better than 0.95 with respect to both objectives were achieved for the Pareto solution closest to the optimum (hereafter referred as the selected Pareto-optimum, blue dot in Figure 1a). The amplitude of periodic terms of TWSC increased markedly in the Pareto solutions when compared to the original model (Figure 2a). Since the narrow uncertainty band given by the Pareto set of solutions does not include the original model time series, the significance of model improvement is substantiated. Although the amplitudes of basin-average TWSC were slightly overestimated by the selected Pareto solution in 2003 and 2006, its root mean square error (RMSE) of the complete (but de-trended) TWSC signal was reduced by 50% compared to the original model version (Table 3). The reduction of RMSE for discharge was even greater, since a phase shift of discharge seasonality could be corrected by the multi-criteria calibration (see Figure 3a). A main reason for the model improvements in the Amazon basin could be attributed to longer residence times of surface water in rivers and floodplains as expressed by lower values for the flow velocity parameter in the Pareto solutions.

Also in the Mississippi basin a very good fit to observations with NSC performances of about 0.9 for both objectives were obtained for the selected Pareto-optimum (Figure 1b). Although the results for river discharge are more uncertain than for TWSC, the improvement compared to the original WGHM is greater for discharge than for TWSC. This is reflected by the reduction of the RMSE of the monthly mean discharge of about 80%, respectively 13 km^3/month (Table 3) for the selected Pareto-optimum. The clear improvement of monthly discharge simulations is also due to the fact that the original model was calibrated for mean annual values and did not take into account the seasonal distribution of discharge as in the present scheme. Therefore, the overestimated peaks of monthly discharge during spring in the standard model version could be corrected for all Pareto solutions (see Figure 3b). The reconstructed calibrated time series of water storage variations shows a slightly shifted phase and an amplitude which is closer to the GRACE time series (Figure 2b). The RMSE of the full de-trended time series of TWSC was improved about 6 mm compared to the original model version (Table 3). This improvement was most likely caused by changes of two model parameters. An increased effective root zone increases the soil storage capacity and an increased snow melt temperature smooths the previously overestimated runoff peaks.

Calibration for the Congo basin resulted in a much wider trade-off between both objectives (note the different scaling of both axes in Figure 1c). The performance of the Pareto solutions varies between 0.0 and 0.8 for discharge and between 0.7 and 0.9 for TWSC (Figure 1c). This trade-off resulted in a wider uncertainty band for the calibrated TWSC periods of the Pareto solutions (Figure 2c). Nevertheless, a small phase shift of TWSC periods was achieved for all Pareto solutions. The RMSE of the full TWSC signal for the selected Pareto-optimum was improved by about 1 mm (Table 3). All other Pareto solutions provide greater RMSE reductions, since they show a higher simulation performance for the significant periods of TWSC, as the selected Pareto-optimum. For discharge, there were slight improvements in the monthly regime (Figure 3c), as indicated by higher peaks during the turns of the year (from October till January) for the re-calibrated hydrograph of the selected Pareto-optimum. While the RMSE for discharge could clearly be decreased by the calibration procedure, the NSC value for the selected Pareto-optimum of 0.76 still indicates only moderate correspondence of simulated and observed river discharge. Though, the rather discontinuous course of the Pareto frontier may imply that a higher number of function evaluations would give better calibration results. These limitations in achieving better discharge and TWSC simulations as well as the wider uncertainty in the calibration of the Congo basin are likely due to the lack of river runoff measurements during the calibration period and complicate the assignment of improved processes for the Congo basin. The particular characteristics of the rainfall distribution in each year will cause substantial deviations from the mean hydrograph that was used for model evaluation in this basin (Figure 3c). This may also point out errors in the model structure, the model input data, or in the parameter space allowed for calibration in the Congo basin and is subject to further studies.

Introduction of further observables to the multi-objective calibration scheme could further reduce the resulting equifinality of parameter sets as expressed by the dense Pareto-Frontier shown for Amazon and Mississippi. In particular, parameter values of storage processes that are represented by these additional observations could be more effectively constrained. For example, surface water storage derived from satellite altimetry and imagery can provide such data sets for an individual storage compartment (Papa et al., 2008). Though, the success will be limited as long as the observables contain high errors (e.g. groundwater, Döll and Fiedler, 2008) or the approach demands sophisticated model modifications to make model state variables match the observables (as for remotely sensed surface soil moisture).

