Interactive comment on "Recent changes in terrestrial water storage in the Upper Nile Basin: an evaluation of commonly used gridded GRACE products" by Mohammad Shamsudduha et al.

Anonymous Referee #1 [*italics*]

<u>Numbered responses are given below each comment (see revised manuscript with track-changes accepted):</u>

This study aims to estimate the TWS change and its individual components in the Upper Nile Basin using GRACE, LSMs and in situ observations. Actually, similar studies have been done in this region by Awange et al. [2013], Awange et al. [2014], and Nanteza et al. [2016]. So, the main point is whether this manuscript can bring enough new knowledge based on new/updated data or methods. Different from previous studies, three different GRACE products (gridded level-3 GRCTellus, JPL mascon and constrained GRGS products) were compared and validated with in situ TWS observations in this study. However, the detailed scaling process used in this study is still unclear for me (see detailed comments below). I also suspect that whether limited 6 monitoring well observations can represent actual large-scale GWS variations in the study region.

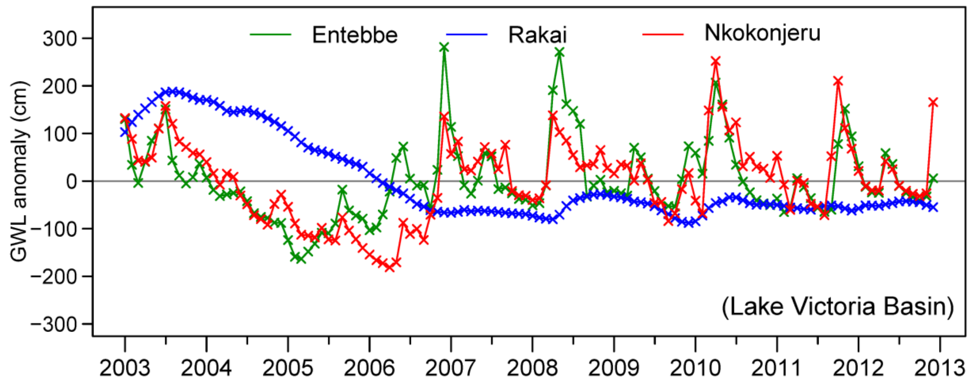
Responses to general comments [G1 to G3]:

[G1] We thank the Anonymous Referee #1 (AR1) for their comments on the manuscript. We are pleased that the reviewer has recognised the central difference between this study and previous studies in the region that include: (1) application of commonly used gridded GRACE products rather than a single GRACE product; and (2) an evaluation of these gridded products to represent the phase and amplitude of changes in terrestrial water storage in the Upper Nile Basin including a large and well-constrained change in surface water storage from 2003 to 2006.

Especially, all three well observations in the LVB are located near the Lake Victoria. The representativeness of these wells is questionable. In addition, there are some obvious typos in the manuscript.

[G2] We agree with AR1 that the representivity of a limited number (6) of monitoring wells in the region is questionable. These daily monitoring records have been selected from a larger database of groundwater-level monitoring records in Uganda on the basis of the completeness and quality of their records from 2003 to 2012. Unfortunately, several time-series records from Uganda were excluded due to unexplained errors and substantial gaps; locations of the several monitoring wells also reside outside of the studied basins. Long time-series records of groundwater levels over the period from 2003 to 2012 from western Kenya, northern Tanzania, Rwanda and Burundi have not been identified despite intensive investigations carried out by *The Chronicles Consortium*, https://www.un-igrac.org/special-project/chronicles-consortium.

In the supplementary information of the revised manuscript, we now include a new figure (**see supplementary Fig. 3**) showing the 6 employed piezometric records that inform in situ ΔGWS (also see Fig. AR1.1). In the Lake Kyoga Basin, piezometric records from 3 sites show consistency in the seasonality and amplitude of groundwater storage changes plotted as monthly groundwater-level anomalies relative to the mean for the period from 2003 to 2012; further details of these oscillations are described by Owor et al. (2009). In the Lake Victoria Basin, groundwater-level records from 2 sites (Entebbe, Nkokonjeru) are similar in their phase and amplitude, and are influenced by changes in the level of Lake Victoria as demonstrated by Owor et al. (2011). The groundwater-level record from Rakai represents local semi-arid conditions that exist within catchment areas (e.g. River Ruizi) draining to the western shore of Lake Victoria in Uganda. Although there are differences in the phase of groundwater-level fluctuations between the semi-arid site at Rakai and both Entebbe and Nkokonjeru (as well as the 3 sites in the Lake Kyoga Basin), amplitudes are similar.



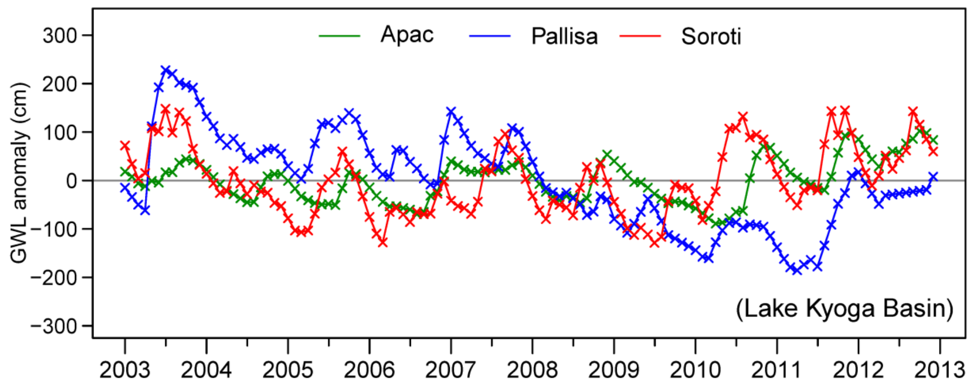


Figure AR1.1. Time-series records of monthly anomaly of groundwater-level monitoring records at three stations in LVB (top), and records at three stations in LKB (bottom).

GWS estimation from GRACE: Based on my understanding on the manuscript, ΔGWS= the rescaled GRACE ΔTWS (sf=1.7 for GRCTellus, sf=? for JPL mascon) minus scale-down ΔSWS (sf=0.11 for GRCTellus and sf=0.39 for JPL mascon) minus simulated ΔSMS. Why so-called a scale down of ΔSWS was used rather than the original ΔSWS (EWH, based on equation 2, Line 317)? In fact, the ΔGWS estimation from GRACE (GRCTellus, JPL mascon and GRGS) was not given in detail. I would suggest the authors explain it in a paragraph in 3.2.2.

[G3] We thank the AR1 for this critical comment on the estimation of ΔGWS derived from GRACE datasets.

First, GRACE ΔTWS time-series records were generated for LVB and LKB following a conventional approach by: (i) selecting $1° \times 1°$ grids within the basin boundary, (ii) applying gridded scaling factors to the corresponding ΔTWS grids; and (iii) taking the average of time-series records of scaled ΔTWS grids over the basin. For *GRCTellus* products (CSR, JPL, GFZ), we applied scaling coefficients derived from CLM4.0 land surface model provided by Landerer and Swenson (2012). Similarly, gridded scaling factors were applied to JPL-Mascons product provided by Wiese et al. (2015). No scaling factors were applied to GRGS GRACE. On the specific question of 'rescaled GRACE ΔTWS', we did not apply a single multiplicative scaling factor of 1.7 to *GRCTellus* ΔTWS (CSR, JPL, GFZ products) to generate a basin-wide time-series data.

Below we explain why we conducted the scaling experiments. Please note, however, that in the revised manuscript, we now only apply an unscaled (original) ∆SWS signal to estimate ∆GWS from GRACE data. To respond to the original query of the reviewer, we now provide a clearer explanation as to why unconventional scaling experiments were conducted (**revised section 3.2.3**). Two separate, unconventional scaling experiments were conducted only for the Lake Victoria Basin (LVB) in order to highlight the discrepancy between *GRCTellus* TWS and in-situ (i.e. now 'bottom-up') TWS and to reconcile *GRCTellus* with in-situ observations since we observe that the amplitude of monthly anomalies of combined ∆SWS+∆SMS signals substantially exceeds GRACE ∆TWS, particularly for the *GRCTellus* GRACE ∆TWS signal (**see Fig. AR1.2 and supplementary Fig. S4**). Under the first experiment, we apply a single multiplicative scaling factor of 1.7, informed by the lowest RMSE, in order to 'scale up'' the *GRCTellus* ensemble mean of ∆TWS data. In the second experiment, we apply a 'scaled down' ΔSWS in the LVB, recognising that ΔSWS is the largest contributor to ΔTWS in the LVB. We apply spatially-averaged scaling factors representative of (1) Lake Victoria and its surrounding grid cells (experiment 1: s=0.71; range 0.02–1.5), and (2) the open water surface of Lake Victoria without surrounding grid cells (experiment 2: s=0.11; range 0.02–0.30)." These experiments suggest that the 'true' *GRCTellus* GRACE ΔTWS signal is weakened during processing and that the combined ΔSWS+ΔSMS signal is greater than ΔTWS, mathematically resulting to a positive estimate of ΔGWS.

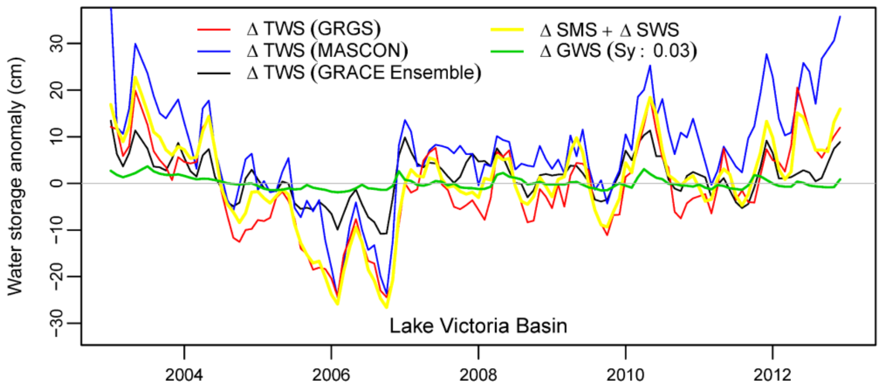


Figure AR1.2. Time-series records of GRACE ∆TWS, sum of in-situ ∆SWS and ensemble mean ∆SMS and in-situ ∆GWS for LVB.

In the revised manuscript, we have also included a new sub-section (**3.2.2. Estimation of ΔGWS from GRACE**) on **lines 422-429** that describes how ∆GWS is estimated from GRACE ∆TWS. In the revised analysis, we apply the original and un-scaled ∆SWS to estimate GRACE-derived ∆GWS in both LVB and LKB (**see Fig. 8 in the revised manuscript and supplementary Fig. S14**).

Responses to specific comments [S1 to S20]:

(1) Line 240, CRS should be CSR.

[S1] Agreed and corrected in revised manuscript [**line 247**]

(2) Line248-249: GRCTellus datasets are provided as 1X1 grids, but ~111 km is not the so-called spatial resolution of GRACE. At least, in some place of the manuscript, the authors should emphasize that the real resolution of GRACE is about 300 km, rather than that provided by these level-3 products.

[S2] We thank AR1 for their point of clarity on the resolution of GRACE products. In the revised manuscript, we have stated clearly the spatial resolution of GRACE footprint of ~200 000 km^2 in several places [**lines 47, 257 and 438**].

(3) Line 254, the citation Geruo et al., 2013 should be A et al. 2013. This is also a mistake in some other papers. Actually, A is his family name and Geruo is his forename.

[S3] Agreed and corrected in revised manuscript [**lines 261, 271 and 730-732**].

(4) Line 287, the citation (CSR, 2016) was not shown in the References. If there is no publication about it, maybe the authors can provide the website link where the information was available.

[S4] The citation of CSR (2016) is now replaced with Tapley et al. (2015) [**lines 301-302, 914-916**].

(5) Line 405, JLP should be JPL. For GRGS, whether scaling factor was applied?

[S5] Agreed, JPL is corrected throughout the revised manuscript [**line 545**]; no scaling factors were applied for GRGS product. Note that gridded scaling factors were applied to *GRCTellus* and JPL-Mascons solutions.

(6) Line 310, Fig.s should be Figs.
[S6] Agreed and corrected in revised manuscript [**line 334**].

(7) Line 394, if I understand it correctly, gridded scale factors from Landerer et al. were not used in this study finally. The authors applied a single scaling (1.7) actually. Based on Figure S1 and the authors' experiment (Fig. S10b), the factors are highly underestimated by Landerer et al. in the LVB.

[S7] We applied the gridded scaling factors, provided by Landerer and Swenson (2012) to generate GRACE-derived ∆TWS time-series data in the two basins (**lines 395-413**). Under an additional scaling experiment applied solely to *GRCTellus*, a single, multiplicative scaling factor of 1.7 was applied in an attempt to reconcile large differences between *GRCTellus* GRACE ∆TWS and the in-situ (now 'bottom-up') ∆TWS.

(8) In 3.2 Methodologies, how to estimate GWS using GRACE in detail? I would suggest the authors explain it in a paragraph in 3.2.2.

[S8] As per response G3, in the revised manuscript, we have included a new sub-section (**3.2.2. Estimation of ΔGWS from GRACE**) in lines **422-429** that describes how ∆GWS is estimated from GRACE ∆TWS.

(9) Line 434, "in both LVB and LKB (see supplementary Figs. S2–S7)." The captions of Figs. S5-S7 are "over the Victoria Nile Basin". Does the Victoria Nile Basin mean the LKB? The caption of Figure S9 also contains "in VNB".

[S9] Agreed and corrected in the revised **supplementary Figs. S8-S10**.

(10) Line 436, simulated ∆SWS should be simulated ∆SMS?

[S10] Agreed and corrected in revised manuscript [**line 498**].

(11) Line 446, "all 5 GRACE ∆TWS and in situ ∆TWS time-series records". There are only 4 curves in each panel of Figure 7.

[S11] Agreed, in the revised manuscript, the text is revised to "all 3 ∆TWS from 5 GRACE products". See **line 508** in revised manuscript.

(12) Line 449, "the period of 2004 to 2006", but in table 3, "2003-2006". This kind of inconsistency occurs several times in the manuscript.

[S12] Agreed and the period of "2004 to 2006" is now corrected as "2003 to 2006" throughout the revised manuscript. Note that the linear trends reported in revise **supplementary Figs. 11 and 12** have changed from previously reported trends which were restricted from 2004 to 2006. We also revised linear trends in the manuscript to reflect changes from the previous values.

(13) Line 464, "see supplementary Table S1". No correlation estimates in table S1 in fact.

[S13] No correlation estimates are provided in the supplementary Table S1, which reports variances explained by linear regression reported as R^2 values of the model fit. Please see the revised **supplementary Table S1** that tabulates now provides variance of individual signals.

(14) Line 465-466, "GRACE ∆TWS is unable to explain natural variability in in situ ∆TWS in LKB though this may be explained by the fact that SWS in Lake Kyoga is influenced by dam releases from LVB". GRACE can detect all mass changes including both natural and anthropogenic variability, but can not disaggregate individual components. If in situ ∆TWS includes all mass change signals, it should be consistent with the GRACE estimate, no matter mass change is natural or anthropogenic. I suspect that the lower correlation in the LKB might be caused by the smaller area of LKB and larger leakage errors from the surrounding regions (including LVB).

[S14] We appreciate that GRACE detects all mass changes whether they are natural or anthropogenic. We also appreciate the explanation suggested by AR1 and provide calculations to support this assertion. In the revised manuscript (**lines 597-605**), we briefly discuss the leakage from Lake Victoria into the adjacent basin but we will expand this discussion and report that our leakage analysis shows that GRACE signal leakage into LKB from LVB, which is 3 times larger in area than LKB, is 3.4 times bigger for both *GRCTellus* GRACE and GRGS products.

(15) The caption of Table S1, no "variablility (i.e., variance, cmˆ2)" in the table. In the caption, what is the meaning of 120 cmˆ2 and 24 cmˆ2? The variances of in situ ∆TWS?

[S15] Yes, the values of 120 and 24 cm^2 are variances in in-situ ∆TWS for LVB and LKB respectively. The caption for the **supplementary Table 1** is correctly in the revised manuscript.

(16) Line 473-477, GRACE-derived ∆TWS was rescaled to recover the actual mass change. But, why the scaling down process was needed to remove ∆SWS for estimating ∆GWS? If rescaled ∆TWS time series was used to estimate ∆GWS, maybe the authors should use in situ ∆SWS (equation 2) rather than scaling down ∆SWS. I also cannot understand the caption of Figure 8. Why a scaling down process of SWS is needed for disaggregating GWS from GRACE-derived rescaled TWS (Line 399-405).

[S16] See above response to the general comment G3.

(17) Line 399-405, were these factors calculated from the product of Landerer and Swenson 2012 (Figure S1)? Note that this product should be used for recovering TWS rather than only for SWS. In line 402, s=0.71 for experiment 1. But in caption of Figure S10, s=0.77 for experiment 1.

[S17] Yes, as per response G3 and S7, gridded scaling factors provided by Landerer and Swenson (2012) were used to generate ∆TWS time-series records for LVB and LKB. However, under the scaling experiments undertaken on the ensemble mean of 3 *GRCTellus* GRACE products (CSR, JPL, GFZ), multiplicative scaling factors were applied to observed ∆SWS time-series data (s=0.71 and s=0.11 in experiment 1) and ∆TWS (s=1.7 in experiment 2), guided by RMSE, in an attempt to reconcile substantial differences between *GRCTellus* GRACE ∆TWS and bottom-up ∆TWS. Please note that these scaling factors under these experiments were not applied to ∆TWS or ∆SWS datasets to finally estimate GRACE-derived ∆GWS (see new section of **3.2.2. Estimation of ΔGWS from GRACE** on lines **422-429**).

(18) Section 3.1.3, GLDAS does not assimilate surface water, which is an important TWS component in the study region. Whether the absence of surface water process will highly affect the accuracy of simulated soil moisture from GLDAS? Maybe the authors can try to use WGHM model which considers the surface water. In Figure S12, the authors compared many LSMs except WGHM, which simulates all TWS components. If the authors removed ∆SMS from WGHM, maybe there will be a better agreement between in situ well observations and GRACE-based ∆GWS, although the representativeness of these wells is also questionable.

[S18] We thank AR1 for their suggestion of the use of WGHM to test GRACE-derived ∆GWS but this is beyond the scope of the current study and we will consider such experiments in future studies.

(19) Line 1117, cantered should be centered.

[S19] Agreed and corrected in the revised manuscript [**line 1234**].

(20) Figure 8, what is the criterion of selecting Sy?

[S20] The explanation for the applied range of *Sy* values is given in the revised manuscript [**lines 378-381**].

References

Landerer, F. W., and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resour. Res., 48, W04531, 2012.

Owor, M., Taylor, R. G., Tindimugaya, C., and Mwesigwa, D.: Rainfall intensity and groundwater recharge: empirical evidence from the Upper Nile Basin, Environmental Research Letters, 1-6, 2009.

Owor, M., Taylor, R. G., Mukwaya, C., and Tindimugaya, C.: Groundwater/surface-water interactions on deeply weathered surfaces of low relief: evidence from Lakes Victoria and Kyoga, Uganda, Hydrogeol. J., 19, 1403-1420, 2011.

Wiese, D. N., Yuan, D.-N., Boening, C., Landerer, F. W., and Watkins, M. M.: JPL GRACE Mascon Ocean, Ice, and Hydrology Equivalent Water Height JPL RL05M.1. Ver. 1, PO.DAAC, CA, USA, 2015.

Interactive comment on "Recent changes in terrestrial water storage in the Upper Nile Basin: an evaluation of commonly used gridded GRACE products" by Mohammad Shamsudduha et al.

Anonymous Referee #2 [*italics*]

<u>Numbered responses are given below each comment (see revised manuscript with track-changes accepted):</u>

[AR2] This study evaluates, for the Upper Nile Basin over the 2003-2012 period, several estimates of terrestrial water storage (TWS) as processed from the Gravity Recovery and Climate Experiment (GRACE) retrievals with in situ and model-derived estimates of its individual terms: surface water storage (SWS), soil moisture storage (SMS), and groundwater storage (GWS).
The authors reach interesting conclusions, namely 1) the pre-processing of GRACE greatly affects estimated annual TWS amplitude and, most notably, reconcilability with bottom-up approaches and 2) uncertainty in GRACE TWS and model-derived prevents a reasonable inference of GWS variation in these aquifers. While I appreciate the scientific value of this work,
I find this manuscript confusing at times in its logic, and lacking rigor regarding how methods and some quantities are defined. Therefore, I recommend resubmission only after the authors have made a substantial rewriting effort to improve the clarity of the presented results.

[G0] We greatly appreciate the critical comments of the Anonymous Referee #2 (AR2) and their recognition of the important conclusions of the manuscript.

General comments [G1-G3]

[AR2] "In situ ΔTWS" is used throughout the manuscript, but this term is quite misleading: as defined in Eq. (1) and then L379-381, this quantity is the sum of ΔSWS, ΔGWS, and ΔSMS estimates. While the two former terms are indeed estimates based on situ measurements, ΔSMS is averaged from simulations with three gridded hydrological models at 0.25 resolution (Sect 3.1.3 and L580-581). This is of particular importance since the whole study is about attempting to reconcile estimates of storage compartments across approaches and scales. I suggest using something like "bottom-up ΔTWS" instead.

[G1] We thank AR2 for their critical comment here. We agree and have adapted the proposed nomenclature, "bottom-up ΔTWS", in the revised manuscript to make the distinction clearer.

[AR2] The method section is rather long, in particular the description of GRACE datasets retrievals and the applied methodology in sections 3.1.2, 3.2.1 and 3.2.2. While I understand the authors want to present the remaining datasets (ΔSWS, ΔSMS . . .) before detailed how ΔTWS is being processed, sect. 3.2.1 and sect 3.2.2, are even frankly confusing at times, e.g., when the ΔTWS scaling methodology is explained (L357-363, see specific comments) and then discussed again (L387-397) so that in the end I am not sure what was used for the study.

[G2] We appreciate that the description of various datasets and the method section are long and keep them separate under two sub-sections: Datasets (3.1) and Methodologies (3.2). The apparent confusion in the application of scaling factors may derive from the fact that we conducted additional scaling experiments only for the ensemble mean ΔTWS of 3

GRCTellus GRACE products (CSR, JPL, GFZ). These additional scaling experiments were conducted in an attempt to reconcile *GRCTellus* GRACE ∆TWS with 'bottom-up ∆TWS'. As per responses S7 and S17 to AR1, we have now clarified the selected methodologies for scaling factors in sections 3.2.1, a new section (**3.2.2. Estimation of ΔGWS from GRACE**) on **lines 422-429**) and 3.2.3 in the revised manuscript.

[AR2] TWS sometimes appears instead of ∆TWS (e.g. L79-86). While this be should a mere technical comment, in some cases TWS would actually be more accurate in the general sense (i.e. the concept of storage), e.g. when discussing reduction in volumetric storage in the whole basin (e.g., L537-539 where "∆TWS" is used).

[G3] We thank AR2 for their comment here and have revised the use of 'TWS' and '∆TWS' accordingly in the revised manuscript (for example, see **lines 20, 30, 675**).

Responses to specific comments [S1 to S3]:

L21-22: *It would be more accurate to say that the authors "test the phase and amplitude of three GRACE ∆TWS estimates derived from 5 commonly-used gridded products [. . .]".*

[S1] We thank AR2 for their critical comment and suggestion here. We agree with AR2 and have employed suggested edits in the revised manuscript [**line 21-22**: we test the phase and amplitude of three GRACE ΔTWS signals from 5 commonly-used gridded products].

L123: *What is the actual time span of the "unintended experiment": 2004-2006 (like stated here)? 2005-2006 (e.g., L553)? 2003-2006 (most of the manuscript)? The authors should delimit this period consistently across the main text, the tables, the figures, and the supplementary materials.*

[S2] Agreed, we have used the time span of 2003-2006 to indicate the "unintended experiment" throughout the revised manuscript.

L169-173: *The authors should comment on the large discrepancy between these two lake area estimates. In addition, why do the authors report the HydroSHEDS area value as being from this study in Table 1?*

[S3] We thank AR2 for their suggestion here and have included in the revised manuscript a statement highlighting the large discrepancy between the delineated area of LVB reported by UNEP (2013) and both Awange et al. (2014) and this study, which employs the *HydroSHEDS* boundary shapefiles for LVB and LKB. In the Table 1 we have now added the reference to *HydroSHEDS* database [**line 1023**].

L357-363: *The authors first state that they spatially aggregate the unscaled ∆TWS signal over the study region in order to have a time series, but then say that the scaling factors are applied to each grid of the GRACE mesh, therefore it is done before spatial aggregation? Please clarify.*

[S4] Yes, gridded scaling factors were applied to corresponding grid cells for ∆TWS before the spatial aggregation over LVB and LKB in order to generate time-series data. We have revised the texts in order to clarify this point [**lines 395-413**].

L395-397: *Along with the regionally-averaged gain factor, why did the authors not also test the third method described L392-394?*

[S5] We do neither possess nor access monthly scaling factors to conduct the third scaling experiment. We have clarified the point that we only employed method (2) in the revised manuscript [**line 400-402**].

