

Introduction

First of all, I want to thank both referees for their constructive comments. In the following document I have answered all comments and listed the changes in our manuscript based on these comments. The new manuscript is added together with the final response. This final response is structured as follows:

- Reaction on comments from referee 1 (C5050)
- Reaction on comments from referee 2 (C5257)
- References

Reaction on comments from referee 1 (C5050):

Comment 1

The overall strategy for this study is not clearly explained. Some people use GRACE data to inform hydrological models in a calibration and/or data assimilation approach (e.g. (Milzow et al., 2011). Others compare GRACE data to hydrological model output to check consistency and identify weaknesses in both the GRACE data and the models. What exactly is the purpose here? This should be explained upfront. If the purpose is to inform the model, then it is essential to document model improvement in an independent validation period.

Reaction

The hypothesis and strategy were indeed not given explicitly in our manuscript and improved the overall strategy and motivation in the new version. The main goal of this study is to show that the water mass changes in a hydrological model and surface water in the region can explain a large part of the mass variations, as observed by GRACE. This is in contrast with other papers like (Voss et al., 2013), which claim that a large part of the water mass changes are caused by anthropogenic groundwater depletion. Therefore, we want to show that it is also possible to explain the water mass variation largely by depletion of lakes and reservoirs and by natural groundwater variations.

The main differences between our study and other studies like (Longuevergne et al., 2013; Voss et al., 2013) are:

- The inclusion of the local geology in our current model, which gives a better insight in the groundwater flows and storages.
- The use of streamflow measurements for at least part of the catchment, which can be used to calibrate our model.
- Exclusion of smaller lakes by earlier researches, which still can have an influence on the total water balance due to large water level variations.

The reason that we included GRACE measurements in the model calibration was because we wanted to show that a hydrologic model can mimic water mass variations from GRACE. However, for an independent comparison between the model and GRACE data, this would indeed need a validation

period, which is not available in our case because a similar drought as the one between 2007 and 2009 did not occur during the lifespan of the GRACE satellite mission.

Therefore, in this version of the manuscript we have chosen to use only streamflow measurements as a calibration for our hydrologic model in first instance, but at the same time we also compare the results with the case where GRACE data is included in the calibration process. This results in almost identical results of both calibration methods for the snow reservoir, unsaturated reservoir and fast runoff reservoir and a difference of about 8 mm EWH for the groundwater reservoir. The total decline from the calibration with only streamflow data was 39 mm EWH and the total decline including GRACE as a calibration parameter is 48 mm EWH. This shows that model results are quite consistent for both methods and lead to the same conclusions, although there is a small difference in water mass decline.

Changes made to manuscript

- Hypothesis/strategy and motivation for research edited in (1 Introduction line 80-90)
- Model approach updated (1 Introduction line 108-119)
- Calibration method explanation added (3.4 Model Calibration line 382-398)
- Method for model selection updated with new calibration method (3.4 Model Calibration line 410-440)
- Comparison two different calibration methods added (4.4 Natural groundwater variations line 524-541)
- Comparison two different calibration methods added (4.6 GRACE and modelled mass variations)
- Model results and figures updated based on new calibration method

Comment 2

GRACE processing: There are basically two ways of processing GRACE data in the literature: Either the level 1 range rate data is inverted for spherical harmonic coefficients or for mass changes on a grid. The second approach is usually termed the MASCON approach (Rowlands, 2005). The method you present here seems to be a hybrid between the two, in the sense that mascon parameters are fitted to level 2 spherical harmonic coefficients instead of range rate data. I know that GRACE processing is not the focus of this paper, but it would still be nice to explain the pros and cons of doing it this way. For instance, I do not understand if you can retrieve the 6 MASCONS shown in Figure 2 independently of the rest of the planet or if you always have to invert for a global set of mascons? Also, why do the MASCONS have circular shape, why not adapt them to the geometry of the basins of interest as for instance in (Krogh et al., 2010) ?

Reaction

To explain the GRACE processing more clearly, the following text is added to the manuscript on line 163-170:

"This method is based of GRACE level-2 data from the Center of Space Research (CSR) and includes modifications for the gravitational flattening term C_{20} from satellite laser ranging. Furthermore the method also considers degree 1 terms associated with geo-center motions as a result of geophysical loading phenomena. The used GRACE method is not a spatial averaging kernel method, instead to obtain equivalent water levels over a region one has to add up the signal from the individual mascons. The signal at these mascons is obtained via a global inversion method."

This means that we use a global set of mascons and that the signal at these mascons was obtained by means of a global inversion method as described in (Schrama et al, 2014). Our original manuscript lacked a description at this point

A comparison between different GRACE methods for mass calculation can be found in (Shepherd et al., 2012). As long as the starting conditions of these methods are the same the differences turn out to be small. The method in (Schrama et al 2014) is an extension of one of the GRACE methods contained in (Shepherd et al 2012), and it shows nearly identical results.

The advantage of the approach in (Schrama et al 2014) is that resulting EWH values are part of a standard GRACE product and will be consistent with other regional solution worldwide. The disadvantage is that we have to adapt our study area to the given mascons. We are thus not able to use approaches like (Krogh et al., 2010), which adapt the mascons to the geometry of the studied basin. However, the changes in EWH due to adaption to the exact study area are small. We tested this using a comparison between the GRACE results for the current dataset and two extended datasets, see Figure 1 in this rebuttal. This shows that the GRACE signal remains largely the same, with only small differences in yearly fluctuation and water depletion over the whole period. If these relatively large changes in study area induce only small changes in the total GRACE signal it is likely that the differences due to adaptation to geometry are negligible.

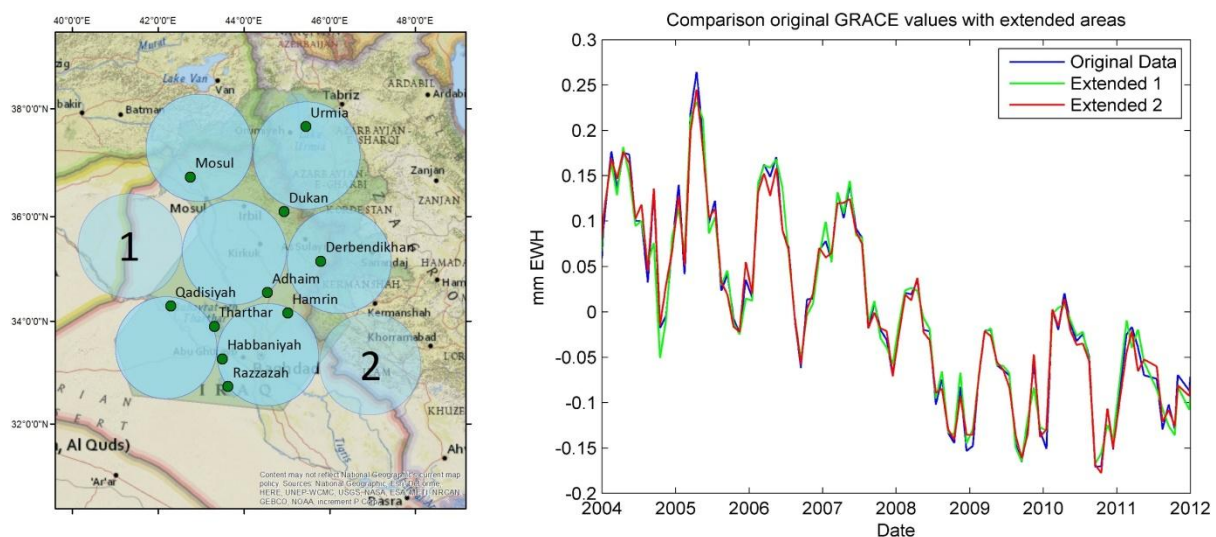


Figure 1 - Comparison of EWH values for current and extended areas.

In the old version a method was given to refine our results (page 11538 line 15-22), but it proved to have negligible influence and was therefore omitted in the new version.

Changes to manuscript

- Description of GRACE processing updated (3.1 GRACE mass calculations line 160-174)
- Comparison between our and other GRACE processing methods added (3.1 GRACE mass calculations line 175-186)

Comment 3

The purpose and basis of the various corrections described on page 11539 are not really clear to me. Why does the contribution of the lakes have to be scaled by 1/2 and 1/3? How did these numbers come about? Why is the GRACE region extended into the southwestern desert, if this is really not part of the study area of interest? All this needs to be motivated and explained much better.

Reaction

The two corrections made to the GRACE results are those based on the extension of our study area to the south west and the correction for lakes at the border of the study area, which are both related to water mass leakage

The water mass depletion of Lake Tharthar is more than 50 % of the total water mass depletion from lakes in our area (figure Figure 2). Therefore we extended the study area to the south-west to be sure that the water mass depletion from this lake was fully captured in our GRACE mass calculations.

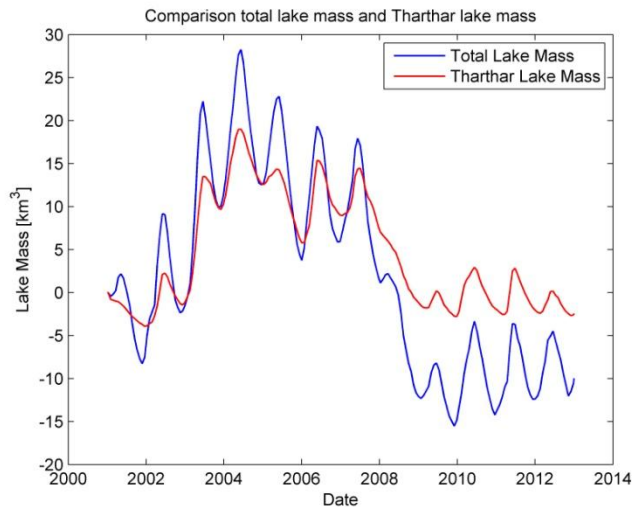


Figure 2 - Mass variation from Lake Tharthar compared with total mass variation from lakes in study area

Because Lake Urmia and lake Razzazah have much less impact on the total lake mass, we have chosen to correct their mass leakage out of our study area based on the method described in (Longuevergne et al., 2013). In this paper it is demonstrated how point masses within a circular mascon contribute to this mascon, based on the size of the mascon and the distance from the mascon center. For this correction we geometry of our whole study area and calculated the correction factors based on the distance of the lakes from the centre of our region. During the revision of this paper we studied the locations of the mascons and lake centres in more detail, which resulted in a slightly larger contribution of lake Urmia of 0.6 and a contribution of Lake Razzazah of 0.65. Corrections for other lakes in the region were also calculated, but the mass changes due to the correction factors of these lakes are negligible.

Changes to manuscript

- More detailed description and motivation of study area extension given (3.1 GRACE mass calculations line 185-240)
- Explanation and motivation of lake mass correction added (3.1 GRACE mass calculations line 198-204 and Derivation of lake mass line 291-297)

Comment 4

It is stated that the aquifers are karstified, and highly transmissive. Water level and storage variations in such aquifers are generally suppressed because of the high transmissivity and water is effectively drained over short time scales. Lines 4-9 on page 11548 seem to suggest the opposite. It would be nice to discuss if the simulated groundwater storage variations are reasonable given the available hydrogeological knowledge and observations from the region.

Reaction

This point indeed deserves an extended elaboration. There is quite some research on this because the water supply in Northern Iraq is dependent on springs that emerge from these aquifers (Ali and Stevanovic, 2010; Ali et al., 2009b). These studies show that the discharge from these karstic aquifers can be split up in two components. The first component is a rapid discharge within a month after rain events, which is related to channels and large fractures in the limestone aquifer. The second component is a much slower discharge, which is generally stable till the new wet season. While the first component is covered by the fast reservoirs in our hydrologic model, the second component is modelled as the slow groundwater reservoir. The recession value of these groundwater reservoirs is based on the recession values of three large springs in the area, the Bestansur, Zalim and Sachinar (Ali and Stevanovic, 2010; Ali et al., 2009a, 2009b), which respective recession coefficients are 0.0049, 0.0032 and 0.0038. The chosen recession bandwidth is chosen between 0.0035 and 0.0045.

Changes in manuscript:

- More elaborate description about groundwater flow added (3.3 Rainfall Runoff model 359-377)

Comment 5

Throughout the manuscript, language and grammar should be checked and clarity of the wording should be improved, see also details listed below. Please always call the same things by the same names. For example, the manuscript sometimes talks about the “Lesser Zab catchment” and sometimes about the “smaller Dukan area”, although, I believe, those two names refer to exactly the same thing.

Reaction

Detailed comments are corrected and paper is checked for grammar and inconsistent wording. Also, tributary names and lake names are added to figure 2 and figure 5 of the manuscript. The Dukan area referred to in the text is a part of the Lesser Zab catchment upstream from Lake Dukan, which is now explained in the text and caption of figure 5 of the manuscript.

Changes in manuscript

- Lake names added to figure 2
- Tributary names added to figure 5

A lot of place names are used in the text, but cannot be found on any of the maps. I think the paper would benefit from a detailed base map that shows and names all places, rivers and lakes referred to in the text.

Reaction

tributary names are added in figure 2 and lake names are added in figure 5

Changes in manuscript

- Lake names added to figure 2
- Tributary names added to figure 5

Comment 7

Please explain how the uncertainty bands for the surface water storage and snow storage in fig 8 were derived.

Reaction

The errors due water level measurements are derived based on the estimated errors in lake levels given by (Crétau et al., 2011; USDA/FAS, 2013). These time series of water level errors are then multiplied with the total area of these lakes and added up:

$$\sigma^t = \sqrt{\frac{\sum_{i=1}^n (\sigma_i^t A_i^t)^2}{n}}$$

Where n is the number of lakes, σ_i^t is the estimated error in water level for one of the lakes at time t and A_i^t and the calculated lake area of a particular lake at time t .

This uncertainty does not include the errors in lake area calculation, but these are difficult to quantify and are likely small due to the use of stage-area curves. For example, the difference in lake size from Lake Mosul and the derived values from (Issa et al., 2013) are about 10 km² but it only causes significant differences in the stage-volume curve for high water levels. Additionally, the relative error for larger lakes is much smaller because the ratio between lake shore length and lake area becomes smaller with increasing lake size. In figure Figure 3 this is illustrated using the resulting stage-area curve from lake Tharthar and Lake Qadasiyah, which have an large contribution to the total water mass depletion. Through the use of a linear/cubic fit of 192 measurements, most of the errors due to lake area and lake level calculations are filtered out.

