Referee #1

I read once more a version of the paper and also the rebuttal letter in response to my first review. However, I am not satisfied with the provided answers on two major points raised in my first review. I am therefore recommending to ask the authors for another revision in order to give them the chance to properly address those concerns.

We thank reviewer 1 for the comments. Below are our responses (in blue) to new comments #1 and #2, including the original comments and replies.

Major Comment # 1

(Original review) I suggest that comparisons with the official ITSG2016 monthly solutions are included in order to demonstrate the added-value of the GC approach over the standard L2 data. Note that comparisons against GRGS or JPL monthly solutions as already (partly) included in the paper will not be sufficient since ITSG2016 is commonly perceived as a GRACE series of particularly high quality.

(Reply) The computation of Delta TWS from ITSG L2 solutions (like other L2 solutions) is subject to the post-processing filters and often followed by the application of empirical scaling factors, obtained from the land surface models. Our study provides a more rigorous way of computing Delta TWS without going through such ad hoc procedures. GRGS and JPL mascon are internally or post-processed and it is not expected users to apply (subjective) post- processing. This is the primary reason we validate our results with GRGS and JPL mascon.

(Second review) I disagree. The results of the new method needs to be compared exactly to conventionally post-processed gravity fields of the very same L2 data source in order to demonstrate the superiority of the proposed method. Conventional postprocessing at least implies restoring spectral deficits (degree 1, 2) and the removal of correlated errors (Kusche et al., 2009; or Swenson and Wahr, 2006). From my point of view, those are in no way just "ad hoc procedures".

(Reply) We added the comparison with the ITSG solutions in the Figure 6 of the revised manuscript. The revised Figure 6 now compares the ΔTWS results from ITSG (a), JPL/Mascon (b), and GRGS solutions (c), and this study GC (e). This particular example of ITSG is from the DDK5 filtering in addition to usual replacement of degree-1 and degree-2/order-0 coefficients. For the reviewer's information, we demonstrate how the different filters (for example, DDK1, DDK5, and DDK8 of Kusche et al., 2009) affect the computation of the storage in Figures S1 and S2.

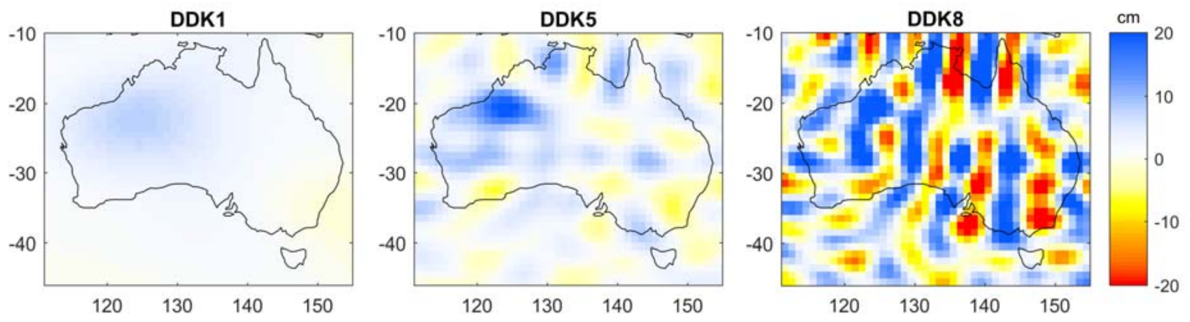

Figure S1: ΔTWS of April 2003, derived from the same ITSG solution with different post-processing filters (DDK1, DDK5, and DDK8) applied.

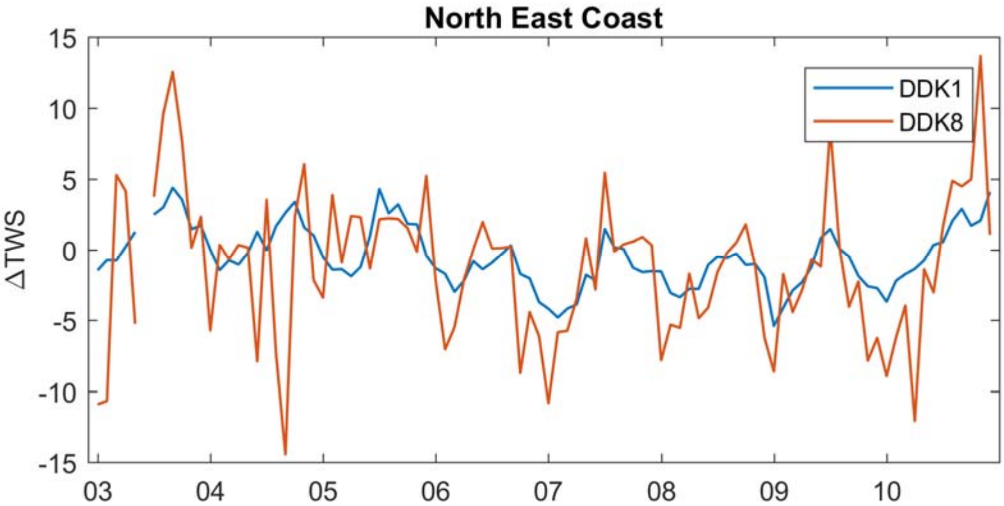

Figure S2: Basin average ΔTWS of the North East Coast basin computed from ITSG-DDK1 and ITSG-DDK8 solutions. The RMS difference between two results is ~ 5.3 cm, greater than the amplitude of DDK1.

Major Comment # 2

(Original review) The GC approach assumes that model errors are normally distributed with zero mean (eq. 1). Authors should provide more evidence that this assumption is indeed justified in their setting.

(Reply) R3: As the reviewer concerned, the GC approach is developed based on the least-square combination, which assumes the uncertainty following the normal distribution with zero mean and covariance C. The derivation and setting of model uncertainty under the given assumption (e.g., zero mean) and its limitation are clarified in Sect. 4.2.
"The GC approach assumes that model errors are normally distributed with zero mean. Any violation of this assumption will yield a bias in the combined solutions. Therefore, the mean value is removed from each ensemble member, $\boldsymbol{\mathcal{H}_R}' = \boldsymbol{\mathcal{H}_R} - \widetilde{\boldsymbol{h}}_{\boldsymbol{R}}$, and the error covariance matrix of the model is empirically computed as"

(Second review) My original request was to provide evidence that the assumption of normal distribution is indeed justified. This was by no means attempted by the presented changes to the manuscript.

(Reply) We thank the reviewer 1 for being patient to clarify the comment once again. The distribution of the model errors is demonstrated in Figure S3 in this letter as well as Figure 2 in the revised manuscript. The figure illustrates the histogram of the 210 ensemble members ($\boldsymbol{\mathcal{H}_R}$) for ΔSM_{top}, ΔSM_{rz}, and ΔGWS, after removing the respective means ($\widetilde{\boldsymbol{h}}_{\boldsymbol{R}}$). The normal distribution of the model error used in the GC approach is grounded on the distribution of $\boldsymbol{\mathcal{H}_R}'$ ($\boldsymbol{\mathcal{H}_R}' = \boldsymbol{\mathcal{H}_R} - \widetilde{\boldsymbol{h}}_{\boldsymbol{R}}$) or the histogram of the ensembles. We revised the manuscript and added this new Figure in Section 4.2.

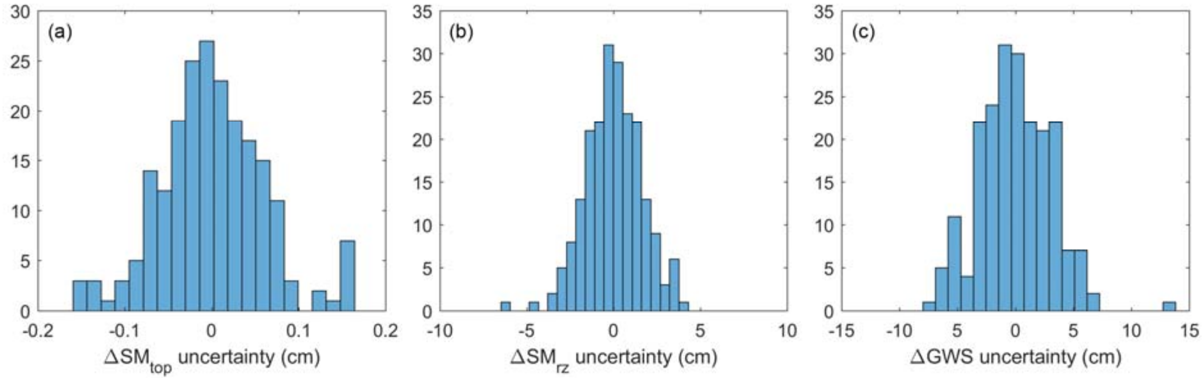

Figure S3: Histograms of the model errors ($\mathcal{H}_R{}'$) computed from 210 ensemble members without the mean. The basin averaged values (from all 10 Australian basins) of January 2003, for example, are used to compute the histogram.

Referee #3

This study introduces an inversion technique to use GRACE L1B data to improve the estimation of soil moisture and groundwater within Australia. As I mentioned in the previous round, I believe this line of research is interesting, however, I am not convinced that the proposed technique is well descried (with the current formulation it is not possible to re-do the work), and also it is not well justified (validation does not prove that the new method works any better than other available techniques). I recommend a reject / major review decision for this contribution.
We thank the reviewer 3 for the comments. Below is our responses (R, in blue):

Major Concerns:
A number of my previous comments are not adequately addressed
• The Methodology section needs to be specified, please add appendices to clearly how the equations are built. I cannot figure out how the normal equation is formulated, whether it includes KBRR and any orbital information? L120–> Please describe how the matrix A is derived and what are the entries. Similarly L128-L130 are unclear.
R1: As written in the original and revised manuscripts, we use the monthly least-square normal equation data from ITSG that are built upon GRACE L1B data products including inter-satellite ranging data. To clarify this further, we revised Section 2 of the manuscript by highlighting the use of the normal equations and moving the observation equation part to the appendix. The GRACE normal equation data are available from, as already given in the manuscript:
https://www.tugraz.at/institute/ifg/downloads/gravity-field-models/itsg-grace2016

• The accuracy of recovery has not been justified a synthetic study should be added to show the framework recovers the introduced gravity signals with a sufficient accuracy, see e.g., https://academic.oup.com/gji/article/158/3/813/2062077
• I cannot accept the following argument: "The approach optimally combines the GRACE's least-squares normal equations with CABLE to improve ΔSM and ΔGWS estimates." Therefore, the accuracy of the estimated result is only compared with the model-only result (please see Fig. 6). The accuracy of our TWS estimate can reach < 2 cm, which is in line with the GRACE accuracy of ~ 2 cm globally (Wahr et al., 2006).
R2: We are afraid that if we understand what is asked by this comment. Of course, our method and numerical codes were already tested with synthetic data sets, before applying it to the real data. We have verified the accurate recovery (within the numerical precision) from the synthetic data. This should be clear from our least-squares development as detailed in the manuscript. In addition, the comprehensive accuracy estimates of different water storage components is presented in Figure 7 of the revised manuscript.