L415-418: *A lag of 2-3 months between lowest rainfall and lowest ∆TWS is also well noticeable, while ∆SMS respond more quickly to rewetting after the driest month (~1 month) and ∆SWS is slower (~4 months lag after minimum rainfall).*

[S6] We appreciate this comment and our explanation is that seasonal hydrological response to rainfall, particularly changes in ∆SWS are also influenced by managed dam operations.

L432-434: *Figs. S5 to S7 are relative to the entire Victoria Nile Basin and not Lake Kyoga Basin, I do not see how the authors can derive the observation that "GRACE-derived ∆TWS signals are strongly correlated in both LVB and LKB (see supplementary Figs. S2–S7)". The same applies L441-444. Maybe the figures were unintentionally swapped with relative to LKB?*

[S7] Agreed and corrected in the revised manuscript [**see supplementary Figs. S8–S10**].

L446-447: *This sentence is misleading since only 3 ∆TWS estimates are used shown, albeit derived from 5 different GRACE products.*

[S8] Agreed, as per responses S11 (AR1) and S1 (AR2), we have revised the text.

L449-456: *The authors might already mention that only ∆GWS shows an increase in 2005-2006, as later discussed in the Discussion section.*

[S9] We thank AR2 for this comment. We have provided an explanation of the apparent rise in ∆GWS in **lines 621-626**.

L457-458: *A support supplementary figure with time series for LKB would help. Is it what Fig. S9 should have been (instead of describing the Victoria Nile Basin)? If so, the authors should add a reference to Fig. S9 here, and replace "[. . .] (see supplementary Figs. S8–S9)." by "[. . .] (see supplementary Figs. S8–S9)." in L456, and caption of Fig. S9 should read "LKB", instead of "VNB".*

[S10] Agreed, LKB is mistakenly labelled as Victoria Nile Basin. We have corrected this in a revised supplementary document [**see supplementary Figs. S8–S10**].

L465-466: *I am not sure what the authors mean, how could the TWS signal miss one of its component, unless it refers to a water transfer within the system? All the more that even if mention of LVB-driven water balance of LKB is given on L175-177, this point is not picked up later in the Discussion section. Is it related to the substantial variability of ∆TWS deriving from ∆SMS in in LKB as compared to LVB? Could the authors expand their idea?*

[S11] We appreciate that GRACE detects all mass changes, whether they are natural or anthropogenic, and regret the confusion caused by our statement, "GRACE ΔTWS is unable to explain natural variability in in situ ΔTWS in LKB though this may be explained by the fact that SWS in Lake Kyoga is influenced by dam releases from LVB". As per response S14 (AR1), further discussion of signal leakage from Lake Victoria into Lake Kyoga is provided in the revised manuscript in which we will report on our leakage analysis showing that GRACE signal leakage into LKB from LVB, which is 3 times larger in area than LKB, is 3.4 times bigger for both *GRCTellus* GRACE and GRGS products [**lines 597-604**].

L476-477: *Why scaling down ΔSWS rather than using the rescaled ΔTWS presented right above (L474-476) to disaggregate ΔGWS?*

[S12] As per response G3 to AR1, we agree that the use of a scaled-down ΔSWS instead of an original signal is confusing and unconventional. Below we explain why we conducted the scaling experiments. Please note, however, that in the revised manuscript, we now only apply an unscaled (original) ΔSWS signal to estimate ΔGWS from GRACE data.

To respond to the original query of AR2, we now provide a clearer explanation as to why unconventional scaling experiments were conducted (**revised section 3.2.3**). Two separate, unconventional scaling experiments were conducted only for the Lake Victoria Basin (LVB) in order to highlight the discrepancy between *GRCTellus* TWS and in-situ (i.e. now 'bottom-up') TWS and to reconcile with in-situ observations. We find that the amplitude of monthly anomalies of combined ΔSWS+ΔSMS signals substantially exceeds GRACE ΔTWS, particularly for the *GRCTellus* GRACE ΔTWS signal (**see Fig. AR2.1 and supplementary Fig. S4**).

Under the first experiment, we apply a single multiplicative scaling factor of 1.7, informed by the lowest RMSE, in order to 'scale up'' the *GRCTellus* ensemble mean of ΔTWS data. In the second experiment, we apply a 'scaled down' ΔSWS in the LVB, recognising that ΔSWS is the largest contributor to ΔTWS in the LVB. We apply spatially-averaged scaling factors representative of (1) Lake Victoria and its surrounding grid cells (experiment 1: s=0.71; range 0.02–1.5), and (2) the open water surface of Lake Victoria without surrounding grid cells (experiment 2: s=0.11; range 0.02–0.30)." These experiments suggest that the 'true' *GRCTellus* GRACE ΔTWS signal is weakened during processing and that the combined ΔSWS+ΔSMS signal is greater than ΔTWS, mathematically resulting to a positive estimate of ΔGWS.

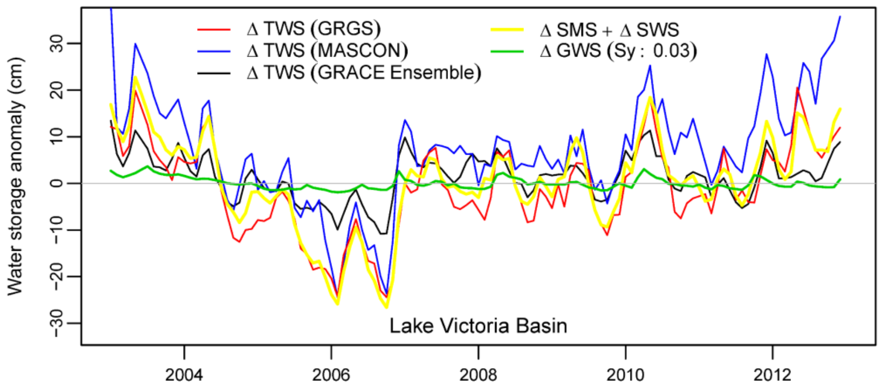


Figure AR2.1. Time-series records of GRACE ΔTWS, sum of in-situ ΔSWS and ensemble mean ΔSMS and in-situ ΔGWS for LVB.

In the revised manuscript, we now include a new sub-section (**3.2.2. Estimation of ΔGWS from GRACE**) in lines **422-429** that describes how ΔGWS is estimated from GRACE ΔTWS. In the revised analysis, we apply the original and un-scaled ΔSWS to estimate GRACE-derived ΔGWS in both LVB and LKB (**see Fig. 8 in the revised manuscript and supplementary Fig. S14**).

L526-527: *This sentence essentially repeats L517-518, with typos (see Technical comments).*

[S13] Agreed, lines 526-527 have been deleted in the revised manuscript [deleted sentence following the **line 592**].

L529: *The measurement error is not necessarily only a bias (systematic) is there are random components; Swenson and Wahr (2006) seem to keep this broader definition.*

[S14] We applied measurement and leakage errors from Landerer and Swenson (2012); reference to Swenson and Wahr (2006) is incorrect and has been deleted in the revised manuscript [**lines 266-268**].

L541-548: *Would not it be more correct to say that the choice of ΔSMS from LSMs contributes to uncertainty in estimating bottom-up ΔTWS (termed in situ in the manuscript, see General Comments), and consequently comparing it to GRACE ΔTWS, rather than uncertainty "GRACE analysis"? In addition, the order of sentences in this paragraph leaves me with the impression that this study did not bring any improvement to estimating bottom-up ΔTWS, while most of the manuscript uses this estimate as a benchmark to test GRACE ΔTWS products. In order to avoid finally leaving the reader with "how reliable is this ΔTWS reconciliation then?", the authors should maybe remind in the discussion that ΔSWS is by far the largest contributor in LVB at least, somewhat limiting the propagation of ΔSMS uncertainty.*

[S15] We agree with this argument of AR2 that ΔSWS is by far the largest contributor to ΔTWS in the LVB and is dominated by an accurately observed ΔSWS signal of 81 km^3, limiting the propagation of ΔSMS uncertainty. We have revised the discussion to reflect this important argument as it relates to statements about uncertainty in GRACE products relative to a 'bottom-up' ΔTWS [**lines 628-630**].

L616-617: *This should probably be stated already in the Discussion.*

[S16] We thank AR2 for this suggestion.

Technical corrections

L101: *SSA is not used anywhere else in the manuscript of supplement.*

[T1] Agreed, "(SSA)" is now deleted in the revised manuscript [**line 105**].

L527: *Likely typos, maybe "[. . .] priori information from LSMs contributes to adding uncertainty to ΔTWS signals".*

[T2] Agreed, this statement is now deleted in the revised manuscript.

Figs. 5 and 6: *What are the dashed vertical lines in the top panels and the horizontal dashed line in the bottom panels?*

[T3] Agreed, we have deleted the vertical line which separates the two periods (2003-2006, 2007-2012). The dashed horizontal line indicates the mean rainfall for the period of 2003-2012; this detail will be made clear in the figure captions (**revised Figs. 5 and 6**) in the revised manuscript. Furthermore, for clarity, we are now showing all 3 SMS lines derived from 3 GLDAS LSMs (CLM, NOAH, VIC).

Recent changes in terrestrial water storage in the Upper Nile Basin: an evaluation of commonly used gridded GRACE products

Mohammad Shamsudduha[1, 2], Richard G. Taylor[2], Darren Jones[3], Laurent Longuevergne[4], Michael Owor[5] and Callist Tindimugaya[6]

[1]Institute for Risk and Disaster Reduction, University College London, UK
[2]Department of Geography, University College London, UK
[3]Centre for Geography, Environment and Society, University of Exeter, UK
[4]CNRS – UMR 6118 Géosciences Rennes, Université de Rennes 1, France
[5]Department of Geology & Petroleum Studies, Makerere University, Uganda
[6]Directorate of Water Resources Management, Ministry of Water & Environment, Uganda

Correspondence to: M. Shamsudduha (m.shamsudduha@ucl.ac.uk)

Abstract

GRACE (Gravity Recovery and Climate Experiment) satellite data monitor large-scale changes in total terrestrial water storage (ΔTWS) providing an invaluable tool where in situ observations are limited. Substantial uncertainty remains, however, in the amplitude of GRACE gravity signals and the disaggregation of ~~Δ~~TWS into individual terrestrial water stores (e.g. groundwater storage). Here, we test the phase and amplitude of three GRACE ΔTWS signals from 5 commonly-used gridded products (i.e., NASA's *GRCTellus*: CSR, JPL GFZ; JPL-Mascons; GRGS GRACE) using in situ data and modelled soil-moisture from the Global Land Data Assimilation System (GLDAS) in two sub basins (LVB: Lake Victoria Basin, LKB: Lake Kyoga Basin) of the Upper Nile Basin. The analysis extends from January 2003 to December 2012 but~~The~~ focuses~~ of this analysis is~~ on a large and accurately observed reduction in ΔTWS of ~~75~~ 83 km^3 from ~~2004~~ 2003 to 2006 in Lake Victoria Basin~~ in the Upper Nile Basin~~. We reveal substantial variability in current GRACE products to quantify the reduction of ΔTWS in Lake Victoria that ranges from ~~68~~ 80 km^3 (JPL-Mascons~~GRGS~~) to

~~50~~ 69 km^3 and ~~26~~31 km^3 for ~~JPL-Mascons~~ GRGS and *GRCTellus*, respectively. Representation of the phase in ~~Δ~~TWS in the Upper Nile Basin by GRACE products varies but is generally robust with GRGS, JPL-Mascons and *GRCTellus* (ensemble mean of CSR, JPL and GFZ time-series data) explaining 90~~1~~ %, 84~~5~~ %, and 75~~7~~ % of the variance, respectively, in 'in-situ' or 'bottom-up' ΔTWS in LVB. Resolution of changes in groundwater storage (ΔGWS) from GRACE ΔTWS is greatly constrained by both uncertainty in ~~modelled~~ changes in soil-moisture storage (ΔSMS) modelled by GLDAS LSMs (CLM, NOAH, VIC) and the low annual amplitudes in ΔGWS (e.g., ~~3.5~~1.8 to ~~4.4~~4.9 cm) observed in deeply weathered crystalline rocks underlying the Upper Nile Basin. Our study highlights the substantial uncertainty in the amplitude of ΔTWS that can result from different data-processing strategies in commonly used, gridded GRACE products.

Keywords: GRACE products; terrestrial water storage; groundwater; hard-rock aquifers; Lake Victoria; Lake Kyoga; Sub-Saharan Africa

1. Introduction

Satellite measurements under the Gravity Recovery and Climate Experiment (GRACE) mission have, since March 2002 (Tapley et al., 2004), enabled remote monitoring of large-scale (i.e., GRACE footprint: ~200 000 km^2), spatio-temporal changes in total terrestrial water storage (ΔTWS) at 10-day to monthly timescales (Longuevergne et al., 2013; Humphrey et al., 2016). Over the last 15 years, studies in basins around the world (Rodell and Famiglietti, 2001; Strassberg et al., 2007; Leblanc et al., 2009; Chen et al., 2010; Longuevergne et al., 2010; Frappart et al., 2011; Jacob et al., 2012; Shamsudduha et al., 2012; Arendt et al., 2013; Kusche et al., 2016) ~~show~~ have demonstrated that GRACE satellites trace natural (e.g., drought, floods, glaciers and ice melting, sea-level rise) and anthropogenic (e.g., abstraction-driven groundwater depletion) influences on ΔTWS. GRACE-derived TWS provides vertically-integrated water storage changes in all water-bearing layers (Wahr et al., 2004; Strassberg et al., 2007; Ramillien et al., 2008) that include (Eq. 1) surface water storage in rivers, lakes, and wetlands (ΔSWS), soil moisture storage (ΔSMS), ice and snow water storage (ΔISS), and groundwater storage (ΔGWS). GRACE measurements have over the last decade become an important hydrological tool for quantifying basin-scale ΔTWS (Güntner, 2008; Xie et al., 2012; Hu and Jiao, 2015) and are increasingly being used to assess spatio-temporal changes in specific water stores (Famiglietti et al., 2011; Shamsudduha et al., 2012; Jiang et al., 2014; Castellazzi et al., 2016; Long et al., 2016; Nanteza et al., 2016) where time-series records of other individual freshwater stores are available (Eq. 1).

$$\Delta TWS_t = \Delta GWS_t + \Delta ISS_t + \Delta SWS_t + \Delta SMS_t \qquad (1)$$

GRACE-derived ΔTWS derive from monthly gravitational fields which can be represented as spherical harmonic coefficients that are noisy as depicted in north-south elongated linear features or "stripes" on monthly global gravity maps (Swenson and Wahr, 2006; Wang et al., 2016). Post-processing of GRACE SH data is therefore required. The most popular GRACE products are NASA's *GRCTellus* land gravity solutions (i.e., spherical harmonics based CSR, JPL and GFZ), which require scaling factors to recover spatially smoothed TWS signals (Swenson and Wahr, 2006; Landerer and Swenson, 2012). Additionally, NASA's new monthly gridded GRACE product, Mass Concentration blocks (i.e., Mascons), estimate terrestrial mass changes directly from inter-satellite acceleration measurements and can be used without further post-processing (Rowlands et al., 2010; Watkins et al., 2015). GRGS GRACE are also spherical harmonic-based products available at a 10-day timestep and can also be used directly since gravity fields are stabilised during the processing of GRACE satellite data (Lemoine et al., 2007; Bruinsma et al., 2010).

Restoration of the amplitude of *GRCTellus* TWS data, dampened by spatial Gaussian filtering with a large smoothing radius (e.g., 300 to 500 km), is commonly achieved using scaling factors that derive from a priori model of freshwater stores, usually a global-scale Land-Surface Model or LSM (Long et al., 2015). However, signal-restoration methods are emerging that do not require hydrological model or LSM (Vishwakarma et al., 2016). Substantial uncertainty nevertheless persists in the magnitude of applied scaling factors (e.g., *GRCTellus*) and corrections (Long et al., 2015). In situ observations provide a valuable and necessary constraint to the scaling of TWS signals over a particular study area as no consistent basis for ground-truthing these factors exists.

The disaggregation of GRACE-derived ΔTWS anomalies into individual water stores (Eq. 1) is commonly constrained by the limited availability of observations of terrestrial freshwater stores (i.e., ΔSWS, ΔSMS, ΔGWS, ΔISS). Indeed, a major source of uncertainty in the attribution of GRACE ΔTWS derives from the continued reliance on modelled ΔSMS derived from LSMs (i.e., CLM, NOAH, VIC, MOSAIC) under the Global Land Data Assimilation System or GLDAS (Rodell et al., 2004) and remote-sensing products (Shamsudduha et al., 2012; Khandu et al., 2016). Further, analyses of GRACE-derived ΔGWS often assume ΔSWS is limited (Kim et al., 2009) yet studies in the humid tropics and engineered systems challenge this assumption showing that it can overestimate ΔGWS (Shamsudduha et al., 2012; Longuevergne et al., 2013). Robust estimates of ΔGWS from GRACE gravity signals have, to date, been developed in locations where ΔSWS is well constrained by in situ observations and groundwater is used intensively for irrigation so that

ΔGWS comprises a significant (>10 %) proportion of ΔTWS (Leblanc et al., 2009; Famiglietti et al., 2011; Shamsudduha et al., 2012; Scanlon et al., 2015). In Sub-Saharan Africa ~~(SSA)~~, intensive groundwater withdrawals are restricted to a limited number of locations (e.g., irrigation schemes, cities) and constrained by low-storage, low-transmissivity aquifers in the deeply weathered crystalline rocks that underlie ~40 % of this region (MacDonald et al., 2012) including the Upper Nile Basin (Fig. 1). Consequently, the ability of low-resolution GRACE gravity signals to trace ΔGWS in these hard-rock environments is unclear. A recent study (Nanteza et al., 2016) applies NASA's *GRCTellus* (CSR GRACE) data over large basin areas (>300 000 km^2) of East Africa and argues that ΔGWS can be estimated with sufficient reliability to characterise regional groundwater systems after accounting for ΔSWS by satellite altimetry and ΔSMS data from the GLDAS LSM ensemble (Rodell et al., 2004).

Here, we exploit a large-scale reduction and recovery in surface water storage that was recorded within Lake Victoria (Fig. 1), the world's second largest lake by surface area (67 220 km^2) (UNEP, 2013) and eighth largest by volume (2 760 km^3) (Awange et al., 2008). This well-constrained reduction in ΔSWS comprises a decline in lake level of 1.2 m between May 2004 and February 2006, equivalent to a lake-water volume (ΔSWS) loss of 81 km^3 that resulted, in part, from excessive dam releases (Fig. 2). We test the ability of current GRACE products to represent the amplitude and phase of this voluminous and well-constrained change in freshwater storage. Our analysis focuses on both the Lake Victoria Basin (hereafter LVB) (256 100 km^2) and Lake Kyoga Basin (hereafter LKB) (79 270 km^2) (Fig. 1). Applying in situ observations of ΔSWS and ΔGWS combined with simulated ΔSMS by the GLDAS LSMs, we assess: (1) the ability of current gridded GRACE products (i.e., *GRCTellus*, JPL-Mascons, GRGS GRACE) to measure a well constrained ΔTWS in the Upper Nile Basin from 2003 to 2012 focusing on the unintended experiment within the LVB from ~~2004~~ 2003 to 2006; and (2) the sensitivity of a disaggregated GRACE ΔTWS signals to trace ΔGWS in a deeply weathered crystalline rock aquifer systems underlying the Upper Nile Basin.

2. The Upper Nile Basin

2.1 Hydroclimatology

The Upper Nile Basin, the headwater area of the ~3 400 000 km^2 Nile Basin (Awange et al., 2014), includes both the Lake Victoria Basin (LVB) and Lake Kyoga Basin (LKB). Mean annual rainfall over the entire basin varies from 650 to 2900 mm (TRMM monthly rainfall; 2003–2012) with an average of 1300 mm ~~(σ=~~and standard deviation of 354 mm~~)~~ (Fig. 3). Mean annual gauged rainfall at different stations, Jinja, Bugondo and Entebbe measured is 1195, 1004 and 1541 mm, respectively (Owor et al., 2011). Rainfall over Lake Victoria is typically 25–30 % greater than that measured in the surrounding catchment (Fig. 3), which is partially explained by the nocturnal 'lake breeze' effect (Yin and Nicholson, 1998; Nicholson et al., 2000; Owor et al., 2011).

Estimates of mean annual evaporation from the surface of Lake Victoria vary from 1260 mm (UNEP, 2013) to 1566 mm (Hoogeveen et al., 2015) whereas mean annual evaporation from the surface of Lake Kyoga is estimated to vary from 1205 mm (Brown and Sutcliffe, 2013) to 1660 mm (Hoogeveen et al., 2015). Evapotranspirative fluxes from the surrounding swamps in Lake Kyoga are estimated to be much higher and approximately 2230 mm yr^{-1} (Brown and Sutcliffe, 2013).

Annual rainfall is predominantly bimodal in distribution (Fig. 4) with two distinct rainy seasons driven by the movement of the Intertropical Convergence Zone (ITCZ) (Awange et al., 2013). Long rains (March to May) and short rains (September to November) account for approximately 40% and 25% of annual rainfall respectively (Basalirwa, 1995; Indeje et al., 2000). The latter rainfalls are particularly influenced by El-Niño Southern Oscillation (ENSO) and Indian Ocean Dipole (IOD). GRACE-derived ΔTWS within the LVB shows a statistical association (R^2) of 0.56 with ENSO and 0.48 with IOD (Awange et al., 2014).

2.2 Lakes Victoria and Kyoga

Located between 31°39' E and 34°53' E longitudes, and 0°20' N and 3°00' S latitudes, Lake Victoria (Fig. 1) is located in Tanzania, Uganda and Kenya where each accounts for 51 %, 43 % and 6 % of lake surface area respectively (Kizza et al., 2012). Lake Victoria is relatively shallow with a mean depth of ~40 m and a maximum depth of 84 m (UNEP, 2013) akin to many shallow, open surface-water bodies as well as permanent and seasonal wetlands occupying low relief plateau across the Great Lakes Region of Africa (Owor et al., 2011). Moreover, the western and northwestern lake bathymetry is characterised by even shallower depths of between 4 and 7 m (Owor, 2010). Hydrologically, lake input is dominated by direct rainfall (84 % of total input); the remainder derives primarily from river inflows as direct groundwater inflow (<1 %) is negligible (Owor et al., 2011). Approximately 25 major rivers flow into Lake Victoria with a total catchment area of ~194 000 km^2; the largest tributary, River Kagera, contributes ~30 % of total river inflows (Sene and Plinston, 1994). Lake Victoria outflow to Lake Kyoga occurs at Jinja (Fig. 1).

Lake Kyoga (Fig. 1), located between 32°10' E and 34°20' E longitudes, and 1°00' N and 2°00' N latitudes, has a mean area of 1 720 km^2 with an estimated mean volume of 12 km^3 (Owor, 2010; UNEP, 2013). According to the recent global *HydroSHEDS* (Hydrological data and maps based on shuttle elevation derivatives at multiple ~~Scales~~scales) database, the Lake

Kyoga has a total surface area of 2 729 km^2 (Lehner et al., 2008). Lake Kyoga comprises lake-zone and flow-through conduit areas. The lake zone in Lake Kyoga is very shallow with a mean depth of 3.5 to 4.5 m (Owor, 2010). Lake Kyoga has a through-flow channel (mean depth 7 to 9 m) where the main Victoria Nile River flows (Owor, 2010) and acts as a linear reservoir with the annual water balance predominantly governed by the discharge of the Victoria Nile from Lake Victoria. Lake Kyoga has a through-flow channel (mean depth 7–9 m) where the main Victoria Nile River flows (Owor, 2010). Whilst numerous rivers flow into Lake Kyoga (e.g. Rivers Mpologoma, Awoja, Omunyal, Abalang, Olweny, Sezibwa and Enget) (Owor, 2010), the majority contributes a fraction of their former volume upon reaching the lake (Krishnamurthy and Ibrahim, 2013) due, in part, to evapotranspirative losses from fringe swamp areas (4 510 km^2) surrounding the lake (UNEP, 2013).

2.3 Hydrogeological setting

The Upper Nile Basin is underlain primarily by deeply weathered crystalline rock aquifer systems that have evolved through long-term, tectonically-driven cycles of deep weathering and erosion (Taylor and Howard, 2000). Groundwater occurs within unconsolidated regoliths or 'saprolite' and, below this, in fractured bedrock, known as 'saprock'. Bulk transmissivities of the saprolite and saprock aquifers are generally low (1 to 20 m^2 d^{-1}) (Taylor and Howard, 2000; Owor, 2010) and field estimates of the specific yield of the saprolite, the primary source of groundwater storage in these aquifer systems, are 2 % based on pumping-tests with tracers (Taylor et al., 2010) and magnetic resonance sounding experiments (Vouillamoz et al., 2014). Borehole yields are highly variable but generally low (0.5 to 20 m^3 h^{-1}) yet are of critical importance to the provision of safe drinking water.