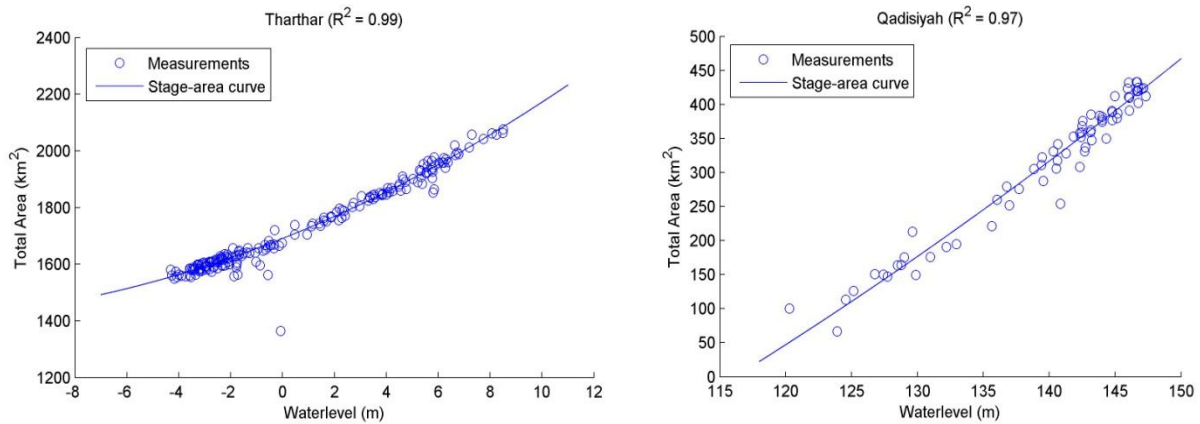


Figure 3 - Stage-area curve Lake Tharthar and Lake Qadasiyah

We replaced the lake mass graph with a new one based on a 95% confidence interval instead of 68%, to ensure consistency with the GRACE graphs. This error does not include the leakage of water mass of lakes outside the study area due to GRACE mass calculations, because these values cannot be clearly quantified.

The uncertainty of the snow storage is based uncertainties in the precipitation from the GLDAS model. These uncertainties were derived from the comparison between monthly rainfall from the GLDAS model and four rainfall stations in the region. In the new version, snow mass is added to the rainfall-runoff model, although differences are small compared to the original method.

Changes in manuscript

- Derivation of lake mass uncertainty added (3.2 Derivation of lake mass line 274-281)

- Description of possible additional uncertainties in lake mass derivation added (3.2 Derivation of lake mass line 282-297)

Comment 8

Why use snow from GLDAS? Why not run a simple snow accumulation and melt routine on top of the hydrological model? At least that would ensure consistent precipitation input. How exactly is the TRMM product corrected for snow from GLDAS?

Reaction

Snow mass from the GLDAS model was used because snow mass in the study area does not have much effect on the total mass decline and it is not likely that a snow accumulation and melt routine would have much effect on the total modelled snow mass. Therefore, the simplest approach was used. Together with the addition of snow mass, the snowfall from GLDAS was subtracted from TRMM precipitation and the snowmelt was added to TRMM precipitation.

Addition of a snow accumulation and melt routine would add to the consistency of the model and was therefore added in the new version. Figure 4 shows a comparison between the snow mass derived from GLDAS with our rainfall-runoff model, which confirms that the results are similar. The added variables are a threshold temperature and a melting coefficient (mm/C/day), which were kept the same for all three geologic zones. Temperature values were derived from GLDAS on a daily basis (Rodell et al., 2004). Because temperature differences are large due to height differences in the basin, it was not possible to model this in a lumped approach. Therefore a semi-distributed model setup is used, with a distributed approach for the snow and unsaturated reservoir.

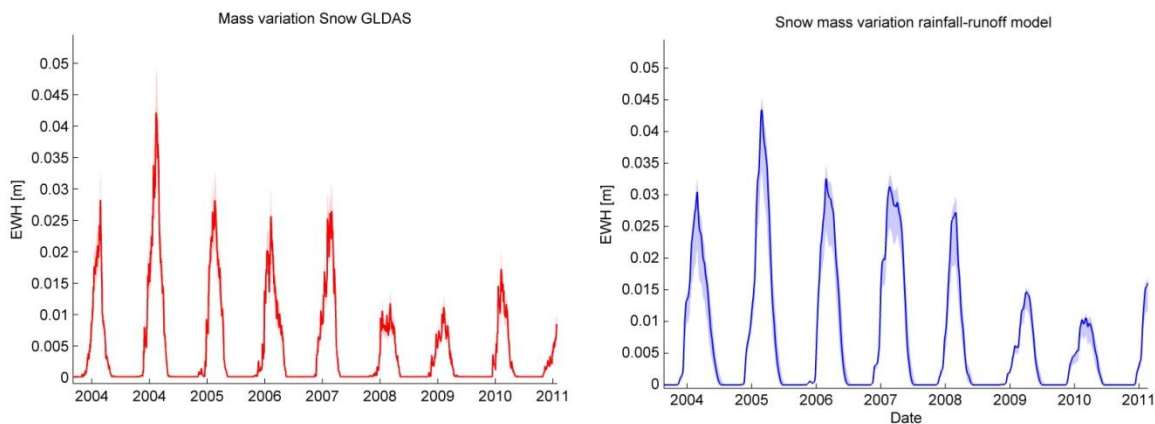


Figure 4 - (left) Snow mass from GLDAS. (right) Snow mass based on a snow accumulation and melting routine

Changes in manuscript:

- Description about snow mass derivation from GLDAS removed

- Description of snow mass reservoir and distributed model approach added (3.3 Rainfall-runoff model line 339-348)

Comment 9

Very little is said about how the uncertainties of the simulated storages have been determined (fig 9). To me, these uncertainty bounds look very narrow. Can they be justified? For instance, were the errors due to parameter transfer from gauged to ungauged parts of the catchment taken into account? Can these uncertainty estimates be hold up against real observations? It would be nice to see the comparison of simulated and observed hydrographs at least for the single available station...

Reaction

The uncertainty band are formed by the ensemble of all models on the pareto front (like in Werth et al., 2009), which is based on the Nash-Sutcliffe and log Nash-Sutcliffe performance. We used the difference between the maximum and minimum values of the ensemble as the uncertainty band. A discharge graph for 2005-2008 is given in figure Figure 5 and shows an ensemble of all model runs from the pareto front. This graph shows that the differences between modelled and measured values can still be substantial, but the baseflow values during the dry season are mainly within the minimum and maximum discharge bands. The pareto ensemble is a measure of uncertainty of the parameter values of the model, but does not represent other uncertainties based on the forcing, up scaling or model structure.

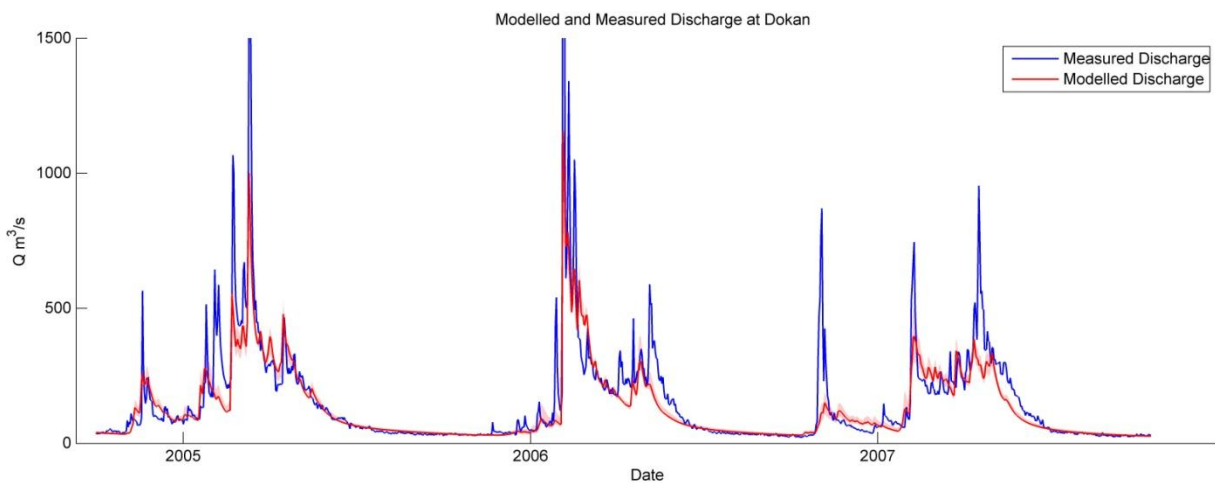


Figure 5 - Modelled and measured discharge at inflow of Lake Dukan

For instance, the effect of errors in rainfall and temperature measurements is not included in this procedure. Also, the spreading of rainfall varies over each TRMM pixel, which could cause lower peak flows in our model. Another influencing factor is the water use upstream from our gauging station, which likely causes a reduction in streamflow, but was not considered in our model. This can result in an

underestimation of the groundwater variation, because baseflow values are likely higher than those used in our model calibration.

Because the mixture of uncertainties between GRACE, lake mass and the model output is questionable the residual graph is removed.

Changes in manuscript:

- Section on model discharge is added (4.3 Modelled discharge Dukan catchment line 500-522)
- Uncertainty bands of rainfall runoff model explained (3.4 Model Calibration line 435-438)
- Added figure 8 about modelled and measured discharge.
- Removed figure about residual water mass
- Discussion about upscaling of catchment added (4.4 Natural groundwater variations line 563-571)

Details

1. Page 11534, line 11: "Corrected for" should probably be "estimated". Total mass variation should include lakes. > **Sentence changed**
2. Page 11534, line 19: "Depletion of geology"? Please re-word. I guess you refer to natural depletion of groundwater. > **Sentence changed**
3. Page 11535, line 6: Is this predicted decrease due to anthropogenic climate change? > **changed to climate models**
4. Page 11535, line 20: "riparian" should be "upstream". > **Changed**
5. Page 11537 line 9: Please give the version of the TRMM 3B42 product used here. Versions 6 and 7 are quite different in other parts of the world, here too? > **V7 added**
6. Page 11537 line 22: Please reword. The lakes influence the GRACE signals. > **Changed**
7. Page 11538 line 9: "dises" should be circles?? > **This part is removed.**
8. Page 11542, line 5: "bias-corrected" may be better than "calibrated" > **Changed**
9. It is not unambiguously stated in the text how GRACE and the model are compared (e.g. page 11546, line 23). It is stated that GRACE and the model are compared, while in fact the comparison is between GRACE minus surface water storage and the model. > **Changed**
10. Figure 1: Both maps should have coordinate systems/scale bars and colorbar legends. > **Added**
11. Figure 2: Needs coordiates/scales > **Added**

Reaction on comments from referee 2 (C5257):

Comment 1

The method should generally be explained in more detail, especially the assumptions/constraints, etc. The way of GRACE processing using the mascon approach is only vaguely given. Moreover, GRACE is used for calibration and data reduction. Does this not imply internal correlations?

Reaction

The assumptions/constraints for several model steps were indeed only shortly discussed. In the new version of our manuscript modelling steps are explained in more detail and assumptions are given explicitly.

To explain the GRACE processing more clearly, the following text is added to the manuscript on line 163-170:

"This method is based of GRACE level-2 data from the Center of Space Research (CSR) and includes modifications for the gravitational flattening term C20 from satellite laser ranging. Furthermore the method also considers degree 1 terms associated with geo-center motions as a result of geophysical loading phenomena. The used GRACE method is not a spatial averaging kernel method, instead to obtain equivalent water levels over a region one has to add up the signal from the individual mascons. The signal at these mascons is obtained via a global inversion method."

The reason that we included GRACE measurements in the model calibration was because we wanted to show that a hydrologic model can mimic water mass variations from GRACE. However, for an independent comparison between the model and GRACE data, this would indeed need a validation period, which is not available in our case because a similar drought as the one between 2007 and 2009 did not occur during the lifespan of the GRACE satellite mission.

Therefore, in this version of the manuscript we have chosen to use only streamflow measurements as a calibration for our hydrologic model in first instance, but at the same time we also compare the results with the case where GRACE data is included in the calibration process. This results in almost identical results of both calibration methods for the snow reservoir, unsaturated reservoir and fast runoff reservoir and a difference of about 8 mm EWH for the groundwater reservoir. The total decline from the calibration with only streamflow data was 39 mm EWH and the total decline including GRACE as a calibration parameter is 48 mm EWH. This shows that model results are quite consistent for both methods and lead to the same conclusions, although there is a small difference in water mass decline.

Changes in manuscript

- Description of GRACE processing updated (3.1 GRACE mass calculations line 163-170)

- Description of GRACE processing updated (3.1 GRACE mass calculations line 160-174)
- Hypothesis/strategy and motivation for research edited in (1 Introduction line 80-90)
- Model approach updated (1 Introduction line 108-119)
- Calibration method explanation added (3.4 Model Calibration line 382-398)
- Method for model selection updated with new calibration method (3.4 Model Calibration line 410-440)
- Comparison two different calibration methods added (4.4 Natural groundwater variations line 524-541)
- Model results and figures updated based on new calibration method

Comment 2

The limitations and error contributions of the various model reductions and assumptions should be given explicitly. Only then the usefulness and quality of the proposed “rainfall-runoff” model can be evaluated.

Did the authors compare their results with independent GRACE-based estimates of mass changes in that region (see, e.g., Sneeuw et al. 2014 for lake Urmia).

Reaction

Uncertainty of the rainfall-runoff model is given explicitly in the new version of the manuscript. Also, limitations and possible errors in the model are added.

The uncertainty band of the rainfall-runoff model are formed by the ensemble of all models on the pareto front (like in Werth et al., 2009), which is based on the Nash-Sutcliffe and log Nash-Sutcliffe performance indexes. We used the difference between the maximum and minimum values of the ensemble as the uncertainty band. A discharge graph for 2005-2008 is given in figure Figure 6 and shows an ensemble of all model runs from the pareto front. This graph shows that the differences between modelled and measured values can still be substantial, but the baseflow values during the dry season are mainly within the minimum and maximum discharge bands. The pareto ensemble is a measure of uncertainty of the parameter values of the model, but does not represent other uncertainties based on the forcing, up scaling or model structure.

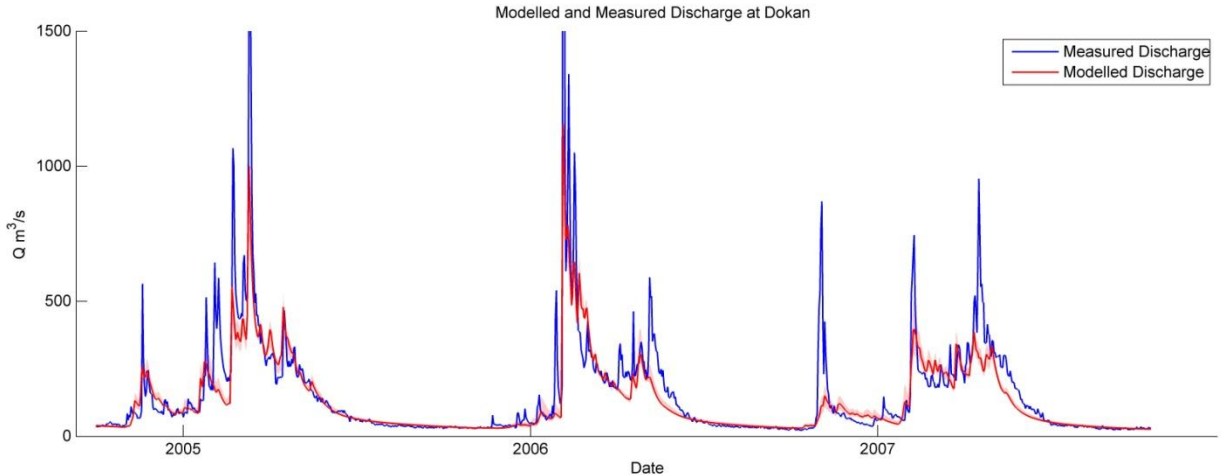


Figure 6 - Modelled and measured discharge at inflow of Lake Dukan

For instance, the effect of errors in rainfall and temperature measurements is not included in this procedure. Also, the spreading of rainfall varies over each TRMM pixel, which could cause lower peak flows in our model. Another influencing factor is the water use upstream from our gauging station, which likely causes a reduction in streamflow, but was not considered in our model. This can result in an underestimation of the groundwater variation, because baseflow values are likely higher than those used in our model calibration. Because the mixture of uncertainties between GRACE, lake mass and the model output is questionable the residual graph is removed.