A large part of the introduction is used to state other methods (e.g., assimilation and decomposition-based inversion) are erroneous. If this is true, reliable evidences must be provided to indicate the presented method works better than other technique
R3: The statements and related references regarding to signal decomposition were removed. For data assimilation, the evidences can be found in the references given (e.g., Girotto et al., 2016; Tian et al., 2017). In this introduction section, we only provide the general background of how ΔGWS has been estimated from GRACE, we do not claim that our presented method works better than other technique.

• L16: (e.g., time-variable gravity fields, i.e., Level 2 data, and ...
R4: The statement is modified.

• L21 remove ', not the post-processed Δ*T*WS grid data.'
R5: Removed.

• L27: ....combination maximizing the strength of the model and observations while suppressing the weaknesses. The approach ..... This is not well justified, in other words, the authors do not show combining the final grace products and models is worse than the proposed joint-inversion.

R6: Revised. "The GRACE-combine (GC) approach is developed for optimal least-squares combination and the approach is applied to estimate…".

• L33-34: ... estimates likely due to limitation of GRACE's temporal and spatial resolution... The way this has been validated in this study does not necessarily back this conclusion up.
R7: Revised. "Significant improvement is found in groundwater storage while marginal improvement is observed in surface soil moisture estimates".

• GRACE information rigorously and negate these limitations, this study uses the fundamental
R8: The statement is incomplete, and we do not take any action on this comment.

• L48-L51, Also L55, 59,etc.: before the references add 'e.g.,' as these references are only examples of existing researches.
R9: Done.

• L61: (Schumacher et al., 2016,2018; Tangdamrongsub et al., 2017)
Schumacher, M, Kusche, J & Döll, P, 2016, A systematic impact assessment of GRACE error correlations on data assimilation in hydrological models. Journal of Geodesy, vol 90., pp. 537-559
Schumacher, M, Forootan, E, van Dijk, A, Schmied, HM, Crosbie, R, Kusche, J & Döll, P, 2018, Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model. Remote Sensing of Environment, 204, pp212-228, doi:10.1016/j.rse.2017.10.029
R10: References are added.

• L80 'incorrect gravity information' --> does it mean all previous researches are wrong?
R11: The application of data tuning alters the original gravity information. This is self-explained.

• L81 ' began to employ' --> have applied the L2's (Schumacher etl al., 2016, 2018; Khaki et al., 2017 a,b)
Khaki, M., Hoteit, I., Kuhn, M., Awange, J., Forootan, E., van Dijk, A., Schumacher, M., Pattiaratchi, C. (2017a), Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model. Advances in Water Resources, 107, pages 301-316, doi:10.1016/j.advwatres.2017.07.001
Khaki, M., Schumacher, M., Forootan, E., Kuhn, M., Awange, J., van Dijk, A. (2017b), Accounting for spatial correlation errors in the assimilation of GRACE into hydrological models through localization. Advances in Water Resources, 108, pages 99-112, doi:10.1016/j.advwatres.2017.07.024
R12: References are added.

• L83 'still affected by the post-processing filter' --> This statement is very negative, and criticizes other processing strategies without providing a real measure. As far as I understand, in an assimilation formulation, incorporating the full co-variance matrix is already very important, useful and sufficient. Other post-processing steps have less impacts on the final assimilation results. See Schumacher et al 2018 and discussions for the Australian case study. If the authors suggest their approach is better than other formulations of signal separation, it should be validated and discussed.
R13: Revised. "Some recent studies began to employ the full variance-covariance information in the data assimilation scheme to enhance the quality of the estimates (Eicker et al., 2014, Schumacher et al., 2016; Tangdamrongsub et al., 2017; Khaki et al., 2017 a,b)".

• L86: It is not clear why this CABLE model has been selected
R14: CABLE provides comprehensive water storage component suitable for our analysis, and particularly the code and model parameter are publicly available. The additional sentence is added to the text to clarify this: "CABLE is a public available land surface model, and can be used to estimate soil moisture and groundwater in terms of volumetric water content …".

• After L92: (Dis-)Similarity of this work with previous studies that use inversion techniques might be addressed, see e.g., https://academic.oup.com/gji/article/158/3/813/2062077

R15: We thank for reviewer suggestion. The reference is addressed in the revised manuscript.

• Please also report on the impact of choices, within the gravity inversion, on the final mass estimation see e.g., http://www.sciencedirect.com/science/article/pii/S0264370716301016?via%3Dihub

R16: It is not clear what the reviewer means by this comment. There are various GRACE processing results around including the one the reviewer refers. It is not the focus of comparing all kinds of GRACE processing in this study, but demonstrating the way of using GRACE data with the actual covariance (normal equation) information.

• After L98: You might add a discussion reflecting the fact GRACE data is sensitive to many processes. You select the signal over Australia to simplify the computation etc.

R17: Revised. "One advantage of the study area is that the state vector can be defined mainly by soil moisture and groundwater as other hydrological components (e.g., glacier) are negligible.".

• L109: 'from a model' -->what does this mean? which model?

R18: Revised. "a land surface model".

• Equation5 --> matrices should be shown, e.g., add a appendix

R19: Revised. Please see R1.

• L162: How is the contribution of the signal from outside removed or corrected?

R20: The contribution of the signal from outside is removed using the GRGS solution. This was clearly described in Sect. 2 and Sect. 3.2 of the revised manuscript.

• Equation11 --> Please discuss the condition of these matrices used in the inversion.

R21: The matrix is well conditioned, and the condition number is around $10^4 – 10^5$.

• Co-variance estimation of the solution and whether it is representative of all measurements errors and model errors should be added.

R22: Fundamentally, the errors estimated from the least-squares combination present both model and observation errors (please see Eq. (7)). We present the covariance estimates (in terms of standard error) in Fig. 7 of the revised manuscript.

• 4.2 Model uncertainty --> From a hydrological point of view, model's errors should contain those uncertainties related to parameters, forcing data and model structure. The current errors do not reflect all these three categories and even the assumptions, that are used to define the distribution of model parameters, are not that sophisticated. Therefore, the impact of over-/under-estimation of the model's co-variance matrix on the final inversion results must be evaluated.

R23: In this study, the model error contains uncertainties related to parameters, forcing data and model structure, as the reviewer mentioned. This is written very clearly in Sect. 4.2. In this paper, the model forcing data (mainly precipitation) and parameters are both perturbed. We estimate the precipitation error based on 7 different products, which we believe it provides more realistic error compared to a simple assumption (e.g., 10-30 % of the value) seen in the previous publications and current practices (e.g., Eicker et al., 2014; Tangdamrongsub et al., 2015). The offline sensitivity study of forcing data is also conducted, and it is found that the water storage estimate is most sensitive to precipitation data, and relatively less sensitive to the change of other forcing components (this is written in Sect. 4.1). This is the main reason the precipitation is mainly perturbed. The parameters are perturbed based on the recommendation of the previous literature and the omission error is also included. In fact, most of previous literature (e.g., Zaitchik et al., 2008; Forman et al., 2012; Eicker et al., 2014; Tian et al., 2017, etc.) adopt a very similar procedure of model error determination we use here.

On the use of GRACE normal equation of intersatellite tracking data for improved estimation of soil moisture and groundwater in Australia

Natthachet Tangdamrongsub [1], Shin-Chan Han [1], Mark Decker [2], In-Young Yeo [1], Hyungjun Kim [3]

[1] School of Engineering, University of Newcastle, Callaghan, New South Wales, Australia
[2] ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney, New South Wales, Australia
[3] Institute of Industrial Science, the University of Tokyo, Tokyo, Japan

Abstract

An accurate estimation of soil moisture and groundwater is essential for monitoring the availability of water supply in domestic and agricultural sectors. In order to improve the water storage estimates, previous studies assimilated terrestrial water storage variation (ΔTWS) derived from Gravity Recovery and Climate Experiment (GRACE) into land surface models. However, the GRACE-derived ΔTWS was generally computed from the high level products (e.g., time-variable gravity fields, i.e., Level 2, and land grid from the Level 3 product). The gridded data products are subjected to several drawbacks such as signal attenuation and/or distortion caused by posteriori filters, and a lack of error covariance information. The post-processing of GRACE data might lead to the undesired alteration of the signal and its statistical property. This study uses the GRACE least-squares normal equation data to exploit the GRACE information rigorously and negate these limitations. Our approach combines the GRACE's least-squares normal equation (obtained from ITSG-Grace2016 product) with the results from the Community Atmosphere Land Exchange (CABLE) model to improve soil moisture and groundwater estimates. This study demonstrates, for the first time, an importance of using the GRACE raw data. The GRACE-combine (GC) approach is developed for optimal least-squares combination and the approach is applied to estimate the soil moisture and groundwater over 10 Australian river basins. The results are validated against the satellite soil moisture observation and the in-situ groundwater data. We demonstrate the GC approach delivers evident improvement of water storage estimates, consistently from all basins, yielding better agreement at seasonal and inter-annual time scales. Significant improvement is found in groundwater storage while marginal improvement is observed in surface soil moisture estimates.

1. Introduction

The changes of Terrestrial Water Storage (ΔTWS) derived from the Gravity Recovery And Climate Experiment (GRACE) data products have been used in the last decade to study global water resources, including groundwater depletion in India and Middle East (Rodell et al., 2009; Voss et al., 2013), water storage accumulation in Canada (Lambert et al., 2013), flood-influenced water storage fluctuation in Cambodia (Tangdamrongsub et al., 2016). The gravity data obtained from GRACE satellites are commonly processed and released in three different product levels (L) that increase in the amount of processing, L1B – satellite tracking data (e.g., Wu et al., 2006), L2 – global gravitational Stokes coefficients (e.g., Bettadpur, 2012), and L3 – global grids (e.g., Landerer and Swenson, 2012). The original (L1B) GRACE information is inevitably altered or sheered due to data processing and successive post-processing filterings, because the error covariance information is not propagated through each post-processing step.

The GRACE-derived ΔTWS has been computed widely from the higher-level products (e.g., L2 and L3) on which various ad hoc post-processing filters were applied (e.g., Gaussian smoothing filter (e.g., Jekeli, 1981), destripe filter (e.g., Swenson and Wahr, 2006)). ΔTWS obtained from these filters lacks proper error covariance information and is attenuated and distorted. To overcome the signal attenuation in GRACE high-level products, empirical approaches have been developed, including the application of scale factors computed from land surface models (Landerer and Swenson, 2012) to the GRACE L3 products. GRACE uncertainty in high level product is usually unknown or assumed. For example, Zaitchik et al. (2008) derived empirically a global average uncertainty that is variable depending on choices of post-processing filters (Sakumura et al., 2014). Furthermore, GRACE error and sensitivity is dependent on latitudes due to the orbit convergence toward poles (Wahr et al., 2006) and any post-processing filters will alter the GRACE data and their error information. Rigorous statistical error information is of equal importance to derivation of ΔTWS for data assimilation and model calibration (Tangdamrongsub et al., 2017; Schumacher et al., 2016, 2018). ΔTWS and its uncertainty estimates should be formulated directly from L1B data considering the complete statistical information.