2.4 An observed reduction in TWS in the LVB

In 1954, the construction of the Nalubaale Dam (formerly Owen Falls Dam) at the outlet of Lake Victoria at Jinja transformed the lake into a controlled reservoir (Sene and Plinston, 1994). Operated as a run-of-river hydroelectric project to mimic pre-dam outflows, the 'Agreed Curve' between Uganda and Egypt dictated dam releases that were controlled on a 10-day basis and generally adhered to, with compensatory discharge releases to minimise any departures, until the construction of the Kiira dam at Jinja in 2002 (Sene and Plinston, 1994; Owor et al., 2011).

The combined discharge of the Nalubaale and Kiira Dams enabled total dam releases (Fig. 2) to substantially exceed the Agreed Curve (Sutcliffe and Petersen, 2007) and between May 2004 and February 2006 the lake level dropped by 1.2 m (equivalent ΔSWS loss of 81 km^3) (Owor et al., 2011). Mean annual releases were 1387 $m^3 s^{-1}$ (+162 % of Agreed Curve) in 2004 and 1114 $m^3 s^{-1}$ (+148 % of Agreed Curve) in 2005. Sharp reductions in dam releases in 2006 helped to arrest and reverse the lake-level decline with lake levels stabilising by early 2007.

3. Data and Methods

3.1 Datasets

We use publicly available time-series records of: (1) GRACE TWS solutions from a number of data--processing strategies and dissemination centres including NASA's *GRCTellus* land solutions ([RL05 for CSR, GFZ (version DSTvSCS1409), RL05.1 for JPL (version DSTvSCS1411),) and JPL-Mascons solution (version RL05M_1.MSCNv01))], andas well as the French National Centre for Space Studies (CNES) GRGS solution (version GRGS RL03-v1); (2) NASA's Global Land Data Assimilation System (GLDAS) simulated soil moisture data from 3 global land surface models (LSMs) (CLM, NOAH, VIC); and (3) monthly precipitation data from NASA's Tropical Rainfall Measuring Mission (TRMM) satellite mission. We also employ in-situ observations of lake levels and groundwater levels from a network of river gauges and monitoring boreholes~~wells~~ operated by the Ministry of Water and Environment in Entebbe (Uganda). Datasets are briefly described ~~briefly~~ below.

3.1.1 Delineation of basin study areas

Delineation of the Lake Victoria Basin (LVB) and Lake Kyoga Basin (LKB) was conducted in Geographic Information System (GIS) environment under ArcGIS (v.10.3.1) environment using the 'Hydrological Basins in Africa' datasets derived from *HydroSHEDS* database (available at http://www.hydrosheds.org/) (Lehner et al., 2006, 2008). Regional water bodies including Lakes Victoria and Kyoga (Fig. 1) were spatially defined by the Inland Water dataset available globally at country scale from DIVA-GIS (Hijmans et al., 2012). Computed areas of the basins and lake surface areas are summarised in Table 1 along with previously estimated figures from other studies.

3.1.2 GRACE-derived terrestrial water storage (TWS)

Twin GRACE satellites provide monthly gravity variations interpretable as ΔTWS (Tapley et al., 2004) with an accuracy of ~1.5 cm (Equivalent Water Thickness or Depth) when spatially averaged (Wahr et al., 2006). In this study, we apply 5 different monthly GRACE solutions for the period of January 2003 to December 2012: post-processed, gridded (1° × 1°) GRACE-TWS time-series records from 3 *GRCTellus* land solutions from CSR~~S~~, JPL and GFZ processing centres (available at http://grace.jpl.nasa.gov/data) (Swenson and Wahr, 2006; Landerer and Swenson, 2012), JPL-Mascons (Watkins et al., 2015; Wiese et al., 2015), and GRGS GRACE products (CNES/GRGS release RL03-v1) (Biancale et al., 2006).

GRCTellus land ~~datasets~~ solutions are post-processed from two versions, RL05 and RL05.1 of spherical harmonics released by the University of Texas at Austin Centre for Space Research (CSR) and the German Research Centre for Geosciences Potsdam (GFZ), and the NASA's Jet Propulsion Laboratory (JPL) respectively. *GRCTellus* gridded datasets are available at monthly timestep at a spatial resolution of 1° × 1° ~~grids~~ (~111 km at equator) though the actual spatial resolution of GRACE footprint is ~450 km or ~200,000 km^2 (Scanlon et al., 2012). Post-processing of *GRCTellus* GRACE datasets primarily involve (i) removal of atmospheric pressure or mass changes based on the European Centre for Medium-Range Weather Forecasts (ECMWF) model; (ii) a glacial isostatic adjustment (GIA) correction based on a viscoelastic 3-D model of the Earth (A et al., 2013); and (iii) an application a destriping filter plus a 300-km Gaussian to minimise the effect of correlated errors (i.e., destriping) manifested by N-S elongated stripes in GRACE monthly maps. However, the use of a large spatial filter and truncation of spherical harmonics leads to energy removal so scaling coefficients or factors are applied to the *GRCTellus* GRACE - derived TWS data in order to restore attenuated signals (Landerer and Swenson, 2012). Dimensionless scaling factors are ~~also~~ provided as 1° × 1° bins (see supplementary Fig.~~ure~~ S1) that derive from the Community Land Model (CLM4.0) (Landerer and Swenson, 2012).

~~*GRCTellus*~~ JPL-Mascons (version RL05M_1.MSCNv01) data processing also involves a glacial isostatic adjustment (GIA) correction based on a viscoelastic 3-D model of the Earth (A et al., 2013). JPL-Mascons applies no spatial filtering as JPL-RL05M directly relates inters-satellite range-rate data to mass concentration blocks or Mascons to estimate global monthly gravity fields in terms of equal area 3° × 3° mass concentration functions to minimise measurement errors. The use of Mascons and the special processing result in better signal-to-noise ratios of the mascon fields compared to the conventional spherical harmonic solutions (Watkins et al., 2015). For convenience, gridded Mascons fields are provided at a spatial sampling of 0.5° in both latitude and longitude (~56 km at the equator). As with *GRCTellus* GRACE datasets the neighbouring grid cells are not 'independent' of each other and cannot be interpreted individually at the 1° or 0.5° grid scale (Watkins et al., 2015). Similar to *GRCTellus* GRACE (CSR, JPL, GFZ) products, dimensionless scaling factors are provided as 0.5° × 0.5° bins (see supplementary Fig. S2) that also derive from the Community Land Model (CLM4.0) (Wiese et al., 2016). The gain factors or scaling coefficients are multiplicative factors that minimize the difference between the smoothed and unfiltered monthly ΔTWS variations from 'actual' land hydrology at a given geographical location (Wiese et al., 2016).

GRGS/CNES GRACE monthly products (version RL03-v1) are processed and made publicly available (http://grgs.obs-mip.fr/grace) by the French Government space agency, National Centre for Space Studies or Centre National d' Études Spatiales (CNES). The post-processing of GRGS data involves taking into account of gravitational variations such as Earth tides, ocean tides, and 3D gravitational potential of the atmosphere and ocean masses (Bruinsma et al., 2010). The remaining signals for time-varying gravity fields therefore represent changes in terrestrial hydrology including snow cover, baroclinic oceanic signals and effects of post-glacial rebound (Biancale et al., 2006; Lemoine et al., 2007). Further details on the Earth's mean gravity-field models can be found on the official website of GRGS/LAGEOS (http://grgs.obs-mip.fr/grace/).

GRACE satellites were launched in 2002 to map the variations in Earth's gravity field over its 5-year lifetime but both satellites are still in operation even after more than 14 years. However, active battery management since 2011 has led the GRACE satellites to be switched off every 5–6 months for 4–5 week durations in order to extend its total lifespan (Tapley et al., 2015). As a result, GRACE ΔTWS time-series data have some missing records that are linearly interpolated (Shamsudduha *et al.*, 2012). In this study, we derive ΔTWS time-series data as equivalent water depth (cm of H_2O) using the basin boundaries (GIS shapefiles) for masking the 1° × 1° grids.

3.1.3 Rainfall data

We apply Tropical Rainfall Measuring Mission (TRMM) (Huffman et al., 2007) monthly product (3B43 version 7) for the period of January 2003 to December 2012 at 0.25° × 0.25° spatial resolution and aggregate to 1° × 1° grids over LVB and LKB. General climatology of the Upper Nile Basin is represented by long-term (2003–2012) mean annual rainfall (Fig. 3) and seasonal rainfall pattern (Fig. 4). TRMM rainfall measurements show a good agreement with limited observational precipitation records (Awange et al., 2008; Awange et al., 2014).

3.1.4 Soil moisture storage (SMS)

NASA's Global Land Data Assimilation System (GLDAS) is an uncoupled land surface modelling system that drives multiple land surface models (GLDAS LSMs: CLM, NOAH, VIC and MOSAIC) globally at high spatial and temporal resolutions (3-hourly to monthly at 0.25° × 0.25° grid resolution) and produces model results in near-real time (Rodell et al., 2004). These LSMs provide a number of output variables which include soil moisture storage (SMS). Similar to the approach applied in the analysis of GRACE-derived ΔTWS analysis in the Bengal Basin (Shamsudduha et al., 2012), we apply simulated monthly ΔSMS records at a spatial resolution of 1° × 1° from 3 GLDAS LSMs: the Community Land Model (CLM, version 2) (Dai et al., 2003), NOAH (version 2.7.1) (Ek et al., 2003) and the Variable Infiltration Capacity (VIC) model (version 2.7.1) (Liang et al., 2003). The respective depths of modelled soil profiles are 3.4 m, 2.0 m, and 1.9 m in CLM (10 vertical layers), NOAH (4 vertical layers), and VIC (version 1.0) (3 vertical layers). Because of the absence of in situ soil moisture data in the study areas we apply an ensemble mean of the aforementioned 3 LSMs-derived simulated ΔSMS time-series records (see Figs. 5 and 6) in order to disaggregate GRACE ΔTWS signals in LVB and LKB.

3.1.5 Surface water storage (SWS)

Daily time-series of ΔSWS are computed from in situ (gauged) lake-level observations at Jinja for Lake Victoria and Bugondo for Lake Kyoga (Figs. 1 and 2) compiled by the Ugandan Ministry of Water and Environment (Directorate of Water Resources Management). Mean monthly anomalies for the period of January 2003 – December 2012 were computed as an equivalent water depth using Eq. (2). Missing data in the time series (2003–2012) records are linearly interpolated. For instance, in case of monthly ΔSWS derived from Lake Kyoga water levels, there is one missing record (December 2005).

$$\Delta SWS = \Delta Lake\ Level \times \left(\frac{Lake\ Area}{Total\ Basin\ Area}\right) \quad (2)$$

3.1.6 Groundwater storage (GWS) from borehole observations

Time series of ΔGWS are constructed from in situ piezometric records from 6 monitoring wells located in LVB and LKB where near-continuous, daily observations exist from January 2003 to December 2012 and have been compiled by the Ugandan Ministry of Water and Environment (Directorate of Water Resources Management) (Owor et al., 2009; Owor et al., 2011). Monitoring boreholes were installed into weathered, crystalline rock aquifers that underlie much of LVB and LKB, and are remote from local abstraction. As such, they represent variations in groundwater storage influenced primarily by climate variability. Mean monthly anomalies of ΔGWS, ~~normalised~~ standardised to mean records from January 2003 to December 2012, were derived from near-continuous, daily observations at Entebbe, Rakai and Nkokonjeru for LVB and at Apac, Pallisa and Soroti for LKB (Fig. 1; Table 2; see supplementary Fig. S3). In the Lake Kyoga Basin, piezometric records from 3 sites show consistency in the seasonality and amplitude of groundwater storage changes plotted as monthly groundwater-level anomalies relative to the mean for the period from January 2003 to December 2012. In the Lake Victoria Basin, groundwater-level records from 2 sites (Entebbe, Nkokonjeru) are similar in their phase and amplitude, and are influenced by changes in the level of Lake Victoria as demonstrated by Owor et al. (2011). The groundwater-level record from Rakai represents local semi-arid conditions that exist within catchment areas (e.g., River Ruizi) draining to the western shore of Lake Victoria in Uganda. Although there are differences in the phase of groundwater-level fluctuations between the semi-arid site at Rakai and both Entebbe and Nkokonjeru (as well as the 3 sites in the Lake Kyoga Basin), annual amplitudes are similar.

~~These~~ The groundwater-level time series data are a sub-set of the total number of available monitoring-well records in the LVB and LKB and selected on the basis of (i) the completeness and quality of the records from 2003 to 2012, and (ii) ~~following a~~ rigorous review of groundwater-level records conducted at a dedicated workshop at the Ministry of Water & Environment in January 2013. These records represent shallow groundwater-level observations within the saprolite that is dynamically connected to surface waters (Owor et al. 2011). ~~The limited spatial coverage in quality-controlled piezometry, especially for the LVB, represents an important limitation in our analysis.~~ Long time-series records of groundwater levels over the period from 2003 to 2012 from western Kenya, northern Tanzania, Rwanda and Burundi have not been identified despite intensive investigations carried out by *The*

Chronicles Consortium[1]. The partial ~~limited~~ spatial coverage in quality-controlled piezometry, especially for the LVB, represents an important limitation in our analysis.

Mean monthly anomalies were translated into an equivalent water depth (Eq. 3) by applying a range of specific yield (S_y) values (1–6 % with an average of 3 %) although estimates of S_y in hard-rock environments are observed to vary from < 2% to 8 % (Taylor et al., 2010; Taylor et al., 2013; Vouillamoz et al., 2014) using Eq. (3). Missing data in the time series were linearly interpolated. In case of monthly ΔGWS that derived from borehole (n=6) observations, missing records range from 1–9 months (120 months in 2003–2012) with three boreholes (Soroti, Rakai and Nkonkonjero) with time-series records ending in June–July 2010.

$$\Delta GWS = \Delta h * S_y * \left(\frac{Land\ Area}{Total\ Basin\ Area}\right) \quad (3)$$

3.2 Methodologies

3.2.1 GRACE ΔTWS estimation

First, the 1° × 1° gridded monthly anomalies of GRACE-derived ΔTWS and GLDAS LSMs derived ΔSMS are masked over the area of LVB and LKB. GRACE ΔTWS along with GLDAS ΔSMS are extracted for the marked 1° × 1° grid cells for LVB and LKB and the grid values are spatially aggregated to form time-series of monthly anomalies ΔTWS and ΔSMS.

~~Firstly,~~ *GRCTellus* GRACE ΔTWS gridded data are ~~generally~~ scaled ~~up~~ using dimensionless, gridded scaling factors. Several GRACE studies (Rodell et al., 2009; Sun et al., 2010; Shamsudduha et al., 2012) have applied scaling factors in three different ways: (1) single scaling factor based on regionally averaged time series, (2) spatially distributed or gridded

[1] The Chronicles Consortium: https://www.un-igrac.org/special-project/chronicles-consortium

scaling factors based on time-series at each grid point, and (3) gridded-gain factors estimated as a function time or of temporal frequency (Landerer and Swenson, 2012; Long et al., 2015).~~that are provided separately and are independent of ΔTWS grids (Landerer and Swenson, 2012)~~. In this study, we apply spatially-distributed scaling approach (method 2 above) to generate basin-averaged ΔTWS time-series records for *GRCTellus* (CSR, JPL, GFZ) products. Scaling factors provided at 1° × 1° grids are applied to each corresponding GRACE ΔTWS grids for NASA's *GRCTellus* products in order to restore attenuated signals during the post-processing (Landerer and Swenson, 2012) using Eq. (4). ~~A number of GRACE studies (Rodell et al., 2009; Sun et al., 2010; Shamsudduha et al., 2012) around the world have applied scaling factors in three different ways: (1) single scaling factor based on regionally averaged time series, (2) spatially distributed or gridded scaling factors based on time-series at each grid point, and (3) gridded-gain factors estimated as a function time or of temporal frequency (Landerer and Swenson, 2012; Long et al., 2015). In this study, we apply the gridded scaling factors approach to adjust ΔTWS time-series records.~~ Similarly, provided scaling factors are applied to JPL-Mascons ΔTWS time-series data but at 0.5° × 0.5° grid resolution. No scaling factors were applied to GRGS GRACE ΔTWS as the monthly gravity solutions have already been stabilised during their generation process.

$$g^1(x, y, t) = g(x, y, t) \times s(x, y) \qquad (4)$$

Here, $g^1(x, y, t)$ represents each un-scaled grid where x represents longitude, y represents latitude, and t represents time (month), and $s(x, y)$ is the corresponding scaling factor.

~~Second, scaling coefficients or factors provided at 1° × 1° grids are applied to each corresponding GRACE ΔTWS grids for NASA's *GRCTellus* products only in order to restore attenuated signals during the post-processing using Eq. (4) (Landerer and Swenson, 2012).~~

For the 3 *GRCTellus* gridded products (i.e., CSR, GFZ, and JPL solutions), ~~W~~we apply an ensemble mean of scaled GRACE ΔTWS ~~of 3 *GRCTellus* gridded products (i.e., CSR, GFZ, and JPL solutions)~~ as our exploratory analyses reveal that ~~the~~ ΔTWS time-series records over the Lake Victoria Basin are highly correlated (r >0.95, p-value <0.001) to each other. Additionally, ~~and the~~ small (ranges from 1.3 to 1.9 cm) Root Mean Square Error (RMSE) ~~is very small (ranges from 1.3 to 1.9 cm)~~ among the ~~time-series records~~GRACE ΔTWS datasets suggests substantial similarities in phase and amplitude.

~~$g^1(x, y, t) = g(x, y, t) \times s(x, y)$ (4)~~

~~Here, $g^1(x, y, t)$ represents each un-scaled grid where x represents longitude, y represents latitude, and t represents time (month), and $s(x, y)$ is the corresponding scaling factor.~~

3.2.2 Estimation of ΔGWS from GRACE

Estimation of groundwater storage changes (ΔGWS) from GRACE measurements is conducted using Eq. (5) in which ΔTWS_t is derived from gridded GRACE products (spatially scaled ΔTWS for *GRCTellus* and JPL-Mascons but unscaled ΔTWS for GRGS), ΔSMS_t is an ensemble mean of 3 GLDAS LSMs (CLM, NOAH, VIC), and ΔSWS_t is area-weighted, in-situ surface water storage estimated from lake-level records using Eq. (2).

$$\Delta GWS_t = \Delta TWS_t - (\Delta SWS_t + \Delta SMS_t) \quad (5)$$

3.2.3 Reconciliation of GRACE ΔTWS disaggregation

Reconciling GRACE-derived TWS with ground-based observations is limited by the paucity of in situ observations of SMS, SWS and GWS in many environments. In addition, direct comparisons between in situ observations of ΔSMS, ΔSWS and ΔGWS and gridded GRACE ΔTWS anomalies are complicated by substantial differences in spatial scales, which need to be considered prior to analysis (Becker et al., 2010). For example, individual groundwater-level monitoring boreholes may represent, depending on borehole depth, a sensing area of several 10s of km^2 (Burgess et al., 2017), whereas the typical GRACE footprint is ~200 000 km^2. The disaggregation of GRACE ΔTWS into individual water store can also propagate errors to disaggregated components. Here, we construct 'in situ' or 'bottom-up'~~in situ~~ ΔTWS (i.e., combined signals of ΔSMS, ΔSWS and ΔGWS) for the Lake Victoria Basin and attempt to reconcile with GRACE-derived ΔTWS. One feature of GRACE ΔTWS among the 3 solutions we apply in this study is the considerable variation in annual amplitudes that exist over the period of 2003 to 2012.

In addition, for the *GRCTellus* products, we conduct unconventional scaling experiments, outlined below~~, to both the ensemble GRACE ΔTWS and in situ ΔSWS~~ in an attempt to reconcile satellite and in situ measures and to shed light on the uncertainty in ΔTWS amplitudes of the *GRCTellus* GRACE products. The ΔTWS signals in CSR, JPL and GFZ products is greatly attenuated due to spatial smoothing and the amplitude is substantially smaller compared to JPL-Mascons and GRGS products. In the first scaling experiment, we apply an additional, basin-averaged, multiplicative scaling factor to ΔTWS ranging from 1.1 to 2.0 and employ RMSE to assess their relative performance. With reference to *GRCTellus* GRACE ΔTWS and bottom-up ΔTWS relationship, the scaling factor producing the lowest

RMSE between the two time series is employed. Secondly, it is observed that in the LVB, ΔSWS is the largest contributor, representing ~50% variance in the in-situ or bottom-up ΔTWS time-series signal. GRACE ΔTWS analyses commonly apply the same scaling factor as ΔTWS to all other individual components (Landerer and Swenson, 2012). Therefore, under the scaling experiment, we apply to in-situ ΔSWS spatially-averaged scaling factors representative of (i) Lake Victoria and its surrounding grid cells (experiment 1: s=0.71; range 0.02–1.5), and (ii) the open-water surface of Lake Victoria without surrounding grid cells (experiment 2: s=0.11; range 0.02–0.30). Furthermore, we find that the amplitude of monthly anomalies of ΔSWS+ΔSMS combined substantially exceed ΔTWS (see supplementary Fig. S4), particularly for the *GRCTellus* GRACE ΔTWS signal that is greatly smoothed due to filtering. This discrepancy is pronounced over the period of 2003–2006, and when applied to estimate GRACE-derived ΔGWS, produces steep, rising trends in the estimated ΔGWS (i.e., GRACE ΔTWS – (ΔSWS+ΔSMS)) whereas borehole observations of groundwater levels show declining trend and of much lower amplitude over the same period.

4. Results

Monthly time-series records (January 2003 to December 2012) are presented in Figures 5 and 6 respectively for Lake Victoria Basin (LVB) and Lake Kyoga Basin (LKB) of (a) GRACE ΔTWS from *GRCTellus* GRACE ΔTWS (ensemble mean of CSR, GFZ, and JPL solutions), GRGS and JPL-Mascons, (b) GLDAS land surface models (LSMs) derived ΔSMS (ensemble mean of 3 LSMs: NOAH, CLM, VIC), (c) in situ ΔSWS from lake levels records, and (d) in situ ΔGWS borehole observations. Monthly rainfall derived from TRMM satellite observations over the same period are shown on the bottom panel (d). Time-series records of all ΔTWS components and rainfall are aggregated for LVB to represent the average seasonal (monthly) pattern of each signal (Fig. 4) that shows an obvious lag (~1 month) between peak rainfall (March–April) and ΔTWS and its individual components.

Mean annual (2003–2012) amplitudes of various GRACE-derived ΔTWS signals, ~~in situ~~bottom-up ΔTWS, ensemble mean of simulated ΔSMS, in situ ΔSWS and ΔGWS time-series records (Figs. 5 and 6) are ~~calculated~~ presented (see supplementary Table S1) for both LVB and LKB. Mean annual amplitude of GRACE ΔTWS ranges from 11~~.7~~ to 21~~0.6~~ cm among *GRCTellus*, GRGS and JPL-~~MASCON~~ Mascons GRACE products in LVB, and from 8.4 to 16.4 respectively in LKB. Mean annual amplitude of in situ ΔSWS is much greater (14.8 cm) in LVB than in LKB (3.8 cm). GLDAS LSMs derived ensemble mean ΔSMS amplitude in LVB is 7.9 cm and 7.3 cm in LKB. The standard deviation in ΔSMS varies substantially in LVB (1.2 cm, 4.2 cm, and 2.9 cm) LKB (1.3 cm, 4.7 cm, and 4.0 cm) for CLM, NOAH, and VIC models respectively. Mean annual amplitude of in situ ΔGWS ranges from 4.4 cm (LVB) to 3.5 cm (LKB).

Time-series correlation (Pearson) analysis over various periods of interests (decadal: 2003–2012; well-constrained SWS reduction or a period of unintended experiment: 2003–2006; controlled dam operation: 2007–2012) reveals that GRACE-derived ΔTWS signals are strongly correlated in both LVB and LKB (see supplementary Figs. S5~~2~~–S10~~7~~). For example, in LVB, in situ ΔSWS shows a statistically significant (*p*-value <0.001) strong correlation (*r*=0.77–0.92) with all GRACE- ΔTWS time-series (2003–2012) records. Similarly, simulated ~~ΔSWS~~ ΔSMS shows statistically significant (*p*-value <0.001) strong correlation (*r*=0.70~~2~~–0.78) with ΔTWS time-series records. In contrast, in situ ΔGWS shows statistically significant (*p*-value <0.001) but moderate correlation (*r*=0.~~4~~63–0.~~5~~69) with ΔTWS time-series records. Correlation among the variables shows similar statistically significant (p-value <0.001) but wide-ranging associations for the periods of unintended experiment (2003–2006) and controlled dam operation (2007–2012). In LKB, however, correlation among in situ ΔSWS and GRACE ΔTWS time-series records is statistically significant (p-value <0.0~~0~~5~~1~~) but poor in correlation strength (r=0.~~28~~22–0.34). In situ ΔGWS shows statistically significant (p-value <0.001) ~~moderate~~ strong correlation (r=0.~~40~~64–0.69~~47~~) with GRACE ΔTWS time-series records.