A validation of the calibrated model was performed for de-trended GRACE signals including non-periodic components and errors from January until December 2007 (see Figure 4). For this year, a simulation run was realized with WGHM using the parameter values that were calibrated for the period 2003-2006. For the Amazon and the Mississippi river basins, simulation results were markedly better for the validation period, when they are compared to the results of the standard model in terms of amplitude, phase and RMSE values. This improvement is similar to what was achieved in the calibration period (see Table 3). This corroborates the model improvement of TWSC that could be achieved by the multi-criterial calibration for these basins. For the Congo river basin, however, the RMSE value increased, indicating that the model performs somewhat worse with the re-calibrated parameter set in the validation period. This confirms the above results that improvements by calibration are difficult to achieve with the present model set up and data availability for this river basin. For further studies it should also be taken into consideration that it might be justified to reduce the weight assigned to the river discharge data during calibration in the Congo basin due to their high uncertainties. This may enable the selection of Pareto optima with higher TWSC-simulation performance (see Figure 1c).

4 Conclusions

The first multi-objective calibration of the global hydrology model WGHM with TWSC data from GRACE and monthly mean river discharge was successfully carried out. By this approach, phase and amplitude differences of periodic water storage variations between GRACE and WGHM could be significantly reduced as compared to earlier versions of WGHM. We could show that the direct integration of GRACE data into the calibration process of WGHM leads to a clear improvement of simulated monthly TWSC signals on a scale of large river basins. At the same time, a better simulation of river discharge could be achieved. This highlights the particular value of multi-objective process analyses. If two observables are considered within the calibration approach, the trade-off in model performance of different hydrological variables is taken into account. Finally, this allows for an improved representation of the water balance as a whole.

It should be pointed out that the calibration approach adopted in this study followed two principles that can be seen as a prerequisite for the successful integration of GRACE water storage data into large-scale hydrological models (Güntner, 2008). First, GRACE and WGHM model data were treated exactly in the same way before comparison and parameter adjustment, i.e., the same methods of filtering and basin-averaging were applied to both data sets. This excludes the risk of poor comparability of the time series if unfiltered model data are compared to filtered GRACE data which may include filter-induced biases. Secondly, with WGHM a hydrological model was used that represents all relevant water storage compartments in the analyzed river basins, including surface water storage. Thus, it is assured that water storage calibrated in the model is consistent with the observation variable, i.e., the integrative nature of GRACE-based TWSC.

A better process understanding in global hydrology is necessary to provide more reliable estimates of changes in the continental water cycle, which constitutes an important input for climate studies or water resources management. In order to get a closer view into the reasons why the model differs from the real world, more accurate input data and improved calibration settings should be applied. The former can be achieved by using up-to-date river discharge data (i.e. for the Congo basin) and better GRACE filter methods. For the latter, technically more extensive model calibrations in terms of the size of parameter set population and of function evaluation are necessary to shift the Pareto frontier towards an even better model performance. Also, the analysis of a posteriori model states and parameter sets will help to uncover potential errors in model structure or input data. In this way, an improved understanding of continental water storage processes may finally be achieved by a stepwise modification of the modelling concept (Fenicia et al., 2008). Especially for regions like the Congo river basin with a very inaccurate or lacking coverage of terrestrial data, the usage of GRACE data is most proliferous concerning model improvement. Longer GRACE time series and the continuing error reduction within GRACE gravity recovery are likely to reduce the uncertainty of GRACE TWSC recovery and therefore the data assimilation into global hydrology modeling in further studies. Additionally, the presented approach is promising for the integration of alternative data sets from remote sensing, such as soil moisture, snow cover or surface water volumes into hydrological models. Furthermore, the methods considered here to achieve consistency of model variables and GRACE observations in terms of, e.g., data filtering and the selection of dominant signals, may similarly apply to other areas of Earth system modelling where GRACE data are to be used as a model constraint, such as for processes of the cryosphere or the Earth's interior.

Acknowledgements

The German Ministry of Education and Research (BMBF) supported these investigations within the geoscientific R+D programme GEOTECHNOLOGIEN "Erfassung des Systems Erde aus dem Weltraum" under grant 03F0424A. Thanks to M. Scheinert for providing the GravTools software.