The GRACE information is not fully exploited in many studies. For example, groundwater storage variation (ΔGWS) is often computed by subtracting the soil moisture variation (ΔSM) component simulated by the land surface model from GRACE-derived ΔTWS data (Rodell et al., 2009, Famiglietti et al., 2011), assuming the model ΔSM is error-free. This may result in the inaccurate ΔGWS and the associated error estimate as the uncertainties of observation and of the land surface model outputs are neglected in the combination (or regression) of two noisy data (e.g., Long et al., 2016). In data assimilation, the GRACE uncertainty is often derived empirically, not necessarily reflecting the actual GRACE error characteristics (e.g., Zaitchik et al., 2008; Tangdamrongsub et al., 2015; Tian et al., 2017). For example, Girotto et al. (2016) used L3 product and showed that it was necessary to adjust GRACE observation and its uncertainty in order to make their water storage estimates more accurate. Similarly, Tian et al. (2017) reported the need of applying a scale factor to GRACE uncertainty (from mascon product) in their GRACE assimilation process. It is apparent that the use of post-processed GRACE products often requires data tuning, leading possibly to an integration of incorrect gravity information into the data assimilation system. Some recent studies began to employ the full variance-covariance information in the data assimilation scheme to enhance the quality of the estimates (Eicker et al., 2014, Schumacher et al., 2016; Tangdamrongsub et al., 2017; Khaki et al., 2017 a,b).

This study aims to use the GRACE information of ΔTWS measurement directly from the raw L1B data. The approach optimally combines the GRACE's least-squares normal equations with the model simulation results from the Community Atmosphere Land Exchange (CABLE, Decker, 2015) to improve ΔSM and ΔGWS estimates. The proposed approach presents three main advantages. Firstly, one can exploit the full GRACE signal and error information by using the normal equation data sets. Secondly, the approach is developed for optimal least-squares combination (e.g., Ramillien et al., 2004), which maximizes the model and observation strength while simultaneously supressing their weaknesses. Finally, the method bypasses empirical, multiple-step post-processing filters.

The main objective of this study is to present the GRACE-combined (GC) approach to estimate improved ΔSM and ΔGWS at regional scales. We demonstrate our approach applied to 10 Australian river basins (Fig. 1a). One advantage of the study area is that the state vector can be defined mainly by ΔSM and ΔGWS as other hydrological components (e.g., snow, glacier) are negligible. We validate the top layer of ΔSM estimates against the satellite soil moisture observation (the Advanced Microwave Scanning Radiometer aboard EOS (AMSR-E), Njoku et al., 2003) over all 10 basins and the ΔGWS estimates against the in-situ groundwater data available over Queensland and Victoria (Fig. 1b, 1c).

This paper is outlined as follows: Firstly, the derivation of GC approach is presented in Sect. 2 while the description of GRACE data processing, including the use of GRACE normal equation, is given in Sect. 3. Secondly, the CABLE modelling is outlined in Sect. 4. This includes the derivation of model uncertainty based on the quality of precipitation data and the model parameter inputs. The processing of validation data is also described in Sect. 4. Thirdly, Sect. 5 presents the result of ΔSM and ΔGWS estimates and comparison to in-situ data. The long-term trends in the Australian mass variation over the last 13 years is also investigated in this section.

2. A method of combining GRACE L1B data with land surface model outputs

The statistical information of ΔTWS computed from a land surface model can be written as:

$$\tilde{\boldsymbol{h}} = \boldsymbol{h} + \boldsymbol{\epsilon}; \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \mathbf{C}), \tag{1}$$

where $\boldsymbol{h}$ is the "truth" (unknown) model state vector while $\tilde{\boldsymbol{h}}$ is the calculated state vector characterized with the model error $\boldsymbol{\epsilon}$. The model error is assumed to have zero mean and covariance $\mathbf{C}$.

The term $\boldsymbol{h}$ is used to represent a vector including global ΔTWS grid, and terms with a subscript R (e.g., $\boldsymbol{h_R}$, $\mathbf{C}_{\boldsymbol{R}}$) is used to represent only a regional set of ΔTWS (for example, in Australia). As such, the observation equation over a region can be rewritten as:

$$\tilde{\boldsymbol{h}}_{\boldsymbol{R}} = \boldsymbol{h_R} + \boldsymbol{\epsilon}; \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \mathbf{C}_{\boldsymbol{R}}). \tag{2}$$

As soil moisture and groundwater are the major components of ΔTWS in Australia (surface water storage being insignificant), the vector $\boldsymbol{h_R}$ can be defined as:

$$\boldsymbol{h_R} = [\Delta \boldsymbol{SM}_{top} \quad \Delta \boldsymbol{SM}_{rz} \quad \Delta \boldsymbol{GWS}]^{\boldsymbol{T}}, \tag{3}$$

where $\Delta \boldsymbol{SM}_{top}$, $\Delta \boldsymbol{SM}_{rz}$, $\Delta \boldsymbol{GWS}$ represent the vectors of top (surface) soil moisture, root zone soil moisture, and groundwater storage variations, respectively.

A least-squares normal equation of GRACE can be written as:

$$\mathbf{N}\,\boldsymbol{x} = \boldsymbol{c} \tag{4}$$

Where $\mathbf{N}$ is a normal matrix, $\boldsymbol{x}$ contains the spherical harmonic coefficients (SHC) of the geopotential, and $\boldsymbol{d}$ is the normal vector. In this study, $\mathbf{N}$ and $\boldsymbol{c}$ can be obtained from the

ITSG-Grace2016 products (Mayer-Gürr et al, 2016; https://www.tugraz.at/institute/ifg/downloads/gravity-field-models/itsg-grace2016, see more details in Sect. 3.1). Eq. (4) can be written in terms of $\boldsymbol{h}$ as follows (see Appendix A for the derivation):

$$(\mathbf{H}^T\mathbf{Y}^T\mathbf{N}\mathbf{Y}\mathbf{H})\widehat{\boldsymbol{h}} = \mathbf{H}^T\mathbf{Y}^T\boldsymbol{c} \qquad (5)$$

where $\mathbf{Y}$ converts ΔTWS to geopotential coefficients considering the load Love numbers (e.g., Wahr et al., 1998) and $\mathbf{H}$ is the operational matrix converting $\Delta \boldsymbol{SM}_{top}$, $\Delta \boldsymbol{SM}_{rz}$, and $\Delta \boldsymbol{GWS}$ to ΔTWS. Eq. (5) is based on the assumption that the GRACE orbital perturbation is a result of ΔTWS variation on the surface. If M is the number of model grid cells, $N_{\max}$ is the maximum degree of the geopotential coefficients, and $L=(N_{\max}+1)^2-4$ is the number of geopotential coefficients from GRACE, the dimension of $\mathbf{Y}$, $\mathbf{H}$, and $\boldsymbol{h}$ are $L\times M$, $M\times 3M$, and $3M\times 1$, respectively. Note that, Eq. (5) is defined with the global grid of $\boldsymbol{h}$. For a regional application, Eq. (5) can be modified as:

$$\left[\mathbf{H}_R^T\mathbf{Y}_R^T \mid \mathbf{H}_o^T\mathbf{Y}_o^T\right] \mathbf{N} \begin{bmatrix}\mathbf{Y}_R\mathbf{H}_R \\ \mathbf{Y}_o\mathbf{H}_o\end{bmatrix} \begin{bmatrix}\widehat{\boldsymbol{h}}_R \\ \widehat{\boldsymbol{h}}_o\end{bmatrix} = \left[\mathbf{H}_R^T\mathbf{Y}_R^T \mid \mathbf{H}_o^T\mathbf{Y}_o^T\right] \boldsymbol{c}, \qquad (6)$$

where the subscript R indicates the grid ΔTWS only in a region of interest, and o for the rest of the globe. If the number of the model grid cells associated with R is J and that of the outside cells is M–J. As such, the dimensions of $\mathbf{Y}_R$, $\mathbf{H}_R$, $\widehat{\boldsymbol{h}}_R$, $\mathbf{Y}_o$, $\mathbf{H}_o$, $\widehat{\boldsymbol{h}}_o$ are $L\times J$, $J\times 3J$, $3J\times 1$, $L\times (M-J)$, $(M-J)\times 3(M-J)$, $3(M-J)\times 1$, respectively. The dimension of $\mathbf{N}$ and $\boldsymbol{c}$ remain unchanged, since they are essentially from the normal equations of the original GRACE L1B data (to be discussed in the following section).

From Eq. (6), the normal equations associated with ΔTWS in the region of interest can then be written as

$$\mathbf{H}_R^T\mathbf{Y}_R^T\mathbf{N}\mathbf{Y}_R\mathbf{H}_R\widehat{\boldsymbol{h}}_R = \mathbf{H}_R^T\mathbf{Y}_R^T\boldsymbol{c} - \mathbf{H}_R^T\mathbf{Y}_R^T\mathbf{N}\mathbf{Y}_o\mathbf{H}_o\widehat{\boldsymbol{h}}_o \qquad (7)$$

or

$$\mathbf{N}_R\widehat{\boldsymbol{h}}_R = \boldsymbol{c}_R \qquad (8)$$

where $\mathbf{N}_R = \mathbf{H}_R^T\mathbf{Y}_R^T\mathbf{N}\mathbf{Y}_R\mathbf{H}_R$ and $\boldsymbol{c}_R = \mathbf{H}_R^T\mathbf{Y}_R^T\boldsymbol{c} - \mathbf{H}_R^T\mathbf{Y}_R^T\mathbf{N}\mathbf{Y}_o\mathbf{H}_o\widehat{\boldsymbol{h}}_o$. As seen, Eq. (7) is the regional representation of Eq. (5) where only the grid cells inside the study region are used, while the contribution from the grid cells outside the region needs to be removed or corrected. Combining the normal equation of Eq. (2) and Eq. (8), the optimal combined solution of $\widehat{\boldsymbol{h}}_R$ can be resolved as follows:

$$\widehat{\boldsymbol{h}}_R = \left(\mathbf{C}_R^{-1} + \mathbf{N}_R\right)^{-1}\left(\mathbf{C}_R^{-1}\widetilde{\boldsymbol{h}}_R + \boldsymbol{c}_R\right) \qquad (9)$$

The computation of model covariance matrix $\mathbf{C}_R$ will be discussed in Sect. 4.2. The posteriori covariance of $\widehat{\boldsymbol{h}}_R$ can be estimated as follows:

$$\widehat{\boldsymbol{\Sigma}} = (\mathbf{C}_R^{-1} + \mathbf{N}_R)^{-1}, \qquad (10)$$

and the uncertainty estimate of $\widehat{\boldsymbol{h}}_R$ is simply calculated as:

$$\sigma_{\hat{h}} = \sqrt{diag(\hat{\Sigma})}, \quad (11)$$

where $diag()$ represents the diagonal element of the given matrix.