Time-series records of all 3 ΔTWS from 5 GRACE products ~~all 5 GRACE ΔTWS~~ and ~~in situ~~bottom-up ΔTWS time-series records in both LVB and LKB are shown in Figure 7 and results of temporal trends are summarised in Table 3. Statistically significant (p-value <0.05) declining trends (–4.1 to –11.0 cm yr^{-1} in LVB; –2.1 to –~~5~~4.6 cm yr^{-1} in LKB) are consistently observed during the period of ~~2004~~ 2003 to 2006. Trends are all positive in GRACE ΔTWS and ~~in situ~~bottom-up ΔTWS time-series records over the recent period of controlled dam operation (2007–2012) in both LVB and LKB. Therefore, the overall, decadal (2003–2012) trends are slightly rising (0.04 to ~~0~~1.00~~79~~ cm yr^{-1}) in LVB but nearly stable (–0.01 cm yr^{-1}) in *GRCTellus* ΔTWS and slightly declining (–0.56 cm yr^{-1}) ~~in situ~~bottom-up ΔTWS over LKB. In addition, short-term volumetric trends (~~2004~~2003–2006) in GRACE and ~~in situ~~bottom-up ΔTWS as well as simulated ΔSMS and in situ ΔSWS are declining whereas in situ ΔGWS and rainfall anomalies show slightly rising trends over the same period in LVB (see supplementary Figs. ~~S8~~S11–S12~~9~~). Similar trends are reported in various signals over LKB but magnitudes are much smaller compared to that of LVB, which is 3 times larger in size than LKB. Volumetric declines in ΔTWS in the LVB for the period ~~2004~~ 2003 to 2006 are: ~~75~~83 km^3 (~~in situ~~bottom-up), ~~68 km^3 (GRGS), 5~~80 km^3 (JPL-Mascons), 69 km^3 (GRGS) and ~~26~~31 km^3 (*GRCTellus* ensemble mean of C~~R~~SR, JPL and GFZ products).

Linear regression reveals that the association between GRACE-derived ΔTWS and ~~in situ~~bottom-up ΔTWS is stronger in LVB (R^2=0.~~77~~75–0.~~91~~90) than in LKB (R^2=0.~~49~~56–0.~~55~~62) (see supplementary Table S1). GRACE ΔTWS is unable to explain natural variability in ~~in situ~~bottom-up ΔTWS in LKB though this may be explained by the fact that SWS in Lake Kyoga is influenced by dam releases from LVB. Multiple linear regression and the ~~analysis~~ Analysis of Variance (ANOVA) reveal~~s~~ that the relative proportion of variability in ~~in situ~~bottom-up ΔTWS time-series record can be explained by ΔSWS (~~88.9~~92.6 %), ΔSMS (~~9.4~~6.5 %) and ΔGWS (~~1.9~~0.66 %) in LVB; and by ~~37.2~~47.9 %, ~~55.9~~48.5 % and ~~6.9~~3.6 % respectively in LKB. These results are indicative only as these percentages can be biased by the presence of strong correlation among variables and the order of these variables listed as predictors in the multiple linear regression models.

Disaggregation of ΔGWS from GRACE ΔTWS time-series record from each product has been carefully considered and estimated following Eq. (5). No further additional scaling factors, as described in the 'scaling experiment' section (see results of scaling experiment in supplementary Fig. 13) are applied in the final disaggregation of ΔGWS from GRACE ΔTWS signals. ~~In case of LVB, we apply a spatially-averaged multiplicative scaling factor (1.7) to *GRCTellus* GRACE-derived ΔTWS dataset to amplify the signal that is better reconciled with in situ ΔTWS (see supplementary Fig. S10). Additionally, for both *GRCTellus* and JPL-Mascons ΔTWS disaggregation to ΔGWS a scaled down signal of in situ ΔSWS is applied.~~ Results of Pearson correlation analysis of the time-series record (2003–2012) of in situ ΔGWS in LVB show ~~Time-series record (2003–2012) of in situ ΔGWS in LVB weakly correlates~~ statistically insignificant and poor correlation (r=0.11~~29~~, p-value ~~<0.001~~0.25) ~~with~~ to JPL-Mascons and an inverse correlation with both the ensemble

GRCTellus (r=–0.55, p-value <0.001) and ~~JPL-Mascons GRACE derived ΔGWS but shows no correlation with~~ GRGS (r=–0.27, p-value=0.003) GRACE-derived estimates of ΔGWS (Fig. 8). In contrast, ~~I~~in LKB, in situ ΔGWS time-series record shows ~~weaker and~~ statistically ~~in~~significant but weak correlations to JPL-Mascons (r=0.~~16–0.19,~~ 34, p-value <0.001~~8~~) ~~with JPL-Mascons~~ and GRGS (r=0.39, p-value <0.001) GRACE-derived ΔGWS but shows ~~no~~ an inverse correlation (r=–0.21, p-value=0.02) ~~with~~ to *GRCTellus* ΔGWS (see supplementary Fig. S14~~1~~). Furthermore, RMSE among various GRACE-derived estimates of ΔGWS and in situ ΔGWS ranges from ~~3.0~~ 7.2 cm (GRACE ensemble), 3.~~7~~ 8 cm (GRGS) to ~~6.4~~ 8.2 cm (JPL-Mascons) in LVB, and from 3.~~4~~ 2 (GRACE ensemble), 5.~~6~~ 3 cm (GRGS) to ~~6.8~~ 5.4 cm (JPL-Mascons) in LKB.

5. Discussion

We apply 5 different gridded GRACE products (*GRCTellus* – CSR, JPL and GFZ; GRGS and JPL-Mascons) to test ΔTWS signals for in the Lake Victoria Basin (LVB) comprising a large and accurately observed reduction (83~~75~~ km^3) in ΔTWS from ~~2004~~ 2003 to 2006. Our analysis reveals that all GRACE products capture this substantial reduction in terrestrial water mass but the magnitude of GRACE ΔTWS among GRACE products varies substantially. For example, *GRCTellus* underrepresents greatly (~~66~~63 %) the reduction of 83 km^3 in ~~in situ~~bottom-up ΔTWS whereas GRGS and JPL-Mascons GRACE products underrepresent~~s~~ ~~slightly (10~~this by -17 % and 4 % respectively~~)~~. Over a longer period (2003–2012) in the Upper Nile Basin, all GRACE products correlate well with ~~in situ~~bottom-up ΔTWS but, similar to the unintended experiment, variability in amplitude is considerable (Fig. 9). The average (2003–2012) annual amplitude of ΔTWS is substantially dampened (i.e., ~~86~~ 45 % less than ~~in situ~~bottom-up ΔTWS) in *GRCTellus* GRACE products relative to GRGS (~~6~~ 4 %) and JPL-Mascons (27 % more than bottom-up ΔTWS) products in the LVB.

The 'true' amplitude in *GRCTellus* ΔTWS signal is generally reduced during the post-processing of GRACE spherical harmonic fields, primarily due to spatial smoothing by a large-scale (e.g., 300 km) Gaussian filter and truncation of gravity fields at a higher (degree 60 = 300 km) spectral degree (Swenson and Wahr, 2006; Landerer and Swenson, 2012). Despite the application of scaling ~~coefficients~~ factors based on CLM v.4.0 to amplify *GRCTellus* ΔTWS amplitudes at individual grids, the basin-averaged (LVB) time-series record represents only ~~77~~ 75 % variability in ~~in-situ~~bottom-up ΔTWS. Scaling experiments conducted here reveal that *GRCTellus* ΔTWS requires an additional multiplicative factor of 1.7 in order to match ~~in-situ~~bottom-up ΔTWS with a minimum RMSE (5.8 cm). On the other hand, NASA's new gridded GRACE product, JPL-Mascons, that applies a priori constraint in space and time to derive monthly gravity fields and undergoes some degree of spatial smoothing (Watkins et al., 2015), represents nearly ~~85~~ 83 % variability in ~~in-situ~~bottom-up ΔTWS. In contrast, the GRGS GRACE product, ~~although applying~~which applies truncation at degree 80 (~250 km) ~~,~~ does not suffer from any large-scale spatial smoothing, and~~,~~ is able to represent well (~~92~~ 90 %) the variability in ~~in-situ~~bottom-up ΔTWS in the LVB.

A priori corrections of *GRCTellus* ensemble mean GRACE signals using a set of LSM-derived scaling factors (i.e., amplitude gain) can lead to substantial uncertainty in ΔTWS (Long et al., 2015). We show that the amplitude of simulated terrestrial water mass over the Upper Nile Basins varies substantially among various LSMs (see supplementary Fig. ~~S12~~S15). Most of these LSMs (GLDAS models: CLM, NOAH, VIC) do not include surface water or groundwater storage (Scanlon et al., 2012). Although CLM (v.4.0 and 4.5) includes a simple representation (i.e., shallow unconfined aquifer) of groundwater (Niu et al., 2007; Oleson et al., 2008), it does not consider recharge from irrigation return flows. In addition, many of these LSMs do not consider lakes and reservoirs and, most critically, LSMs are not reconciled with in situ observations. ~~As a result, methods of rescaling the amplitude of GRACE signals based on a priori information from LSMs contribute uncertainty to TWS signals.~~

The combined measurement and leakage errors, $\sqrt{(bias^2 + leak^2)}$ (Swenson and Wahr, 2006) for *GRCTellus* ΔTWS based on CLM4.0 model for LVB and LKB are 7.2 cm and 6.6 cm respectively. These values, however, do not represent mass leakage from the lake to the surrounding area within the basin itself. A sensitivity analysis of *GRCTellus* and GRGS signals reveal that signal leakage occurs from lake to its surrounding basin area as well as between basins. For instance, GRACE signal leakage into LKB from LVB, which is 3 times larger in area than LKB, is 3.4 times bigger for both GRCTellus GRACE and GRGS products. Furthermore, ~~for leakage from the lake into the basin area~~ the analysis shows that leakage from Lake Victoria to LVB for *GRCTellus* is substantially greater than GRGS product by a factor of ~2.6. In other words, 1 mm change in the level of Lake Victoria represents an equivalent change of 0.12 mm in ΔTWS in LVB for *GRCTellus* compared to 0.32 mm for GRGS. Consequently, changes in the amplitude of GRGS ΔTWS are much greater (~38 %) than *GRCTellus*. During the observed reduction in ΔTWS (~~75~~ 83 km^3) from ~~2004~~ 2003 to 2006, the computed ~~amplitude~~ volumetric reduction for GRGS is ~~68~~ found to be 69 km^3 whereas it is ~~26~~ 31 km^3 for *GRCTellus*.

Another source of uncertainty that contributes toward ΔTWS anomalies in GRACE analysis is the choice of simulated ΔSMS from various global-scale LSMs (e.g., Shamsudduha et al., 2012; Scanlon et al., 2015). For example, the mean annual (2003–2012) amplitudes in simulated ΔSMS in GLDAS LSMs (CLM, NOAH, VIC) vary substantially in LVB (3.5 cm, 10.2 cm, and 10.5 cm) and LKB (3.7 cm, 10.6 cm, and 7.7 cm) respectively. Due to an absence of a dedicated monitoring network for soil moisture in the Upper Nile Basin, this study like many other GRACE studies, is resigned to applying simulated ΔSMS from multiple LSMs arguing that the use of an ensemble mean minimises the error associated with ΔSMS (Rodell et al., 2009).

Computed contributions of ΔGWS to ΔTWS in the Upper Nile Basins are low (<10 %). GRACE-derived estimates of ΔGWS from all three products (*GRCTellus*, GRGS and JPL-Mascons) correlate very weakly with in situ ΔGWS in both LVB and LKB. One curious observation in LVB during the unintended experiment (~~2005~~2003–2006) is that in situ ΔGWS rises whereas in situ ΔSWS and simulated ΔSMS decline. The available evidence in groundwater-level records (e.g., Entebbe, Uganda) suggests that rainfall-generated groundwater recharge led to an increased in ΔGWS while dam releases exceeding the "Agreed Curve" continued to reduce ΔSWS (Owor et al., 2011).

Uncertainties in the estimation of GRACE-derived ΔGWS remain in: (i) accurate representation of the largest individual signal of in-situ ΔSWS in the disaggregation~~the choice of scaling factors applied to in situ ΔSWS associated with the disaggregation~~ of GRACE ΔTWS signal as it can limit the propagation of uncertainty in simulated ΔSMS~~of ΔTWS from JPL-Mascons and *GRCTellus* GRACE products~~, (ii) simulated ΔSMS by GLDAS land surface models, (iii) the very limited spatial coverage in piezometry to represent in situ ΔGWS, and (iv) applied S_y (3 % with range from 1 % to 6 %) to convert in situ groundwater levels to ΔGWS. The lack of any strong correlation in ~~GRGS~~ GRACE-derived ΔGWS and in situ ΔGWS time-series records indicates that the magnitude of uncertainty is larger than the overall variability in ΔGWS in low-storage, low-transmissivity weathered crystalline aquifers within the Upper Nile Basin.~~-~~ Furthermore, statistically significant but negative correlations in both LVB and LKB arise from a positive change in GRACE-derived ΔGWS when in situ ΔGWS is declining (e.g., 2003 to 2006 in LVB; 2008 to 2010 in LKB). This inconsistency suggests that the 'true' GRACE ΔTWS signal is weakened during processing and that the combined ΔSWS+ΔSMS signal is greater than ΔTWS, mathematically resulting to a positive estimate of ΔGWS. In contrast to the assertions of Nanteza et al. (2016) applying the *GRCTellus* CSR solution, we find that this uncertainty prevents robust resolution of ΔGWS from GRACE ΔTWS in these complex hydrogeological environments of East Africa. Despite substantial efforts to improve groundwater-level monitoring and to collate existing groundwater-level records across Africa, we recognise that understanding of in situ ΔGWS remains greatly constrained by limitations in current observational networks and records. Since present uncertainties and limitations identified in the Upper Nile Basin occur in many of the weathered hard-rock aquifer environments that underlie 40% of Sub-Saharan Africa (MacDonald et al., 2012), tracing of ΔGWS using GRACE in these areas is unlikely to be robust until these uncertainties and limitations are better constrained.

6. Conclusions

The analysis of a large, accurately recorded reduction of 1.2 m in the water level of Lake Victoria, equivalent to ΔSWS decline of 81 km^3~~in the volume of Lake Victoria (ΔSWS=81 km^3)~~ from 2004 to 2006 exposes substantial variability among commonly-used 5 gridded GRACE products (*GRCTellus* CSR, JPL, GFZ; GRGS; JPL-Mascons) to quantify the amplitude of changes in terrestrial water storage (ΔTWS). ~~For~~ Around this event, we estimate an overall decline in 'in situ' or 'bottom-up' ΔTWS (i.e., in situ ΔSWS and ΔGWS; simulated ΔSMS) over the Lake Victoria Basin (LVB) of ~~75~~ 83 km^3 from 2003 to 2006. This value compares favourably with JPL-Mascons GRACE ΔTWS ~~GRGS GRACE ΔTWS~~ (~~68~~ 80 km^3), is underrepresented by GRGS GRACE ΔTWS~~JPL-Mascons GRACE ΔTWS~~ (~~50~~ 69 km^3), and is substantially underrepresented by the ensemble mean of *GRCTellus* GRACE ΔTWS (~~26~~ 31 km^3). Attempts to better reconcile *GRCTellus* GRACE ΔTWS to ~~in situ~~bottom-up ΔTWS through scaling techniques are unable to represent adequately the observed amplitude in ΔTWS but highlight the uncertainty in the amplitude of gridded GRACE ΔTWS datasets generated by various processing strategies.

From 2003 to 2012, GRGS, JPL-Mascons and *GRCTellus* GRACE products trace well the phase in ~~in situ~~bottom-up ΔTWS in the Upper Nile Basin that comprises both the LVB and Lake Kyoga Basin (LKB). In the LVB for example, each explains ~~91~~ 90 % (GRGS), ~~85~~ 83 % (JPL-Mascons), and ~~77~~ 75 % (*GRCTellus* ensemble mean of CSR, JPL and GFZ) of the variance, respectively, in ~~in situ~~bottom-up ΔTWS. The relative proportion of variability in ~~in situ~~bottom-up ΔTWS (variance 120 cm^2 LVB, 24 cm^2 LKB) is explained by in situ ΔSWS (~~89~~ 93 % LVB; ~~37~~ 49 % LKB), GLDAS ensemble mean ΔSMS (~~9~~ 6 % LVB; ~~56~~ 48 % LKB) and in situ ΔGWS (~~2~~ ~1 % LVB; ~~7~~ 4 % LKB); these percentages are indicative and can vary as individual TWS components are strongly correlated and the order of explanatory variables in regression equation can affect the Analysis of Variance (ANOVA). In situ ΔGWS contributes minimally to ΔTWS and is only moderately associated with GRACE ΔTWS (strongest correlation of r=0.~~57~~39, p-value <0.001). Resolution of ΔGWS from GRACE ΔTWS in the Upper Nile Basin relies upon robust measures of ΔSWS and ΔSMS; the former is observed in situ whereas the latter is limited by uncertainty in simulated ΔSMS, represented here and in many GRACE studies by an ensemble mean of GLDAS LSMs. Mean annual amplitudes in observed ΔGWS (2003–2012) from limited piezometry for the low-storage and low-transmissivity aquifers in deeply weathered crystalline rocks that underlie the Upper Nile Basin are small (1.8 to 4.9 3.5 to 4.4 cm for S_y= 0.03) and, given the current uncertainty in simulated ΔSMS, are beyond the limit of what can be reliably quantified using current GRACE satellite products.

Our examination of a large, mass-storage change (20034 to 2006) observed in the Lake Victoria Basin highlights substantial variability in the measurement of ΔTWS using different gridded GRACE products. Although the phase in ΔTWS is generally well recorded by all tested GRACE products, substantial differences exist in the amplitude of ΔTWS that also influence the disaggregation of individual terrestrial stores (e.g., groundwater storage) and estimation of trends in TWS and individual, disaggregated freshwater stores. We note that the stronger filtering of the large-scale (~300 km) gravity signal associated with *GRCTellus* results in greater signal leakage relative to GRGS and JPL-Mascons. As a result, greater rescaling is required to resurrect signal amplitudes in *GRCTellus* relative to GRGS and JPL-Mascons and these scaling factors depend upon uncertain and incomplete a priori knowledge of terrestrial water stores derived from large-scale land-surface or hydrological models, which generally do not consider the existence of Lake Victoria, the second largest lake by area in the world.

Author contribution

RT conceived this study for which preliminary analyses were carried out by DJ and MS. MS and DJ have processed GRACE and all observational datasets and conducted statistical analyses and GIS mapping. LL conducted the analysis of spatial leakage and bias in GRACE signals. CT, RT and MO helped to establish, collate and analyse groundwater-level data; CT provided dam release data. MS and RT wrote the manuscript and LL, DJ, MO and CT commented on draft manuscripts.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We kindly acknowledge NASA's MEaSUREs Program (http://grace.jpl.nasa.gov) for the freely available gridded *GRCTellus* and JPL-MASCON GRACE data and French National Centre for Space Studies (CNES) for GRGS GRACE data. NASA's Precipitation Processing Centre and NASA's Hydrological Sciences Laboratory and the Goddard Earth Sciences Data and Information Services Centre (GES DISC) are duly acknowledged for TRMM rainfall and soil moisture data from GLDAS Land Surface Models. We kindly acknowledge the Directorate of Water Resources Management in the Ministry of Water and Environment (Uganda) for the provision of piezometric and lake-level data. Support from the UK government's UPGro Programme, funded by the Natural Environment Research Council (NERC), Economic and Social Research Council (ESRC) and the Department For International Development (DFID) through the *GroFutures: Groundwater Futures in Sub-Saharan Africa* catalyst NE/L002043/1) and consortium (NE/M008932/1) grant awards, is gratefully acknowledged.

References

A, G., Wahr, J., and Zhong, S.: Computations of the viscoelastic response of a 3-D compressible Earth to surface loading: an application to Glacial Isostatic Adjustment in Antarctica and Canada, Geophys. J. Int., 192, 557-572, doi:10.1093/gji/ggs030, 2013.

Arendt, A. A., Luthcke, S. B., Gardner, A. S., O'Neel, S., Hill, D., Moholdt, G., and Abdalati, W.: Analysis of a GRACE global mascon solution for Gulf of Alaska glaciers, Journal of Glaciology, 59, 913-924, doi:10.3189/2013JoG12J197, 2013.

Awange, J. L., Sharifi, M. A., Ogonda, G., Wickert, J., Grafarend, E., and Omulo, M.: The falling Lake Victoria water levels: GRACE, TRIMM and CHAMP satellite analysis of the lake basin, Water Resources Management, 22, 775-796, 2008.

Awange, J. L., Anyah, R., Agola, N., Forootan, E., and Omondi, P.: Potential impacts of climate and environmental change on the stored water of Lake Victoria Basin and economic implications, Water Resour. Res., 49, 8160-8173, 2013.

Awange, J. L., Forootan, E., Kuhn, M., Kusche, J., and Heck, B.: Water storage changes and climate variability within the Nile Basin between 2002 and 2011, Advances in Water Resources, 73, 1-15, 2014.

Basalirwa, C. P. K.: Delineation of Uganda into climatological rainfall zones using the method of Principle Component Analysis, International Journal of Climatology, 15, 1161-1177, 1995.

Becker, M., LLovel, W., Cazenave, A., Güntner, A., and Crétaux, J.-F.: Recent hydrological behaviour of the East African great lakes region inferred from GRACE, satellite altimetry and rainfall observations, Comptes Rendus Geoscience, 342, 223-233, 2010.

Biancale, R., Lemoine, J.-M., Balmino, G., Loyer, S., Bruisma, S., Perosanz, F., Marty, J.-C., and Gégout, P.: 3 Years of Geoid Variations from GRACE and LAGEOS Data at 10-day Intervals from July 2002 to March 2005, CNES/GRGS, 2006.

Brown, E., and Sutcliffe, J. V.: The water balance of Lake Kyoga, Uganda, Hydrological Sciences Journal, 58, 342-353, doi: 10.1080/02626667.2012.753148, 2013.

Bruinsma, S., Lemoine, J.-M., and Biancale, R.: CNES/GRGS 10-day gravity field models (release 2) and their evaluation Adv. Space Res., 45, 587-601, 10.1016/j.asr.2009.10.012, 2010.

Burgess, W. G., Shamsudduha, M., Taylor, R. G., Zahid, A., Ahmed, K. M., Mukherjee, A., Lapworth, D. J., and Bense, V. F.: Terrestrial water load and groundwater fluctuation in the Bengal Basin, Scientific Reports, 10.1038/s41598-017-04159-w, 2017.

Castellazzi, P., Martel, R., Galloway, D. L., Longuevergne, L., and Rivera, A.: Assessing Groundwater Depletion and Dynamics Using GRACE and InSAR: Potential and Limitations, Ground Water, doi:10.1111/gwat.12453, 2016.

Chen, J. L., Wilson, C. R., and Tapley, B. D.: The 2009 exceptional Amazon flood and interannual terrestrial water storage change observed by GRACE, Water Resour. Res., 46, W12526, 2010.

Dai, Y., Zeng, X., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., Denning, A. S., Dirmeyer, P. A., Houser, P. R., Niu, G., Oleson, K. W., Schlosser, C. A., and Yang, Z.-L.: The common land model (CLM), Bull. Am. Meteorol. Soc., 84, 1013-1023, 2003.

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., and Tarpley, J. D.: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, J. Geophys. Res., 108(D22), 8851, 10.1029/2002JD003296, 2003.

Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K. J., Syed, T. H., Swenson, S. C., Linage, C. R. d., and Rodell, M.: Satellites measure recent rates of groundwater depletion in California's Central Valley, Geophys. Res. Lett., 38, L03403, 10.1029/2010GL046442, 2011.

Frappart, F., Ramillien, G., and Famiglietti, J. S.: Water balance of the Arctic drainage system using GRACE gravimetry products, International Journal of Remote Sensing, 32, 431-453, doi:10.1080/01431160903474954, 2011.

Güntner, A.: Improvement of Global Hydrological Models Using GRACE Data, Surveys in Geophysics, 29, 375-397, 2008.

Hoogeveen, J., Faurès, J.-M., Peiser, L., Burke, J., and Giesen, N. v. d.: GlobWat – a global water balance model to assess water use in irrigated agriculture, Hydrol. Earth Syst. Sci., 19, 3829-3844, 2015.

Hu, L., and Jiao, J. J.: Calibration of a large-scale groundwater flow model using GRACE data: a case study in the Qaidam Basin, China, Hydrogeol. J., 23, 1305-1317, 2015.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolff, D. B.: The TRMM multi-satellite precipitation analysis: quasi-global, multi-year, combined-sensor precipitation estimates at fine scale, J. Hydrometeorol., 8 (1), 38-55, 2007.

Humphrey, V., Gudmundsson, L., and Seneviratne, S. I.: Assessing Global Water Storage Variability from GRACE: Trends, Seasonal Cycle, Subseasonal Anomalies and Extremes, Surveys in Geophysics, 37, 357-395, doi:10.1007/s10712-016-9367-1, 2016.

Indeje, M., Semazzi, F. H. M., and Ogallo, L. J.: ENSO signals in East African rainfall seasons, International Journal of Climatology, 20, 19-46, 2000.

Jacob, T., Wahr, J., Pfeffer, W. T., and Swenson, S.: Recent contributions of glaciers and ice caps to sea level rise, Nature, 482, 514-518, 2012.