References

Andersen, O., Seneviratne, S. I., Hinderer, J., Viterbo, P., 2005. GRACE-derived terrestrial water storage depletion associated with the 2003 european heat wave. Geophys. Res. Lett. 32, doi:10.1029/2005GL023574.

Chen, J. L., Tapley, B. D., Wilson, C. R., 2006. Alaskan mountain glacial melting observed by satellite gravimetry. Earth Planet. Sci. Let. 248 (1-2), 368–378.

Chen, J. L., Wilson, C. R., Tapley, B. D., Grand, S., 2007. GRACE detects coseismic and postseismic deformation from the sumatra-andaman earthquake. Geophys. Res. Lett. 34 (13), doi:10.1029/2007GL030356.

Deb, K., Agrawal, S., Pratap, A., Meyarivan, T., 2000. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. Tech. rep., Kanpur Genetic Algorithms Laboratory.

Dirmeyer, P., 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. Amer. Meteor. Soc. 87, 1381–1397.

Döll, P., Fiedler, K., 2008. Global-scale modeling of groundwater recharge. Hydrol. Earth Syst. Sci. 12, 863–885.

Döll, P., Kaspar, F., Lehner, B., 2003. A global hydrological model for deriving water availability indicators: model tuning and validation. J. Hydr. 270 (1-2), 105–134.

Fenicia, F., Savenije, H. H. G., Matgen, P., Pfister, L., 2008. Understanding catchment behavior through stepwise model concept improvement. Water Resour. Res. 44, W01402, doi:10.1029/2006WR005563.

Flechtner, F., 2007. GFZ level-2 processing standards document for level-2 product release 0004. GRACE 327-743 (GR-GFZ-STD-001) Rev. 1.0, February 19, 2007.

Frappart, F., Ramillien, G., Biancamaria, S., Mognard, N., Cazenave, A., 2006. Evolution of high-latitude snow mass derived from the GRACE gravimetry mission (2002-2004). Geophys. Res. Lett 33 (2), doi:10.1029/2005GL024778.

Güntner, A., 2008. Improvement of global hydrological models using GRACE data. Surv. Geoph., doi:10.1007/s10712–008–9038–y.

Güntner, A., Stuck, J., Werth, S., Döll, P., Verzano, K., Merz, B., 2007. A global analysis of temporal and spatial variations in continental water storage. Wat. Resour. Res. 43 (5), W05416, doi:10.1029/2006WR005247.

Gupta, H. V., Sorooshian, S., Yapo, P. O., 1998. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resour. Res. 34 (4), 751–763.

Hornberger, G. M., Spear, R. C., 1981. An approach to the preliminary analysis of environmental systems. J. Eviron. Mgmt. 12 (1), 7–18.

Hunger, M., Döll, P., 2008. Value of river discharge data for global-scale hydrological modeling. Hydrol. Earth Syst. Sci. 12 (3), 841.

Jekeli, C., 1981. Alternative methods to smooth the earth's gravity field. Tech. Rep. 329, Dep. of Geod. Sci. and Surv., Ohio State Univ., Columbus.

Kaspar, F., 2004. Development and uncertainty analysis of a global hydrological model. Kassel University Press, Kassel, Germany Ph.D. thesis (in German) (ISBN 3-89958-071-0).

Kollat, J., Reed, P., 2006. Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design. Adv. Water Res. 29 (6), 792–807.

Lettenmaier, D. P., Famiglietti, J. S., 2006. Hydrology: Water from on high. Nature 444 (7119), 562–563.

Luthcke, S. B., Zwally, H. J., Abdalati, W., Rowlands, D. D., Ray, R. D., Nerem, R. S., Lemoine, F. G., McCarthy, J. J., Chinn, D. S., 2006. Recent greenland ice mass loss by drainage system from satellite gravity observations. Science 314 (5803), 1286.

Nash, J. E., Sutcliffe, J. V., 1970. River flow forecasting through conceptual models. part-1 a discussion of principles. J. Hydr. 10 (3), 282–290.

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 Res. Res. 43, W04427, doi:10.1029/2006WR004941.

Niu, G.-Y., Seo, K.-W., Yang, Z.-L., Wilson, C., Su, H., Chen, J., Rodell, M., 2007. Retrieving snow mass from GRACE terrestrial water storage change with a land surface model. Geophys. Res. Lett. 34, L15704, doi:10.1029/2007GL030413.