3. GRACE data

3.1 GRACE least-squares normal equations

In this study, the least-squares normal equations are obtained from the ITSG-Grace2016 products between January 2003 and March 2016. All L1B data including KBR inter-satellite tracking data, attitude, accelerometer, GPS based kinematic orbit data and AOD1B corrections are reduced in terms of the normal equations. These data products are usually used to compute the Earth's geopotential field to the maximum harmonic degree and order of 90, or at a spatial resolution of ~220 km. The products contain the information of the normal matrix $\mathbf{N}$ and the vector $\boldsymbol{c}$ (as shown in Eq. (4)) as well as the a-priori time-varying gravity field coefficients predicted with the GOCO05s solution (Mayer-Gürr et al., 2015). Note that the solution of the ITSG-Grace2016 normal equation is the anomalous geopotential coefficient vector ($\Delta\boldsymbol{x}$), which is referenced to the a-priori time-varying gravity field ($\boldsymbol{x_0}$), through:

$$\mathbf{N}\,\Delta\boldsymbol{x} = \boldsymbol{d} \quad (12)$$

where $\boldsymbol{d}$ and $\boldsymbol{x_0}$ are given. To obtain a complete gravity field variation between the study period ($\boldsymbol{x}$ term in in Eq. (4)), the a-priori time-varying gravity field, $\boldsymbol{x_0}$ is firstly restored to Eq. (12), and the mean gravity field ($\overline{\boldsymbol{x}}_0$) computed from all $\boldsymbol{x_0}$ between January 2003 and March 2016 is then removed as follows:

$$\mathbf{N}\,(\Delta\boldsymbol{x} + \boldsymbol{x_0} - \overline{\boldsymbol{x}}_0) = \boldsymbol{d} + \boldsymbol{N}(\boldsymbol{x_0} - \overline{\boldsymbol{x}}_0) \quad (13)$$

$$\mathbf{N}\,\boldsymbol{x} = \boldsymbol{d} + \boldsymbol{N}(\boldsymbol{x_0} - \overline{\boldsymbol{x}}_0) \quad (14)$$

Therefore, in Sect. 2 (e.g., Eq. (5)), the matrix $\mathbf{N}$ remains unchanged while the vector $\boldsymbol{c}$ can be simply replaced by $\boldsymbol{c} = \boldsymbol{d} + \boldsymbol{N}(\boldsymbol{x_0} - \overline{\boldsymbol{x}}_0)$.

3.2 GRACE-derived ΔTWS products

Three monthly GRACE-derived ΔTWS products are also used, the ITSG-Grace2016 DDK5 solution (ITSG-DDK5 for short, http://icgem.gfz-potsdam.de/series/99_non-iso/ITSG-Grace2016), the CNES/GRGS Release 3 (RL3) (GRGS for short, Lemoine et al., 2015; http://grgs.obs-mip.fr/grace/variable-models-grace-lageos/grace-solutions-release-03) and the JPL RL05M mascon-CRI version 2 product (mascon for short, Watkins et al., 2015; Wiese et al., 2016; http://grace.jpl.nasa.gov/data/get-data/jpl_global_mascons). The ITSG-DDK5 product is the post-processed version of the ITSG L2 solution where the non-isotropic filter DDK5 (Kusche et al., 2009) is applied. The DDK5 solution is empirically selected here to be a good balance between the over-smoothed (e.g., DDK1) and noisy (e.g., DDK8) solutions. The GRGS solution provides ΔTWS at 1°×1° globally, derived from the Earth's geopotential coefficients up to the maximum degree and order 80, and no filter nor scale factor is applied (L2 data product). Mascon provides ΔTWS at equal-area 3° spherical cap grid globally. In contrast to the ITSG-DDK5 and GRGS solutions, the mascon uses a gain factor derived from the land surface model (LSM) to restore mitigated signals and reduce leakage errors (L3 data products) (Watkins et al., 2015; Wiese et al., 2016). Additionally, mascon provides the ΔTWS uncertainty together with the solution. The uncertainty is computed based on several geophysical models (see Watkins et al. (2015) and Wiese et al. (2016) for more details). The uncertainty information is not available in the ITSG-DDK5 or GRGS product.

The GRACE data are obtained between January 2003 and March 2016. After retrieval, the long-term mean value between January 2003 and March 2016 is computed and subtracted from the monthly products. To be consistent with CABLE grid spacing (see Sect. 4), the ΔTWS is computed using 0.5° spatial resolution. The coarse scale datasets (e.g., mascon, GRGS) are resampled to 0.5°×0.5° using the nearest grid values.

In this study, the independent GRACE solutions are used for two main reasons:

1. To obtain the ΔTWS values outside Australia. As shown in Eq. (7), the $\boldsymbol{\hat{h}_o}$ vector needs to be known, which can be from the GRACE-derived ΔTWS solution. We use the GRGS solutions as the GRGS solution is not subject to the filter choice and it provides ΔTWS at a spatial resolution comparable to the normal equation data.
2. To compare with the ΔTWS estimates from our approaches. All solutions are used to compare and validate our ΔTWS estimates.

4. Hydrology model and validation data

4.1 Model setup

The extensive description of the CABLE model is given in Decker (2015) and Ukkola et al. (2016). This section describes the model setup and specific changes applied to this study. CABLE is a public available land surface model and can be used to estimate soil moisture and groundwater in terms of volumetric water content every 3 hours at a 0.5°×0.5° spatial resolution. The soil moisture and groundwater storage can be simply computed by multiplying the estimates with thicknesses of various layers. For soil moisture, the thickness of 6 soil layers is 0.022, 0.058, 0.154, 0.409, 1.085, and 2.872 m, from top to bottom, respectively. The thickness of the groundwater layer is modeled to be 20 m uniformly. Recalling Eq. (3), $\boldsymbol{\Delta SM_{top}}$ is defined as the soil moisture storage variation at the top 0.022 m thick layer, while $\boldsymbol{\Delta SM_{rz}}$ is the variation accumulated over the second to the bottom soil layers (depth between 0.022 m and 4.6 m).

CABLE is initially forced with the data from the Global Soil Wetness Project Phase 3 (GSWP3), which is currently available until December 2010 (http://hydro.iis.u-tokyo.ac.jp/GSWP3, https://doi.org/10.20783/dias.501). We replace GSWP3 forcing data with GLDAS data (Rodell et al., 2004) to compute the water storage changes to 2016. The forcing data used in CABLE are precipitation, air temperature, snowfall rate, wind speed, humidity, surface pressure, and short-wave and long-wave downward radiations. To investigate the impact of different forcing data, the offline sensitivity study is conducted by comparing the water storage estimates computed using:

1. All 8 forcing data components of GSWP3,

2. GSWP3 data with replacing one component obtained from GLDAS forcing data.

It is found that the water storage estimate is most sensitive to the replacement of precipitation data, as expected, and relatively less sensitive to the change of other forcing components. We use the GLDAS forcing data in this study and also further test 7 different precipitation data products (see more details in Sect. 4.2). The forcing data are up/down sampled to a 0.5°×0.5° spatial grid to reconcile with the CABLE spatial resolution.

4.2 Model uncertainty

In this study, the CABLE uncertainty is derived from 210 ensemble estimates associated with different forcing data and model parameters. The 7 different precipitation products (see Table 1) are used to run the model independently. Most products are available to present day while GSWP3, Princeton, and MERRA are only available until December 2010, December 2012, and February 2016, respectively. For each precipitation forcing, 30 ensembles are generated by perturbing the model parameters within +/– 10% of the nominal values. The perturbed size of 10% is similar to Dumedah and Walker (2014). Based on the CABLE structure, the $\Delta\boldsymbol{SM}$ and $\Delta\boldsymbol{GWS}$ estimates are most sensitive to the model parameters listed in Table 2. For example, the fractions of clay, sand, and silt (f_{clay}, f_{sand}, f_{silt}) are used to compute soil parameters including field capacity, hydraulic conductivity, and soil saturation which mainly affect soil moisture storage. Similarly, the drainage parameters (e.g., q_{sub}, f_p) control the amount of subsurface runoff, which has a direct impact on root zone soil moisture and groundwater storages.

From ensemble generations, total $K = 210$ sets of the ensemble water storage estimates ($\boldsymbol{h_e}$) are obtained:

$$\boldsymbol{\mathcal{H}_R} = [\boldsymbol{h_e}|_{k=1} \quad \boldsymbol{h_e}|_{k=2} \quad \boldsymbol{h_e}|_{k=3} \quad \dots \quad \boldsymbol{h_e}|_{k=K}] \tag{15}$$

and the mean value of $\boldsymbol{\mathcal{H}_R}$ is computed as follows:

$$\tilde{\boldsymbol{h}}_{\boldsymbol{R}} = \frac{1}{K}\sum_{k=1}^{K}\boldsymbol{h_e}|_k \tag{16}$$

Note that due to the absence of GSWP3, Princeton, and MERRA data, the number of ensembles reduces to $K = 180$ after December 2010, $K = 150$ after December 2012, and $K = 120$ after February 2016, respectively. The GC approach assumes that model errors are normally distributed with zero mean. Any violation of this assumption will yield a bias in the combined solutions. Therefore, the mean value is removed from each ensemble member, $\boldsymbol{\mathcal{H}_R}' = \boldsymbol{\mathcal{H}_R} - \tilde{\boldsymbol{h}}_{\boldsymbol{R}}$, and the error covariance matrix of the model is empirically computed as:

$$\mathbf{C}_{\boldsymbol{R}} = \boldsymbol{\mathcal{H}_R}'(\boldsymbol{\mathcal{H}_R}')^T/(K-1) \tag{17}$$

The $\tilde{\boldsymbol{h}}_{\boldsymbol{R}}$ (Eq. (16)) and $\mathbf{C}_{\boldsymbol{R}}$ (Eq. (17)) terms can be directly used in Eq. (9). The distribution of model errors is demonstrated in Fig. 2. The figure illustrates the histogram of model errors ($\boldsymbol{\mathcal{H}_R}'$) computed using 210 ensemble members of the model estimated $\Delta\boldsymbol{SM}$ and $\Delta\boldsymbol{GWS}$ in Jan 2003. The histogram indicates that the model error may be approximately described by a normal distribution as introduced in Eq. (1).

Furthermore, in practice, the sampling error caused by finite sample size might lead to spurious correlations in the model covariance matrix (Hamill et al., 2001). The effect can be reduced by applying an exponential decay with a particular spatial correlation length to $\mathbf{C}_{\boldsymbol{R}}$. In this study, the correlation length is determined based on the empirical covariance of model estimated ΔTWS. The covariance function of ΔTWS is firstly assumed isotropic, and it is computed empirically based on the method given in Tscherning and Rapp (1974). The distance where the maximum value of the variance decreases to half is defined as the correlation length. The obtained values vary month-to-month, and the mean value of 250 km is used in this study.

It is emphasized that the model omission error caused by imperfect modeling of hydrological process within the LSM is not taken into account in the above description. The omission error may increase the model covariance and introduce a bias as well. We account for the omission error by increasing 20% of the model covariance. (i.e., multiplying $\mathbf{C}_{\boldsymbol{R}}$ by 1.2). We determine such omission error based on trial-and-error such that it increases the model error (due to the omission error) but not exceeds the model error value reported by Dumedah and Walker (2014). We acknowledge that this is only a simple practical way of accounting for the omission error into the total model error.

4.3 Validation data

4.3.1 Satellite soil moisture observation

The satellite observed surface soil moisture data is obtained from the Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E) using the Land Parameter Retrieval Model (Njoku et al., 2003). The observation is used to validate our estimates of top soil moisture changes (ΔSM_{top}). The AMSR-E product provides volumetric water content in the top layer derived from a passive microwave data (from NASA EOS Aqua satellite) and forward radiative transfer model. In this study, the level 3 product, available daily between June 2002 and June 2011 at 0.25°×0.25° spatial resolution is used (Owe et al., 2008). The measurements from ascending and descending overpasses are averaged for each frequency band (C and X). Then, the monthly mean value is computed by averaging the daily data within a month. To obtain the variation of the surface soil moisture, the long-term mean between June 2002 and June 2011 is removed from the monthly data. Regarding the different depth measured in CABLE and AMSR-E, the CDF-matching technique (Reichle and Koster, 2004) is used to reduce the bias between the top soil moisture model and the observation. The CDF is built using the 2003-2004 data, and it is used for the entire period. There is no satellite observed or ground measured root zone soil moisture data for meaningful comparison with our results, particularly at continental scale. Validation of ΔSM_{rz} at regional and continental scales is currently unachievable due to a complete lack of observations at this spatial scale.