Jiang, D., Wang, J., Huang, Y., Zhou, K., Ding, X., and Fu, J.: The Review of GRACE Data Applications in Terrestrial Hydrology Monitoring, Advances in Meteorology, Article ID 725131, 2014.

Khandu, Forootan, E., Schumacher, M., Awange, J. L., and Schmied, H. M.: Exploring the influence ofprecipitation extremes and humanwater use on total water storage (TWS)changes in the Ganges-Brahmaputra-Meghna River Basin, Water Resour. Res., 52, 2240-2258, doi:10.1002/2015WR018113, 2016.

Kim, H., Yeh, P. J.-F., Oki, T., and Kanae, S.: Role of rivers in the seasonal variations of terrestrial water storage over global basins, Geophys. Res. Lett., 36, L17402, doi:10.1029/2009GL039006, 2009.

Kizza, M., Westerberg, I., Rodhe, A., and Ntale, H.: Estimating areal rainfall over Lake Victoria and its basin using ground-based and satellite data, Journal of Hydrology, 464-465, 401-411, 2012.

Krishnamurthy, K. V., and Ibrahim, A. M.: Hydrometeorological Studies of Lakes Victoria, Kyoga, and Albert, in: Unintended Lakes: Their Problems and Environmental Effects, edited by: Ackermann, W. C., White, G. F., Worthington, E. B., and Ivens, J. L., American Geophysical Union, Washington D.C., 2013.

Kusche, J., Eicker, A., Forootan, E., Springer, A., and Longuevergne, L.: Mapping probabilities of extreme continental water storage changes from space gravimetry, Geophys. Res. Lett., 43, 8026-8034, doi:10.1002/2016GL069538, 2016.

Landerer, F. W., and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resour. Res., 48, W04531, 2012.

Leblanc, M. J., Tregoning, P., Ramillien, G., Tweed, S. O., and Fakes, A.: Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia, Water Resour. Res., 45, W04408, doi:10.1029/2008WR007333, 2009.

Lehner, B., Verdin, K., and Jarvis, A.: HydroSHEDS technical documentation, World Wildlife Fund, Washington D.C., 2006.

Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation data, Eos, Transactions American Geophysical Union, 89, 93-94, 2008.

Lemoine, J.-M., Bruisma, S., Loyer, S., Biancale, R., Marty, J.-C., Perosanz, F., and Balmino, G.: Temporal gravity field models inferred from GRACE data, Adv. Space Res., 39, 1620-1629, doi: 10.1016/j.asr.2007.03.062, 2007.

Liang, X., Xie, Z., and Huang, M.: A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model, J. Geophys. Res., 108(D16), 8613, 10.1029/2002JD003090, 2003.

Long, D., Longuevergne, L., and Scanlon, B. R.: Global analysis of approaches for deriving total waterstorage changes from GRACE satellites, Water Resour. Res., 51, 2574–2594, doi:10.1002/2014WR016853, 2015.

Long, D., Chen, X., Scanlon, B. R., Wada, Y., Hong, Y., Singh, V. P., Chen, Y., Wang, C., Han, Z., and Yang, W.: Have GRACE satellites overestimated groundwater depletion in the Northwest India Aquifer?, Nature Scientific Reports, 6, 24398, doi:10.1038/srep24398, 2016.

Longuevergne, L., Scanlon, B. R., and Wilson, C. R.: GRACE hydrological estimates for small basins: evaluating processing approaches on the High Plains Aquifer, USA, Water Resour. Res., 46, W11517, 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, 4817-4830, doi:10.5194/hess-17-4817-2013, 2013.

MacDonald, A. M., Bonsor, H. C., Dochartaigh, B. E. O., and Taylor, R. G.: Quantitative maps of groundwater resources in Africa, Environ. Res. Lett., 7, doi:10.1088/1748-9326/1087/1082/024009, 2012.

Nanteza, J., de Linage, C. R., Thomas, B. F., and Famiglietti, J. S.: Monitoring groundwater storage changes in complex basement aquifers: An evaluation of the GRACE satellites over East Africa, Water Resour. Res., 52, doi:10.1002/2016WR018846, 2016.

Nicholson, S. E., Yin, X., and Ba, M. B.: On the feasibility of using a lake water balance model to infer rainfall: an example from Lake Victoria, Hydrological Science Journal, 45, 75-95, 2000.

Niu, G.-Y., Yang, Z.-L., Dickinson, R. E., Gulden, L. E., and Su, H.: Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, J. Geophys. Res., 112, D07103, doi:10.1029/2006JD007522, 2007.

Oleson, K. W., Niu, G.-Y., Yang, Z.-L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., Stockli, R., Dickinson, R. E., Bonan, G. B., Levis, S., Dai, A., and Qian, T.: Improvements to the Community Land Model and their impact on the hydrological cycle, J. Geophys. Res., 113, G01021, doi:10.1029/2007JG000563, 2008.

Owor, M., Taylor, R. G., Tindimugaya, C., and Mwesigwa, D.: Rainfall intensity and groundwater recharge: empirical evidence from the Upper Nile Basin, Environmental Research Letters, 1-6, 2009.

Owor, M.: Groundwater - surface water interactions on deeply weathered surfaces of low relief in the Upper Nile Basin of Uganda, Ph.D., Geography, University College London, London, 271 pp., 2010.

Owor, M., Taylor, R. G., Mukwaya, C., and Tindimugaya, C.: Groundwater/surface-water interactions on deeply weathered surfaces of low relief: evidence from Lakes Victoria and Kyoga, Uganda, Hydrogeol. J., 19, 1403-1420, 2011.

Ramillien, G., Famiglietti, J. S., and Wahr, J.: Detection of Continental Hydrology and Glaciology Signals from GRACE: A Review, Surv. Geophys., 29, 361-374, 2008.

Rodell, M., and Famiglietti, J. S.: Terrestrial Water Storage Variations over Illinois: Analysis of Observations and Implications for GRACE, Wat. Resour. Res., 37, 1327-1340, 2001.

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, 381-394, 2004.

Rodell, M., Velicogna, I., and Famiglietti, J. S.: Satellite-based estimates of groundwater depletion in India, Nature, 460, 999-1003, doi:10.1038/nature08238, 2009.

Rowlands, D. D., Luthcke, S. B., McCarthy, J. J., Klosko, S. M., Chinn, D. S., Lemoine, F. G., Boy, J.-P., and Sabaka, T. J.: Global mass fluxsolutions from GRACE: A comparison of parameter estimation strategies-Mass concentrations versus stokes coefficients, J. Geophys. Res., 115, B01403, doi:10.1029/2009JB006546, 2010.

Scanlon, B. R., Longuevergne, L., and Long, D.: Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA, Water Resour. Res., 48, W04520, 2012.

Scanlon, B. R., Zhang, Z., Reedy, R. C., Pool, D. R., Save, H., Long, D., Chen, J., Wolock, D. M., Conway, B. D., and Winester, D.: Hydrologic implications of GRACE satellite data in the Colorado River Basin, Water Resour. Res., 51, 9891-9903, doi:10.1002/2015WR018090, 2015.

Sene, K. J., and Plinston, D. T.: A review and update of the hydrology of Lake Victoria in East Africa, Hydrological Sciences Journal, 39, 47-63, 1994.

Shamsudduha, M., Taylor, R. G., and Longuevergne, L.: Monitoring groundwater storage changes in the highly seasonal humid tropics: validation of GRACE measurements in the Bengal Basin, Water Resour. Res., 48, W02508, doi:10.1029/2011WR010993, 2012.

Strassberg, G., Scanlon, B. R., and Rodell, M.: Comparison of seasonal terrestrial water storage variations from GRACE with groundwater-level measurements from the High Plains Aquifer (USA), Geophys. Res.Lett., 34, L14402, 10.1029/2007GL030139, 2007.

Sun, A. Y., Green, R., Rodell, M., and Swenson, S.: Inferring aquifer storage parameters using satellite and in situ measurements: Estimation under uncertainty, Geophys. Res. Lett., 37, L10401, 10.1029/2010GL043231, 2010.

Sutcliffe, J. V., and Petersen, G.: Lake Victoria: derivation of a corrected natural water level series, Hydrological Sciences Journal, 52, 1316-1321, 2007.

Swenson, S., and Wahr, J.: Post-processing removal of correlated errors in GRACE data, Geophys. Res. Lett., 33, L08402, doi:10.1029/2005GL025285, 2006.

Tapley, B., Flechtner, F., Watkins, M., and Bettadpur, S.: GRACE mission: status and prospects, 2015 GRACE Science Team Meeting, Austin, Texas, September 20-22, 2015, 2015.

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., and Watkins, M. M.: GRACE measurements of mass variability in the Earth system, Science, 305, 503-505, 2004.

Taylor, R., Tindimugaya, C., Barker, J., MacDonald, D., and Kulabako, R.: Convergent radial tracing of viral and solute transport in gneiss saprolite, Ground Water, 48, 284-294, 2010.

Taylor, R. G., and Howard, K. W. F.: A tectonic geomorphic model of the hydrogeology of deeply weathered crystalline rock: evidence from Uganda, Hydrogeol. J., 8, 279-294, 2000.

Taylor, R. G., Todd, M. C., Kongola, L., Maurice, L., Nahozya, E., Sanga, H., and MacDonald, A. M.: Evidence of the dependence of groundwater resources on extreme rainfall in East Africa, Nature Climate Change, 3, 374-378, doi:10.1038/nclimate1731, 2013.

UNEP: Adaptation to Climate-change Induced Water Stress in the Nile Basin: A Vulnerability Assessment Report, Division of Early Warning and Assessment (DEWA). United Nations Environment Programme (UNEP), Nairobi, Kenya, 2013.

Vishwakarma, B. D., Devaraju, B., and Sneeuw, N.: Minimizing the effects of filtering on catchment scale GRACE solutions, Water Resour. Res., 52, 5868-5890, doi:10.1002/2016WR018960, 2016.

Vouillamoz, J. M., Lawson, F. M. A., Yalo, N., and Descloitres, M.: The use of magnetic resonance sounding for quantifying specific yield and transmissivity in hard rock aquifers: The example of Benin, Journal of Applied Geophysics, 107, 16-24, 2014.

Wahr, J., Swenson, S., Zlotnicki, V., and Velicogna, I.: Time-variable gravity from GRACE: First results, Geophys. Res. Lett., 31, L11501, doi:10.1029/2004GL019779, 2004.

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

Wang, L., Davis, J. L., Hill, E. M., and Tamisiea, M. E.: Stochasticfiltering for determining gravityvariations for decade-long timeseries of GRACE gravity, J. Geophys.Res. Solid Earth, 121, 2915-2931, doi:10.1002/2015JB012650, 2016.

Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., and Landerer, F. W.: Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons, J. Geophys. Res. Solid Earth, 120, 2648–2671, doi:10.1002/2014JB011547, 2015.

Wiese, D. N., Yuan, D.-N., Boening, C., Landerer, F. W., and Watkins, M. M.: JPL GRACE Mascon Ocean, Ice, and Hydrology Equivalent Water Height JPL RL05M.1. Ver. 1, PO.DAAC, CA, USA, 2015.

Wiese, D. N., Landerer, F. W., and Watkins, M. M.: Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution, Water Resour. Res., 52, 7490-7502, doi:10.1002/2016WR019344, 2016.

Xie, H., Longuevergne, L., Ringler, C., and Scanlon, B. R.: Calibration and evaluation of a semi-distributed watershed model of Sub-Saharan Africa using GRACE data, Hydrol. Earth Syst. Sci., 16, 3083-3099, 2012.

Yin, X., and Nicholson, S. E.: The water balance of Lake Victoria, Hydrol. Sci. J., 43, 789-811, 1998.

Figure Captions

Figure 1. Map of the study area encompassing the Lake Victoria Basin (LVB) and Lake Kyoga Basin (LKB), and location of the in situ monitoring stations. The Upper Nile Basin is marked by a rectangle (red) within the entire Nile River Basin shown as a shaded relief index map. Piezometric monitoring (red circles) and lake-level gauging (dark blue squares) stations are shown on the map.

Figure 2. Observed daily total dam releases (blue line) and the agreed curve (red line) at the outlet of Lake Victoria in Jinja from November 2007 to July 2009 (Owor et al., 2011).

Figure 3. Mean annual rainfall for the period of 2003–2012 derived from TRMM satellite observations. Greater annual rainfall is observed over much of the Lake Victoria and northeastern corner of the Lake Victoria Basin.

Figure 4. Seasonal pattern (monthly mean from January 2003 to December 2012) of TRMM-derived monthly rainfall, various GRACE-derived ΔTWS signals [GRCTellus~~E~~=ensemble mean of CSR, ~~GFZ, and~~ JPL and GFZ; GRGS and JPL-Mascons (MSCN) products], the bottom-up TWS; GLDAS LSMs ensemble mean ΔSMS, in situ ΔSWS and borehole-derived estimate of ΔGWS over the Lake Victoria Basin.

Figure 5. Monthly time-series datasets for the Lake Victoria Basin (LVB) from January 2003 to December 2012: (a) *GRCTellus* GRACE-derived ΔTWS (ensemble mean of CSR, GFZ, and JPL), GRGS and JPL-Mascons ΔTWS time-series data; (b) GLDAS-derived ΔSMS (individual signals as well as an ensemble mean of NOAH, CLM, and VIC); (c) lake-level-derived ΔSWS; and (d) borehole-derived ΔGWS time-series data. Note that monthly rainfall records derived from TRMM satellite are plotted on panel (d) where the dashed horizontal line represents the mean monthly rainfall for the period of January 2003 to December 2012.

Figure 6. Monthly time-series datasets for the Lake Kyoga Basin (LKB) from January 2003 to December 2012: (a) *GRCTellus* GRACE-derived ΔTWS (ensemble mean of CSR, GFZ, and JPL), GRGS and J~~L~~PL-Mascons ΔTWS time-series data; (b) GLDAS-derived ΔSMS (individual signals as well as an ensemble mean of NOAH, CLM, and VIC); (c) lake-level-

derived ΔSWS; and (d) borehole-derived ΔGWS time-series data. Note that monthly rainfall records derived from TRMM satellite are plotted on panel (d) where the dashed horizontal line represents the mean monthly rainfall for the period of January 2003 to December 2012.

Figure 7. Comparison among time-series records of ΔTWS from *GRCTellus* (ensemble mean of CSR, GFZ, and JPL), GRGS and JPL-~~MASCON~~ Mascons GRACE products and ~~in situ~~bottom-up ΔTWS for the Lake Victoria Basin (LVB) (a), and Lake Kyoga Basin (LKB)~~,~~ (b) for the period of January 2003 to December 2012. The vertical grey lines represent monthly rainfall anomalies in LVB and LKB.

Figure 8. Estimates of in situ ΔGWS and GRACE-derived ΔGWS time-series records (January 2003 to December 2012~~2003–2012~~) in LVB show a substantial variations among themselves. ~~Note that an adjusted~~An ensemble mean ΔSMS (GLDAS 3 LSMs: CLM, NOAH and VIC) and an unscaled ΔSWS ~~(scaling factor of 0.11)~~ ~~is~~ are applied in the disaggregation of ΔGWS using *GRCTellus* GRACE (ensemble mean of CSR, GFZ, and JPL) ~~product; similarly, an adjusted ΔSWS (scaling factor of 0.39) is applied for the~~and JPL-Mascons products.

Figure 9. Taylor diagram shows strength of statistical association, variability in amplitudes of time-series records and agreement among the reference data, ~~in situ~~bottom-up ΔTWS and *GRCTellus* GRACE-derived ΔTWS (ensemble mean of CSR, GFZ, and JPL, GRGS and JPL-Mascons ΔTWS time-series records), simulated ΔSMS (ensemble mean of NOAH, CLM~~,~~ and VIC), in situ ΔSWS, and in situ ΔGWS over the LVB. The solid arcs around the reference point (black square) indicate c~~ea~~ent~~e~~red Root Mean Square (RMS) differences among ~~in situ~~bottom-up ΔTWS and other variables, and the dashed arcs from the origin of the diagram indicate variability in time-series records. Data for Lake Victoria Basin (LVB) are only shown in this diagram.

Table 1. Estimated areal extent (km^2) of the Lake Victoria Basin (LVB), Lake Kyoga Basin (LKB), Lake Victoria and Lake Kyoga.

Basin/Lake	**This study [*HydroSHEDS* database]**	**UNEP (2013)**	**Awange et al. (2014)**
Lake Victoria Basin	256 100	184 000	258 000
Lake Victoria	67 220	68 800	-
Lake Kyoga Basin	79 270	75 000	75 000
Lake Kyoga	2 730	1 720	-

Table 2. Details of groundwater and lake level monitoring stations located in Lake Victoria Basin and Lake Kyoga Basin.

Monitoring Station	Basin	Parameter	Longitude	Latitude	Depth (m bgl)
Apac	LKB	Groundwater level	32.50	1.99	15.0
Pallisa	LKB	Groundwater level	33.69	1.20	46.2
Soroti	LKB	Groundwater level	33.63	1.69	66.0
Bugondo	LKB	Lake level	33.20	0.45	-
Entebbe	LVB	Groundwater level	32.47	0.04	48.0
Rakai	LVB	Groundwater level	31.40	–0.69	53.0
Nkokonjeru	LVB	Groundwater level	32.91	0.24	30.0
Jinja	LVB	Lake level	33.23	1.59	-

Table 3. Linear trends (cm yr^{-1}) in GRACE ΔTWS and ~~in situ~~bottom-up ΔTWS in Lake Victoria Basin and Lake Kyoga Basin over various time periods (statistically significant trends, p values <0.05 are marked by an asterisk).

Period	**GRACE Ensemble**	**GRGS**	**JPL-Mascons**	**~~In situ~~Bottom-up TWS**
		Lake Victoria Basin (LVB)		
2003–2006	–4.10*	–9.00*	–~~7.70~~10.0*	–11.00*
2007–2012	–0.31	1.50*	2~~1~~.7~~9~~0*	1.10*
2003–2012	0.04	0.58	~~0.79~~1.00*	0.54*
		Lake Kyoga Basin (LKB)		
2003–2006	–2.10*	–4.60*	–3~~5~~.~~6~~50*	–2.80*
2007–2012	0.22	2.00*	1~~2~~.5~~2~~0*	0.48
2003–2012	–0.01	0.54*	0.54~~5~~*	–0.56*

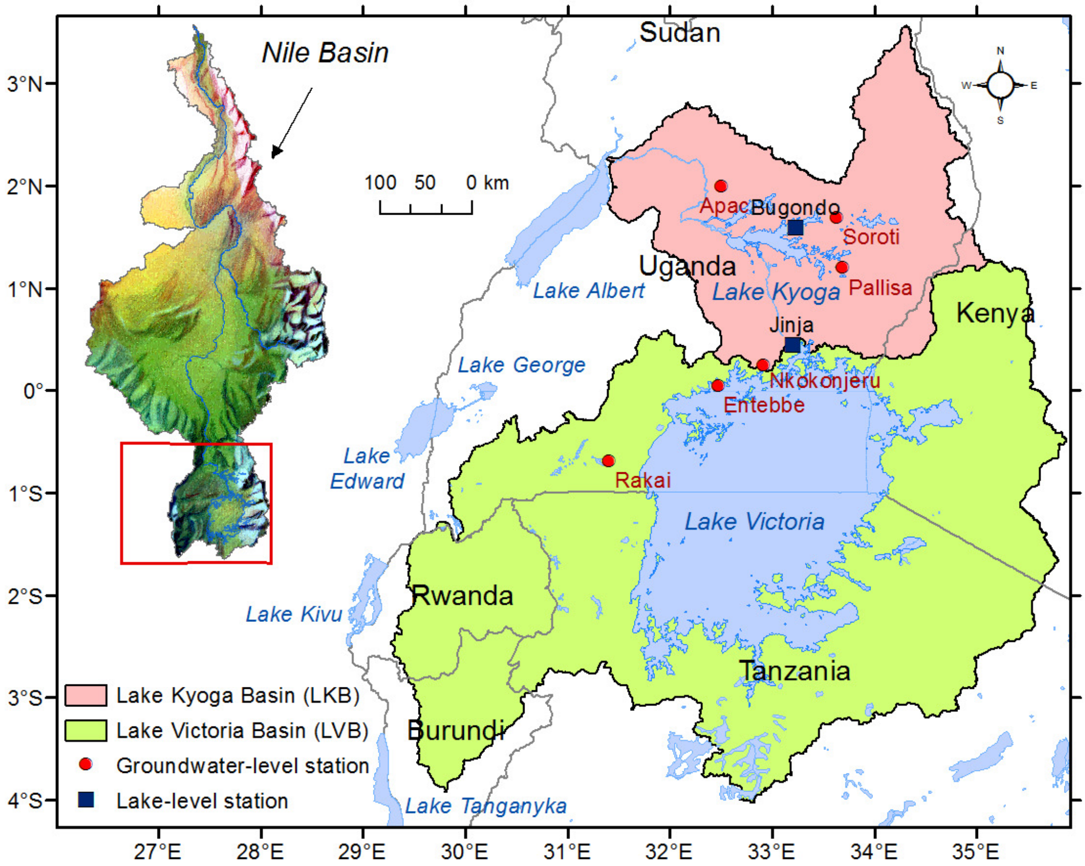


Figure 1. Map of the study area encompassing the Lake Victoria Basin (LVB) and Lake Kyoga Basin (LKB), and location of the in situ monitoring stations. The Upper Nile Basin is marked by a rectangle (red) within the entire Nile River Basin shown as a shaded relief index map. Piezometric monitoring (red circles) and lake-level gauging (dark blue squares) stations are shown on the map.

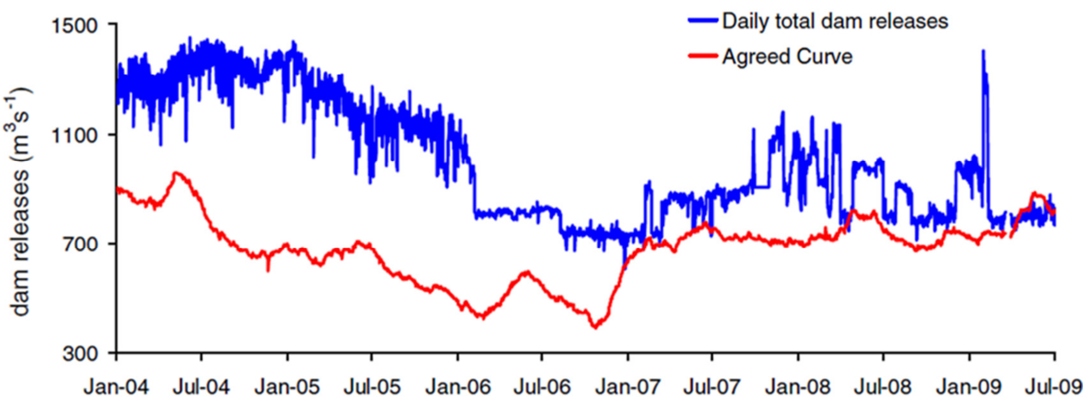


Figure 2. Observed daily total dam releases (blue line) and the agreed curve (red line) at the outlet of Lake Victoria in Jinja from November 2007 to July 2009 (Owor et al., 2011).

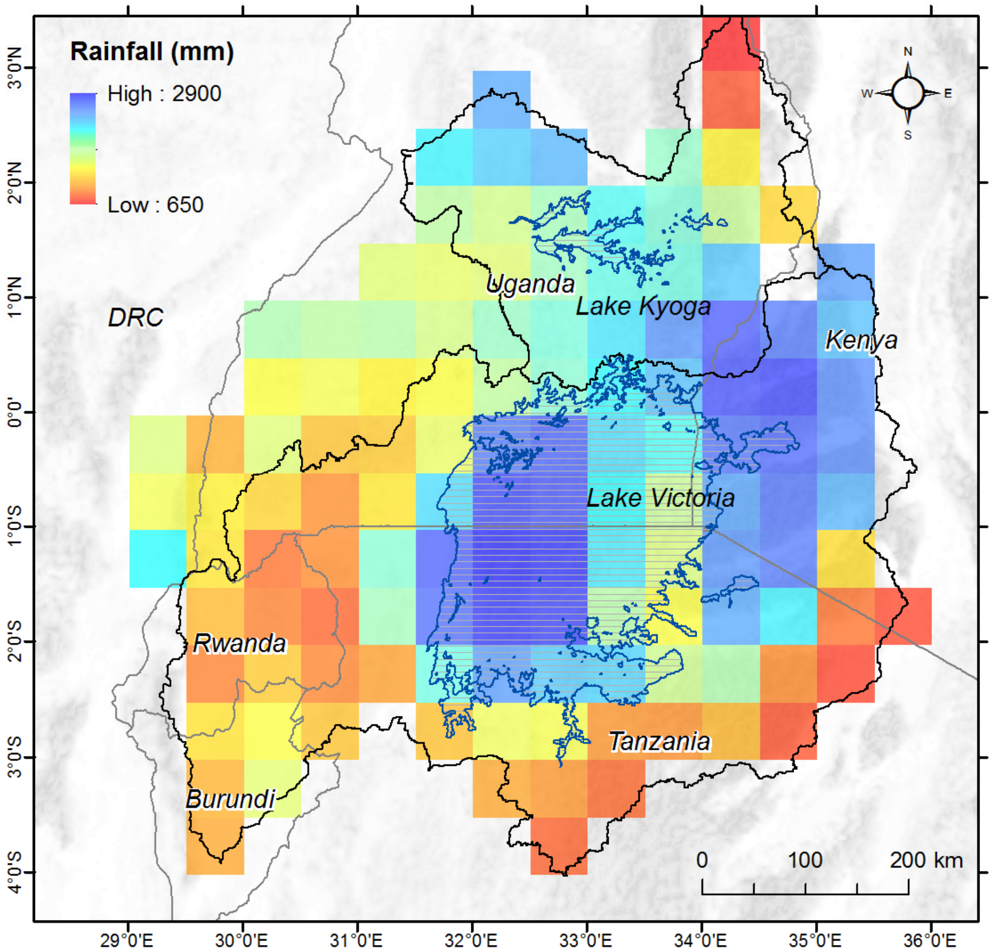


Figure 3. Mean annual rainfall for the period of 2003–2012 derived from TRMM satellite observations. Greater annual rainfall is observed over much of the Lake Victoria and northeastern corner of the Lake Victoria Basin.