We did compare our GRACE result with different other GRACE-based estimates from (Crétau et al., 2011; Longuevergne et al., 2013; Voss et al., 2013). These studies obtain similar results, but the extent of our study area is not exactly the same. Also, comparison between our method and other common methods by (Shepherd et al., 2012) showed that differences are generally small. Comparison with results from (Tourian et al., 2015) ,which represents a part of our study area were not made. Figure Figure 2 gives a comparison between the resulting lake mass variation from our study and from (Tourian et al., 2015). This shows that although yearly variations between those models differ (especially 2006-2008), the trend in mass decline is almost the same. Differences between the results from different study are possibly related with the used satellite data, which was mainly from Envisat in (Tourian et al., 2015) and Jason 1&2 in our study.

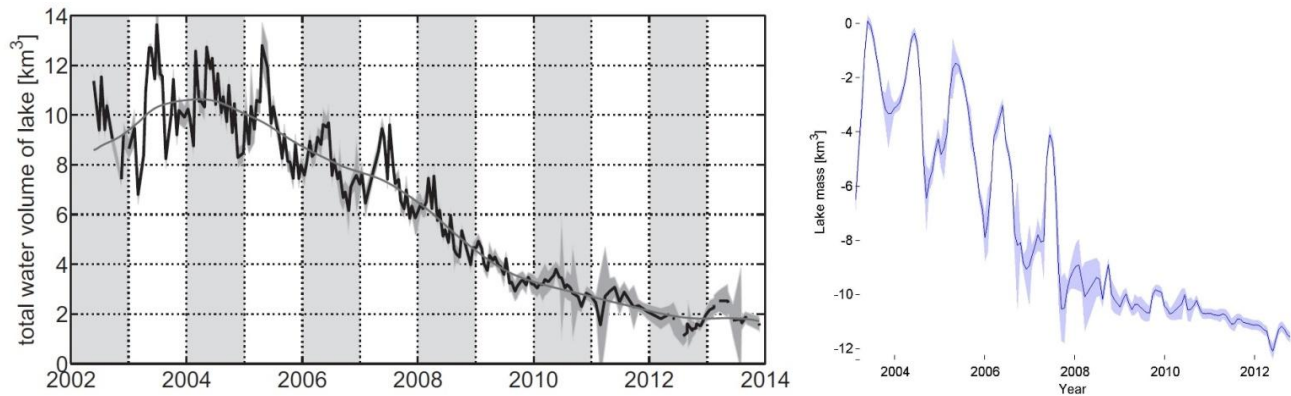


Figure 7 - Modelled water mass from Sneeuw and this study

Changes in manuscript

- Section on modelled discharge is added. In this section the derivation of the uncertainty band of the rainfall runoff model is given. Also other possible sources of uncertainty are discussed. (4.3 Modelled discharge Dukan catchment line 500 - 522)
- Uncertainty bands of rainfall runoff model explained (3.4 Model Calibration line 435-438)
- Discussion about upscaling of catchment added (4.4 Natural groundwater variations line 563-571)
- Added figure 8 about modelled and measured discharge.
- Removed figure about residual water mass

Comment 3

Throughout the paper starting with the abstract, the authors use different (partly incomplete) units for representing mass variations. For example, mass loss is represented sometimes in mm and sometimes in km³, for the first probably mm in EWH is meant, where the second is a volume change. It should be used consistently.

Reaction

Mass variations are always given as mm EWH in the new manuscript. Masses in km³ are also given in some cases, but always together with a value in mm EWH.

Comment 4

On page 11539 (lines 8-9), the reason for selecting the weight 1/2 for the lake mass of Urmia and 1/3 for lake Razazzah is not obvious. Do these coefficients come from some empirical model?

Reaction

The correction for the lake masses of Lake Urmia and Lake Razzazah are based on work from (Longuevergne et al., 2013). In this paper is demonstrated how point masses within a circular mascon contribute to this mascon, based on the size of the mascon and the distance from the mascon center. For this correction we geometry of our whole study area and calculated the correction factors based on the distance of the lakes from the centre of our region. During the revision of this paper we studied the locations of the mascons and lake centres in more detail, which resulted in a slightly larger contribution of lake Urmia of 0.6 and a contribution of Lake Razzazah of 0.65. Corrections for other lakes in the region were also calculated, but the mass changes due to the correction factors of these lakes are negligible.

Changes in manuscript

- Explanation and motivation of lake mass correction added (3.1 GRACE mass calculations line 195-205 and 3.2 Derivation of lake mass line 291-297)

Comment 5ⁱ

Are the estimated lake mass variations reliable and accurate?

Reaction

The errors due water level measurements are derived based on the estimated errors in lake levels given by (Crétau et al., 2011; USDA/FAS, 2013). These time series of water level errors are then multiplied with the total area of these lakes and added up:

$$\sigma^t = \sqrt{\frac{\sum_{i=1}^n (\sigma_i^t A_i^t)^2}{n}}$$

Where n is the number of lakes, σ_i^t is the estimated error in water level for one of the lakes at time t and A_i^t and the calculated lake area of a particular lake at time t .

The estimated errors are based on the footprint from the different satellites and comparison between in-situ measurements and satellite altimetry . Because the error estimates came with the data from (Crétau et al., 2011; USDA/FAS, 2013), it is quite difficult for us to give an idea of the accuracy of the estimated errors. For example, the graphs in figure Figure 2 show that the water levels difference of two separate satellite products can be much larger than the estimated errors in those products.

The presented uncertainty does not include the errors in lake area calculation, because these are difficult to quantify and are assumed to be relatively small due to the use of stage-area curves. For example, the difference in lake size from Lake Mosul and the derived values from (Issa et al., 2013) are about 10 km² but it only causes significant differences in the stage-volume curve for high water levels. Additionally, the relative error for larger lakes is much smaller because the ratio between lake shore length and lake area becomes smaller with increasing lake size. In figure Figure 3 this is illustrated using the resulting stage-area curve from lake Tharthar and Lake Qadasiyah, which have an large contribution

to the total water mass depletion. Through the use of a linear/cubic fit of 192 measurements, most of the errors due to lake area and lake level calculations are filtered out.

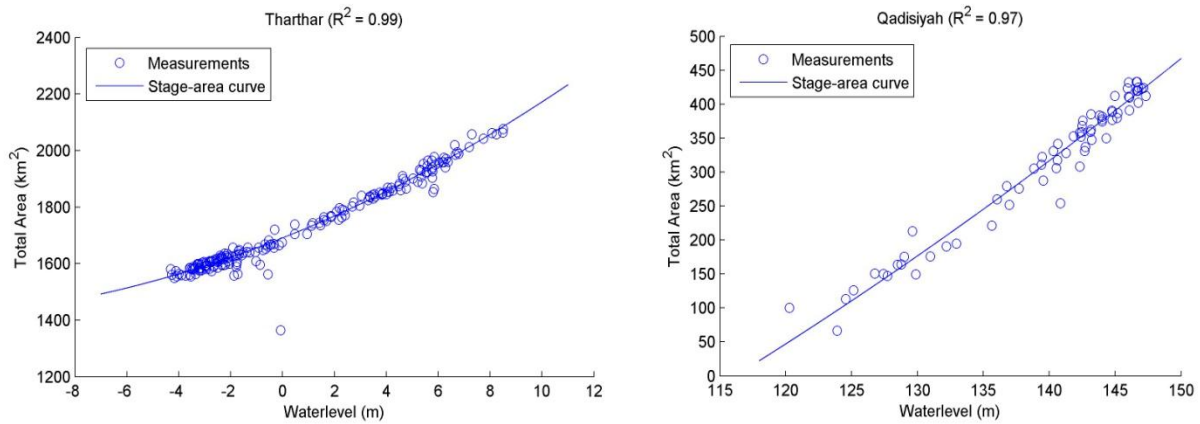


Figure 8 - Stage-area curve Lake Tharthar and Lake Qadisiyah

We replaced the lake mass graph with a new one based on a 95% confidence interval instead of 68%, to ensure consistency with the GRACE graphs. This error does not include the leakage of water mass of lakes outside the study area due to GRACE mass calculations, because these values cannot be clearly quantified.

Changes in manuscript

- Derivation of lake mass uncertainty added (3.2 Derivation of lake mass line 274-281)
- Description of possible additional uncertainties in lake mass derivation added (3.2 Derivation of lake mass line 282-297)

Comment 6

For the snowfall and snowmelt calculations, the authors used the GLDAS model. How reliable is that model for such calculations?

Reaction

Snow mass from the GLDAS model was used because snow mass in the study area does not have much effect on the total mass decline and it is not likely that a snow accumulation and melt routine would have much effect on the total modelled snow mass. Therefore, the simplest approach was used.

Addition of a snow accumulation and melt routine would add to the consistency of the model and was therefore added in the new version. Figure Figure 4 shows a comparison between the snow mass derived from GLDAS with our rainfall-runoff model, which confirms that the results are similar. The added variables are a threshold temperature and a melting coefficient (mm/C/day), which were kept the same for all three geologic zones. Temperature values were derived from GLDAS on a daily basis (Rodell et al., 2004).

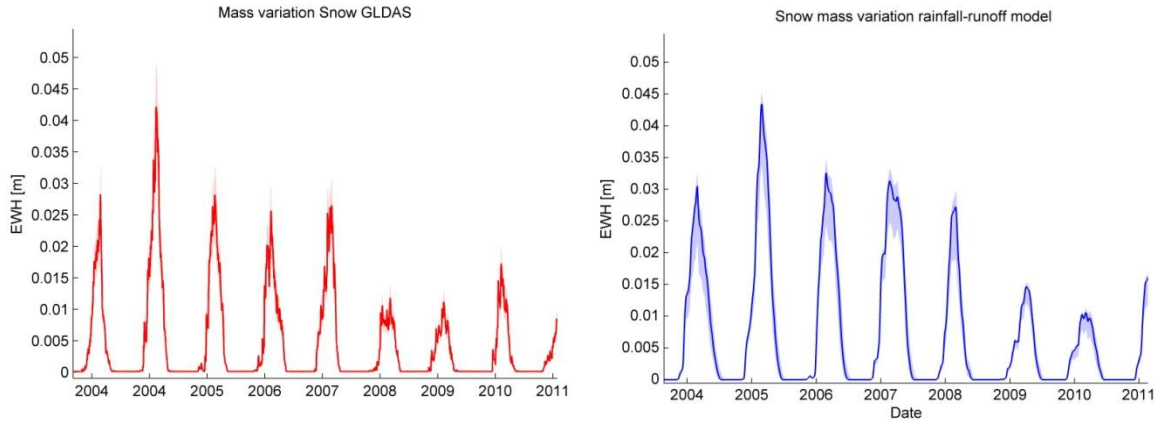


Figure 9 - (left) Snow mass from GLDAS. (right) Snow mass based on a snow accumulation and melting routine

Changes in manuscript

- Description about snow mass derivation from GLDAS removed
- Description of snow mass reservoir and distributed model approach added (3.3 Rainfall-runoff model line 339-348)

Comment 7

For groundwater level estimation, how many stations are used and how reliable are the data?

Reaction

Groundwater levels are based on the groundwater reservoirs in our rainfall-runoff model, which was stated on line 20-25 of page 11545 and in the caption of figure 9.

Uncertainty of these groundwater levels are therefore based on the bandwidth of the pareto ensemble.

Details

1. All abbreviations (e.g. GRACE) should be explained at the first time of appearance. For example, GRACE is explained more often, see page 11535, line 28. But other abbreviations were never defined, e.g., WGHM, GGP and SD, etc. > **Checked and changed**
2. All data used, incl. background models should be summarized in a table. > **Added**
3. On page 11537 (line 12), there is one more “and” that should be removed. > **Changed**
4. The word “River” is sometimes written in capital and sometimes in small letters. > **Checked and changed**
5. In the section 3.1, the title “GRACE mass calculations” should be changed to “GRACE mass variation calculation” or to something similar. In addition, it should be said which GSM model from which analysis center has been used in the GRACE calculations. > **Changed and added**

6. On page 11543 in the formula section, punctuation should be used at the end of the formulas. >
This comment is not clear for me. Should I add a dot to end the formula?

Figures

1. In Fig. 1, the legend for the colours should be included to specify the range of rainfall and topography variations. > **Changed**
2. For Fig. 6, it should be explained how to read and how to understand what is shown there. What can be learnt from such a representation?
The pareto front is a measure for the uncertainty in model parameters, based on different performance indicators. Description of the pareto front/ensemble is added in section: 4.3 Modelled discharge Dukan catchment.
3. The x-axis of Fig. 8 is not labelled and has to be corrected for the starting year (year 2004 is used two times). > **Changed**
4. The words “Left” and “Right” in the caption of Fig. 9 are differently used as in Fig. 5. > **Changed**
5. The residuals in terms of EWH that are represented in Fig. 10 are rather big. Any explanation for this? Did the authors consider soil moisture at all levels down to the depth of 2 m?
Differences in soil moisture could indeed be the reason of these residuals as is stated in section GRACE and modelled values. Other causes could be errors in rainfall data, model structure, anthropogenic activities and errors the mass leakage from lakes in the region.