Niu, G.-Y., Yang, Z.-L., 2006. Assessing a land surface model's improvements with GRACE estimates. Geophys. Res. Lett. 33, doi:10.1029/2005GL025555.

Papa, F., Güntner, A., Frappart, F., Prigent, C., Rossow, W. B., 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.

Ramillien, G., Bouhours, S., Lombard, A., Cazenave, A., Flechtner, F., Schmidt, R., 2008. Land water storage contribution to sea level from GRACE geoid data over 2003-2006. Glob. Planet. Change 60 (3-4), 381.

Ramillien, G., Frappart, F., Cazenave, A., Güntner, A., 2005. Time variations of land water storage from an inversion of 2 years of GRACE geoids. Science 235, 283–301.

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. Wat. Resour. Res. 42, W10403, doi:10.1029/2005WR004331.

Rodell, M., Chen, J., Kato, H., Famiglietti, J., Nigro, J., Wilson, C., 2007. Estimating groundwater storage changes in the mississippi river basin (USA) using GRACE. Hydrogeol. J. 15 (1), 159.

Rodell, M., Famiglietti, J. S., Chen, J., Seneviratne, S. I., Viterbo, P., Holl, S., Wilson, C. R., 2004. Basin scale estimates of evapotranspiration using GRACE and other observations. Geophys. Res. Lett. 31 (20), doi:10.1029/2004GL020873.

Schmidt, R., Flechtner, F., Güntner, A., Knig, R., Meyer, U., Neumayer, K.-

H., Reigber, C., Rothacher, M., Petrovic, S., Zhu, S., 2007. GRACE time-variable gravity accuracy assessment. In: Dynamic planet: Monitoring and understanding a dynamic planet with geodetic and oceanographic tools. IAG Symposium, Cairns, Australia, 22-26 August. 2005, Springer, 237–243.

Schmidt, R., Flechtner, F., Meyer, U., Neumayer, K.-H., Dahle, C., König, R., Kusche, J., 2008a. Hydrological signals observed by the GRACE satellites. Surv. Geoph., doi: 10.1007/s10712-008-9033-3.

Schmidt, R., Petrovic, S., Güntner, A., Barthelmes, F., Wünsch, J., Kusche, J., 2008b. Periodic components of water storage changes from GRACE and global hydrology models. J. Geophys. Res. 113, B08419, doi:10.1029/2007JB005363.

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

Seitz, F., Schmidt, M., Shum, C. K., 2008. Signals of extreme weather conditions in central europe in GRACE 4-d hydrological mass variations. Earth Planet. Sci. Let. 268 (1-2), 165.

Strassberg, G., Scanlon, B., Rodell, M., 2007. Comparison of seasonal terrestrial water storage variations from GRACE with groundwater-level measurements from the high plains aquifer (USA). Geophys. Res. Lett. 34 (14), doi:10.1029/2007GL030139.

Swenson, S., Wahr, J., 2002. Methods for inferring regional surface-mass anomalies from gravity recovery and climate experiment (GRACE) measurements of time-variable gravity. J. Geophys. Res. 107 (B9), 2193, doi:10.1029/2001JB000576.

Swenson, S., Wahr, J., 2006. Post-processing removal of correlated errors in

GRACE data. Geophys. Res. Lett. 33, L08402, doi:10.1029/2005GL025285.

Syed, T. H., Famiglietti, J. S., Zlotnicki, V., Rodell, M., 2007. Contemporary estimates of pan-arctic freshwater discharge from GRACE and reanalysis. Geophys. Res. Lett. 34 (19), doi:10.1029/2007GL031254.

Tamisiea, M. E., Mitrovica, J. X., Davis, J. L., 2007. GRACE gravity data constrain ancient ice geometries and continental dynamics over laurentia. Science 316 (5826), 881–883.

Tang, Y., Reed, P., Wagener, T., 2006. How effective and efficient are multi-objective evolutionary algorithms at hydrologic model calibration? Hydrol. Earth Syst. Sci. 10, 289–307.

Tapley, B. D., Bettadpur, S., Watkins, M., Reigber, C., 2004. The gravity recovery and climate experiment: Mission overview and early results. Geophys. Res. Lett. 31, L09607, doi:10.1029/2004GL019920.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., Sorooshian, S., 2003. Effective and efficient algorithm for multiobjective optimization of hydrologic models. Water Resour. Res. 39 (8), doi:10.1029/2002WR001746.