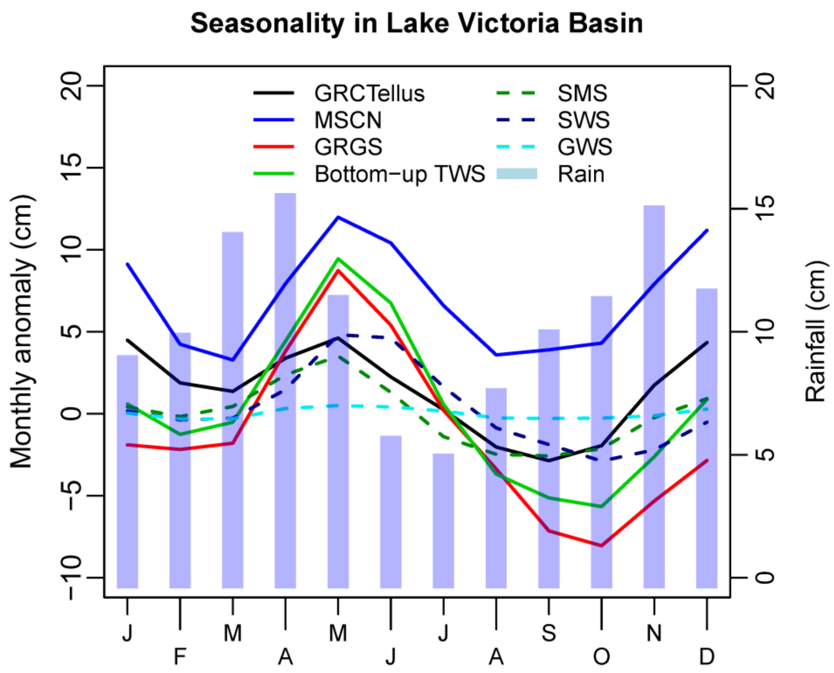


Figure 4. Seasonal pattern (monthly mean from January 2003 to December 2012) of TRMM-derived monthly rainfall, various GRACE-derived ΔTWS signals [GRCTellus~~E~~=ensemble mean of CSR, ~~GFZ, and~~ JPL and GFZ; GRGS and JPL-Mascons (MSCN) products], the bottom-up TWS; GLDAS LSMs ensemble mean ΔSMS, in situ ΔSWS and borehole-derived estimate of ΔGWS over the Lake Victoria Basin.

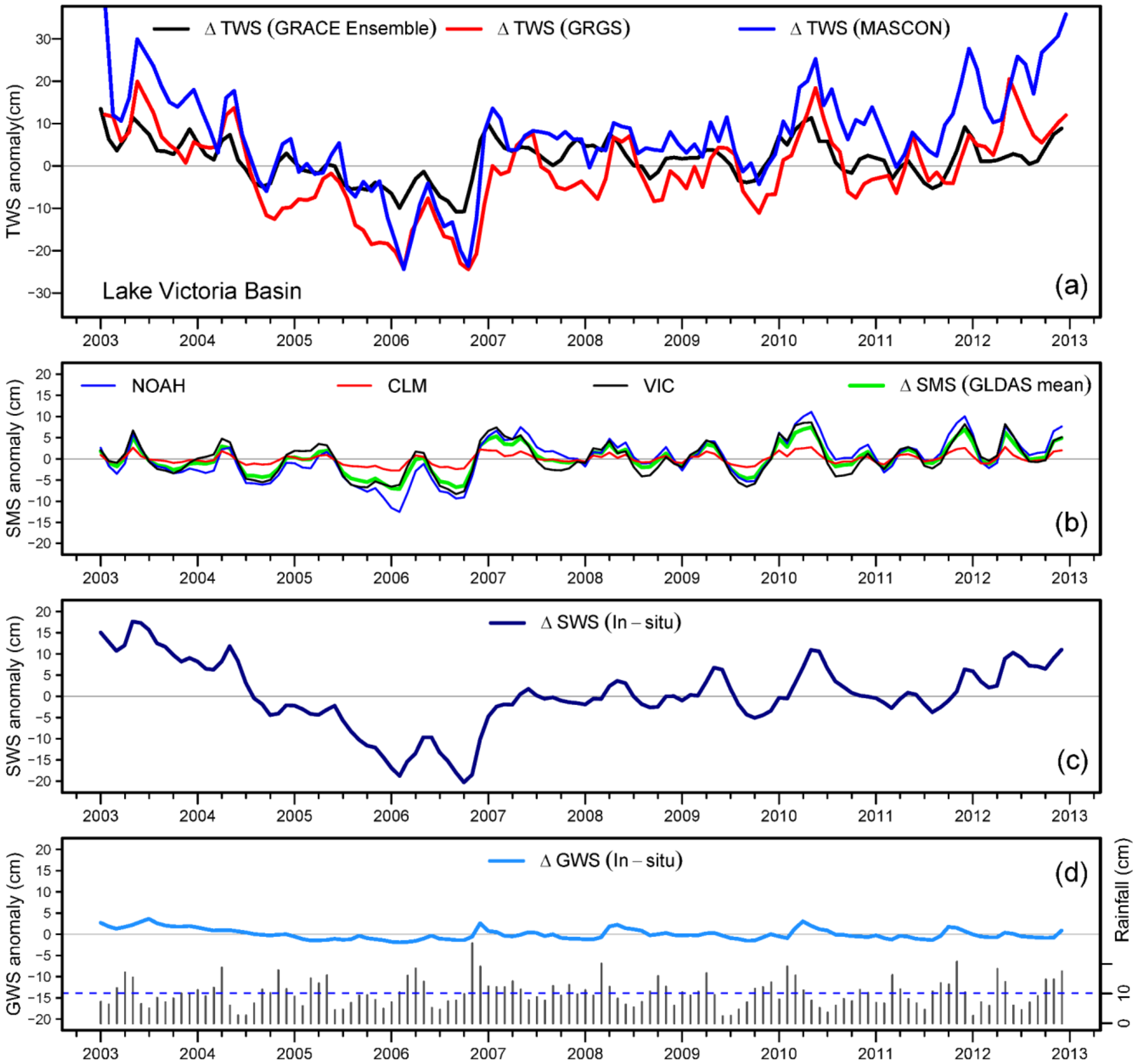


Figure 5. Monthly time-series datasets for the Lake Victoria Basin (LVB) from January 2003 to December 2012: (a) *GRCTellus* GRACE-derived ΔTWS (ensemble mean of CSR, GFZ, and JPL), GRGS and JPL-Mascons ΔTWS time-series data; (b) GLDAS-derived ΔSMS (individual signals as well as an ensemble mean of NOAH, CLM, and VIC); (c) lake-level-derived ΔSWS; and (d) borehole-derived ΔGWS time-series data. Note that monthly rainfall records derived from TRMM satellite are plotted on panel (d) where the dashed horizontal line represents the mean monthly rainfall for the period of January 2003 to December 2012.

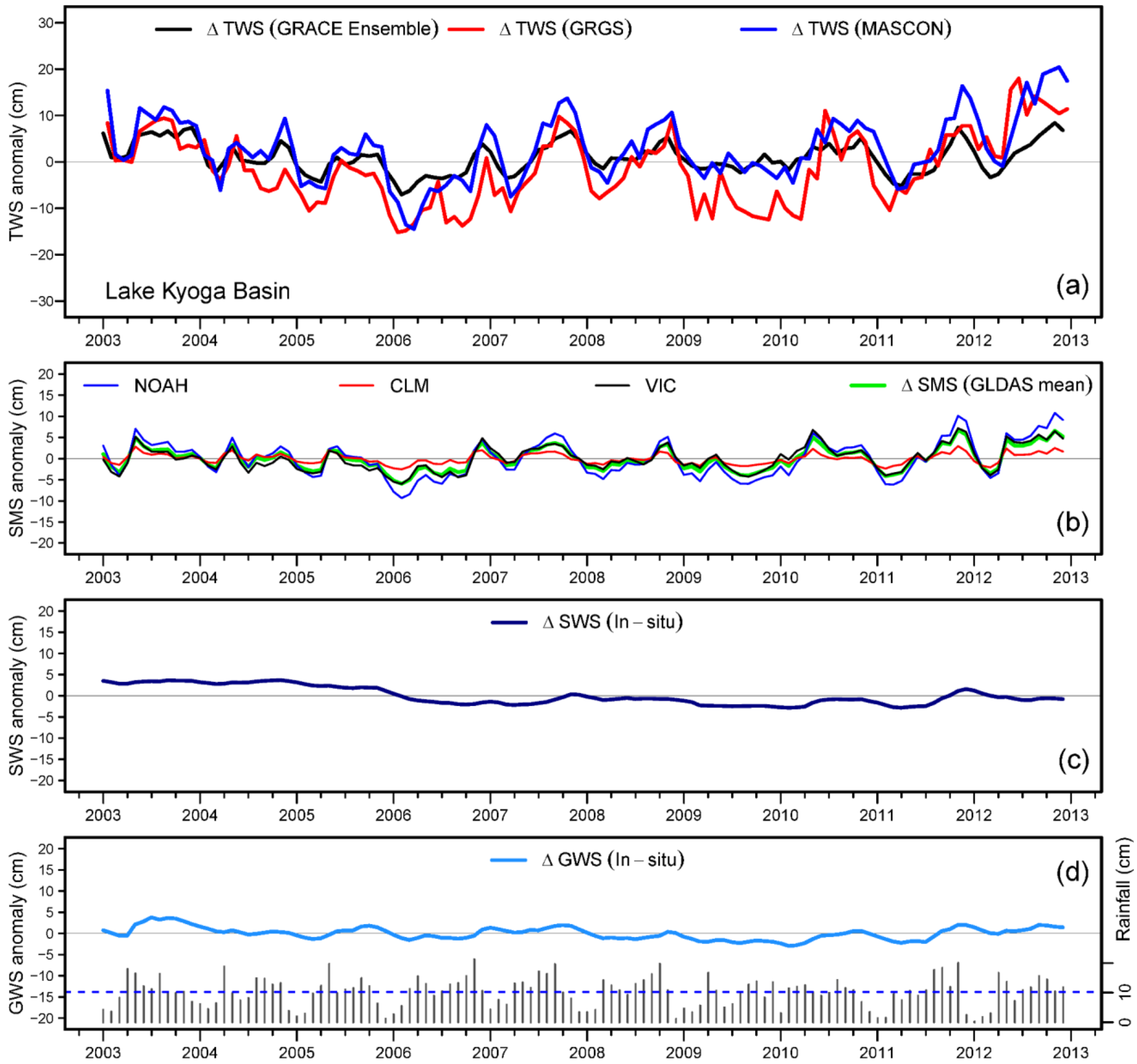


Figure 6. Monthly time-series datasets for the Lake Kyoga Basin (LKB) from January 2003 to December 2012: (a) *GRCTellus* GRACE-derived ΔTWS (ensemble mean of CSR, GFZ, and JPL), GRGS and J~~L~~PL-Mascons ΔTWS time-series data; (b) GLDAS-derived ΔSMS (individual signals as well as an ensemble mean of NOAH, CLM, and VIC); (c) lake-level-derived ΔSWS; and (d) borehole-derived ΔGWS time-series data. Note that monthly rainfall records derived from TRMM satellite are plotted on panel (d) where the dashed horizontal line represents the mean monthly rainfall for the period of January 2003 to December 2012.

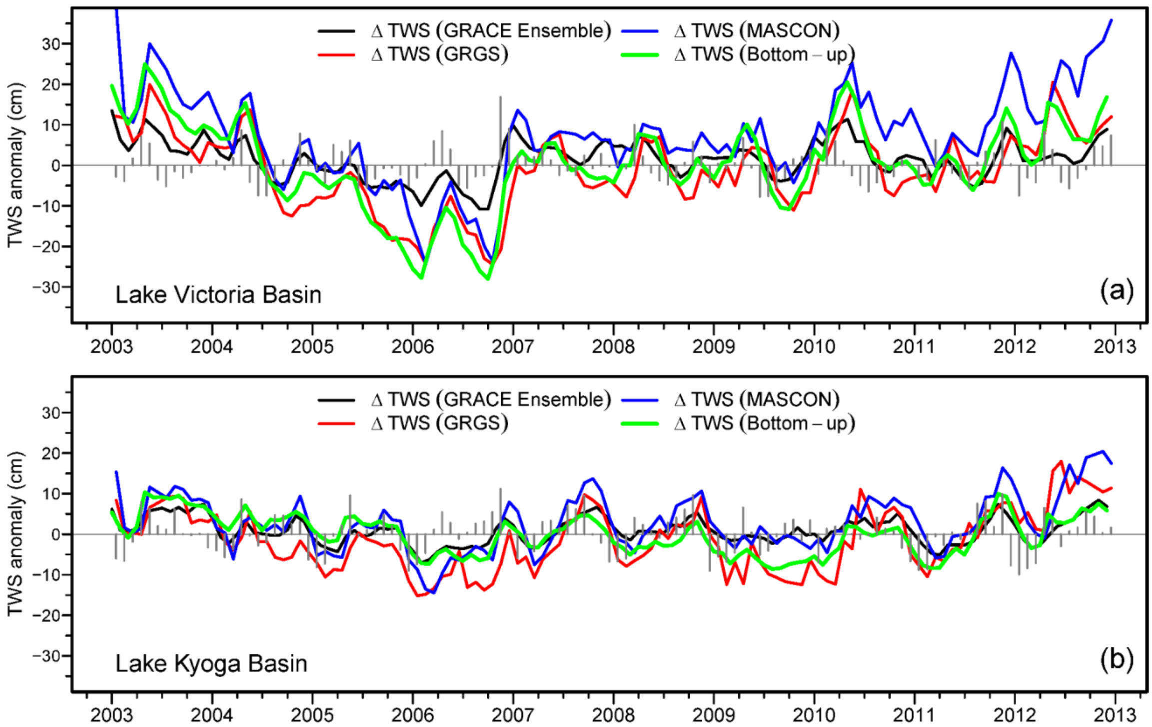


Figure 7. Comparison among time-series records of ΔTWS from *GRCTellus* (ensemble mean of CSR, GFZ, and JPL), GRGS and JPL-~~MASCON~~ Mascons GRACE products and ~~in situ~~bottom-up ΔTWS for the Lake Victoria Basin (LVB) (a), and Lake Kyoga Basin (LKB)~~,~~ (b) for the period of January 2003 to December 2012. The vertical grey lines represent monthly rainfall anomalies in LVB and LKB.

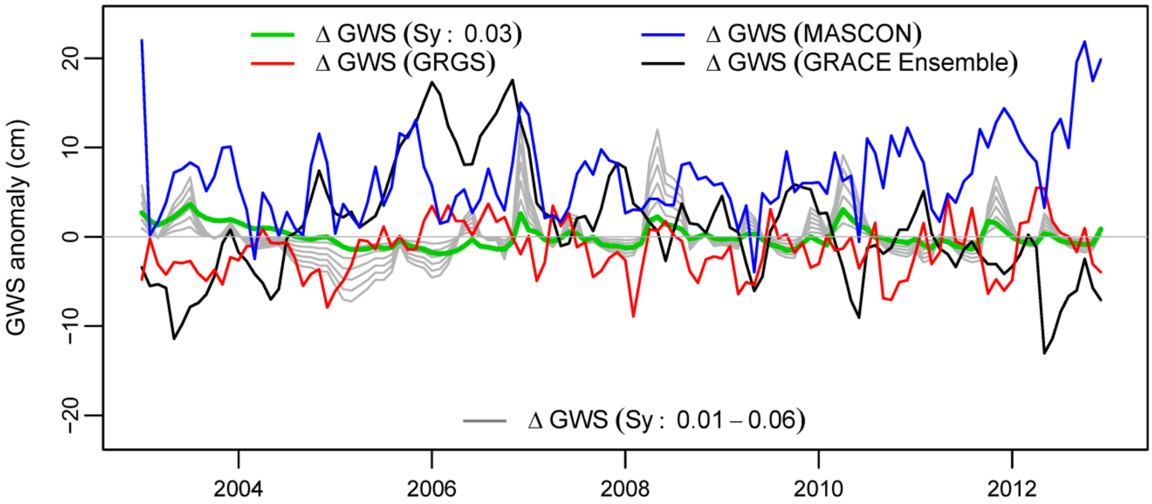


Figure 8. Estimates of in situ ΔGWS and GRACE-derived ΔGWS time-series records (January 2003 to December 20122003–2012) in LVB show a substantial variations among themselves. Note that an adjustedAn ensemble mean ΔSMS (GLDAS 3 LSMs: CLM, NOAH and VIC) and an unscaled ΔSWS (scaling factor of 0.11) is are applied in the disaggregation of ΔGWS using *GRCTellus* GRACE (ensemble mean of CSR, GFZ, and JPL) product; similarly, an adjusted ΔSWS (scaling factor of 0.39) is applied for theand JPL-Mascons products.

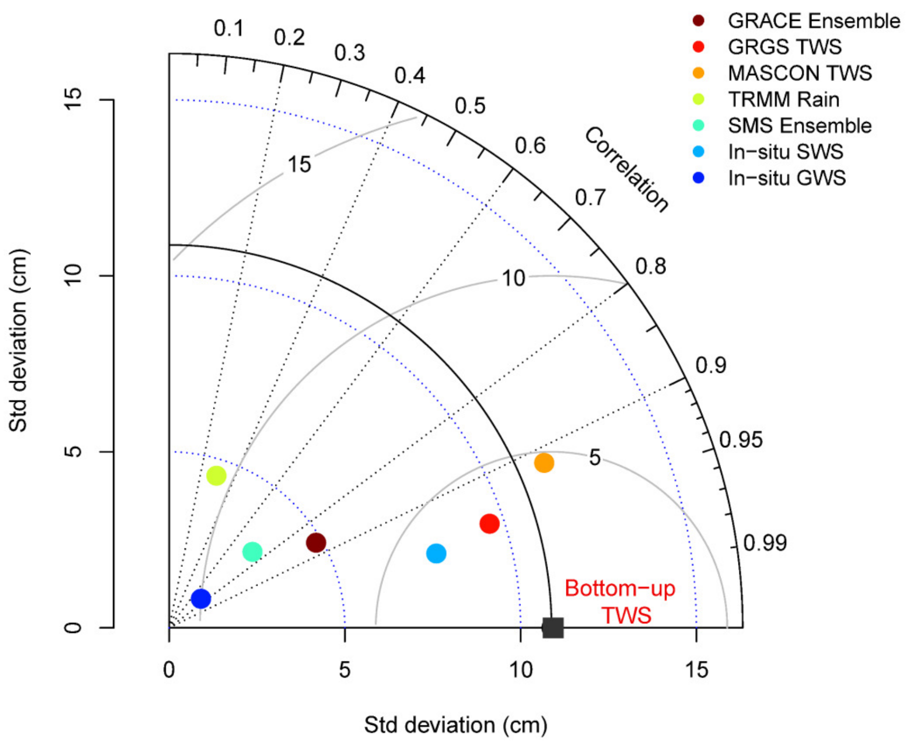


Figure 9. Taylor diagram shows strength of statistical association, variability in amplitudes of time-series records and agreement among the reference data, ~~in situ~~bottom-up ΔTWS and *GRCTellus* GRACE-derived ΔTWS (ensemble mean of CSR, GFZ, and JPL, GRGS and JPL-Mascons ΔTWS time-series records), simulated ΔSMS (ensemble mean of NOAH, CLM~~,~~ and VIC), in situ ΔSWS, and in situ ΔGWS over the LVB. The solid arcs around the reference point (black square) indicate ceantered Root Mean Square (RMS) differences among ~~in situ~~bottom-up ΔTWS and other variables, and the dashed arcs from the origin of the diagram indicate variability in time-series records. Data for Lake Victoria Basin (LVB) are only shown in this diagram.

Recent changes in terrestrial water storage in the Upper Nile Basin: an evaluation of commonly used gridded GRACE products

Mohammad Shamsudduha[1, 2], Richard G. Taylor[2], Darren Jones[3], Laurent Longuevergne[4], Michael Owor[5] and Callist Tindimugaya[6]

[1]Institute for Risk and Disaster Reduction, University College London, London, UK
[2]Department of Geography, University College London, London, UK
[3]Centre for Geography, Environment and Society, University of Exeter, Exeter, UK
[4]CNRS – UMR 6118 Géosciences Rennes, Université de Rennes 1, Rennes, France
[5]Department of Geology & Petroleum Studies, Makerere University, Kampala, Uganda
[6]Directorate of Water Resources Management, Ministry of Water & Environment, Uganda

Correspondence to: M. Shamsudduha (m.shamsudduha@ucl.ac.uk)

Supplementary Material

Supplementary Table:

Table S1. In each sub sections (LVB and LKB) of the table below, ~~M~~mean annual (2003–2012) amplitudes (first row) and ~~variability (i.e.,~~ variance~~, cm^2)~~ (second row) in ~~in situ~~bottom-up ΔTWS (time-series variance in LVB: 120 cm^2 and LKB: 24 cm^2) explained in linear regression (R^2) by individual signals (~~second~~ third row) such as various GRACE-derived ΔTWS, simulated ΔSMS, in situ ΔSWS and in situ ΔGWS in both ~~LVB (120 cm^2) and LKB (24 cm^2)~~ and the relative proportion of variability (~~third~~ fourth row) in ~~in situ~~bottom-up ΔTWS as explained by ΔSMS, ΔSWS and ΔGWS.

Lake Victoria Basin (LVB)					
GRACE Ensemble ΔTWS	GRGS ΔTWS	JPL-MASCON ΔTWS	ΔSMS	ΔSWS	ΔGWS
11.7 cm	20.6 cm	~~20~~27.~~5~~3 cm	7.9 cm	14.8 cm	~~4~~2.~~4~~9 cm
23.5 cm^2	92.6 cm^2	136.9 cm^2	10.4 cm^2	62.8 cm^2	1.5 cm^2
~~77~~75%	~~91~~90%	~~85~~83%	~~59~~54%	~~89~~92%	~~32~~54%
-	-	-	~~9~~6.~~4~~5%	~~88~~92.6%	~~1.9~~0.66%
Lake Kyoga Basin (LKB)					
8.4 cm	16.2 cm	16.~~4~~5 cm	7.3 cm	3.8 cm	~~3~~2.~~5~~9 cm
11.2 cm^2	58.3 cm^2	47.8 cm^2	7.4 cm^2	4.5 cm^2	2.1 cm^2
~~55~~62%	~~49~~56%	~~53~~57%	~~56~~62%	~~51~~48%	~~64~~76%
-	-	-	~~55.9~~48.5%	~~37.2~~47.9%	~~6.9~~3.6%

Supplementary Figures:

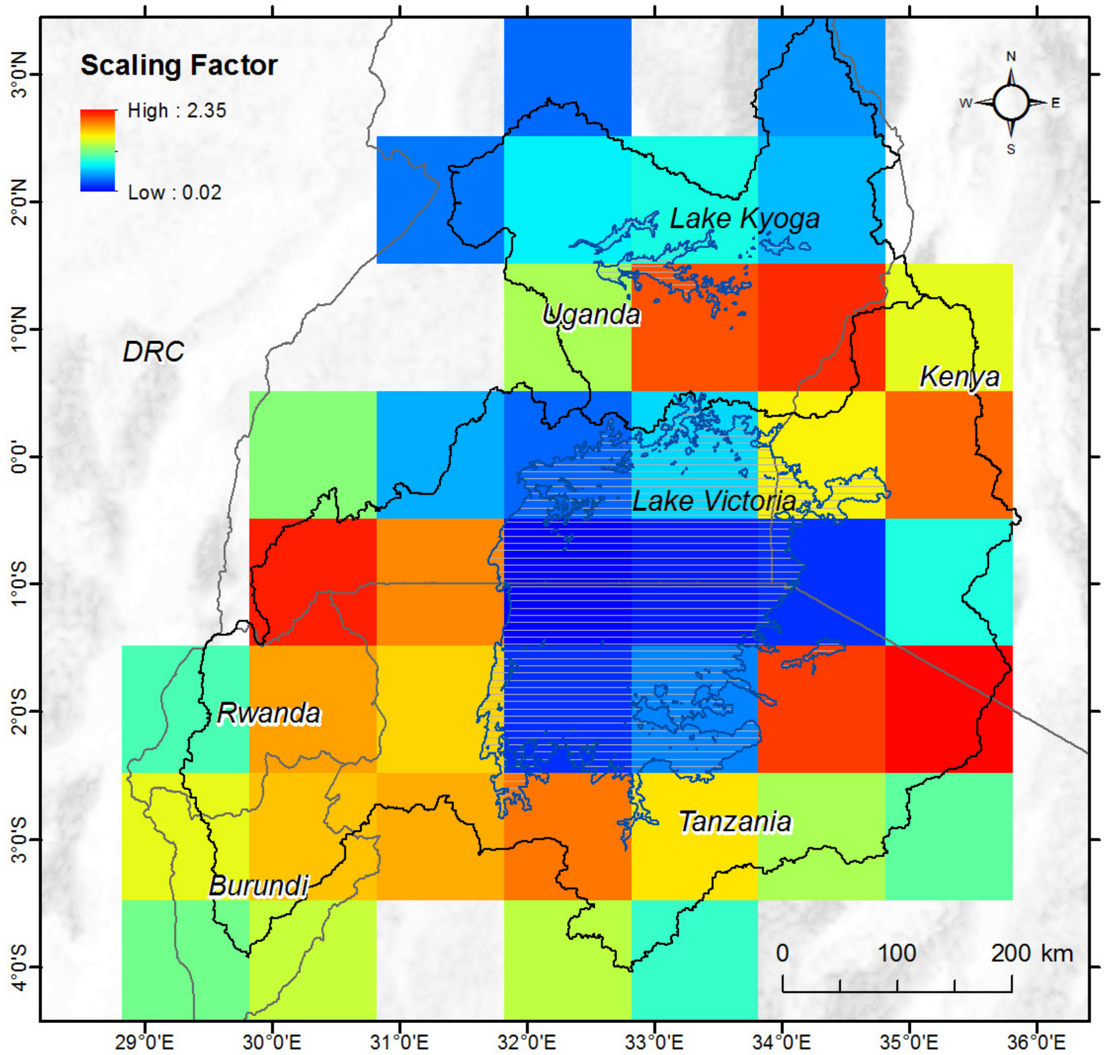


Figure S1. The general outline of Lake Victoria Basin (LVB) and Lake Kyoga Basin (LKB) within the Upper Nile Basin and the gridded (1° × 1°) scaling coefficients for *GRCTellus* solutions derived from CLM4.0 land surface model (Landerer and Swenson, 2012).