References:

- Ali, S. S. and Stevanovic, Z.: Time Series Analysis of Saraw Springs - SE of Sulaimaniya, Iraqi Kurdistan Region, in *Advances in Research in Karst Media*, edited by B. Andreo, F. Carrasco, J. J. Durán, and J. W. LaMoreaux, pp. 89–94, Springer, Leipzig., 2010.
- Ali, S. S., Stevanovic, Z. and Al-Jabbari, M. H.: Auto and Cross Correlation Analysis of Bestansur spring- Sharazoor Plain Iraqi Kurdistan Region, in *Proceeding of 3rd scientific conference of the College of Science, University of Baghdad*, pp. 1941–1948., 2009a.
- Ali, S. S., Stevanovic, Z. and Jemcov, I.: The Mechanism and Influence on Karstic Spring Flow - Sachinar Spring Case Example, Sulaimaniya, NE Iraq, *Iraqi Bull. Geol. Min.*, 5(2), 87–100, 2009b.
- Crétaux, J.-F., Jelinski, W., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., Gennero, M.-C., Nino, F., Abarca Del Rio, R., Cazenave, A. and Maisongrande, P.: SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data, *Adv. Sp. Res.*, 47(9), 1497–1507, doi:10.1016/j.asr.2011.01.004, 2011.
- Issa, I. E., Al-Ansari, N. and Knutsson, S.: Sedimentation and new operational curves for Mosul Dam, Iraq, *Hydrol. Sci. J.*, 58(7), 1456–1466, doi:10.1080/02626667.2013.789138, 2013.
- Krogh, P. E., Andersen, O. B., Michailovsky, C. I. B., Bauer-Gottwein, P., Rowlands, D. D., Luthcke, S. B. and Chinn, D. S.: Evaluating terrestrial water storage variations from regionally constrained GRACE mascon data and hydrological

models over Southern Africa – preliminary results, *Int. J. Remote Sens.*, 31(14), 3899–3912, doi:10.1080/01431161.2010.483483, 2010.

Longuevergne, L., Wilson, C. R., Scanlon, B. R. and Crétaux, J. F.: GRACE water storage estimates for the Middle East and other regions with significant reservoir and lake storage, *Hydrol. Earth Syst. Sci.*, 17(12), 4817–4830, doi:10.5194/hess-17-4817-2013, 2013.

Milzow, C., Krogh, P. E. and Bauer-Gottwein, P.: Combining satellite radar altimetry, SAR surface soil moisture and GRACE total storage changes for hydrological model calibration in a large poorly gauged catchment, *Hydrol. Earth Syst. Sci.*, 15(6), 1729–1743, doi:10.5194/hess-15-1729-2011, 2011.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin*, J. K., Walker, J. P., Lohmann, D. and Toll, D.: The Global Land Data Assimilation System, *Bull. Am. Meteorol. Soc.*, 85(3), 381–394, doi:10.1175/BAMS-85-3-381, 2004.

Rowlands, D. D.: Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements, *Geophys. Res. Lett.*, 32(4), L04310, doi:10.1029/2004GL021908, 2005.

Shepherd, A., Ivins, E. R., a, G., Barletta, V. R., Bentley, M. J., Bettadpur, S., Briggs, K. H., Bromwich, D. H., Forsberg, R., Galin, N., Horwath, M., Jacobs, S., Joughin, I., King, M. a, Lenaerts, J. T. M., Li, J., Ligtenberg, S. R. M., Luckman, A., Luthcke, S. B., McMillan, M., Meister, R., Milne, G., Mouginot, J., Muir, A., Nicolas, J. P., Paden, J., Payne, A. J., Pritchard, H., Rignot, E., Rott, H., Sørensen, L. S., Scambos, T. a, Scheuchl, B., Schrama, E. J. O., Smith, B., Sundal, A. V., van Angelen, J. H., van de Berg, W. J., van den Broeke, M. R., Vaughan, D. G., Velicogna, I., Wahr, J., Whitehouse, P. L., Wingham, D. J., Yi, D., Young, D. and Zwally, H. J.: A reconciled estimate of ice-sheet mass balance., *Science*, 338(6111), 1183–9, doi:10.1126/science.1228102, 2012.

Tourian, M. J., Elmi, O., Chen, Q., Devaraju, B., Roohi, S. and Sneeuw, N.: A spaceborne multisensor approach to monitor the desiccation of Lake Urmia in Iran, *Remote Sens. Environ.*, 156, 349–360, doi:10.1016/j.rse.2014.10.006, 2015.

USDA/FAS: Global Reservoirs and Lake Monitor, [online] Available from: http://www.pecad.fas.usda.gov/cropexplorer/global_reservoir/ (Accessed 25 November 2014), 2013.

Voss, K. a, Famiglietti, J. S., Lo, M., Linage, C., Rodell, M. and Swenson, S. C.: Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region., *Water Resour. Res.*, 49(2), 904–914, doi:10.1002/wrcr.20078, 2013.

Werth, S., Güntner, A., Petrovic, S. and Schmidt, R.: Integration of GRACE mass variations into a global hydrological model, *Earth Planet. Sci. Lett.*, 277(1-2), 166–173, doi:10.1016/j.epsl.2008.10.021, 2009.

Identifying water mass depletion in Northern Iraq observed by GRACE

G. Mulder¹, T.N. Olsthoorn¹, D.A.M.A. Al-Manmi², E.J.O. Schrama¹, and E.H. Smidt¹

¹Delft University of Technology, Delft, The Netherlands

²University of Sulaimani, Department of Geology, Sulaymaniyah, Iraq

Correspondence to: Gert Mulder (gert.mulder@gmail.com)

Abstract. Observations acquired by Gravity Recovery And Climate Experiment (GRACE) mission indicate a mass loss 146 ± 6 mm equivalent water height (EWH) in Northern Iraq between 2007 and 2009. This data is used as an independent validation of a hydrologic model of the region including lake mass variations. We developed a rainfall-runoff model for five tributaries of the Tigris River, based on local geology and climate conditions. Model inputs are precipitation from Tropical Rainfall Measurement Mission (TRMM) observations, and climatic parameters from Global Land Data Assimilation Systems (GLDAS) model parameters. Our model includes a representation of the karstified aquifers and is calibrated with observed river discharge. Lake mass variations were derived from Moderate Resolution Imaging Spectroradiometer (MODIS) in combination with satellite altimetry and some in-situ data. Our rainfall-runoff model confirms that Northern Iraq suffered a drought between 2007 and 2009 and shows the same patterns as the observed GRACE data. Also, GRACE observed the annual cycle predicted by the rainfall-runoff model. The total mass depletion seen by GRACE between 2007 and 2009 is mainly explained by a lake mass depletion of 75 ± 3 mm EWH and a natural groundwater depletion of 39 ± 8 mm EWH. Our findings indicate that anthropogenic groundwater extraction has a minor influence in this region, while a decline in lake mass and natural depletion of groundwater play a key role.

of water mass in the region (UN-ESCWA and BGR, 2013; Voss et al., 2013). Decrease in rainfall and water availability directly affected the water supply of towns and villages (Michel et al., 2012) and caused a strong decline in crop yields in Northern Iraq (Trigo et al., 2010). About 100,000 people have left their homes in Northern Iraq as a consequence of depleted water sources (McLeman, 2011). From 2009 onwards, rainfall rates have been rising, but are still lower than before the drought period. Moreover, a permanent decrease of rainfall rates in the region is predicted, based on climate models (Gibelin and Déqué, 2003; Giorgi and Lionello, 2008; Mariotti et al., 2008).

Concurrent with decreasing water availability in the region, water demands are fast increasing due to population growth and increase of irrigated agriculture (Altinbilek, 2004; Beaumont, 1998). Especially in Turkey, water demands increase rapidly due to the Southeastern Anatolia Project (GAP), which includes the construction of dams and irrigation schemes in the upstream Tigris catchment. At this moment, about 42,000 ha of this irrigation scheme is operational, with 53,400 ha under development and another 500,000 ha planned in future years (Altinbilek, 1997). Additionally, several dams and irrigations projects are under construction in the Iranian headwaters of the Tigris, which will reduce river flows in Northern Iraq permanently (Ali, 2007).

This means that especially Northern Iraq has to cope with a permanent decrease of its water resources due to lower rainfall and lower river flows from upstream countries. Because agreements on water between riparian countries are either hardly effective or non-existent, there are no guarantees that Iraq will ever receive as much water as before (UN-ESCWA and BGR, 2013; Al-Manmi, 2009).

Several hydrologic studies of the region exist (Chenoweth et al., 2011; Kavvas et al., 2011), but they are generally coarse due to the lack of ground truth and do not yield spe-

1 Introduction

During 2007 till 2009, Northern Iraq suffered a severe drought, with rainfall rates 40% below normal levels (Trigo et al., 2010; Fadhil, 2011). In the same period, discharge of large springs and rivers decreased substantially and data from the GRACE satellite mission indicated a permanent loss

cific information on hydrology and groundwater storages. 120
 Mass observations from GRACE provide a valuable tool to
 70 give more insight in the terrestrial water storages and are
 widely used as a validation of global hydrologic models like
 GLDAS, WaterGAP Global Hydrology Model (WGHM) and
 Global Geodynamics Project (GGP) (Voss et al., 2013; Werth
 et al., 2009; Awange et al., 2011; Hinderer et al., 2006; 125
 75 Schmidt et al., 2008; Ngo-Duc et al., 2007; Llubes et al.,
 2004). For the Euphrates and Tigris region, a comparison be-
 tween GRACE and the GLDAS model was made by Voss
 et al. (2013), which showed a large difference in both yearly
 and long-term mass variations. This could indicate that an- 130
 thropogenic groundwater extraction is a main cause of water
 depletion in the region, but the inability of the GLDAS model
 to capture yearly water mass variations makes the outcomes
 doubtful. Moreover, the GLDAS model does not contain a
 groundwater reservoir, which is important to track natural 135
 85 groundwater variations.

In this study, independent mass variations obtained from
 GRACE data are compared with water mass variation
 from lakes/reservoirs and a newly developed rainfall-runoff
 model, to show that these are likely the main causes of water 140
 mass depletion in the region. GRACE mass variation is de-
 rived using a mascon approach from Schrama et al. (2014)
 and the rainfall-runoff model is based on the general hydrology
 and geology of the region using the topo flex approach
 from Savenije (2010); Fenicia et al. (2011). Forcing data for 145
 95 this model is based on climatic parameters from the GLDAS
 model, daily rainfall from TRMM and local rainfall stations.
 Calibration of the model is done using local discharge mea-
 surements. MODIS surface reflectance data was used in com-
 bination with altimetry data from the Environmental Satellite 150
 100 (Envisat), Jason 1&2 and GEOSAT Follow-On (GFO) satel-
 lite missions (Crétaux et al., 2011) to find lake mass varia-
 tions. Local hydrologic and geologic data were obtained dur-
 ing fieldwork in cooperation with local water experts. In-situ
 data includes discharge data of one of the Tigris River tribu-
 105 taries (Directorate Dokan Dam, unpublished data) and rain-
 fall data from four stations in the region (Meteorological de-
 partment Kurdistan, unpublished data).

The following modelling steps will be used to com-
 110 pare GRACE data with water mass variations from
 lakes/reservoirs and the rainfall-runoff model: Firstly, the to-
 tal mass variation from an extended study area is derived
 from GRACE using a mascon approach. Secondly, the sur-
 face water mass is calculated for the same area and extracted 160
 from GRACE data, to obtain an approximate soil moisture
 and groundwater mass variation. Thirdly, the study area is
 115 reduced to Northern Iraq only, using soil moisture data from
 GLDAS. Finally, the natural variation in soil moisture and
 groundwater mass for Northern Iraq is calculated using a 165
 newly developed rainfall-runoff model.

2 Study Area

Most of Northern Iraq is part of the upstream catchment of
 the Tigris River, which originates in Turkey and flows south-
 wards to the Persian gulf. The total yearly flow of the Tigris
 at Baghdad is about $50 \text{ km}^3 \text{ yr}^{-1}$, of which half originates
 from upstream catchments in Turkey and half from tribu-
 taries in Northern Iraq (Brooks, 1997; Altinbilek, 2004). Fig-
 ure 1 gives an elevation map of the total study area, which in-
 cludes five main tributaries of the Tigris River, of which four
 have headwaters in Turkey or Iran. Figure 1 also presents a
 map of the mean yearly rainfall rates between 2002 and 2012,
 based on TRMM 3B42 (V7) data (Huffman et al., 2007).
 This map illustrates the large climatic variations in this area.
 While the south-western part of the catchment has a desert
 climate with rainfall rates of about 200 mm yr^{-1} , the north-
 eastern part consists of a mountain range with a considerably
 colder climate with rainfall rates up to 1000 mm yr^{-1} . The
 mountainous region in the north and northeast of the catch-
 ment is the main source of water of the Tigris River, while
 the arid areas in the southwest are totally dependent on up-
 stream river water (Beaumont, 1998; Brooks, 1997). Beside
 spatial differences in climates, there are also large seasonal
 and yearly variations in rainfall and temperatures. Almost all
 rain falls during the winter period, between November and
 March, and mean rainfall rates can drop by fifty percent in
 dry years.

To ensure water supply in the region during dry periods,
 many reservoirs exist and several are currently under con-
 struction. Northern Iraq includes the lakes of Mosul, Dukan,
 Derbendikhan, Adhaim and Hamrin, which all show a water
 mass decline during the 2007-2009 drought. But also water
 mass decline from lakes close to our study area, like Lake
 Tharthar, Habbaniyah, Qadisiyah, Razzazah, Urmia and Raz-
 zazah, are considered in this research (fig. 2).

3 Methods

3.1 GRACE mass variation calculations

During the last years, several methods have been developed
 to calculate mass change based on GRACE data like Swen-
 son and Wahr (2006); Schrama and Wouters (2011). Most
 methods comprise different processing steps, mainly to re-
 duce noise, apply geophysical corrections, add consistency
 and improve the ability of GRACE to see spatial details. This
 study uses a mascon approach, based on circular mascons
 with a radius of approximately one degree, which are evenly
 distributed over the Earth's surface (Schrama et al., 2014).
 This method is based of GRACE level-2 data from the Cen-
 ter of Space Research (CSR) and includes modifications for
 the gravitational flattening term C20 from satellite laser rang-
 ing. Furthermore the method also considers degree 1 terms
 associated with geo-center motions as a result of geophys-

ical loading phenomena. The used GRACE method is not a spatial averaging kernel method, instead to obtain equivalent water levels over a region one has to add up the signal from the individual mascons. The signal at these mascons is obtained via a global inversion method. The advantage of this approach is that resulting EWH values are part of a standard GRACE product and will be consistent with other regional solution worldwide. The disadvantage is that we have to adapt our study area to the given mascons. We are thus not able to use approaches like Krogh et al. (2010), which adapt the mascons to the geometry of the studied basin. However, a simple test where adjacent mascons were added to our study area, showed that small changes in study area do not have significant effects on the resulting GRACE values. Also, comparison between our method and other common methods by Shepherd et al. (2012) showed that differences are generally small.

To calculate water mass decline from GRACE data, an extended study area is used. Figure 2 shows the used area and mascon coverage for the GRACE mass calculations. This area is about $260 \cdot 10^3 \text{ km}^2$ and includes a large part of the catchment of Lake Urmia ($\pm 45 \cdot 10^3 \text{ km}^2$) and a part of the desert to the south-west of North Iraq ($\pm 95 \cdot 10^3 \text{ km}^2$). The desert area was added because we want to make sure that the mass decline from Lake Tharthar (fig. 2) is fully captured by our GRACE mass calculation and can be corrected for later on. This lake showed a decline in water mass of about 48 mm EWH for the extended study area during the 2007-2009 drought, which is more than 50 % of the total lake mass decline. Other lakes like Lake Razazzah and Lake Urmia are still close to the border, but these lakes have shown much less water mass depletion between 2007-2009 (tab. 1). This induces mass leakage of these lakes outside our study area (Longuevergne et al., 2013), which is corrected as explained in section 3.2.