4.3.2 In-situ groundwater

The in-situ groundwater level from bore measurements are obtained from 2 different ground observation networks (see Fig. 1). The data in Queensland are obtained from Department of Natural Resources and Mines (DNRM) while the data in Victoria is from Department of Environment and Primary Industries (DWPI). More than 10,000 measurements are available from each network, but the data gap and outliers are present. Therefore, the bore measurement is firstly filtered by removing the sites that present no data or data gap longer than 30 months during the study period.

To obtain the monthly mean value, the hourly or daily data are averaged in a particular month. The outliers are detected and fixed using the Hampel filter (Pearson, 2005) where the remaining data gaps are filled using the cubic spline interpolation. To obtain the groundwater level variation, the long-term mean groundwater level computed between the study period is removed from the monthly values. The groundwater level variation (ΔL) is then converted to ΔGWS using $\Delta GWS = S_y \cdot \Delta L$, where S_y is specific yield. Based on Chen et al. (2016), $S_y = 0.1$ is used for the Victoria network. Specific yields of Queensland's network have been found ranging from 0.045 (Rassam et al., 2013) to 0.06 (Welsh 2008), and an averaged $S_y = 0.05$ is used in this study. Finally, the mean value computed from all data (in each network) is used to represent the in-situ data of the network.

5. Results

5.1 Model-only performance

We study the model ΔTWS changes under different meteorological forcing and land parameterization. Total 210 estimates of monthly TWS (sum of SM_{top}, SM_{rz}, and GWS) are obtained between January 2003 and March 2016 from the ensemble run based on 7 different precipitation inputs. Then, the averaged values of the TWS estimates are computed from the 30 precipitation-associated ensemble members. This results in 7 sets of monthly mean TWS estimates from 7 different precipitation data. For each set, the monthly ΔTWS is computed by removing the long-term mean computed between January 2003 and March 2016.

The precipitation-based ΔTWS are then compared with the GRACE-mascon solution (see Sect. 3.2) over 10 different Australian basins. The comparison is carried out between January 2003 and March 2016. Due to the availability of the data, the periods used are shorter in cases of GSWP3, Princeton, and MERRA precipitation (see Table 1). The metric used to evaluate a goodness of fit between CABLE run and GRACE mascon estimates is the Nash-Sutcliff (NS) coefficient (see Eq. (B1)) (Fig. 3).

Figure 3 demonstrates CABLE ΔTWS varies noticeably by precipitation as well as locations. The area-weighted average values (see Eq. (B2)) computed from Princeton, GSWP3, and TRMM yields the model ΔTWS reasonably agreeing with GRACE by giving the NS coefficient greater than 0.45, while MERRA, PERSIANN, and GLDAS show NS = ~0.3. The less agreement is mainly due to the quality of rainfall estimates over Australia. The NS of ECMWF is around 0.4.

All model ensembles are consistent with the GRACE data over the Timor Sea and inner parts of Australia (e.g., LKE, MRD, NWP) where the NS value can reach as high as 0.9 (see, e.g., TRMM over TIM). On the contrary, the less agreement is found mostly over the coastal basins. Very small or even negative NS values indicate the misfit between CABLE and GRACE mascon solutions, and they are observed over the Indian Ocean (see GLDAS), North East Coast (see GSWP3, PERSIANN, TRMM), South East Coast (see MERRA, TRMM), South West Coast (see GSWP3, GLDAS, MERRA), and South West Plateau (see MERRA).

By averaging all ΔTWS estimates from seven different precipitation datasets, the mean-ensemble estimate (MN) delivers the best agreement with GRACE as seen by the highest average NS value (MN of AVG = 0.55) among all ensembles. Particularly, NS values are greater than 0.4 in all basins and no negative NS values are presented in MN. In average, it can be clearly seen that using the mean value (MN) is a viable option to increase the overall performance of the ΔTWS estimates. Therefore, only CABLE MN result will be used in further analyses. The comparison with the GRGS GRACE solution was also evaluated (not shown here) and the overall results are similar to Fig. 3.

5.2 Impact of GRACE on storage estimates

5.2.1 Contribution of GRACE

This section investigates the impact of the GC approach on the estimates of various water storage components. The ΔTWS estimate obtained from the GC approach is demonstrated in Sect. 5.1, by comparing with the independent GRACE mascon solution. Figure 3 shows the GC result yields the highest NS values in all basins, outperforming all other CABLE runs. In average (AVG), the NS value increases by ~35% (0.55 to 0.74) from the MN case. The similar behaviour is also seen when compared with the GRGS GRACE solution (not shown); the average NS value increases from 0.50 to 0.74. This is not surprising as the GC approach uses the fundamental GRACE tracking data as GRACE mascon and GRGS solutions do. Improvement of NS coefficient indicates merely the successfulness of integrating GRACE data and the model estimates.

Figures 4 and 5 show the GC results of ΔTWS as well as ΔSM_{top}, ΔSM_{rz}, and ΔGWS in different basins. The monthly time-series and the de-seasonalized time-series are shown. In general, GRACE tends to increase ΔTWS when the model ΔTWS (MN) is predicted to be underestimated (see e.g., LKE, MRD, NWP, SWP, TIM between 2011 and 2012) and by decrease ΔTWS when determined to be overestimated (see all basins between 2008 and 2010). A clear example is seen over Gulf of Carpentaria (Fig. 4d), where CABLE overestimates ΔTWS and produces phase delay between 2008 and 2010. The over estimated amplitude and phase delay seen in CABLE ΔGWS during this above period (Fig. 4c) is caused by an overestimation of soil and groundwater storage. The positively biased soil and groundwater storage causes a phase delay by increasing the amount of time required for the subsurface drainage (baseflow) to reduce to soil and groundwater stores. The overestimation of water storage is the result of overestimated precipitation or underestimated evapotranspiration. The amplitude and phase of the water storage estimate are adjusted toward GRACE observation in the GC approach.

The impact of GRACE varies across the individual storage as well as across the geographical location (climate regime). In general, the major contributors to ΔTWS are ΔSM_{rz} and ΔGWS. Due to a small store size (only ~2 cm thick), ΔSM_{top} contributes only ~2 % to ΔTWS. As such, ΔSM_{rz}, and ΔGWS have greater variations, which commonly lead to greater uncertainty compared to ΔSM_{top}, and therefore, the stores anticipate greater shares from the GRACE update. This behaviour is seen over all basins where the differences between CABLE-simulated and GC ΔSM_{rz}, and ΔGWS estimates are greater (compared to ΔSM_{top}).

Furthermore, the impact of GRACE on ΔSM_{rz}, and ΔGWS is different across the continent. For example, over central and southern Australia (see e.g., LKE, MRD, NWP, SWP), the dry climate is responsible for a small amount of groundwater recharge and most of the infiltration is stored in soil compartments. In this climate condition, ΔSM_{rz} amplitude is significantly larger than ΔGWS and it plays a greater role in ΔTWS, and consequently, the GRACE contribution is mostly seen in ΔSM_{rz} component. Different behaviour is seen over the northern Australia (GOC, NEC, TIM) where ΔGWS amplitude are greater (~40 % of ΔTWS) compared to other basins (only ~17 % of ΔTWS). This is due to the sufficient amount of rainfall over the wet climate region, replenishing groundwater recharges and resulting in greater variability in ΔGWS. Therefore, compared to the dry climate basin, the GRACE contributes to ΔGWS over these basins by the larger amount.

5.2.2 Impact on long-term trend estimates

The spatial patterns of the long-term trends of water storage changes over January 2003 and March 2016 are analysed before and after applying the GC approach (Fig. 6). For comparison, the long-term trends of ΔTWS derived from the ITSG-DDK5, mascon, and GRGS solutions are also shown (Fig. 6a, 6b, 6c). From Fig. 6e, GRACE effectively changes the long-term trend estimates in most basins in a way the spatial pattern of the ΔTWS trend of the GC solution consistent to the mascon and GRGS solutions, while satisfying the model processes and keeping the spatial resolution. The trend of ΔSM_{top} is insignificant (Fig. 6f) and the GC approach does not change (Fig. 6g). The largest adjustment is seen in ΔSM_{rz} and ΔGWS components, to be consistent with the GRACE data in most basins (Fig. 6i, 6k).

GRACE shows significant changes in the ΔTWS trend estimates particularly over the northern and western parts of the continent. The model estimates around the Gulf of Carpentaria basin show a strong negative trend that is inconsistent from the GRACE data. It is found that underestimated precipitation after 2012 is likely the cause of such an incompatible negative trend (see Fig. 4d). Applying the GC approach clearly improves the trend (Fig. 6d vs. 6e). The other example is seen over the western part of the continent (see rectangular area in Fig. 6d, 6e) where the averaged long-term trend of ΔTWS was predicted to be –0.4 cm/year but changed to be –1.2 cm/year (see also Sect. 5.4) by the GC approach. The precipitation over the western Australia is understood to be overestimated after 2012, evidently seen by that the model ΔTWS is always greater than the GC solution (see e.g., Fig. 4h, 5d, 5p). The GC approach reveals that the water loss over the western Australia is at least twice greater than what has predicted by the CABLE model.

In addition, the shortage of water storage in the south-eastern part of the continent from the millennium drought (McGrath et al., 2012) has been recovered (seen as a positive water storage trend in Fig. 6) after the rainfall between 2009 and 2012, while the western part is still drying out (seen as negative trends). The trend estimates in terms of mass change are discussed in more detail in Sect. 5.4.

5.2.3 Reduction of uncertainty

Influenced by climate pattern, the uncertainty of water storage estimates significantly varies across Australia. The uncertainty of the model estimate is computed from the variability induced by different precipitation and model parameters while the uncertainty of GC solution is computed using Eq. (11). As expected, larger uncertainties are observed in ΔSM_{rz} and ΔGWS than in ΔSM_{top} (an order of magnitude smaller) since ΔSM_{top} is smaller than others (Fig. 7). Over the wet basins, larger amplitude of the water storage leads to larger uncertainty, seen over Gulf of Carpentaria, North East Coast, South East Coast, and Timor Sea where the CABLE-simulated ΔTWS uncertainty is approximately 28 % larger than other basins. The smaller uncertainty is found over the dry regions (e.g., LKE, SWP). In most basins, the uncertainty of ΔSM_{rz} is larger than the ΔGWS, except the wet basins (e.g., GOC, NEC, TIM) where the greater groundwater recharge leads to a larger uncertainty of ΔGWS.

Figure 7 demonstrates how much the formal error of each of storage components is reduced by the GC approach. Overall, the estimated CABLE uncertainties averaged over all basins (AVG) are 0.2, 4.0, 4.0, and 5.7 cm for ΔSM_{top}, ΔSM_{rz} , ΔGWS, and ΔTWS, respectively. With the GC approach, the uncertainties of ΔSM_{top}, ΔSM_{rz} , ΔGWS, and ΔTWS decrease by approximately 26%, 35%, 39%, and 37%, respectively.