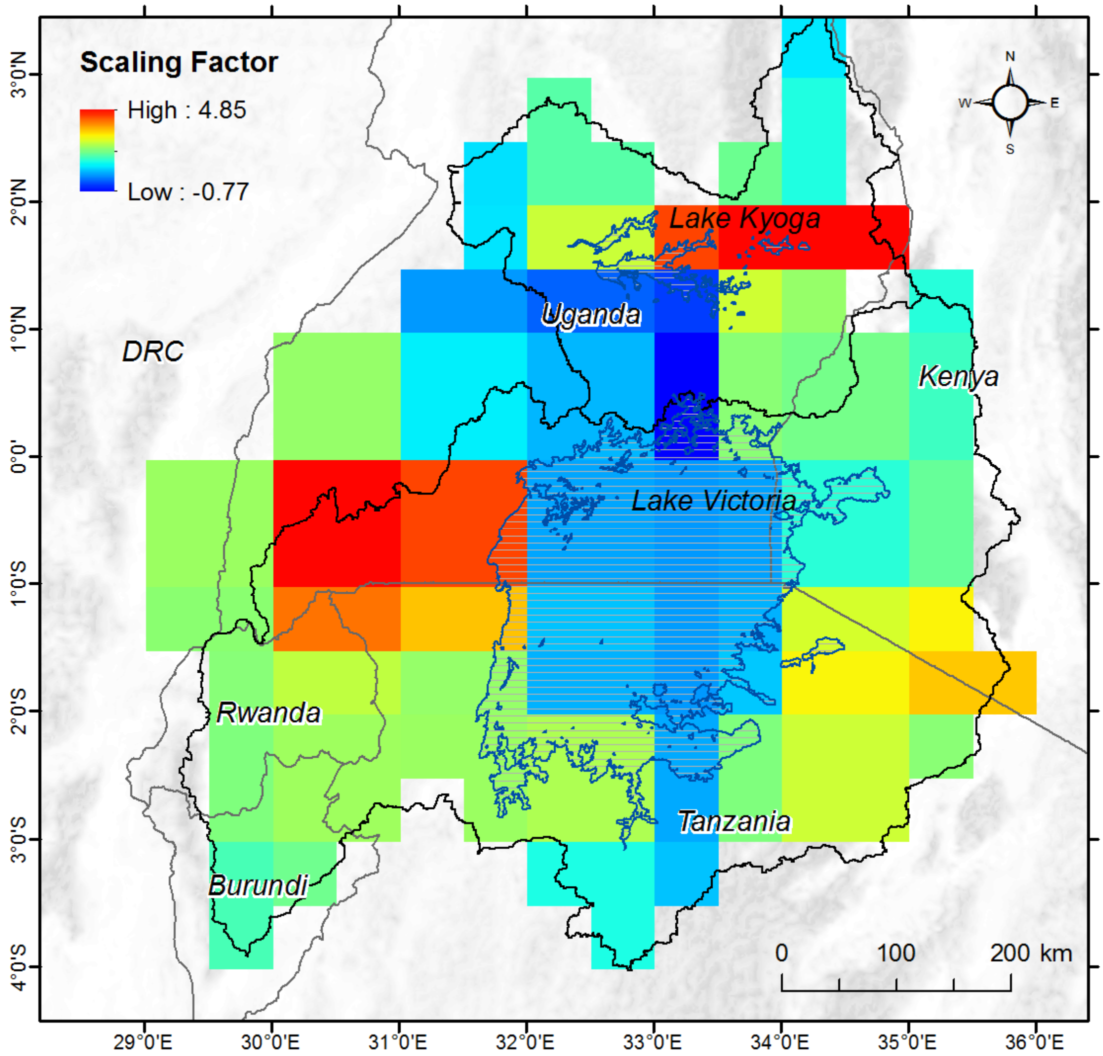
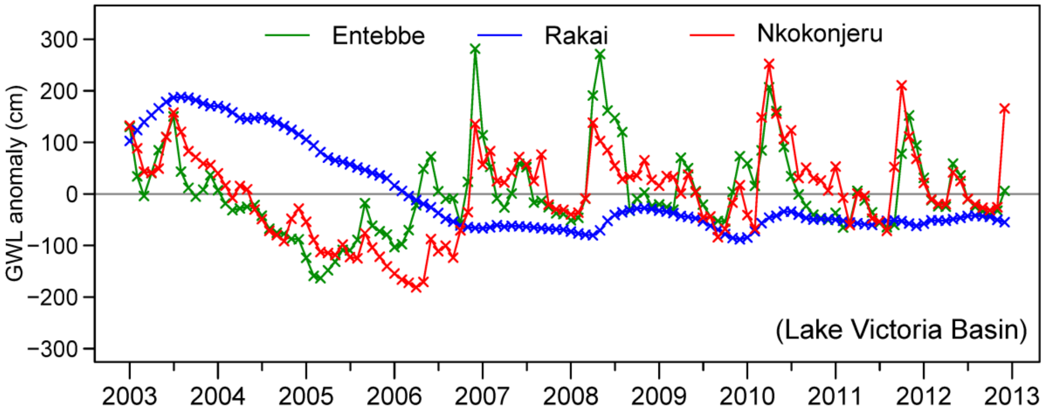


Figure S2. The general outline of Lake Victoria Basin (LVB) and Lake Kyoga Basin (LKB) within the Upper Nile Basin and the gridded (0.5° × 0.5°) scaling coefficients for JPL-Mascons derived from CLM4.0 land surface model (Wiese et al., 2016).



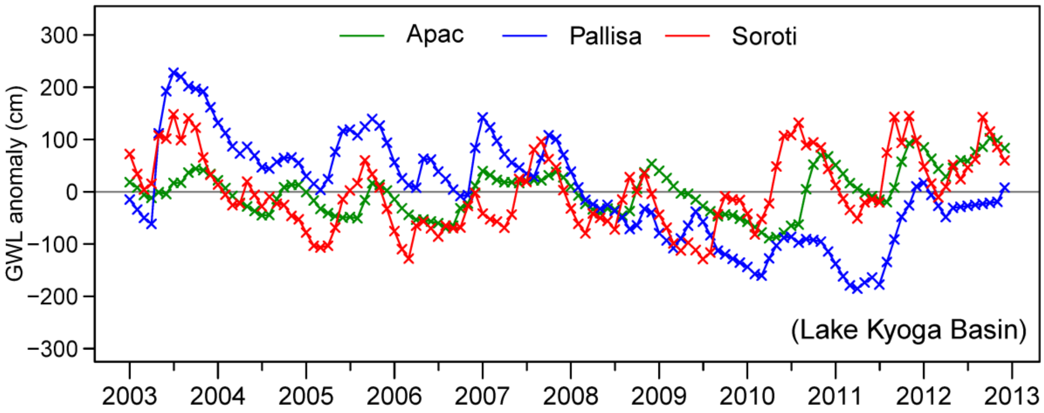


Figure S3. Observed groundwater-level monitoring records (January 2003 to December 2012) at 6 monitoring boreholes: Entebbe, Rakai and Nkokonjeru from Late Victoria Basin (LVB) and Apac, Palissa and Soroti from Lake Kyoga Basin (LKB).

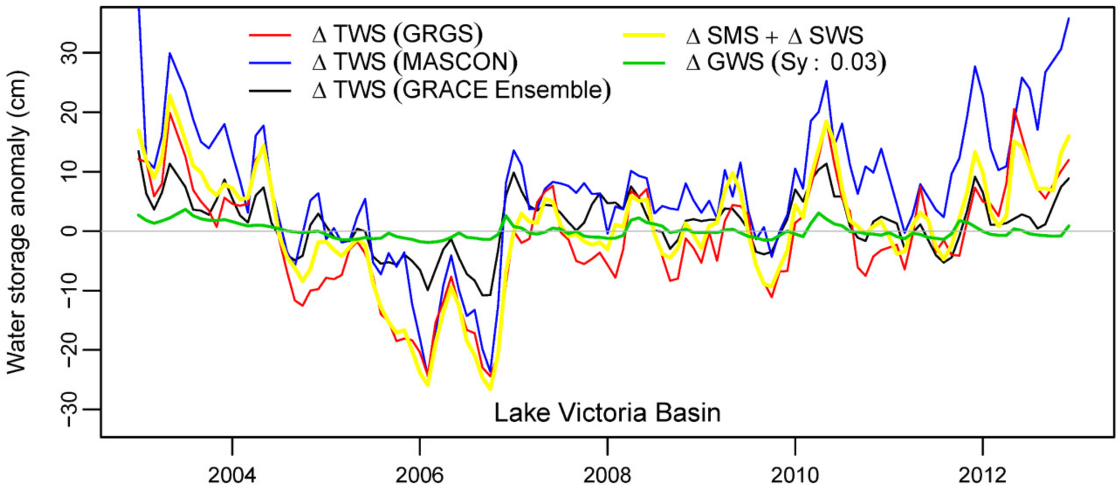


Figure S4. Time-series records of various GRACE ΔTWS solutions, sum of in-situ ΔSWS and ensemble mean of ΔSMS signals and in-situ ΔGWS for Lake Victoria Basin. The graph illustrates that combined signals of ΔSWS+ΔSMS clearly exceed GRACE ΔTWS anomalies (positive and negative sides of the y-axis) in several monthly instances over the period of 2003 to 2012.

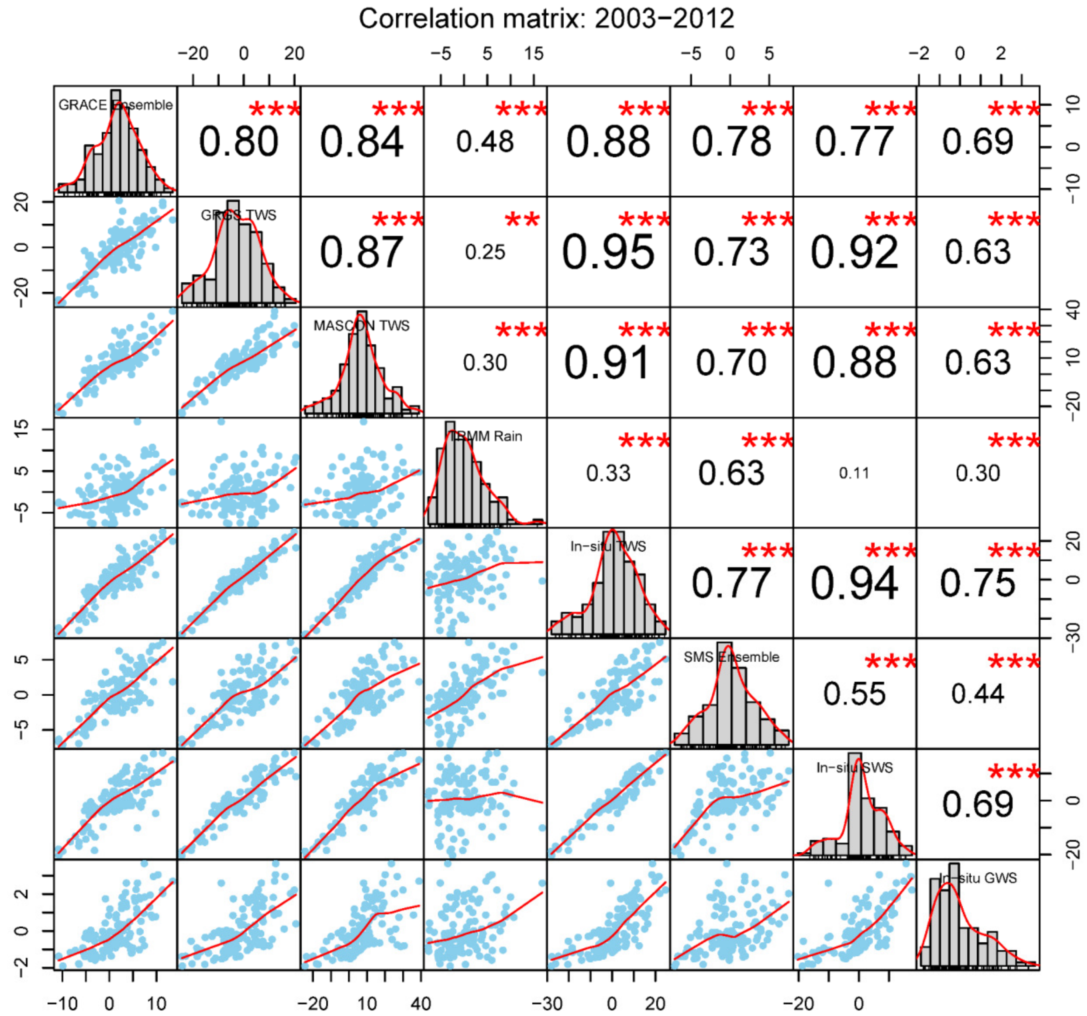


Figure ~~S2~~S5. Pearson correlation coefficients among the time-series variables collated over the Lake Victoria Basin for the period of 2003 to 2012. Statistically significant correlated variables are marked with asterisks where the significance asterisks represent *p*-values < 0.05 (1 asterisk), <0.01 (2 asterisks) and <0.001 (3 asterisks). Histograms with kernel density overlays and bivariate scatterplots of variables are shown. The fitted curve lines in bivariate scatterplots represent locally-weighted polynomial regression (i.e., Lowess) lines.

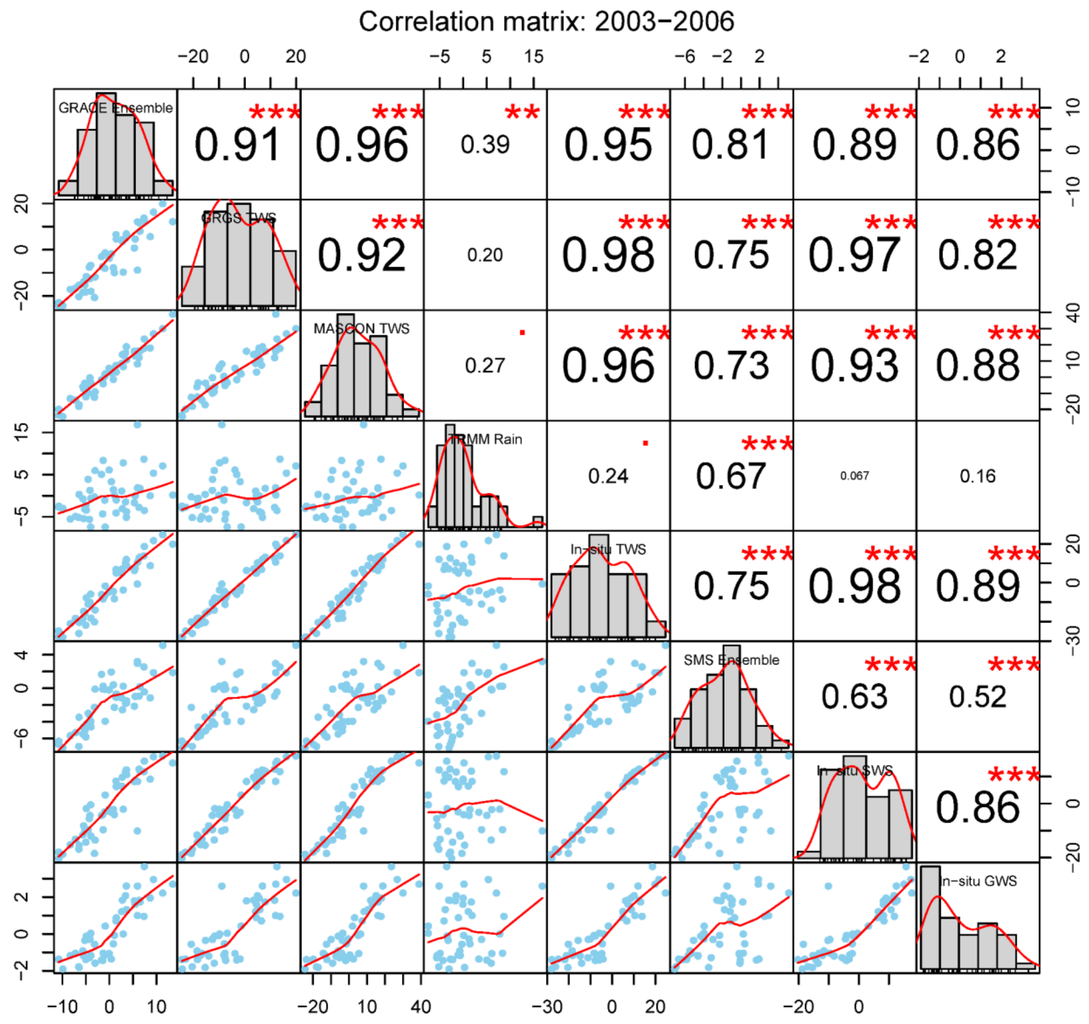


Figure ~~S3~~S6. Pearson correlation coefficients among the time-series variables collated over the Lake Victoria Basin for the period of 2003 to 2006. Statistically significant correlated variables are marked with asterisks where the significance asterisks represent *p*-values < 0.05 (1 asterisk), <0.01 (2 asterisks) and <0.001 (3 asterisks). Histograms with kernel density overlays and bivariate scatterplots of variables are shown. The fitted curve lines in bivariate scatterplots represent locally-weighted polynomial regression (i.e., Lowess) lines.

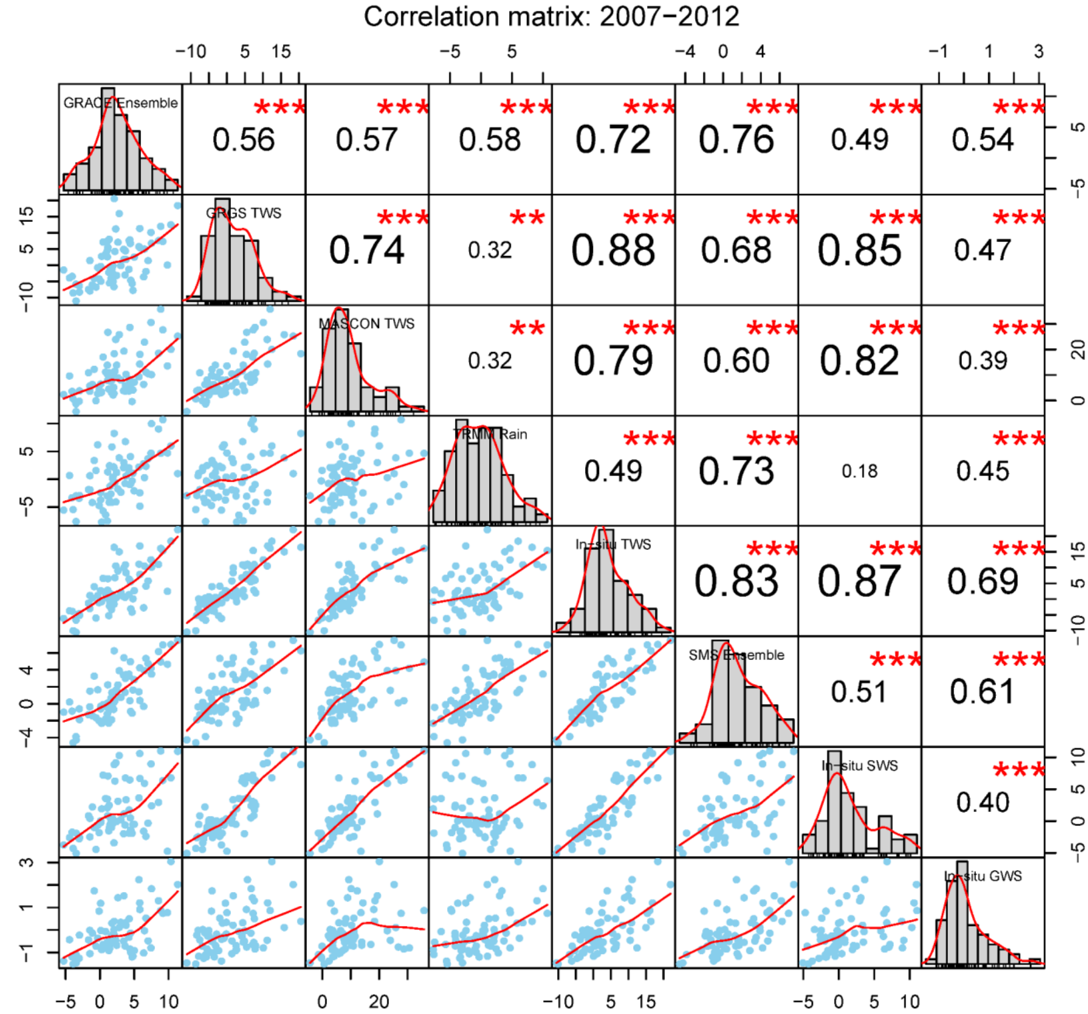


Figure ~~S4~~S7. Pearson correlation coefficients among the time-series variables collated over the Lake Victoria Basin for the period of 2007 to 2012. Statistically significant correlated variables are marked with asterisks where the significance asterisks represent *p*-values < 0.05 (1 asterisk), <0.01 (2 asterisks) and <0.001 (3 asterisks). Histograms with kernel density overlays and bivariate scatterplots of variables are shown. The fitted curve lines in bivariate scatterplots represent locally-weighted polynomial regression (i.e., Lowess) lines.