To compare the total water mass variation of the rainfall-runoff of Northern Iraq with GRACE, the derived GRACE mass was corrected for soil moisture and groundwater mass variation of the extended area to the southwestern desert. Groundwater pumping in these areas is generally small due to high salt content or deep groundwater tables (Krásný et al., 2006) and is estimated around $30\text{-}35 \cdot 10^{-3} \text{ km}^3 \text{ yr}^{-1}$ (UN-ESCWA and BGR, 2013) Because recharge and discharge rates of the aquifers are also very low in this area, we do not expect significant groundwater variations in this region. Therefore, we assumed that the soil moisture profile from GLDAS Rodell et al. (2004) can be used as a measure for the total water mass variation in the area. The new GRACE values for Northern Iraq then becomes:

$$H = \frac{H_t A_t - G_d A_d}{A_t - A_d} \quad (1)$$

Where H is the equivalent water height, A is the total area and G is the soil moisture variation in water depth from the

GLDAS model. Subscripts t and d indicate total and desert areas.

This correction assumes a pristine area, which is not true for the irrigated areas around Baghdad in Central Iraq. This likely causes an underestimation of the yearly water mass variation due to the growing season of the crops in this region. However, water mass decline due to anthropogenic groundwater extraction is small because of the high salt content of groundwater in this region Krásný et al. (2006).

Because the influence of groundwater variation is large in the catchment of Lake Urmia (fig. 2), the correction used for the Desert area cannot be applied here Tourian et al. (2015). Instead, the hydrology of the region is comparable with Northern Iraq. Therefore, the water mass variation in mm EWH from the catchment of Lake Urmia are neglected and assumed to be the same as in Northern Iraq. This introduces an error in the model on a small time scale due to local rainfall events, but on a longer timescale these errors will be minor due the similarities between the regions.

3.2 Derivation of lake mass

The total lake mass variations play an important role in the water balance of the Tigris region (Voss et al., 2013; Longuevergne et al., 2013). Figure 2 gives an oversight of the important lakes in the region. To obtain total lake mass contributions in our study area, time series of both lake level and lake area were calculated. The lake levels are derived from satellite altimetry by Crétaux et al. (2011) and USDA/FAS (2013), which includes data from the Envisat, Jason 1&2 and GFO satellite missions. Area calculations were based on MODIS satellite data to detect water areas, combined with digital elevation maps (DEM) from the Shuttle Radar Topography Mission (SRTM) to distinguish different water bodies. The actual method uses a $250 \times 250 \text{ m}$ grid and comprises of three steps: First, the possible extend of the lake was calculated from the DEM using a minimum and maximum elevation. Secondly, the larger lake areas were selected using MODIS reflectance band 5 (1240 nm, 500 m resolution). Finally, the exact lake borders were defined using MODIS reflectance band 2 (858 nm, 250 m resolution).

From the lake level and lake-area time series stage-area curves were created using linear or cubic regression. In fig. 3, a comparison is given of the derived stage-area curve from our model and a survey of Lake Mosul using sonar by Issa et al. (2013). To decide whether a cubic regression gives a significant improvement, the F-test was used for a 95 % interval. Total volume change over time was then derived from the stage-volume curve, which is the integrated stage-area curve. Table 1 gives the derived water mass decline between 2007 and 2009 of the main lakes in the region. This water mass decline is derived by subtracting the average water mass in 2009 from the average water mass in 2006.

Uncertainties in water mass calculations from lakes are derived from estimated errors in water level measurements

275 given by Crétaux et al. (2011) and USDA/FAS (2013). The 325
total error is given by:

$$\sigma^t = \sqrt{\frac{\sum_{i=1}^n (\sigma_i^t A_i^t)^2}{n}} \quad (2)$$

Where σ_i^t is the estimated standard deviation of the lake wa- 330
ter level from lake i at time t , n is the number of lakes and
280 A_i^t is the lake area at the same moment.

Another source of uncertainty are the errors in lake area 335
calculations, but these are generally small and are difficult to
quantify. For example, the difference in lake size from lake
285 Mosul and the derived values from Issa et al. (2013) (fig. 3)
are about 10 km² but it only causes significant differences in
the stage-volume curve for high water levels. Additionally,
the relative error for larger lakes is much smaller because
the ratio between lake shore length and lake area becomes
smaller with increasing lake size. 340

290 To give an approximation of mass leakage from lakes in
our study area results from Longuevergne et al. (2013) are
used. This led to a correction for Lake Razzazah of 0.65 and
a correction for Lake Urmia of 0.6, because these lakes are
located close to the border and far from the centre of the 345
study area. Corrections for other lakes were not significant
and were therefore omitted.

3.3 Rainfall-Runoff model

In this study, a rainfall-runoff model was used based on the 350
topo flex approach as proposed by Savenije (2010); Feni-
cia et al. (2011) and a simplified snow routine based on
300 Lindström et al. (1997). This resulted in a semi-distributed
model structure based on the geology and topography of 355
the five main tributaries of the Tigris River in Northern Iraq
(fig. 5). Forcing parameters of the rainfall-runoff model are
calibrated daily precipitation data from TRMM (Huffman
305 et al., 2007), daily temperature values from GLDAS and
daily reference evaporation derived from GLDAS climatic 360
parameters (Rodell et al., 2004; Allen et al., 1998). TRMM
data was bias-corrected by linear regression with monthly
precipitation from four gauging stations in Sulaymaniyah,
310 Dokan, Derbendikhan and Penjwen (Meteorological depart-
ment Kurdistan, unpublished data). 365

The rainfall-runoff model is identical for all five tribu-
taries, and is based on three geologic zones in accordance
315 with geologic maps of Stevanovic and Iurkiewicz (2008).
The three zones are:

1. Infiltrative or karstified zone: About one third of the 370
surface area of the mountainous zone consists of kars-
tified limestone and is, therefore, highly infiltrative.
320 These limestones have infiltration rates of more than
50 percent and transmissivities ranging from 9 to 8000
 m^2/day (Krásný et al., 2006). 375
2. Non-infiltrative zone: This zone consists of the other
mountainous areas, which are characterised by fast

runoff due to shallow soil layers, steep slopes and im-
permeable underlying formations.

3. Alluvial zone: Most of the soils in the dry south west-
ern part consist of clay and silt sediments. In this re-
gion flash floods are common during the scarce rainfall
events.

For all five tributaries of the Tigris River in Northern Iraq
and the remaining area close to the Tigris River the same
model setup is used, with the total areas of the three geologic
zones as the only difference. For example the flow regime
of the Adhaim River is governed by the alluvial part of the
model, while the Greater Zab is mainly fed by water from the
infiltrative and non-infiltrative zones.

Figure 4 gives a detailed oversight of the model reservoirs
and parameters. The snow routine and unsaturated reservoirs
are modelled in a distributed way, using the 0.25 degree grid
from TRMM as a basis. When grid cells are part of two or
more different basins or geologic zones the cell is split up
in different parts, which contribute to their respective zones
or catchments. Parameter values of different grid cells were
kept the same within the geologic zones, while input pre-
cipitation from TRMM and climatic values from GLDAS
were separately assigned to each cell. From the unsaturated
zone, water is either routed via the fast runoff or groundwater
reservoir to the river. A lag function was added to the model
to simulate the routing of water through streams and rivers
to the catchment outflow. The fast runoff and deep ground-
water reservoirs of the infiltrative and non-infiltrative zones
were combined, because the topography and top soils are
comparable and share the same underlying aquifers (Krásný
et al., 2006). A total number of eighteen parameters was
used in the model, which were restricted to minimum
and maximum bounds during calibration to prevent equi-
finality and ensure realism of the model. Especially the pa-
rameters for the karstified/infiltrative groundwater reservoir
have a strong influence on the modelled water mass varia-
tions. The parameter value for this reservoir is based on the
recession curves of spring discharges, which emerge from the
karstified aquifer in Northern Iraq. Generally, the discharge
from the karstic aquifers in Northern Iraq can be split up in
two components (Stevanovic and Iurkiewicz, 2008; Ali and
Stevanovic, 2010). The first component is a rapid discharge
within a month after major rain events, which is related to
channels and large fractures in the limestone aquifer. The
second component is a much slower discharge with a stable
recession coefficient during the whole dry season. This
coefficient is about 0.004 per day on average, based on dis-
charge time series of several large springs in the region (Ali
and Stevanovic, 2010; Ali et al., 2009b, a). The first compo-
nent is covered by the fast reservoir in our hydrologic model,
while the second component is modelled as a slow ground-
water reservoir. The bounds of the storage and infiltration
rates in the unsaturated reservoirs were based on fieldwork
and personal communication with local hydrologists.

3.4 Model Calibration

In literature generally two methods are used to integrate GRACE mass into the calibration and validation process. Either GRACE information is used to inform the model as a calibration parameter like in Werth et al. (2009), or to validate of the model like in Syed et al. (2008). In our case it would be best to use the GRACE data to inform the model, because we want to show that a hydrologic model is able to mimic the mass depletion observed by GRACE. However, during the lifespan of the GRACE satellite only one drought like in 2007-2009 did occur in our region, which makes it impossible to use a separate calibration and validation period. Therefore, we have chosen to use the GRACE data only as a validation of the model, but include also the results for the case that GRACE was used to inform the model. This approach shows that the resulting water depletion will increase due to the inclusion of GRACE data, although it does only have relatively small impacts on model behaviour.

The primary source for model calibration is river discharge of the Lesser Zab (Directorate Dukan Dam, unpublished data) at the inflow of Lake Dukan (see fig. 5). The most convenient data to calibrate the model on would be the discharge from the total study area at Baghdad, but this data was not available for us. Additionally, the discharge at those points is not suitable for rainfall-runoff modelling, because it is strongly influenced by the operation of upstream dams and reservoirs. Therefore, the performance of the model was first evaluated for the upstream area of Lake Dukan and expanded to the whole of Northern Iraq, to allow comparison with GRACE. This expansion is possible due to the geologic similarities in the region, which are explained in section 3.3.

Model calibration was done using a Monte Carlo simulation with randomly chosen parameter values for every model run, within given parameter bounds. Total discharge for the Dukan area was evaluated by the Nash-Sutcliffe efficiency (NS) for medium and high flows and by the log Nash-Sutcliffe efficiency (logNS) for low flows:

$$NS_Q = 1 - \frac{\sum_{t=1}^T (Q_o^t - Q_m^t)^2}{\sum_{t=1}^T (Q_o^t - \overline{Q_o})^2} \quad (3)$$

$$\log NS_Q = 1 - \frac{\sum_{t=1}^T (\log(Q_o^t) - \log(Q_m^t))^2}{\sum_{t=1}^T (\log(Q_o^t) - \log(\overline{Q_o}))^2} \quad (4)$$

Where Q_o represents the observed daily discharges, $\overline{Q_o}$ is the average observed daily discharge and Q_m the modelled daily discharge. In the case that GRACE mass was also used as a calibration parameter, model mass performance

was evaluated using the Nash Sutcliffe efficiency for water mass variation:

$$NS_M = 1 - \frac{\sum_{t=1}^T (M_o^t - M_m^t)^2}{\sum_{t=1}^T (M_o^t - \overline{M_o})^2} \quad (5)$$

Where M_o represents the observed mass variations from GRACE, $\overline{M_o}$ is the average observed GRACE mass and M_m the modelled mass variation from lakes and the rainfall-runoff model. Figure 9 gives the pareto front for the average NS_Q and $\log NS_Q$ of the river discharge. In the same figure the NS_Q and $\log NS_Q$ values are given of the additional models from the pareto front if NS_M would be added as a third objective parameter. The presented optimal solution was chosen based on the following condition:

$$\max(NS_Q + \log NS_Q) \quad (6)$$

In the result and discussion section, the model ensemble of the models from the pareto front are given as an uncertainty band, beside the optimal solution.

4 Results and discussion

4.1 GRACE

Figure 6 shows the resulting GRACE values in terms of equivalent water height, with an estimated 95 % confidence band of 20 mm (Schrama et al., 2007). Water mass depletion between 2007 and 2009 is 146 ± 6 mm EWH. These values are based on average GRACE values before and after the drought, which show more or less constant values. Yearly variation is 286 ± 24 mm based on the yearly minimum and maximum values between 2003 and 2011. In the lower graph of fig. 6, the average monthly rainfall is given for the same period to show the relation between GRACE values and rainfall. The periods with more than average rainfall generally coincide with increasing GRACE values, due to of accumulation of rainwater in the catchment areas. The drought period between 2007 and 2009 coincides with an overall decrease in water mass, while the water mass is more or less stable during the periods before and after the drought. The magnitude of the yearly variations in water mass follows the yearly rainfall trend and is therefore largest before 2007 and smallest during 2007-2009. The GRACE data after 2009 is more or less stable, with almost no difference in average water mass. This could indicate that a new equilibrium state is reached after the drought, where the outflow of the system is in balance with lower precipitation values.

The given GRACE values could also be influenced by leakage of mass over the borders of our study area, due to

465 large soil water or groundwater mass variation just inside or
outside our study area. But because the GRACE signal for a
larger region, as given by Voss et al. (2013) are comparable
to the signal we found, the effect of this mass leakage will be
limited.

520

470 4.2 Lakes and Reservoirs

To compare the total lake mass change with GRACE, all lake
mass variations were added up and divided by the total area
of the region. Figure 7 shows the surface water mass variation
in terms of EWH, with a 95 % confidence interval. The total
475 lake mass accounts for 75 ± 3 mm of the water mass deple-
tion between 2007 and 2009 and 55 ± 6 mm of the yearly wa-
ter mass variation during 2003-2011. This means that more
than 50 % of the total water mass depletion is caused by a
decline of surface water mass. Such declines in surface wa-
480 ter mass were already claimed for the whole Euphrates and
Tigris basin by Longuevergne et al. (2013), but their values
are higher than those given by Voss et al. (2013). Different
results between our and other researches are likely caused by
differences in the size of the studied basins, but exclusion of
485 mostly smaller lakes can also play an important role. For ex-
ample, lakes Hamrin, Adhaim, Dukan, Habbaniyah and Mo-
sul, had a total volume decline of about 13 mm EWH during
the study period, but are not considered by Longuevergne
et al. (2013) and Voss et al. (2013). Beside exclusion of sev-
490 eral lakes, also mass leakage from lakes in the study area can
lead to different results.

Figure 7 also shows that large lake mass variations are not
uncommon in this region, which is due to large differences in
rainfall from year to year and recurring droughts Trigo et al.
495 (2010). However, increasing water use in upstream countries
like Turkey and Iran will hinder replenishment of lakes and
reservoirs in Northern Iraq Ali (2007); Beaumont (1998).