Wahr, J., Molenaar, M., Bryan, F., 1998. Time variability of the earth's gravity field: Hydrological and oceanic effencts and their possible detection using GRACE. J. Geophys. Res. 108 (B12), 30,205–30,229.

Wahr, J., Swenson, S., Velicogna, I., 2006. Accuracy of GRACE mass estimates. Geophys. Res. Lett. 33 (6), L06401, doi:10.1029/2005GL025305.

Widen-Nilsson, E., Halldin, S., Xu, C.-y., 2007. Global water-balance modelling with WASMOD-M: Parameter estimation and regionalisation. J. Hydr. 340 (1-2), 105.

Zaitchik, B. F., Rodell, M., Reichle, R. H., 2008. Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the mississippi river basin. J. Hydromet., 535–548.

List of Figures

List of Tables

Table 1
(a) Re-calibrated river basins with (b) corresponding area and (c) discharge station. (d) Discharge source and time series for computation of monthly means. (e) Number of WGHM parameters from different processes (S: Soil, SW: Surface water, GW: groundwater, ER: Evaporation and Radiation, SN: Snow, IN: Interception) derived from a sensitivity study against TWSC and river discharge. The underlined process includes the most sensitive parameter. (f) Calibration parameter in corresponding order to row (e) (MCWH: maximum canopy water height, PT: Priestley-Taylor). (g) Significant GRACE derived TWSC periods P_n of basin averages with associated amplitudes A_n and phases ϕ_n, with t_0 =01.01.2005. (h) Cumulative proportion of the significant periods in the full GRACE signal variability.

(a)	Amazon	Mississippi	Congo
(b)	5.9 Mio km^2	3.0 Mio km^2	3.6 Mio km^2
(c)	Obidos	Tarbert Landing	Kinshasa
	1.9°S, 55.5°E	31.6°N, 91.5°W	4.3°S, 15.3°W
(d)	ORE HYBAM	US ACE	GRDC
	2003-2006	2003-2006	1903-1983
(e)	3 SW, 1 GW, 1 S,	1 SW, 1 S, 2 ER	2 ER, 1 S, 1 GW,
	1 IN	1 SN, 1 IN	2 SW
(f)	runoff coefficients	runoff coefficient	radiation proportion
	river velocity	root depth	PT coefficient
	wetland depth	radiation proportion	rooting depth
	GW baseflow coeff.	PT coefficient	GW baseflow coeff.
	rooting depth	snow melt temperature	wetland depth
	MCWH	MCWH	SW baseflow coeff.
(g)	$P_1 = 0.9833$ a	$P_1 = 0.9826$ a	$P_1 = 0.9881$ a
	$A_1 = 146$ mm	$A_1 = 33$ mm	$A_1 = 30$ mm
	$\phi_1 = 3.82$ mon	$\phi_1 = 2.99$ mon	$\phi_1 = 1.82$ mon
	$P_2 = 2.5297$ a	$P_2 = 2.4824$ a	$P_2 = 0.5022$ a
	$A_2 = 22$ mm	$A_2 = 22$ mm	$A_2 = 15$ mm
	$\phi_2 = 19.29$ mon	$\phi_2 = 29.59$ mon	$\phi_2 = 4.75$ mon
(h)	99%	75%	73%

Table 2
Calibration parameter values and their ranges for the calibration work (GW: groundwater, MCWH: maximum canopy water height, PT: Priestley-Taylor, SW: surface water).

Parameter	Standard value and unit	Minimum	Maximum
GW baseflow coefficient	0.01 / day	0.006	0.1
MCWH	0.3 mm	0.1	1.4
PT coefficient	1.26	0.885	1.65
radiation proportion	0.25	0.08	0.54
river velocity	1 m/s	0.05	2.0
root depth mult.	1	0.5	2.0
runoff coefficient mult.	1	0.5	2.0
snow melt temperature	0°C	-3.75	3.75
SW baseflow coefficient	0.01 / day	0.001	0.1
wetland depth	2 m	1.0	5.0

Table 3
RMSE of simulated versus detected hydrological states for the calibration period 01/2003-12/2006: monthly mean river discharge (col. 2-3) and de-trended TWSC signal with non-periodic components (col. 4-5). RMSE is given for the original WGHM (original) and for the selected Pareto solution of the re-calibrated (re-cal.) WGHM version. For the validation period 01/2007-12/2007 RMSE of TWSC is given in col. 6-7.