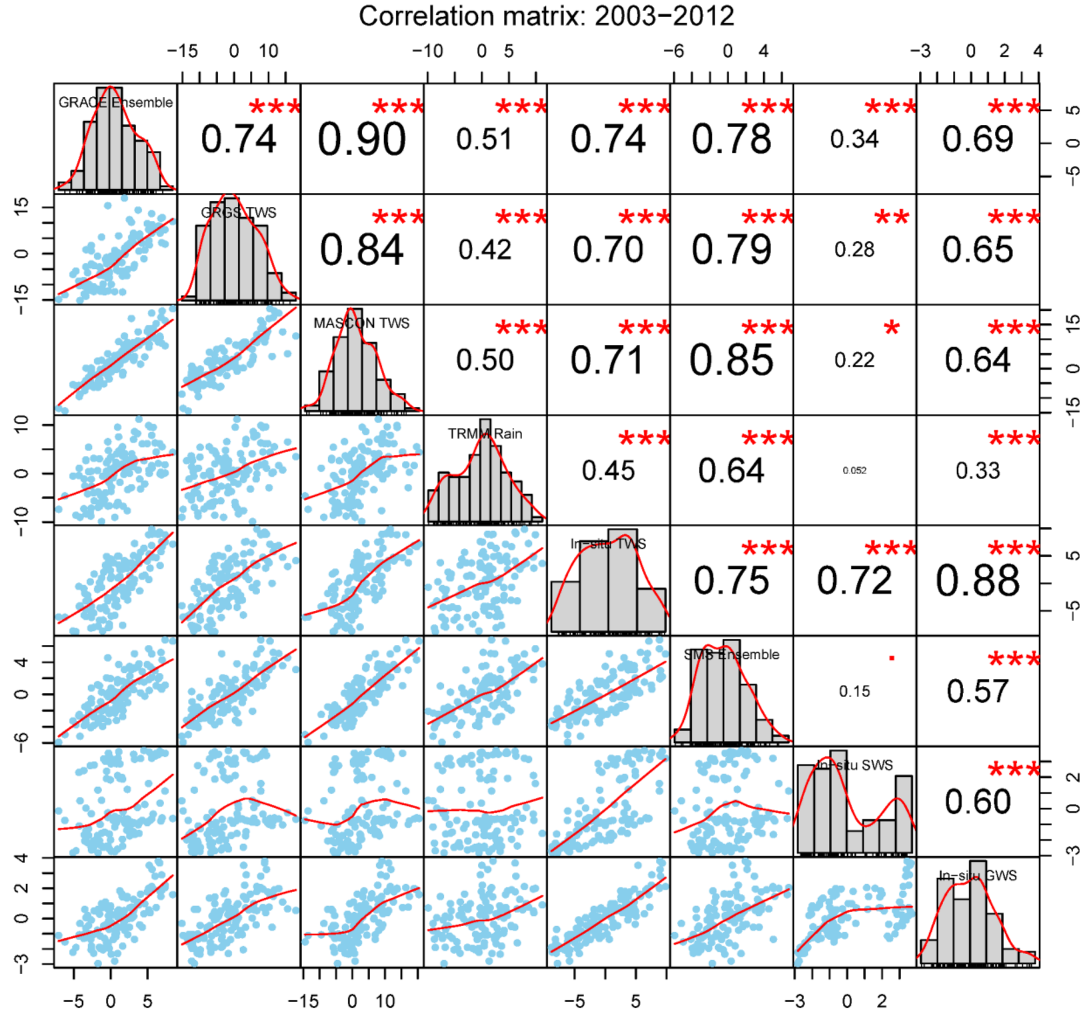


Figure ~~S5~~S8. Pearson correlation coefficients among the time-series variables collated over the ~~Victoria Nile~~Lake Kyoga Basin for the period of 2003 to 2012. Statistically significant correlated variables are marked with asterisks where the significance asterisks represent *p*-values < 0.05 (1 asterisk), <0.01 (2 asterisks) and <0.001 (3 asterisks). Histograms with kernel density overlays and bivariate scatterplots of variables are shown. The fitted curve lines in bivariate scatterplots represent locally-weighted polynomial regression (i.e., Lowess) lines.

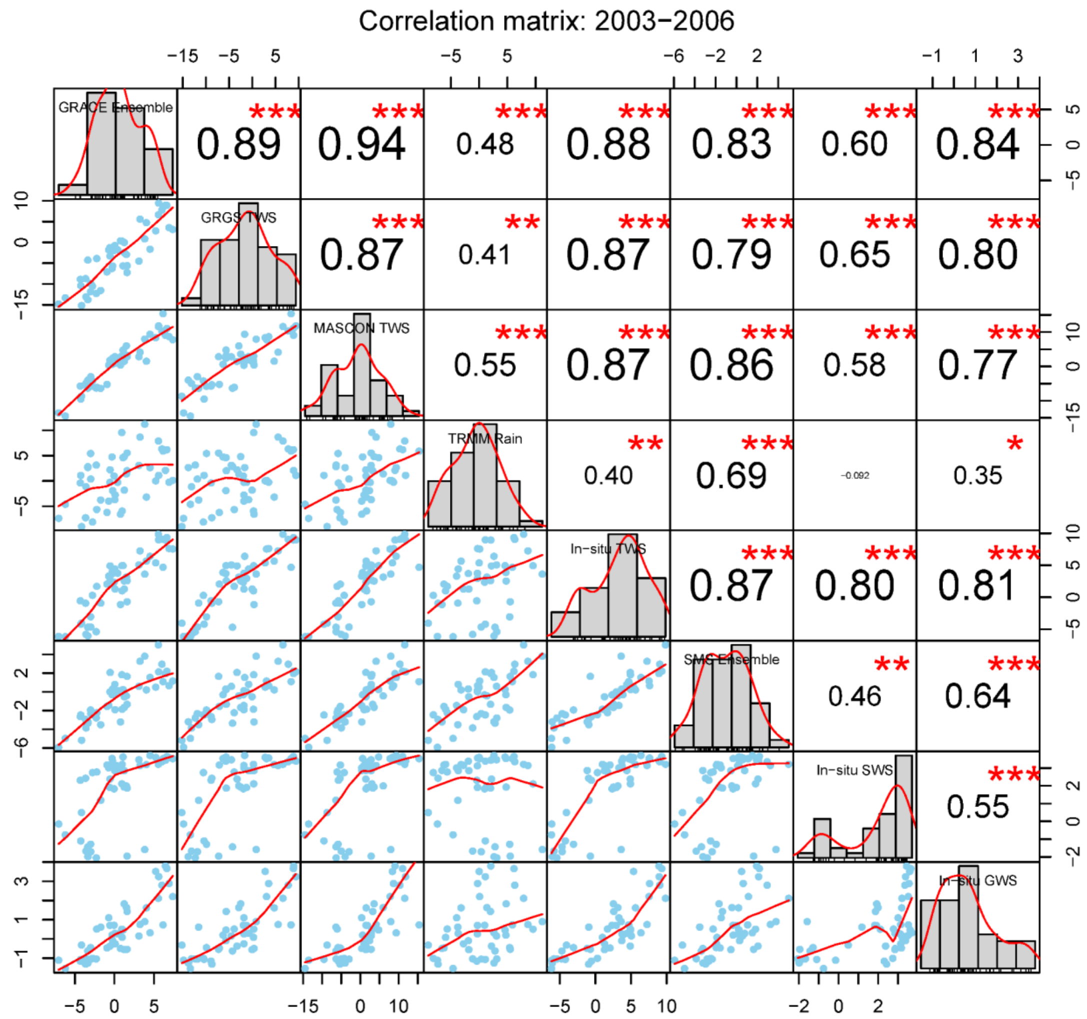


Figure ~~S6~~S9. Pearson correlation coefficients among the time-series variables collated over the Lake Kyoga~~Victoria Nile~~ Basin for the period of 2003 to 2006. Statistically significant correlated variables are marked with asterisks where the significance asterisks represent *p*-values < 0.05 (1 asterisk), <0.01 (2 asterisks) and <0.001 (3 asterisks). Histograms with kernel density overlays and bivariate scatterplots of variables are shown. The fitted curve lines in bivariate scatterplots represent locally-weighted polynomial regression (i.e., Lowess) lines.

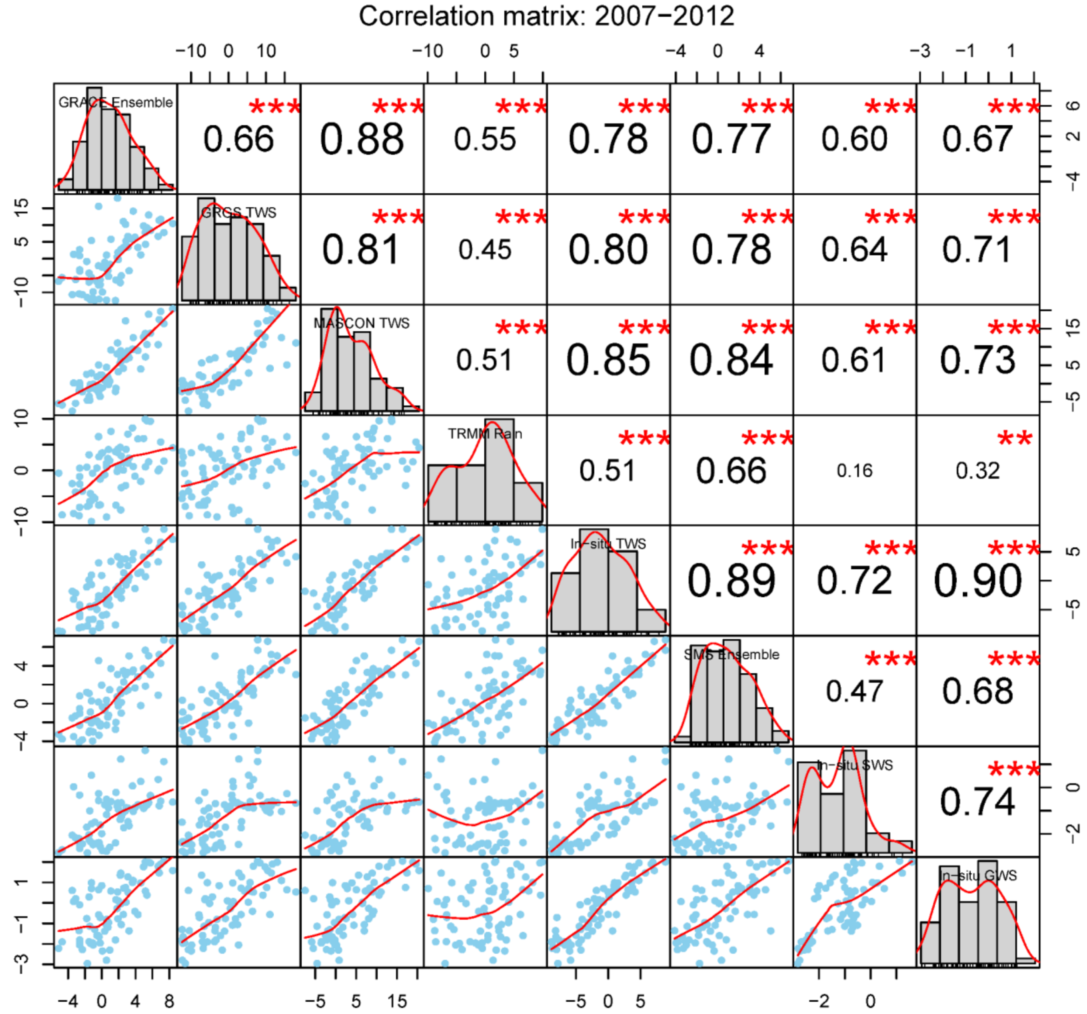


Figure ~~S7~~S10. Pearson correlation coefficients among the time-series variables collated over the Lake Kyoga~~Victoria Nile~~ Basin for the period of 2007 to 2012. Statistically significant correlated variables are marked with asterisks where the significance asterisks represent *p*-values < 0.05 (1 asterisk), <0.01 (2 asterisks) and <0.001 (3 asterisks). Histograms with kernel density overlays and bivariate scatterplots of variables are shown. The fitted curve lines in bivariate scatterplots represent locally-weighted polynomial regression (i.e., Lowess) lines.

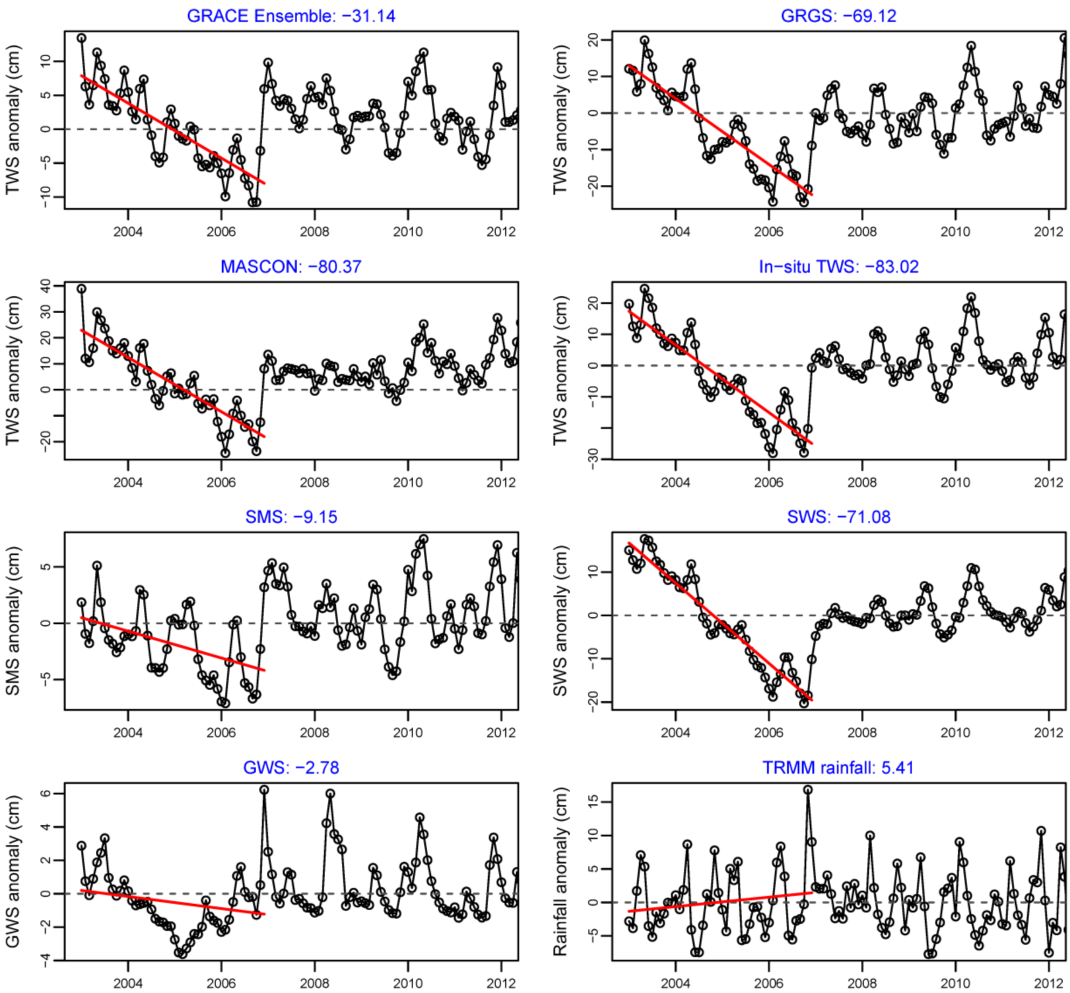


Figure ~~S8~~S11. Time-series records of various GRACE ΔTWS signals, in-situ ΔTWS, in-situ ΔSWS, simulated ΔSMS and in-situ ΔGWS in LVB and linear trends (red line) for the period of 20034 to 2006. Figure (blue) on top of each panel indicates the estimated storage change in km^3.

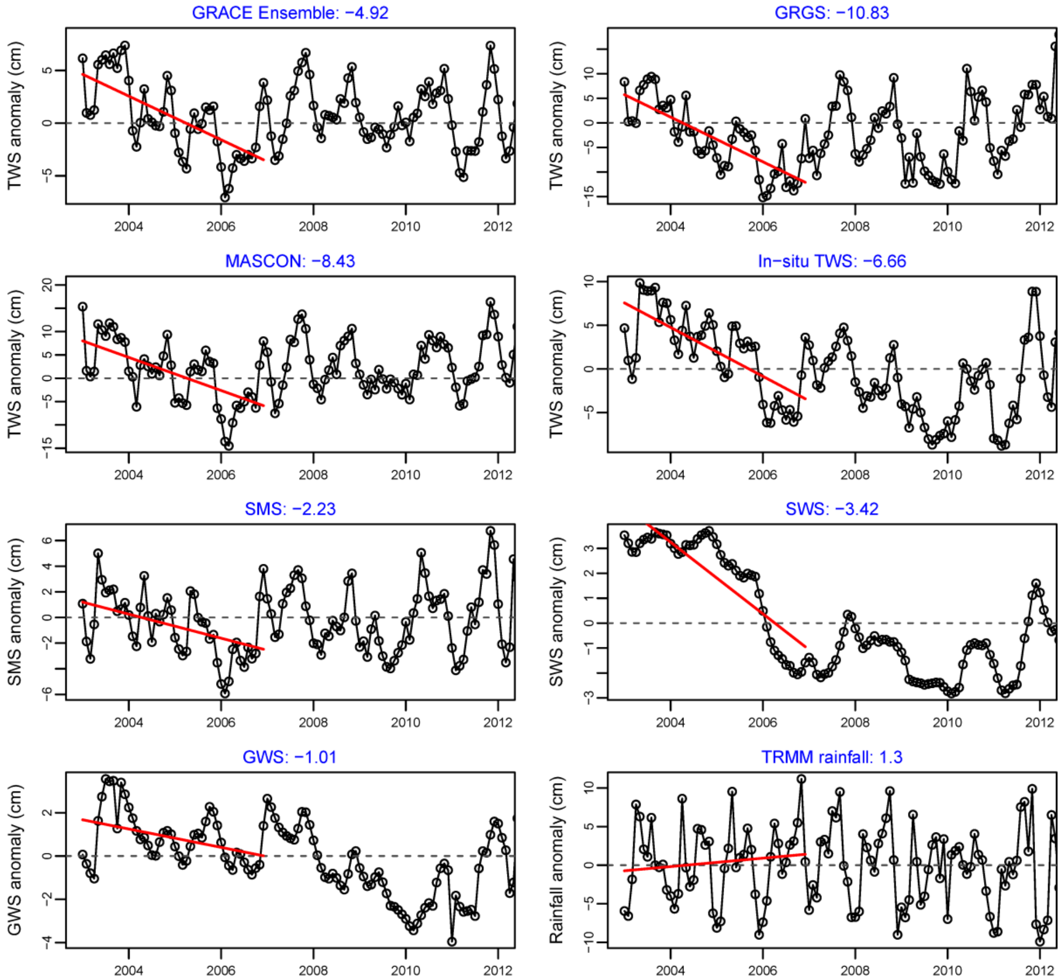


Figure S9S12. Time-series records of various GRACE ΔTWS signals, in-situ ΔTWS, in-situ ΔSWS, simulated ΔSMS and in-situ ΔGWS in VNB and linear trends (red line) for the period of 20034 to 2006. Figure (blue) on top of each panel indicates the estimated storage change in km^3.

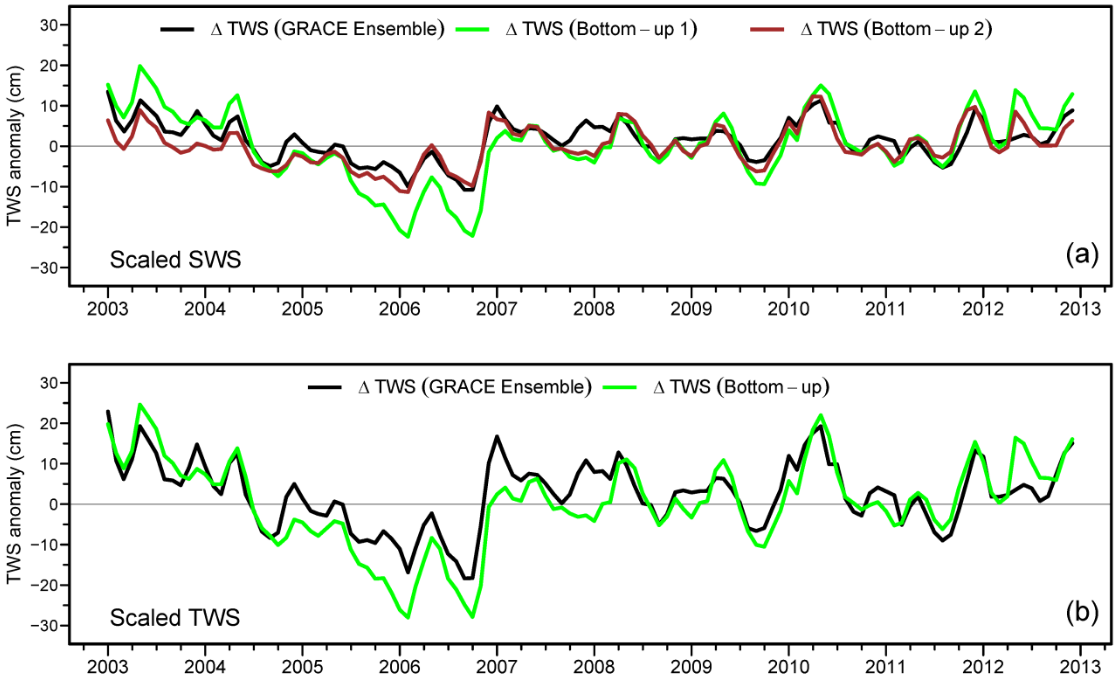


Figure ~~S10~~S13. Results of scaling experiments on *GRCTellus* GRACE and in situ ΔTWS over LVB. Panel (a) shows the comparison between *GRCTellus* GRACE-derived ΔTWS and ~~in situ~~bottom-up ΔTWS where a scaled down (scaling factor of 0.77 ~~in situ~~bottom-up ΔTWS-1; scaling factor of 0.11 ~~in situ~~bottom-up ΔTWS-2) ΔSWS signal is applied; on bottom panel (b) shows comparison between *GRCTellus* GRACE-derived ΔTWS and ~~in situ~~bottom-up ΔTWS where the GRACE-ΔTWS signal is scaled up by a factor of 1.7 ~~that shows~~based on the lowest RMSE of 5.76 cm with the bottom-up ΔTWS.

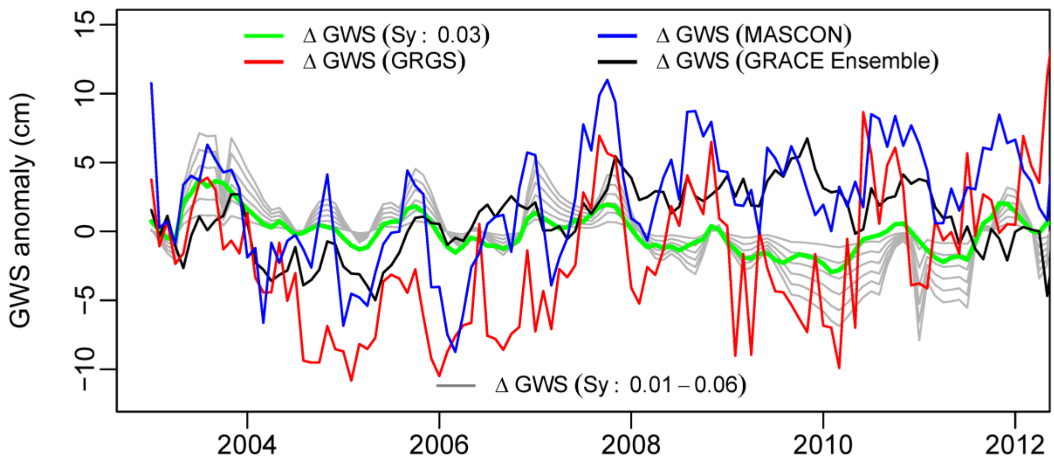


Figure ~~S11~~S14. Estimates of in situ ΔGWS and GRACE-derived ΔGWS time-series records (2003–2012) in LKB show a substantial variations among themselves. No scaling experiments were applied for LKB in the disaggregation of ΔGWS using *GRCTellus* (ensemble mean of CSR, GFZ, and JPL) and JPL-Mascons GRACE products.

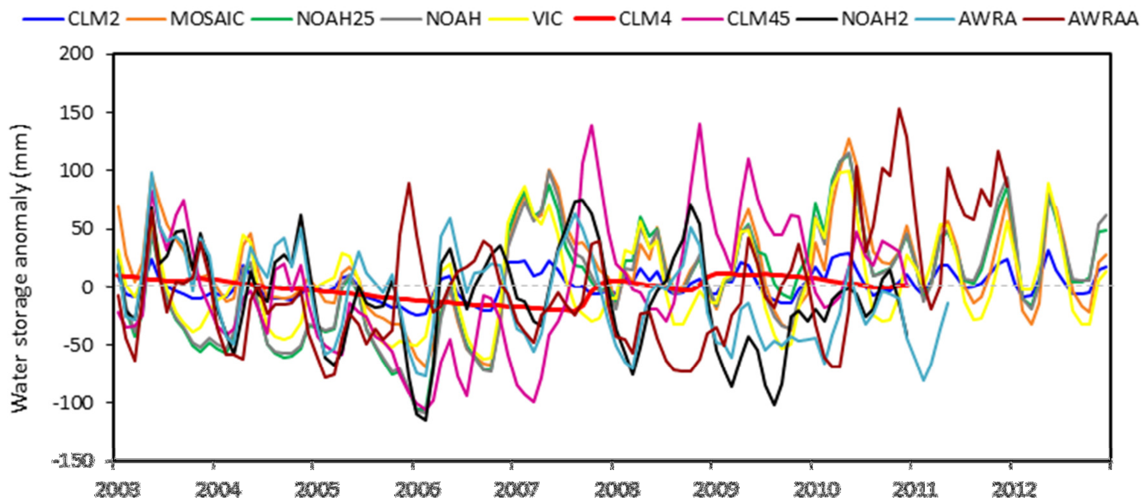


Figure ~~S12~~S15. Simulated terrestrial water storage anomaly from 10 LSMs for the Lake Victoria Basin. Note that not all LSMs simulate groundwater storage; for example, latest versions of the Community Land Model (CLM4.0 and 4.5) simulate groundwater storage.