4.3 Modelled discharge Dukan catchment

550

Figure 8 shows the modelled and measured discharge at the
inflow of Lake Dukan from 2005 till 2007. Generally, the low
flows are captured quite well with discharge values within
the given bands, while the modelled high flows deviate much
555 more from the measured ones. This is likely related with fast
and slow runoff mechanisms in the model, represented by the
fast runoff and groundwater reservoirs respectively. The fast
runoff reservoir is mainly fed by heavy or local rain events
on short time scales, which are not well represented by the
560 TRMM data. The slow runoff or groundwater reservoir has a
much longer timescale and relies more on seasonal rainfall,
which is captured much better by TRMM Almazroui (2011).

The modelled results are given as a pareto ensemble based
on the pareto optimal solution for the NS and logNS perfor-
565 mance indicators (fig. 9). This ensemble is mainly a measure
for the uncertainty of the model due to the model parameters.
515 Uncertainty of the model due to forcings like rainfall or po-

tential evaporation were not evaluated. Further, we assume a
pristine catchment, but in reality the flows will be somewhat
higher because of water use upstream. This mainly affects the
baseflow and will result in higher groundwater variations. It
is therefore more likely that we underestimate the water mass
variation in the Dukan catchment.

4.4 Natural groundwater variations

The natural groundwater mass variations derived from the
rainfall-runoff model is given in fig. 10. This figure shows
the optimal solution based on formula 6 together with the
pareto ensemble for the NS_Q and $\log NS_Q$ performances
and the pareto ensemble based on the NS_Q , $\log NS_Q$ and
 NS_M performances. Note that the pareto ensemble where
GRACE mass is included also covers the other pareto ensem-
ble and includes both the red and the blue bands. Modelled
groundwater variations from the NS_Q and $\log NS_Q$ ensem-
ble contribute 42 ± 5 mm EWH to the yearly water mass
variations and 26 ± 8 mm EWH to the water mass depletion
between 2007 and 2009. Modelled groundwater variations
from the NS_Q , $\log NS_Q$ and NS_M ensemble contribute
49 ± 7 mm EWH to the yearly water mass variations and
 34 ± 14 mm EWH to the water mass depletion between 2007
and 2009. This shows that modelled water mass depletion
does increase when GRACE data is used as a calibration pa-
rameter, but model results are consistent. The contribution
of the natural groundwater mass explains why those values
where not reproduced by the GLDAS model, which does not
include groundwater storage. The time series for the ground-
water reservoir as given in fig. 10 comprises a period with
slowly decreasing groundwater levels till 2007, followed by
a period with a strong groundwater depletion from 2007 till
2009, and a period with slowly increasing groundwater lev-
els from 2010 till 2012. Groundwater from the karstified
aquifers plays an important role in the groundwater deple-
tion between 2007 and 2009, because of its high recharge
during wet periods and fast discharge through springs during
dry periods. While the aquifers still discharge water through
springs during dry years, there is much less replenishment
of the groundwater and groundwater levels will drop. In the
governorates of Sulaymaniyah and Duhok alone, about 1.5
 km^3/year water emerges every year from springs (Stevanovic
and Markovic, 2004; UN-ESCWA and BGR, 2013). After
2009, modelled groundwater levels remained almost stable
because rainfall rates were still below average, but a se-
quence of years with higher rainfall will result in a rise of
groundwater levels and a revival of spring discharge.

To convert the groundwater mass variation from the Dukan
area only to the whole of Northern Iraq, the model was
extended using the similarities between geological regions.
These assumption is based on the identical geological back-
ground of these regions Stevanovic et al. (2009). We do not
know exactly whether the related aquifers developed in the
same way, but because of the occurrence of springs in the

region with similar characteristics we think this assumption
can be made.

4.5 Mass variations model reservoirs

Beside the groundwater reservoir the model consist of three additional reservoirs to model the snow layer, unsaturated zone and fast runoff (fig. 11). These reservoirs mainly contribute to the yearly mass variation of 122 ± 7 mm EWH and have only a small contribution to the water mass decline with 10 ± 1 mm EWH. The unsaturated reservoirs show a recurring pattern every year where the soil layers get saturated during the wet season and dry out during the dry periods. The yearly variation of these reservoirs is 65 ± 8 mm EWH. The fast runoff reservoir, which represents overland flow and interflow in the basin, shows peaks up to about 60 mm EWH during and shortly after intense rainfall events. The average yearly variation of the fast reservoir is 38 ± 7 mm EWH and the decline after the drought is negligible. The snow reservoirs show the largest differences between dry and wet years, because it accumulates precipitation for a whole rainy season. Still, almost all snow melts away during summer due to the strong temperature differences between seasons. The average yearly variation of the snow mass is 25 ± 4 mm EWH and the average decline 6 ± 1 mm EWH.

4.6 GRACE and modelled values

Figure 12 compares the total water mass variation from GRACE and the lakes plus the rainfall-runoff model. GRACE values indicate a mass depletion 146 ± 6 mm EWH between 2007 and 2009 and a yearly mass variation of 286 ± 24 mm EWH. The combined water mass variation of lakes and the rainfall-runoff model result in a water mass decline of 114 ± 9 mm EWH and an average yearly variation of 225 ± 9 mm EWH. The two graphs differ mainly in the winter and summer peaks, but also a part of the water mass decline remains unexplained.

Possible causes for the differences in yearly water mass variations are an underestimation of accumulated snow water or random errors in rainfall rates from TRMM data, which are 23 % on average. Also the impact of anthropogenic activities can cause higher peaks. This can be due to surface water irrigation and groundwater use, which is replenished during the wet season. Other possible causes are additional water storage in depressions or water mass variability of smaller lakes. In theory, we could have changed the maximum storage of the unsaturated zone to fit the given curves better, but this would create a large difference between the modelled and the literature values.

Possible causes for the differences in mass decline are anthropogenic groundwater extraction or lakes, which were not included in the model. But also model uncertainties for due to mass leakage or the groundwater reservoirs can be a main cause.

5 Conclusions

5.1 Water masses in Northern Iraq

The presented approach offers the possibility to quantify different hydrologic processes in the region as well as the shares of surface water, soil moisture and groundwater in the total water mass variation. More importantly, the overall model shows that natural variation of groundwater, snow depth and soil moisture have a share of about 25 % of the total water mass decline. With a depletion of 39 ± 8 mm EWH it also explains more than half of the remaining mass decline if lake masses are extracted from GRACE results. This shows that natural groundwater variation has to be taken into account when GRACE mass values are used to determine overdraft of aquifers.

Especially, in the limestone aquifers of Northern Iraq, strong groundwater variations are common due to extensive karst networks with high transmissivities and infiltration rates, feeding numerous springs in the region. Therefore, overpumping of these aquifers is unlikely, as the groundwater table can vary strongly and the regional water supply is mainly supported by surface water. Additionally, almost all irrigation schemes in those areas are directly linked to large reservoirs.

The dependency of this region on surface water is also reflected by the large water mass variations of the surface water, which contributed about 75 mm out of 146 mm EWH observed by GRACE. With decreasing water availability and increasing water demands from riparian countries in the Tigris River catchment, the need for reliable water management tools is growing, including transboundary models. The developed model helps to give insights in the available water resources and water flows between concerned countries and can be used as a base for water allocation and water agreements. Contrary to other studies like (Chenoweth et al., 2011; Kavvas et al., 2011; Voss et al., 2013), main aquifers and water storages were modelled separately. Results are, therefore, more useful to water managers. Moreover, the model is based and calibrated on both satellite and in-situ data, which enhances its reliability and predictive power.

5.2 Model structure and input data

This research showed that GRACE can be an important data source in rainfall-runoff models because it gives direct measurements of the total water balance of a larger region. Especially in Northern Iraq, where water resources and data on water resources are scarce, this is valuable information. However, it is not possible to determine what causes the water variations in these regions without additional data on precipitation, geology and river discharges. This data can partly be obtained from satellites, but the use of in-situ data is still of vital importance. In our situation, there was only little data available, but it could be used for both model structure and

calibration. Additionally, knowledge from local water experts and field observations gave important information on governing hydrologic processes.

In our case, the rainfall and discharge stations covered only a part of the region, which resulted in increased model uncertainties. For example, the uncertainties in the total groundwater values are mainly caused by the alluvial groundwater reservoirs, which have only a small contribution to the total flow at lake Dukan. Inclusion of discharge data series from other tributaries would reduce these uncertainties and give a better insight in the spatial variability of the region at the same time.

References

- Al-Manmi, D. A. M. A.: Sirwan Transboundary River basin management, Tech. rep., Stockholm, 2009.
- Ali, M. H.: Transboundary waterways and streams along the Iraq-Iran border lines. . . .the reality and future, Tech. rep., Baghdad University, Baghdad, 2007.
- Ali, S. S. and Stevanovic, Z.: Time Series Analysis of Saraw Springs - SE of Sulaimaniya, Iraqi Kurdistan Region, in: *Advances in Research in Karst Media*, edited by Andreo, B., Carrasco, F., Durán, J. J., and LaMoreaux, J. W., pp. 89–94, Springer, Leipzig, doi:10.1007/978-3-642-12486-0, 2010.
- Ali, S. S., Stevanovic, Z., and Al-Jabbari, M. H.: Auto and Cross Correlation Analysis of Bestansur spring- Sharazoor Plain Iraqi Kurdistan Region, in: *Proceeding of 3rd scientific conference of the College of Science, University of Baghdad*, pp. 1941–1948, 2009a.
- Ali, S. S., Stevanovic, Z., and Jemcov, I.: The Mechanism and Influence on Karstic Spring Flow - Sachinar Spring Case Example, Sulaimaniya, NE Iraq, *Iraqi Bulletin of Geology and Mining*, 5, 87–100, 2009b.
- Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: *FAO Irrigation and Drainage Paper No. 56*, Tech. Rep. 56, FAO, Rome, Italy, <http://www.fao.org/docrep/x0490e/x0490e00.htm>, 1998.
- Almazroui, M.: Calibration of TRMM rainfall climatology over Saudi Arabia during 1998-2009, *Atmospheric Research*, 99, 400–414, doi:10.1016/j.atmosres.2010.11.006, 2011.
- Altinbilek, D.: Development and management of the Euphrates-Tigris basin, *International Journal of Water Resources Development*, 20, 15–33, doi:10.1080/07900620310001635584, 2004.
- Altinbilek, H. D.: Water and Land Resources Development in Southeastern Turkey, *International Journal of Water Resources Development*, 13, 311–332, doi:10.1080/07900629749719, 1997.
- Awange, J. L., Fleming, K. M., Kuhn, M., Featherstone, W. E., Heck, B., and Anjasmara, I.: On the suitability of the 4x4 degrees GRACE mascon solutions for remote sensing Australian hydrology, *Remote Sensing of Environment*, 115, 864–875, doi:10.1016/j.rse.2010.11.014, 2011.
- Beaumont, P.: Restructuring of Water Usage in the Tigris-Euphrates Basin : The Impact of Modern Water Management Policies, *Middle Eastern Natural Environments*, 103, 168–186, 1998.
- Brooks, D. B.: Between the Great Rivers : Water in the Heart of the Middle East, *International Journal of Water Resources Development*, 13, 291–310, doi:10.1080/07900629749700, 1997.
- Chenoweth, J., Hadjinicolaou, P., Bruggeman, A., Lelieveld, J., Levin, Z., Lange, M. a., Xoplaki, E., and Hadjikakou, M.: Impact of climate change on the water resources of the eastern Mediterranean and Middle East region: Modeled 21st century changes and implications, *Water Resources Research*, 47, W06 506, doi:10.1029/2010WR010269, 2011.
- Crétaux, J.-F., Jelinski, W., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., Gennero, M.-C., Nino, F., Abarca Del Rio, R., Cazenave, A., and Maisongrande, P.: SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data, *Advances in Space Research*, 47, 1497–1507, doi:10.1016/j.asr.2011.01.004, <http://www.legos.obs-mip.fr/en/soa/hydrologie/hydroweb/>, 2011.
- Fadhil, A. M.: Drought mapping using Geoinformation technology for some sites in the Iraqi Kurdistan region, *International Journal of Digital Earth*, 4, 239–257, doi:10.1080/17538947.2010.489971, 2011.
- Fenicia, F., Kavetski, D., and Savenije, H. H. G.: Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development, *Water Resources Research*, 47, W11 510, doi:10.1029/2010WR010174, 2011.
- Gibelin, A. and Déqué, M.: Anthropogenic climate change over the Mediterranean region simulated by a global variable resolution model, *Climate Dynamics*, 20, 327–339, doi:10.1007/s00382-002-0277-1, 2003.
- Giorgi, F. and Lionello, P.: Climate change projections for the Mediterranean region, *Global and Planetary Change*, 63, 90–104, doi:10.1016/j.gloplacha.2007.09.005, <http://linkinghub.elsevier.com/retrieve/pii/S0921818107001750>, 2008.
- Hinderer, J., Andersen, O., Lemoine, F., Crossley, D., and Boy, J.-P.: Seasonal changes in the European gravity field from GRACE: A comparison with superconducting gravimeters and hydrology model predictions, *Journal of Geodynamics*, 41, 59–68, doi:10.1016/j.jog.2005.08.037, 2006.
- Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, *Journal of Hydrometeorology*, 8, 38–55, doi:10.1175/JHM560.1, 2007.
- Issa, I. E., Al-Ansari, N., and Knutsson, S.: Sedimentation and new operational curves for Mosul Dam, Iraq, *Hydrological Sciences Journal*, 58, 1456–1466, doi:10.1080/02626667.2013.789138, 2013.
- Kavvas, M., Chen, Z., Anderson, M., Ohara, N., Yoon, J., and Xiang, F.: A study of water balances over the Tigris-Euphrates watershed, *Physics and Chemistry of the Earth, Parts A/B/C*, 36, 197–203, doi:10.1016/j.pce.2010.02.005, 2011.
- Krásný, J., Alsam, S., and Jassim, S. Z.: *Hydrogeology*, in: *Geology of Iraq*, edited by Jassim, S. Z. and Goff, J. C., pp. 251–287, Publishers Dolin, Prague, first edn., 2006.
- Krogh, P. E., Andersen, O. B., Michailovsky, C. I. B., Bauer-Gottwein, P., Rowlands, D. D., Luthcke, S. B., and Chinn, D. S.: Evaluating terrestrial water storage variations from regionally constrained GRACE mascon data and hydrological models over Southern Africa – preliminary results, *International Journal of Remote Sensing*, 31, 3899–3912, doi:10.1080/01431161.2010.483483, <http://www.tandfonline.com/doi/abs/10.1080/01431161.2010.483483>, 2010.