	Discharge [$kg3/month$]		TWSC [mm]		TWSC [mm] (2007)	
Basin	original	re-cal.	original	re-cal.	original	re-cal.
Amazon	126.6	28.1	49.6	24.7	64.0	34.1
Mississippi	17.2	4.0	21.8	16.1	18.9	13.4
Congo	22.0	13.5	24.7	23.6	25.4	30.5

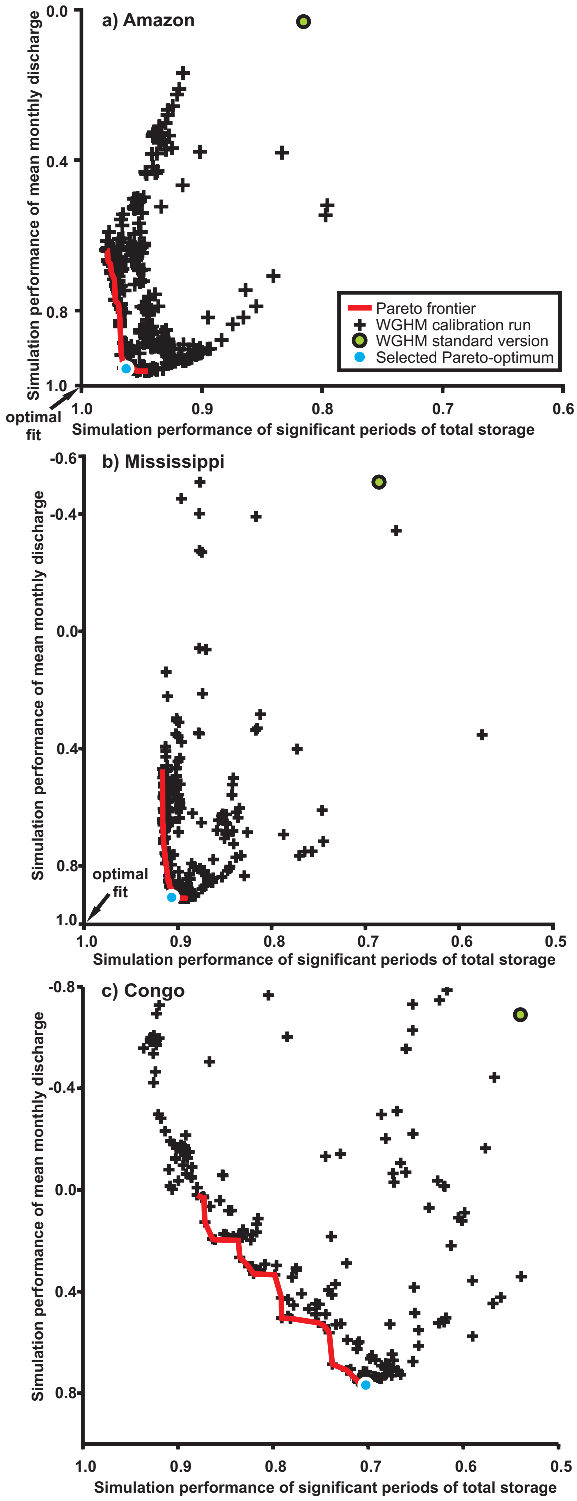

Fig. 1. Calibration results in terms of NSC indicating the simulation performance for a) the Amazon, b) the Mississippi and c) the Congo river basin.