It is worth mentioning that the model uncertainty is mainly influenced by the meteorological forcing data. The uncertainty of precipitation derived from seven different precipitation products is shown in Fig. 7e. The spatial pattern of the precipitation uncertainty is correlated with the uncertainty of water storage estimates. The larger water storage uncertainty is deduced from the larger precipitation uncertainty. The quality of precipitation forcing data is found to be an important factor to determine the accuracy of water storage computation.

5.3 Comparison with independent data

5.3.1 Soil moisture

The ΔSM_{top} estimates are compared with the AMSR-E derived soil moisture. The processing of AMSR-E data is described in Sect 4.3.1. The performance is assessed using Nash-Sutcliff coefficients, given in Table 3. In general, CABLE (MN) shows a good performance in the top soil moisture simulation showing NS value of >0.4 for most of the basins. The top soil moisture estimate shows slightly better agreement with the C-band measurement of the AMSR-E product. This is likely caused by the greater emitting depth of the C-band measurement (~1 cm), which is closer to the depth of the top soil layer (~2 cm) used in this study (Njoku et al., 2003).

The GC approach leads to a small bit of improvement of the top soil estimate consistently from C- and X-band measurements and from all basins. No degradation of the NS value is observed in the GC solutions. The largest improvement is seen over LKE and NEC, where NS increases by 10 – 15%. For other regions, the change in the NS coefficient may be incremental.

5.3.2 Groundwater

The ΔGWS estimates from the model and the GC method are compared with the in situ data obtained from 2 different ground networks in Queensland and Vitoria. For each network, all ΔGWS data inside the groundwater network boundary (see polygons in Fig. 1) are used to compute the average ΔGWS time series. From the comparison given in Fig. 8, it is found that the GC solutions of ΔGWS follows the overall inter-annual pattern of CABLE but with a considerably larger amplitude. This results in a better agreement with the in situ ΔGWS data seen from both networks. The NS coefficient of ΔGWS between the estimates and the in situ data are given in Table 4. The CABLE ΔGWS performs significantly better in Queensland (NS = ~0.5) than Victoria (NS = ~0.3). Significant improvement is found from the GC solutions in both networks, where the NS value increases from 0.5 to 0.6 (~ 22 %) in Queensland and from 0.3 to 0.6 (~85 %) in Victoria. Even greater improvement is seen when the inter-annual patterns are compared. The NS value increase from 0.5 to 0.7 (~ 32 %), and 0.4 to 0.8 (~93 %) in Queensland and Victoria, respectively.

The comparison of the long-term trend of ΔGWS is also evaluated. The estimated trends in Queensland and Victoria are given in Table 4. Beneficially from the GC approach, the ΔGWS trend is improved by approximately 20 % (from 0.4 to 0.6, compared to 1.6 cm/year) in Queensland. Increasing of ΔGWS is mainly influenced by the large amount of rainfall during the 2009 – 2012 La Niña episodes (see Fig. 8a). In Victoria, significant improvement of ΔGWS trend by about 76 % (from 0.1 to –0.2, compared to –0.3 cm/year) is observed. Similar improvement of long-term trend estimates is seen in de-seasonalized time series (improves by ~15 % in Queensland and by ~74 % in Victoria). Decreasing of ΔGWS in Victoria is mainly due to the highly-demanded groundwater consumption by agriculture and domestic activities (van Dijk et al., 2007; Chen et al., 2016). As the groundwater consumption is not parameterized in CABLE, the decreasing of ΔGWS estimate cannot properly captured in the model simulation. Applying GC approach effectively reduces the model deficiency and improves the quality of the groundwater estimations.

5.4 Assessment of mass variation in the past 13 years

Australia experiences significant climate variability; for example, the millennium drought starting from late '90 (Van Dijk et al., 2013) and extremely wet condition during several La Niña episodes (Trenberth 2012; Han 2017). These periods are referred as "Big Dry" and "Big Wet" (Ummenhofer et al., 2009; Xie et al., 2016). To understand the total water storage (mass) variation influenced by these two distinct climate variabilities, the water storage change obtained from the GC approach during Big Dry and Big Wet is separately investigated over 10 basins. The time window between January 2003 and December 2009 is defined as the Big Dry period while between January 2010 and December 2012 is defined as the Big Wet period following Xie et al. (2016). In each period, the long-term trends of GC estimates of ΔTWS, ΔSM_{top}, ΔSM_{rz} , and ΔGWS are firstly calculated. Then, the total water storage variation (in meter) is simply obtained by multiplying the long-term trend (in m/year) with the number of years in the specific period, 7 years for Big Dry and 3 years for Big Wet. To obtain the mass variation, the water storage variation is multiplied by the area of the basin and the density of water (1000 kg/m^3). The estimated mass variations during Big Dry and Big Wet are displayed in Fig. 9. The long-term mass variation of the entire period (January 2003 – March 2016) is also shown.

During Big Dry (2003 – 2009), a significant loss of total storage (40 – 60 Gton over 7 years) is observed over LKE, MRD, NWP, and SWP basins. The largest groundwater loss of >20 Gton is found from LKE and MRD. No significant change is observed over the tropical climate regions (e.g., GOC, NEC). The mass loss mostly occurs in the root zone and groundwater compartments where the sum of ΔSM_{rz} and ΔGWS explains more than 90% of the ΔTWS value. The mass loss is also observed in ΔSM_{top} but >10 times smaller than ΔSM_{rz} and ΔGWS.

During Big Wet (2010 – 2012), the basins like LKE, MRD and TIM exhibit the significant total storage gain of >100 Gton. The gain is particularly larger in ΔSM_{rz} over the basins that experienced the significant loss during Big Dry. For example, over LKE and MRD, the gain of ΔSM_{rz} is approximately 2 – 3 times greater than ΔGWS. It implies that most of the infiltration (from the 2009 – 2012 La Niña rainfall) is stored as soil moisture through the long drought period, and that the groundwater recharge is secondary to the ΔSM_{rz} increase.

The opposite behaviour is observed over the basins (such as NEC and GOC) that experienced mass gain during Big Dry. The water storage gain is greater in ΔGWS compared to ΔSM_{rz}. In NEC, ΔGWS gain is ~8 times larger than ΔSM_{rz} during Big Wet. The soil compartment may be saturated during Big Dry and additional infiltration from the Big Wet precipitation leads to an increased groundwater recharge. The ΔSM_{rz} loss observed over GOC is simply caused by the timing selection of Big Wet period, which ends earlier (~2011) in GOC than in other basins. The ΔSM_{rz} gain becomes ~26 Gton if the Big Wet period is defined as 2008 – 2011. During the post-Big Wet period (2012 and afterwards), the decreasing trend of water storage is observed from all basins (see Fig. 4, 5). This is mainly caused by the decrease in precipitation after 2012 and by gradual water loss through evapotranspiration (Fasullo et al., 2013).

The overall water storage change in the last 13 years demonstrates that the severe water loss from most basins during Big Dry (the millennium drought) is balanced with the gain during Big Wet (the La Niña). The negative ΔTWS estimated during Big Dry becomes positive in LKE, MRD, and SEC and less negative in TIM, and the greatest gain is observed from NEC by ~50 Gton during 13 year-period (see Fig. 9c). However, the water mass loss is still detected over the western basins (e.g., IND, NWP, SWP, SWC), and their magnitudes are even larger than the mass loss during Big Dry. For example, the greatest ΔTWS loss of ~79 Gton is observed over NWP, which is ~25 Gton greater than the loss during Big Dry (see Fig. 9a and 9c). The basin is less affected by the La Niña, and the rainfall during Big Wet is clearly inadequate to support the water storage recovery in the basin. Rainfall deficiency also reduces the groundwater recharge, resulting in even more decreasing of ΔGWS, compared to the Millennium Drought period (see Fig. 9j and 9l). The continual decrease in water storage over western basins is likely caused by the interaction of complex climate patterns like El Niño Southern Oscillation, Indian Ocean Dipole, and Southern Annular Mode cycles (Australian Bureau of Meteorology, 2012; Xie et al., 2016).

5.5 Comparison of GC approach with alternatives

The simplest approach to estimate ΔGWS is to subtract the model soil moisture component from GRACE ΔTWS data, without considering uncertainty in the model output, as used in Rodell et al. (2009) and Famiglietti et al. (2011). This method is called Approach 1 (App 1). In Approach 2 (App 2) as in Tangdamrongsub et al. (2017), by accounting for the uncertainty of model outputs and GRACE data, the water storage states are updated through a Kalman filter:

$$\widehat{\boldsymbol{h}}_{\boldsymbol{R}} = \widetilde{\boldsymbol{h}}_{\boldsymbol{R}} + \mathbf{HC}_{\boldsymbol{R}}^{\boldsymbol{T}}(\mathbf{HRH}^{\boldsymbol{T}} + \mathbf{C}_{\boldsymbol{R}})^{-1}(\boldsymbol{b} - \mathbf{H}\widetilde{\boldsymbol{h}}_{\boldsymbol{R}}) \quad (18)$$

where $\widetilde{\boldsymbol{h}}_{\boldsymbol{R}}$, $\mathbf{H}$, $\mathbf{C}_{\boldsymbol{R}}$ are described in Sect. 2, $\boldsymbol{b}$ is an observation vector containing GRACE-derived ΔTWS, and **R** is an error variance-covariance matrix of the observation. The GRACE-derived ΔTWS and its error information is obtained from the mascon solution. The matrix **R** is a (diagonal) error variance matrix since no covariance information is given in the mascon product. Note that the model uncertainty remains the same as in GC approach (Sect. 4.2). The different results from App1 and App2 are mainly attributed to the different estimates of the uncertainty.

The ΔGWS estimates from App1, App2 and GC in Queensland and Victoria are shown in Fig. 10. It is clearly seen that ΔGWS from App1 are overestimated while the one from App2 fits the ground data significantly better. This behaviour was also seen in Tangdamrongsub et al. (2017) that the water storage estimates tend to be overestimated when error components such as spatial correlation error were neglected as in App1. ΔGWS from App2 shows clear improvements in terms of NS coefficients in both networks. Considering the de-seasonalized ΔGWS estimates, in Queensland, the trend increases from 0.39 ± 0.03 to 0.42 ± 0.03 cm/year (improves by 1.5%), and the NS value increases from 0.46 to 0.53. In Victoria, the trend decreases from 0.73 ± 0.10 to 0.46 ± 0.05 cm/year (improves by 27%), and the NS value increases from –0.89 to 0.30. Although App2 is not yet as good as the GC solution based on the most comprehensive error propagation, this simple test demonstrates an important of considering the uncertainty. The reason of App2 being less accurate than GC is likely due to too simplified error information implemented in App2.

6. Conclusion

This study presents an approach of combining the raw GRACE observation with model simulation to improve water storage estimates over Australia. Distinct from other methods, we exploit the fundamental GRACE satellite tracking data and the full data error variance-covariance information to avoid alteration of signal and measurement error information present in higher level data products.

We compare groundwater storage estimates from GC approach and two other approaches, subject to inclusion of GRACE uncertainty in the ΔGWS calculation. Validating three results of ΔGWS against the in situ groundwater data, we find that the GC approach delivers the most accurate groundwater estimate, followed by the approach based on incomplete information of GRACE's data error. The poorest estimate of groundwater storage is seen when the GRACE uncertainty is completely ignored. This confirms the critical value of using the complete GRACE signal and error information at the raw data level.