- Lindström, G., Johansson, B., and Persson, M.: Development and test of the distributed HBV-96 hydrological model, *Journal of Hydrology*, 201, 272–288, <http://www.sciencedirect.com/science/article/pii/S0022169497000413>, 1997. ⁸⁴⁵
- Llubes, M., Florsch, N., Hinderer, J., Longuevergne, L., and Amalvict, M.: Local hydrology, the Global Geodynamics Project and CHAMP/GRACE perspective: some case studies, *Journal of Geodynamics*, 38, 355–374, doi:10.1016/j.jog.2004.07.015, 2004. ⁷⁹⁰
- Longuevergne, L., Wilson, C. R., Scanlon, B. R., and Crétaux, J. F.: GRACE water storage estimates for the Middle East and other regions with significant reservoir and lake storage, *Hydrology and Earth System Sciences*, 17, 4817–4830, doi:10.5194/hess-17-4817-2013, 2013. ⁷⁹⁵
- Mariotti, A., Zeng, N., Yoon, J.-H., Artale, V., Navarra, A., Alpert, P., and Li, L. Z. X.: Mediterranean water cycle changes: transition to drier 21st century conditions in observations and CMIP3 simulations, *Environmental Research Letters*, 3, 044 001, doi:10.1088/1748-9326/3/4/044001, 2008. ⁸⁰⁰
- McLeman, R. A.: Settlement abandonment in the context of global environmental change, *Global Environmental Change*, 21, S108–S120, doi:10.1016/j.gloenvcha.2011.08.004, 2011. ⁸⁰⁵
- Michel, D., Pandya, A., Hasnain, S. I., Sticklor, R., and Panuganti, S.: Water Challenges and Cooperative Response in the Middle East and North Africa, in: U.S.-Islamic World Forum, pp. 11–15, U.S. World Forum Papers, <http://www.brookings.edu/~media/Research/Files/Papers/2012/11/iwfpapers/Waterweb.pdf>, 2012. ⁸¹⁰
- Ngo-Duc, T., Laval, K., Ramillien, G., Polcher, J., and Cazenave, A.: 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, W04 427, doi:10.1029/2006WR004941, 2007. ⁸¹⁵
- Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin*, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, *Bulletin of the American Meteorological Society*, 85, 381–394, doi:10.1175/BAMS-85-3-381, <http://ldas.gsfc.nasa.gov/gldas/>, 2004. ⁸²⁰
- Savenije, H. H. G.: HESS opinions "Topography driven conceptual modelling (FLEX-Topo)", *Hydrology and Earth System Sciences*, 14, 2681–2692, doi:10.5194/hess-14-2681-2010, 2010. ⁸²⁵
- Schmidt, R., Petrovic, S., Güntner, A., Barthelmes, F., Wunsch, J., and Kusche, J.: Periodic components of water storage changes from GRACE and global hydrology models, *Journal of Geophysical Research*, 113, B08 419, doi:10.1029/2007JB005363, 2008. ⁸³⁰
- Schrama, E. J., Wouters, B., and Rietbroek, R.: A mascon approach to assess ice sheet and glacier mass balances and their uncertainties from GRACE data, *Journal of Geophysical Research: Solid Earth*, 119, 6048–6066, doi:10.1002/2013JB010923, 2014. ⁸³⁵
- Schrama, E. J. O. and Wouters, B.: Revisiting Greenland ice sheet mass loss observed by GRACE, *Journal of Geophysical Research*, 116, B02 407, doi:10.1029/2009JB006847, 2011. ⁸⁴⁰
- Schrama, E. J. O., Wouters, B., and Lavallée, D. A.: Signal and noise in Gravity Recovery and Climate Experiment (GRACE) observed surface mass variations, *Journal of Geophysical Research*, 112, B08 407, doi:10.1029/2006JB004882, 2007. ⁹⁰⁰
- Shepherd, A., Ivins, E. R., A. G., Barletta, V. R., Bentley, M. J., Bettadpur, S., Briggs, K. H., Bromwich, D. H., Forsberg, R., Galin, N., Horwath, M., Jacobs, S., Joughin, I., King, M. a., Lenaerts, J. T. M., Li, J., Ligtenberg, S. R. M., Luckman, A., Luthcke, S. B., McMillan, M., Meister, R., Milne, G., Mouginot, J., Muir, A., Nicolas, J. P., Paden, J., Payne, A. J., Pritchard, H., Rignot, E., Rott, H., Sørensen, L. S., Scambos, T. a., Scheuchl, B., Schrama, E. J. O., Smith, B., Sundal, A. V., van Angelen, J. H., van de Berg, W. J., van den Broeke, M. R., Vaughan, D. G., Velicogna, I., Wahr, J., Whitehouse, P. L., Wingham, D. J., Yi, D., Young, D., and Zwally, H. J.: A reconciled estimate of ice-sheet mass balance., *Science (New York, N.Y.)*, 338, 1183–9, doi:10.1126/science.1228102, <http://www.ncbi.nlm.nih.gov/pubmed/23197528>, 2012.
- Stevanovic, Z. and Iurkiewicz, A.: Groundwater management in northern Iraq, *Hydrogeology Journal*, 17, 367–378, doi:10.1007/s10040-008-0331-0, 2008.
- Stevanovic, Z. and Markovic, M.: Hydrogeology of Northern Iraq, Food and Agriculture Organisation of the United Nations, Rome, 2004.
- Stevanovic, Z., Iurkiewicz, A., and Maran, A.: New Insights Into Karst and Caves of Northwestern Zagros Northern Iraq, *Acta carologica*, 38, 83–96, doi:10.1007/978-3-642-12486-0, 2009.
- Swenson, S. and Wahr, J.: Post-processing removal of correlated errors in GRACE data, *Geophysical Research Letters*, 33, L08 402, doi:10.1029/2005GL025285, <http://doi.wiley.com/10.1029/2005GL025285>, 2006.
- Syed, T. H., Famiglietti, J. S., Rodell, M., Chen, J., and Wilson, C. R.: Analysis of terrestrial water storage changes from GRACE and GLDAS, *Water Resources Research*, 44, n/a–n/a, doi:10.1029/2006WR005779, <http://doi.wiley.com/10.1029/2006WR005779>, 2008.
- Tourian, M., Elmi, O., Chen, Q., Devaraju, B., Roohi, S., and Sneeuw, N.: A spaceborne multisensor approach to monitor the desiccation of Lake Urmia in Iran, *Remote Sensing of Environment*, 156, 349–360, doi:10.1016/j.rse.2014.10.006, <http://linkinghub.elsevier.com/retrieve/pii/S0034425714004027>, 2015.
- Trigo, R. M., Gouveia, C. M., and Barriopedro, D.: The intense 2007–2009 drought in the Fertile Crescent: Impacts and associated atmospheric circulation, *Agricultural and Forest Meteorology*, 150, 1245–1257, doi:10.1016/j.agrformet.2010.05.006, 2010.
- UN-ESCWA and BGR: Inventory of Shared Water Resources in Western Asia, Tech. rep., United Nations Economic and Social Commission for Western Asia; Bundesanstalt für Geowissenschaften und Rohstoffe, Beirut, <http://waterinventory.org/>, 2013.
- USDA/FAS: Global Reservoirs and Lake Monitor, http://www.pecad.fas.usda.gov/cropexplorer/global_reservoir/, 2013.
- Voss, K. a., Famiglietti, J. S., Lo, M., Linage, C., Rodell, M., and Swenson, S. C.: Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region., *Water resources research*, 49, 904–914, doi:10.1002/wrcr.20078, 2013.
- Werth, S., Güntner, A., Petrovic, S., and Schmidt, R.: Integration of GRACE mass variations into a global hydrological model, *Earth and Planetary Science Letters*, 277, 166–173, doi:10.1016/j.epsl.2008.10.021, 2009.

Table 1. Oversight of water mass decline of lakes within the extended study area in km³ and mm EWH. Note that the mass decline of Lake Urmia and Lake Razzazah is part of a more gradual mass decline, while the water mass decline of other lakes mainly occurred during the 2007-2009 drought due to lake management.

Lake	Mass Decline 2007-2009	
	km ³	mm EWH
Tharthar	12.45 ± 0.04	47.9 ± 0.14
Habbaniyah	0.66 ± 0.04	2.53 ± 0.15
Razzazah	1.18 ± 0.10	4.57 ± 0.39
Hamrin	0.55 ± 0.02	2.11 ± 0.08
Adhaim	0.19 ± 0.10	0.73 ± 0.38
Dukan	0.94 ± 0.06	3.60 ± 0.21
Qadisisyah	5.08 ± 0.05	19.5 ± 0.2
Urmia	4.06 ± 0.14	15.6 ± 0.6
Mosul	0.97 ± 0.06	3.74 ± 0.22

Table 2. Summary of data used in this study

Variable	Dataset	Product	Resolution		Period
			Spatial	Temporal	
Water mass variation	GRACE CSR	-	-	1 mo	2003-2012
Precipitation	TRMM	3B42 V7	0.25° x 0.25°	1 d	1999-2012
Precipitation	Data Metereological department Kurdistan	-	-	1 d	2001-2012
Streamflow	Data Directorate Dukan Dam	-	-	1 d	2001-2012
Temperature	GLDAS	NOAH V2.7	0.25° x 0.25°	3 h	2001-2012
Climatic parameters	GLDAS	NOAH V2.7	0.25° x 0.25°	3 h	2001-2012
Soil moisture	GLDAS	NOAH V2.7	0.25° x 0.25°	3 h	2001-2012
Lake levels	ENVISAT	-	-	35 d	2002-2010
	Jason 1&2	-	-	10 d	2002-2015
Lake area	MODIS	MOD09Q1	250 x 250 m	8 d	2002-2012
Lake area	MODIS	MOD09A1	500 x 500 m	8 d	2002-2012

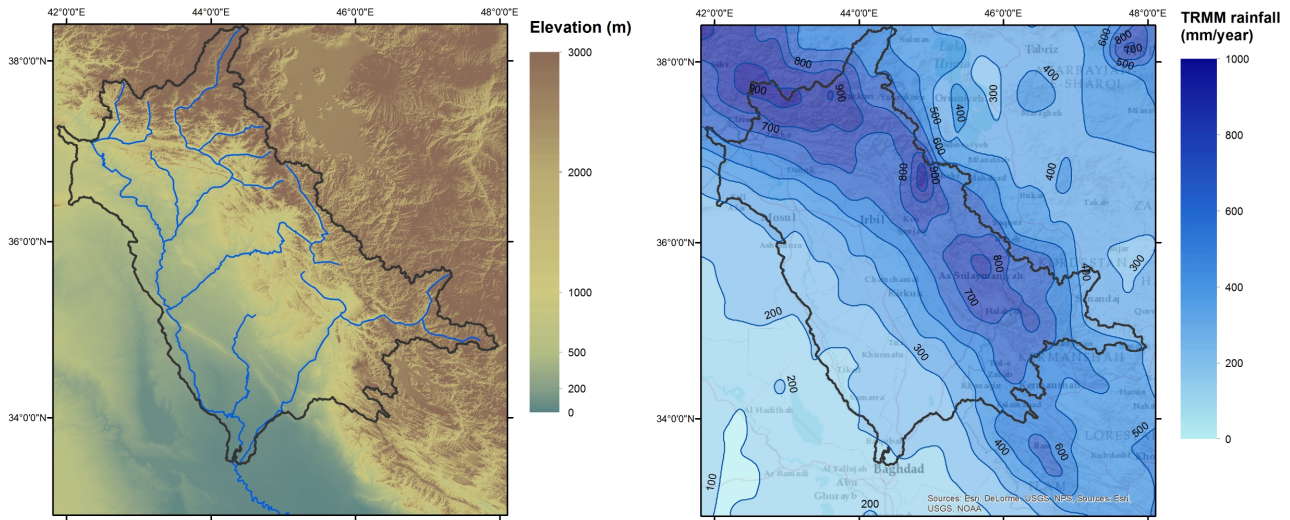


Figure 1. (left) Topographic map of Northern Iraq based on SRTM data and (right) the average yearly rainfall between 2002 and 2012 (mm yr^{-1}) based on TRMM data

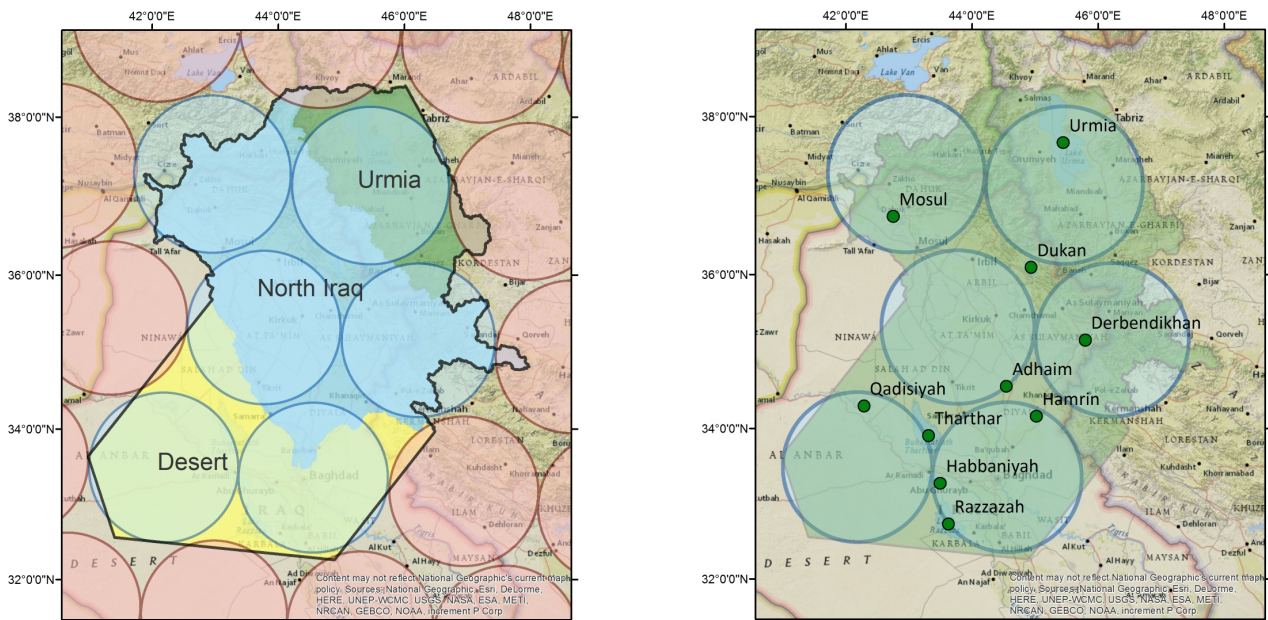


Figure 2. (left) Mascon coverage area for GRACE calculations. In the south-west the included desert area and in the north-east the included Urmia catchment. The blue circles show the coverage of the used mascons and the red circles mascons outside the study area. (right) Map of the total study area including the used mascons and lakes.

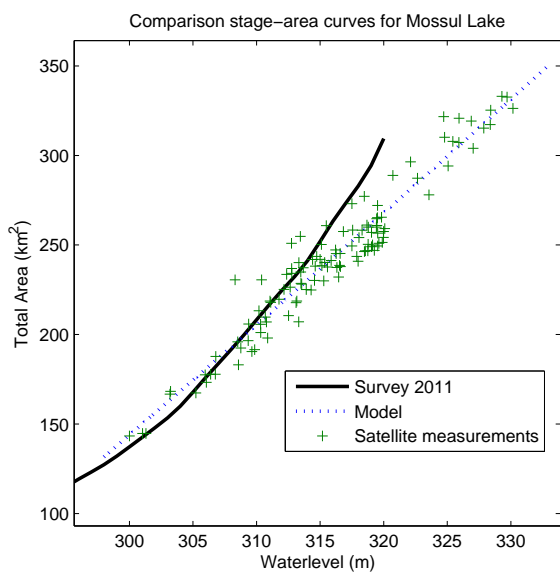


Figure 3. Linear regression stage-area curve for lake Mossul. The curve is compared with a survey in 2011 using sonar by Issa et al. (2013).