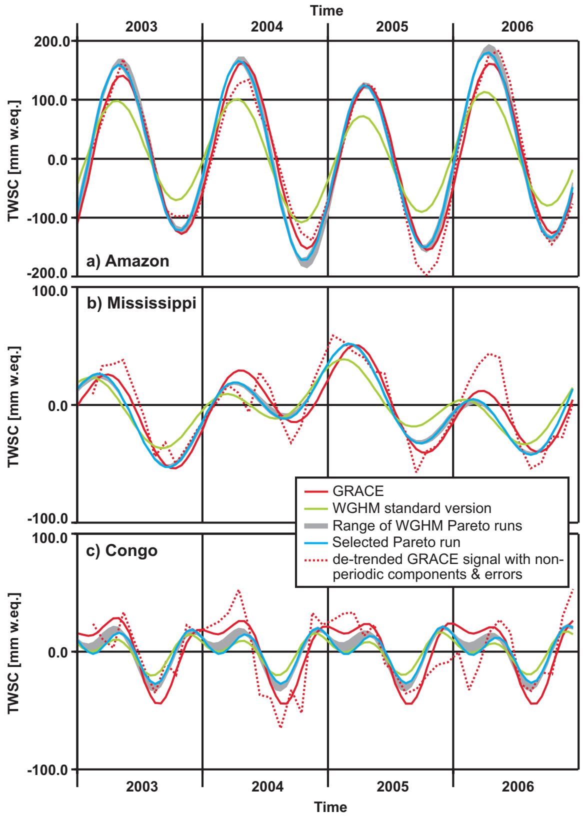

Fig. 2. Calibration results in terms of time series of TWSC from reconstructed periodic terms for a) the Amazon, b) the Mississippi and c) the Congo river basin.

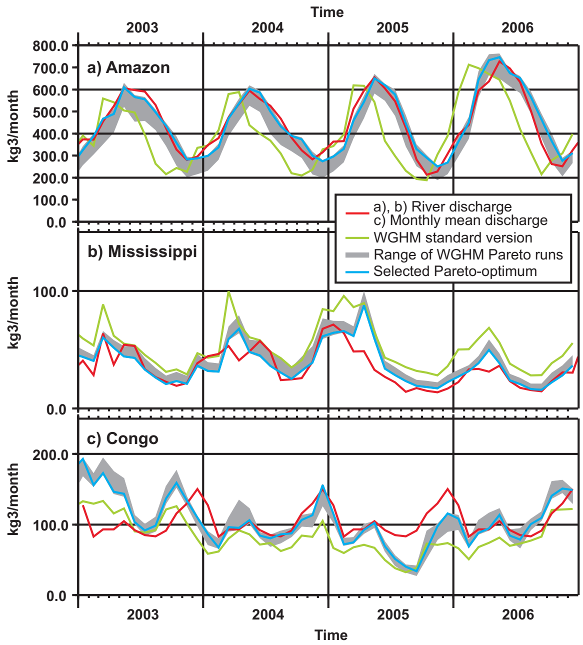

Fig. 3. Calibration results in terms of measured time series of monthly river discharge for a) the Amazon, b) the Mississippi and of mean monthly river discharge for c) the Congo river basin. See Table 1 for detailed sources of discharge measurements.

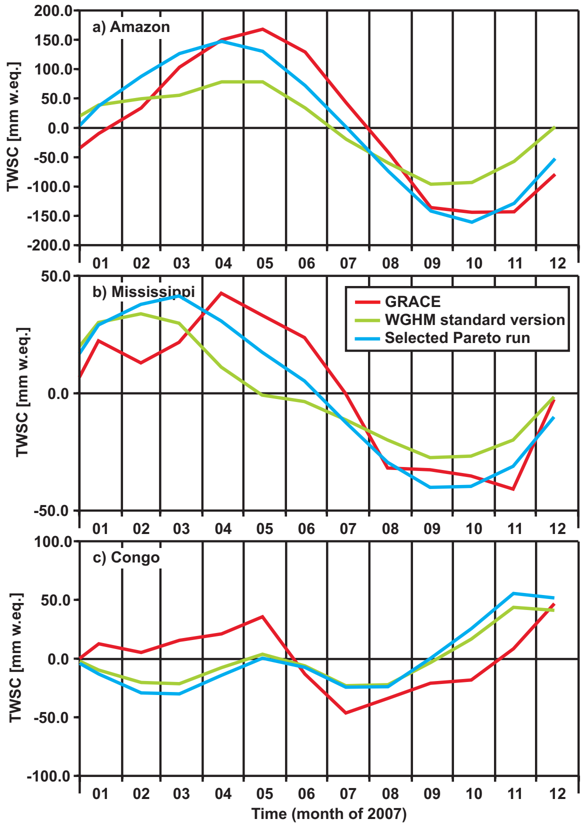

Fig. 4. Validation of de-trended TWSC by model simulations during the period 01-12/2007 (which was not used for calibration) for a) the Amazon, b) the Mississippi and c) the Congo river basins.