The analysis of water storage change between 2003 and 2016 reveals that half of the continent (5 out of 10 basins) is still not fully recovered from the Millennium Drought. The TWS decrease in Western Australia has been most characteristic, and the GC approach finds that the water loss mainly occurs in groundwater layer. Rainfall inadequacy is attributed to the continual dry condition, leading to a greater decreasing of groundwater recharge and storage over Western Australia.

The land surface model we used is deficient in anthropogenic groundwater consumption. The model calibration will never help, and the groundwater consumption must be brought in by external sources. On the contrary, the statistical approach like our GC approach may be useful to fill in the missing component and lead to a more comprehensive water storage inventory.

However, it is difficult to constrain different water storage components by only using total storage observation like GRACE. In addition, it is challenging to improve surface soil moisture varying rapidly in time, using a monthly mean GRACE observation. Tian et al. (2017) utilized the satellite soil moisture observation from the Soil Moisture and Ocean Salinity (SMOS, Kerr et al., 2001) in addition to GRACE data for their data assimilation and showed a clear improvement in the top soil moisture estimate. The GC approach with complementary observations at higher temporal resolution should be considered particularly to enhance the surface soil moisture computation.

Furthermore, the GC approach can be simply extended for GRACE data assimilation. Assimilating the raw GRACE data into land surface models like CABLE enables the model state and parameter to be adjusted with the realistic error information, allowing reliable storage computation. The GC data assimilation will be developed in our future study.

Acknowledgement

This work is funded by The University of Newcastle to support NASA's GRACE and GRACE Follow-On projects as an international science team member to the missions. MD was supported by ARC Centre of Excellence for Climate Systems Science. HK was supported by Japan Society for the Promotion of Science KAKENHI (16H06291). We thank Torsten Mayer-Gürr for GRACE data products in the form of the least-squares normal equations. We also thank three anonymous reviewers for helping us improve the manuscript.

Appendix A: Least-squares normal equation of GRACE

A linearized GRACE satellite-tracking observation equation is formulated as:

$$\boldsymbol{y} = \mathbf{A}\boldsymbol{x} + \boldsymbol{e}; \boldsymbol{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \quad \text{(A1)}$$

where $\boldsymbol{y}$ is the observation vector containing various kinds of L1B data including the inter-satellite ranging data, $\mathbf{A}$ is the design (partial derivative) matrix relating the data and the Earth gravity field variations, $\boldsymbol{x}$ contains the Stokes coefficients of time-varying geopotential fields (e.g., Wahr et al., 1998), and $\boldsymbol{e}$ is the L1B data noise, which has zero mean and covariance $\boldsymbol{\Sigma}$. Eq. (A1) can be modified explicitly in terms of soil moisture and groundwater storage variations as:

$$\boldsymbol{y} = \mathbf{A}\mathbf{S}\bar{\mathbf{Y}}\mathbf{H}\boldsymbol{h} + \boldsymbol{e}; \boldsymbol{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \quad \text{(A2)}$$

where $\mathbf{S}$ contains a factor used to convert ΔTWS to geopotential coefficients considering the load Love numbers (e.g., Wahr et al., 1998), $\bar{\mathbf{Y}}$ converts the gridded data into the corresponding spherical harmonic coefficients. For convenience, the term $\mathbf{Y} = \mathbf{S}\bar{\mathbf{Y}}$ is used in the further derivation. A least-squares solution of Eq. (A2) is given as:

$$(\mathbf{H}^T\mathbf{Y}^T\mathbf{A}^T\boldsymbol{\Sigma}^{-1}\mathbf{A}\mathbf{Y}\mathbf{H})\hat{\boldsymbol{h}} = \mathbf{H}^T\mathbf{Y}^T\mathbf{A}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y}. \quad \text{(A3)}$$

It can be simplified as:

$$\mathbf{H}^T\mathbf{Y}^T\mathbf{N}\,\mathbf{Y}\mathbf{H}\,\hat{\boldsymbol{h}} = \mathbf{H}^T\mathbf{Y}^T\boldsymbol{c}, \quad \text{(A4)}$$

where $\mathbf{N} = \mathbf{A}^T\boldsymbol{\Sigma}^{-1}\mathbf{A}$ and $\boldsymbol{c} = \mathbf{A}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y}$. Eq. (A4) is identical to Eq. (5).

Appendix B: Nash-Sutcliff coefficient and area-weighted average

Nash-Sutcliff coefficient (NS) is computed as follows:

$$NS = 1 - \frac{\sum_{i=1}^{N}(\boldsymbol{y}_i - \hat{\boldsymbol{x}}_i)^2}{\sum_{i=1}^{N}(\boldsymbol{y}_i - \bar{\boldsymbol{y}})^2} \quad \text{(B1)}$$

where $\boldsymbol{y}$ is an observation vector, $\bar{\boldsymbol{y}}$ is the mean of the observation, $\hat{\boldsymbol{x}}$ is a vector containing the simulated result, i is the index of observation, and N is the number of observation.

Area-weighted average ($\bar{Z}$) is compute as follows:

$$\bar{Z} = \frac{\sum_{j=1}^{M} w_j \bar{z}_j}{\sum_{j=1}^{M} w_j} \quad \text{(B2)}$$

where w is the area size, $\bar{z}$ is the mean value inside the considered area, j is the area index, and M is the number of considered area.

References

Australian Bureau of Meteorology (2012) Record-breaking La Niña events: An analysis of the La Niña life cycle and the impacts and significance of the 2010–11 and 2011–12 La Niña events in Australia, National Climate Centre, Bureau of Meteorology, http://www.bom.gov.au/climate/enso/history/La-Nina-2010-12.pdf (last accessed: 5 January 2017).

Bettadpur, S.: CSR Level-2 Processing Standards Document for Product Release 05, GRACE 327-742, Center for Space Research, The University of Texas at Austin, 2012.

Chen, J. L., Wilson, C. R., Tapley, B. D., Scanlon, B., Güntner, A.: Long-term groundwater storage change in Victoria, Australia from satellite gravity and in situ observations, Glob. Planet. Change, 139, 56–65, doi: http://dx.doi.org/10.1016/j.gloplacha.2016.01.002, 2016.

Decker, M.: Development and evaluation of a new soil moisture and runoff parameterization for the CABLE LSM including subgrid-scale processes, J. Adv. Model. Earth Syst., 7, 1788–1809, doi:10.1002/2015MS000507, 2015.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haiberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, 2011.

Dumedah, G., and Walker, J. P.: Intercomparison of the JULES and CALBE land surface models through assimilation of remote sensed soil moisture in southeast Australia, Adv. Wat. Res., 74, 231 – 244, doi:hmattp://dx.doi.org/10.1016/j.advwatres.2014.09.011, 2014.

Eicker, A., Schumacher, M., Kusche, J., Döll, P., and Müller Schmied, H.: Calibration data assimilation approach for integrating GRACE data into the WaterGAP Global Hydrology Model (WGHM) using an Ensemble Kalman Filter: First Results, Surv. Geophys., 35(6), 1285-1309, doi:10.1007/s10712-014-9309-8, 2014.

Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, Ocean Dyn., 53(4), 343-367, doi:10.1007/S10236-003-0036-9, 2003.

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

Fasullo, J. T., Boening, C., Landerer, F. W., and Nerem, R. S.: Australia's unique influence on global sea level in 2010–2011, Geophys. Res. Lett., 40, 4368–4373, doi:10.1002/grl.50834, 2013.

Forman, B. A., Reichle, R. H., and Rodell, M.: Assimilation of terrestrial water storage from GRACE in a snow-domained basin, Water Resour. Res., 48, W01507, doi:10.1029/2011WR011239, 2012.

Girotto, M., De Lannoy, G. J. M., Reichle, R. H., and Rodell, M.: Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model, Water Resour. Res., 52(5), 4164–4183, doi:10.1002/2015WR018417, 2016.

Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter, Mon. Weather Rev., 129, 2776–2790, 2001.

Han, S.-C.: Elastic deformation of the Australian continent induced by seasonal water cycles and the 2010-11 La Niña determined using GPS and GRACE, Geophys. Res. Lett., 44, doi: 10.1002/2017GL072999, 2017.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolf, D. B.: The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, J. Hydrometeor., 8, 38–55, doi:10.1175/JHM560.1, 2007.

Jekeli, C.: Alternative methods to smooth the Earth's gravity field, Rep., 327, Dept. of Geod. Sci. and Surv., Ohio State Univ., Columbus, 1981.

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Martinuzzi, J.-M., Font, J., and Berger, M.: Soil moisture retrieval from space: The soil moisture and ocean salinity (SMOS) mission, IEEE Trans. Geosci. Remote Sens., 39(8), 1729–1735, 2001.

Khaki, M., Hoteit, I., Kuhn, M., Awange, J., Forootan, E., van Dijk, A., Schumacher, M., Pattiaratchi, C., Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model. Adv. Wat Res., 107, 301–316, doi:10.1016/j.advwatres.2017.07.001, 2017a.

Khaki, M., Schumacher, M., Forootan, E., Kuhn, M., Awange, J., van Dijk, A., Accounting for spatial correlation errors in the assimilation of GRACE into hydrological models through localization. Adv. Wat Res., 108, 99–112, doi:10.1016/j.advwatres.2017.07.024, 2017b.

Kusche, J., Schmidt, R., Petrovic, S. and Rietbroek, R.: Decorrelated GRACE time-variable gravity solutions by GFZ, and their validation using a hydrological model, J. Geod., 83(10), 903–913, doi:10.1007/s00190-009-0308-3, 2009.

Lambert, A., Huang, J., van der Kamp, G., Henton, J., Mazzotti, S., James, T. S., Courtier, N., and Barr, A. G.: Measuring water accumulation rates using GRACE data in areas experiencing glacial isostatic adjustment: The Nelson River basin, Geophys. Res. Lett., 40, 6118–6122, doi:10.1002/2013GL057973, 2013.

Landerer, F. W., and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resour. Res., 48, W04531, doi:10.1029/2011WR011453, 2012.

Leblanc, M., Tweed, S., Van Dijk, A., Timbal, B.:, A review of historic and future hydrological changes in the Murray-Darling Basin, Global and Planetary Change, 80 – 81, 226 – 246, doi:10.1016/j.gloplacha.2011.10.012, 2012.

Lemoine, J. M., Bourgogne, S., Bruinsma, S., Gégout, P., Reinquin, F., Biancale R.: GRACE RL03-v2 monthly time series of solutions from CNES/GRGS, EGU2015-14461, EGU General Assembly 2015, Vienna, Austria, 2015.

Mayer-Gürr, T., Behzadpour, S., Ellmer, M., Kvas, A., Klinger, B., Zehentner, N.: ITSG-Grace2016 - Monthly and Daily Gravity Field Solutions from GRACE. GFZ Data Services. http://doi.org/10.5880/icgem.2016.007, 2016.

Mayer-Gürr T., Pail R., Gruber T., Fecher T., Rexer M., Schuh W.-D., Kusche J., Brockmann J.-M., Rieser D., Zehentner N., Kvas A., Klinger B., Baur O., Höck E., Krauss S., Jäggi A.: The combined satellite gravity field model GOCO05s, EGU 2015, Vienna, 2015.