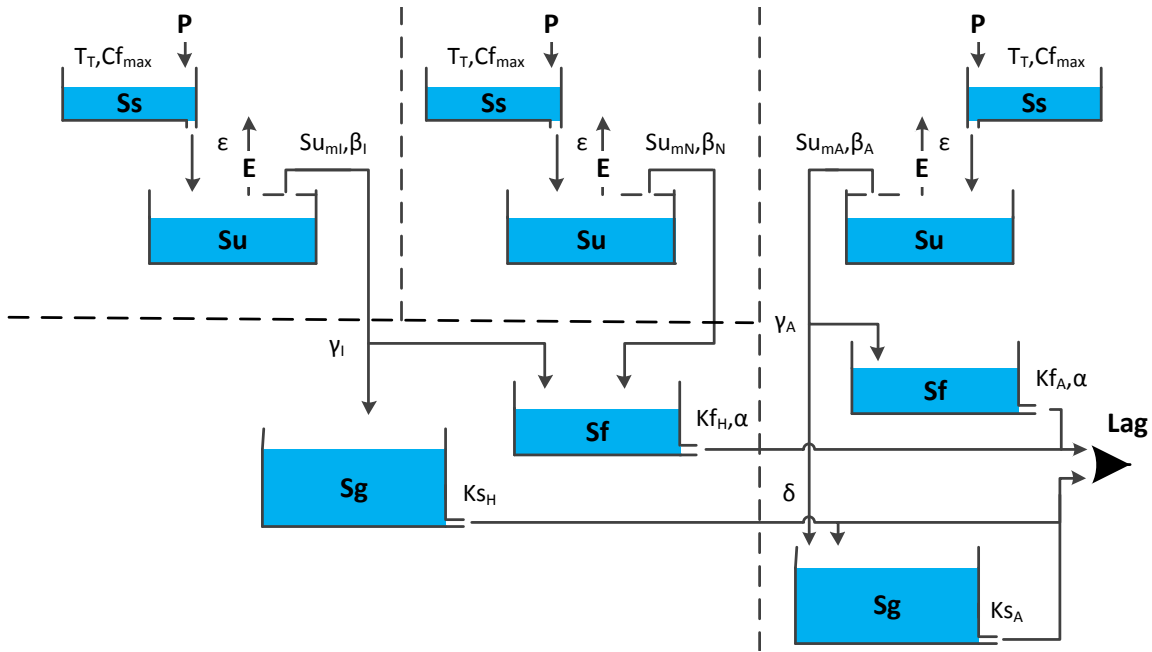


Figure 4. Setup of the rainfall-runoff model based on the three main land classes in Northern Iraq. The upper three reservoirs (Ss) represent the snow accumulation in the basin based in precipitation and temperature. The second layer of three reservoirs (Su) represent the water storage in the unsaturated zone and routes runoff to the fast runoff (Sf) and groundwater (Sg) reservoirs. The third layer of two reservoirs (Sf) represent the water storage related to fast runoff processes, which consist of overland flow and interflow. The two bottom reservoirs represent the groundwater storage, which is the main focus in this study. The water fluxes, indicated with arrows, are calculated based on reservoir levels and model parameters.

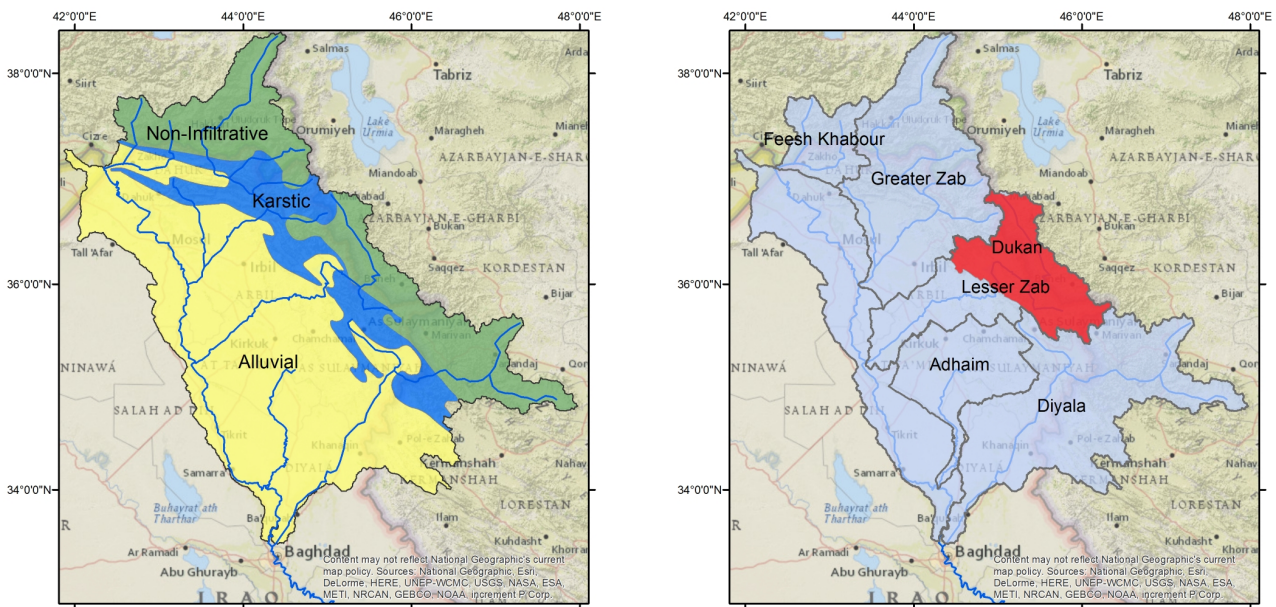


Figure 5. (left) Approximate division of Northern Iraq into three geologic zones, mainly based on (Stevanovic and Jurkiewicz, 2008). (right) Boundaries of main tributaries of the Tigris in Northern Iraq. The Dukan catchment, which is the upper part of the Lesser Zab catchment, is indicated in red. Calibration on streamflow is based on measurements from the Dukan area and calculation of water mass for the whole of Northern Iraq is based on all catchments.

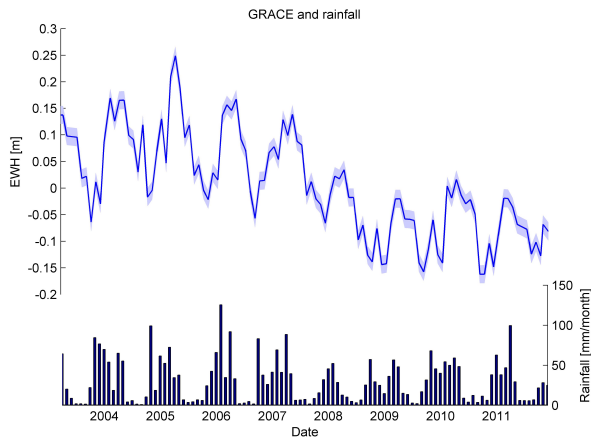


Figure 6. GRACE values and monthly precipitation for extended study area. During the wet winter periods, water accumulates in the region and total water mass increases. Largest water mass depletion occurred during seasons of 2007/2008 and 2008/2009.

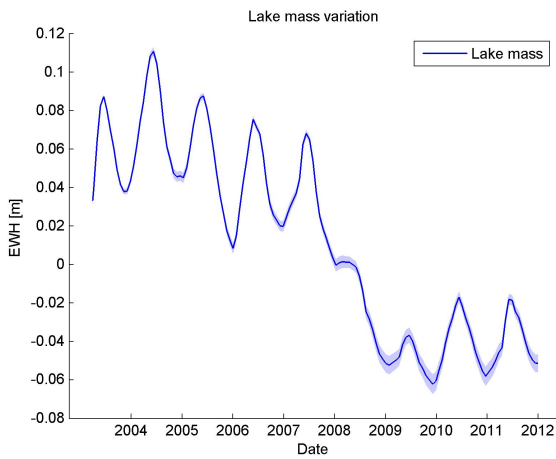


Figure 7. Lake mass variation.

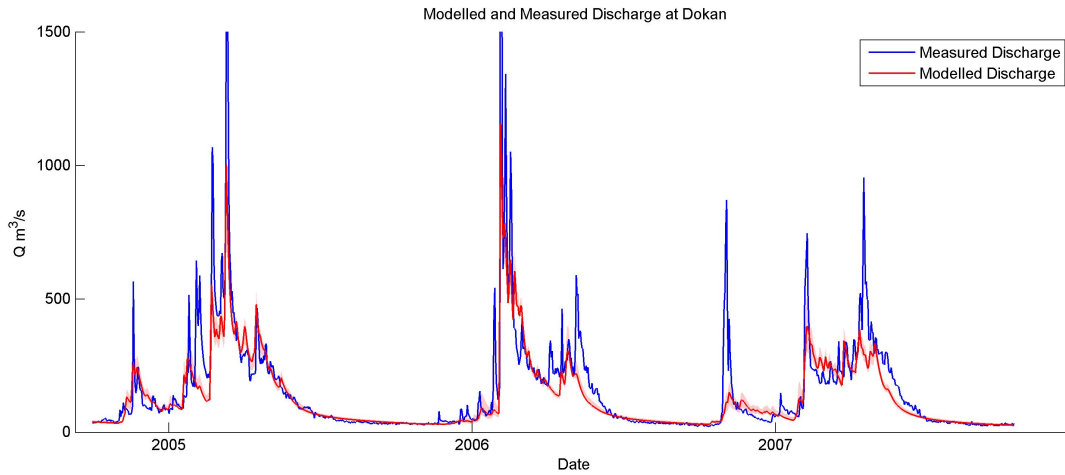


Figure 8. Discharge curves for the Dukan catchment between [Oct 2005 and Oct 2007]. (left) Measured discharge in blue against modelled discharge in red. The red line gives the optimal solution and the bandwith represents solutions within the pareto ensemble. (right) Identical to left graph but here on a log scale, which gives a better view on the low flows during the dry season.

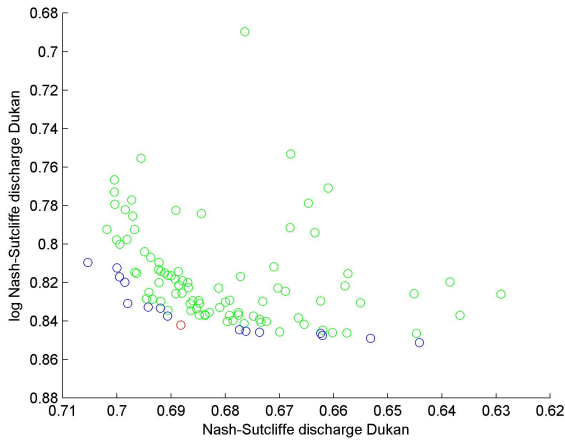


Figure 9. Pareto front for the performance indicators for river flow at the inflow of lake Dukan. The NS model performance is given on the y -axis and the logNS model performance x -axis. The blue dots represents the models which form the pareto front and the red dot the chosen optimal solution. These solution are used to calculate pareto ensembles for model mass and discharge. The green dots represent the models which would be added to the current pareto front if the model is also calibrated using the NS performance of GRACE. The relative little spread of the resulting performances indexes shows that the model produces consistent results for low/high discharges and GRACE values.

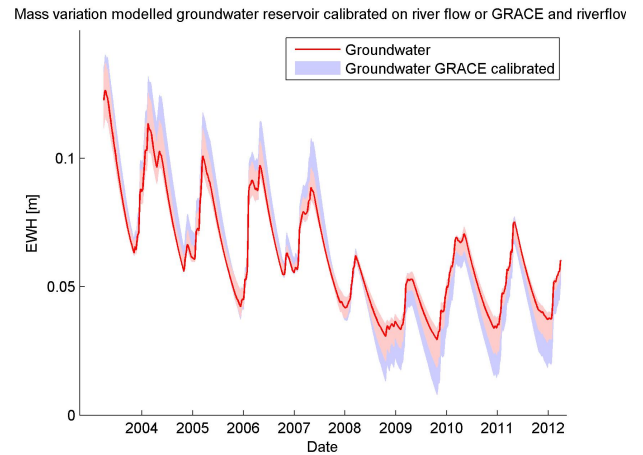


Figure 10. Water mass of groundwater reservoir smoothed over 10 days, showing a permanent decline of groundwater between 2007 and 2009. The red line and bandwith represent the pareto ensemble based on the NS and logNS of streamflow, while the blue bandwith shows the extension of the pareto ensemble when the NS for water mass is added as a third objective.

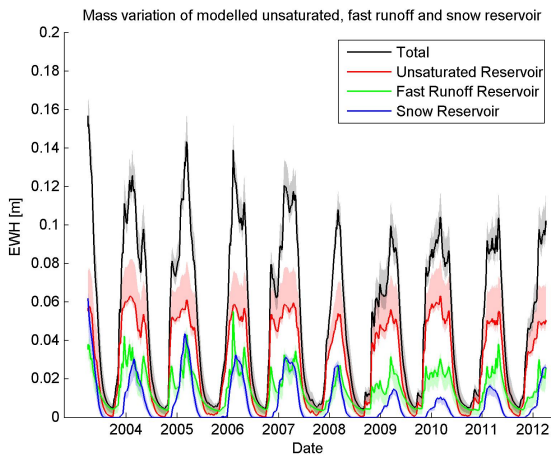


Figure 11. Water mass of the snow, unsaturated and fast runoff reservoir smoothed over 10 days. Lines represent optimal model and the bandwidth with the pareto ensemble based on NS and logNS of streamflow.

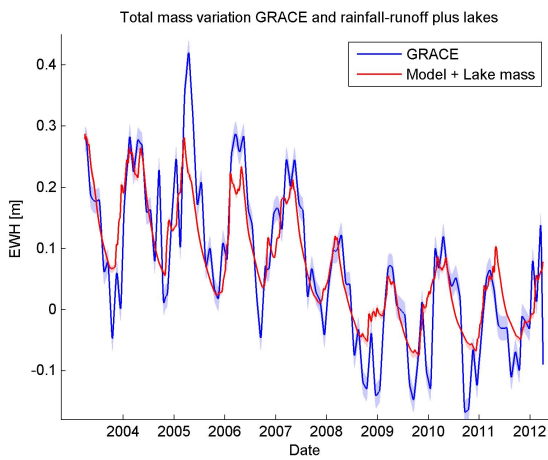


Figure 12. Comparison between resulting mass variation from GRACE and mass variation from lakes and rainfall-runoff model.