McGrath, G. S., Sadler, R.,Fleming, K., Tregoning, P., Hinz, C., and Veneklaas, E. J.: Tropical cyclones and the ecohydrology of Australia's recent continental-scale drought, Geophys. Res. Lett., 39, L03404, doi:10.1029/2011GL050263, 2012.

Njoku, E. G., Jackson, T. L., Lakshmi, V., Chan, T., Nghiem, S. V.: Soil Moisture Retrieval from AMSR-E, IEEE T. Geosci. Remote, 41 (2): 215-229, 2003.

Owe, M.,de Jeu, R., Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, J. Geophys. Res., 113, F01002, 17 pp., doi:10.1029/2007JF000769, 2008.

Pearson, E. K.: Mining imperfect data: Dealing with contamination and incomplete records, ProSanos Corporation, Harrisburg, Pennsylvania, ISBN: 978-0-89871-582-8, doi: http://dx.doi.org/10.1137/1.9780898717884, 2005.

Ramillien, G., Cazenave, A., and Brunau, O., Global time variations of hydrological signals from GRACE satellite gravimetry, Geophys. J. Int., 158, 813–826, 2004.

Rassam, D. W., Peeters, L., Pickett, T., Jolly, I., Holz, L.: Accounting for surfaceegroundwater interactions and their uncertainty in river and groundwater models: A case study in the Namoi River, Australia, Environ. Modell. Softw., 50, 108-119, http://dx.doi.org/10.1016/j.envsoft.2013.09.004, 2013.

Reichle, R. H., and Koster, R. D.: Bias reduction in short records of satellite soil moisture, Geophys. Res. Lett., 31, L19501, doi:10.1029/2004GL020938, 2004.

Rienecker, M. M., Suarez, M. J., Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich MG, Schubert SD, Takacs L, Kim G-K, Bloom S, Chen J, Collins D, Conaty A, da Silva A, Gu W, Joiner J, Koster RD, Lucchesi R, Molod A, Owens T, Pawson S, Pegion P, Redder CR, Reichle R, Robertson FR, Ruddick AG, Sienkiewicz M, Woollen J. 2011. MERRA—NASA's Modern-Era Retrospective Analysis for Research and Applications. J. Climate, DOI: 10.1175/JCLI-D-11-00015.1.

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. Amer. Meteor. Soc., 85(3), 381–394, 2004.

Rodell, M., Velicogna, I., Famiglietti, J. S.: Satellite-based estimates of groundwater depletion in India, Nature. 460, 999-1002, doi:10.1038/nature08238, 2009.

Sakumura, C., Bettadpur, S., and Bruinsma, S.: Ensemble prediction and intercomparison analysis of GRACE time-variable gravity field models, Geophys. Res. Lett., 41, 1389–1397, doi:10.1002/2013GL058632, 2014.

Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-yr high-resolution global dataset of meteorological forcings for land surface modeling, J. Climate, 19 (13), 3088-3111, 2005.

Schumacher, M., Kusche, J., and Döll, P.: A Systematic Impact Assessment of GRACE Error Correlation on Data Assimilation in Hydrological Models, J. Geod., 90(6), 537–559. doi:10.1007/s00190-016-0892-y, 2016.

Schumacher, M, Forootan, E, van Dijk, A, Schmied, HM, Crosbie, R, Kusche, J., and Döll, P, Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model. Remote Sens. Environ., 204, 212–228, doi:10.1016/j.rse.2017.10.029, 2018.

Sorooshian, S., Hsu, K., Gao, X., Gupta, H. V., Imam, B., and Braithwaite, D.: Evaluation of PERSIANN System Satellite-Based Estimates of Tropical Rainfall, Bulletin of the American Meteorological Society, Vol. 81, No. 9, 2035-2046, 2000.

Swenson. S. C.: GRACE monthly land water mass grids NETCDF RELEASE 5.0. Ver. 5.0. PO.DAAC, CA, USA, http://dx.doi.org/10.5067/TELND-NC005, 2012. (last accessed: 5 January 2017).

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

Tangdamrongsub, N., Ditmar, P. G., Steele-Dunne, S. C., Gunter, B. C., Sutanudjaja, E. H. (2016) Assessing total water storage and identifying flood events over Tonlé Sap basin in Cambodia using GRACE and MODIS satellite observations combined with hydrological models, Remote Sens. Environ., 181, 162–173, doi:http://dx.doi.org/10.1016/j.rse.2016.03.030.

Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., and Weerts, A. H.: Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin, Hydrol. Earth Syst. Sci., 19, 2079–2100, doi:10.5194/hess-19-2079-2015, 2015.

Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., Sutanudjaja, E. H., Xie, T, Wang, Z.: Improving estimates of water resources in a semi-arid region by assimilating GRACE data into the PCR-GLOBWB hydrological model, Hydrol. Earth Syst. Sci., 21, 2053 – 2074, doi:10.5194/hess-21-2053-2017, 2017.

Tian, S., Tregoning, P., Renzullo, L. J., van Dijk, A. I. J. M., Walker, J. P., Pauwels, V. R. N., Allgeyer, S.: Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals, Water Resour. Res., 53, doi:10.1002/2016WR019641, 2017.

Trenberth, K. E.: Framing the way to relate climate extremes to climate change, Climatic Change, 115 283–290 , doi:10.1007/s10584-012-0441-5, 2012.

Tscherning, C. C. and Rapp R. H.: Closed covariance expressions for gravity anomalies, geoid undulations, and deflections of the vertical implied by anomaly degree variance models, Rep. 208, Dep. of Geod. Sci. and Surv., Ohio State Univ., Columbus, 1974.

Ukkola, A. M., Pitman, A. J., Decker, M., De Kauwe, M. G., Abramowitz, G., Kala, J., and Wang, Y.-P.: Modelling evapotranspiration during precipitation deficits: identifying critical processes in a land surface model. Hydrol. Earth Syst. Sci., 20, 2403–2419, doi:10.5194/hess-20-2403-2016, 2016.

Van Dijk, A., Beck, H. E., Crosbie, R. S., De Jeu, E. A. M., Liu, Y. Y., Podger, G. M., Timbal, B., Viney, N. R: The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society, Water Resour. Res., 49 (2), 1040 - 1057, doi:10.1002/wrcr.20123, 2013.

Van Dijk, A., Podger, G., Kirby, M.: Integrated water resources management in the Murray-Darling Basin. In: Schumann, A., Pahlow, M. (Eds.), Increasing demands on decreasing supplies, in Reducing the Vulnerability of Societies to Water Related Risks at the Basin Scale, IAHS Publ. 24–30, 2007.

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

Wahr, J., Molenaar, M., and Bryan, F.: Time variability of the Earth’s gravity field: Hydrological and oceanic effects and their possible detection using GRACE, J. Geophys. Res., 103(B12), 30205–30229, 1998.

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

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.

Weerts, A. H. and El Serafy G. Y. H.: Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, Water Resour. Res., 42, W09403, doi:10.1029/2005WR004093, 2006.

Welsh, W.D.: Water balance modelling in Bowen, Queensland, and the ten iterative steps in model development and evaluation, Environ. Modell. Softw., 23 (2), 195-205, 2008.

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.

Wu, S. C., Kruizinga, G., and Bertiger, W.: Algorithm theoretical basis document for GRACE Level-1B data processing V1.2, JPL D-27672, Jet Propul. Lab., Pasadena, Calif, 2006.

Xie, Z., Huete, A., Restrepo-Coupea, N., Maa, X., Devadasa, R., Caprarellib, G.: Spatial partitioning and temporal evolution of Australia's total water storage under extreme hydroclimatic impacts, Remote Sens. Environ., 183, 43–52, http://dx.doi.org/10.1016/j.rse.2016.05.017, 2016.

Zaitchik, B. F., Rodell, M., and Reichle, E. H.: Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi basin, Amer. Meteor. Soc., J. Hydrometeor, 9, 535–548, 2008.

Table 1. Precipitation data from 7 different products used in this study, the Global Soil Wetness Project Phase 3 (GSWP3), the Global Land Data Assimilation System (GLDAS), the Tropical Rainfall Measuring Mission (TRMM), the Modern-Era Retrospective Analysis for Research and Applications (MERRA), the European Centre for Medium-Range Weather Forecasts (ECMWF), the Princeton's Global Meteorological Forcing Dataset (Princeton), and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN). The temporal resolution of all products is 3 hours. Most products are available to present while GSWP3, MERRA, and Princeton terminate earlier.

Product	Availability	Spatial resolution	References
GSWP3	1901/01 – 2010/12	0.5°×0.5°	http://hydro.iis.u-tokyo.ac.jp/GSWP3
GLDAS (NOAH025SUBP 3H)	2000/03 – present	0.25°×0.25°	Rodell et al. (2004)
TRMM (3B42)	1998/01 – present	0.25°×0.25°	Huffman et al. (2007)
MERRA (MSTMNXMLD.5.2.0)	1980/01 – 2016/02	0.5°×0.67°	Rienecker et al. (2011)
ECMWF (ERA-Interim)	1979/01 – present	0.75°×0.75°	Dee et al. (2011)
Princeton (V2 0.5°)	1987/01 – 2012/12	0.5°×0.5°	Sheffield et al. (2005)
PERSIANN (3 hr)	2002/03 – present	0.25°×0.25°	Sorooshian et al. (2000)

Table 2. Model parameters that are sensitive to SM and GWS estimates. The following parameters were perturbed using the additive noise with the boundary conditions given in the last column. The further parameter description can be found in Decker (2015) and Ukkola et al. (2016).

Parameter	Name	Spatial variability	Perturbed range
f_{clay}, f_{sand}, f_{silt}	Fraction of clay, sand, and silt	Yes	0 – 1
f_{sat}	Fraction of grid cell that is saturated	No	810 – 990
q_{sub}	Maximum rate of subsurface drainage assuming a fully saturated soil column	No	0.009 – 0.01
f_p	Tuneable parameter controlling drainage speed	No	1.9 – 2.2

Table 3. NS coefficients between top soil moisture estimates and the satellite soil moisture observations from AMSR-E products over 10 different Australian basins. The area-weighted average value (AVG) is also shown.

	C-band		X-band	
	CABLE	GC	CABLE	GC
GOC	0.67	0.68	0.58	0.60
IND	0.53	0.54	0.41	0.41
LKE	0.48	0.53	0.36	0.42
MRD	0.77	0.80	0.75	0.78
NEC	0.34	0.39	0.14	0.19
NWP	0.33	0.36	0.38	0.42
SEC	0.68	0.68	0.69	0.71
SWC	0.85	0.85	0.89	0.89
SWP	0.55	0.56	0.46	0.48
TIM	0.44	0.45	0.16	0.16
AVG	0.53	0.56	0.47	0.50

Table 4. NS coefficient and long-term trend of ΔGWS estimated from the model-only and GC solutions in Queensland and Victoria groundwater network. The long-term trend of the in-situ data is also shown.

	Queensland			Victoria		
	In-situ	CABLE	GC	In-situ	CABLE	GC
Original time-series						
NS [-]	-	0.49	0.60	-	0.34	0.63
Trend [cm/year]	1.60 ± 0.05	0.39 ± 0.02	0.63 ± 0.05	–0.27 ± 0.05	0.10 ± 002	–0.18 ± 0.03
De-seasonalized time-series						
NS [-]	-	0.50	0.66	-	0.43	0.83
Trend [cm/year]	1.60 ± 0.05	0.39 ± 0.02	0.57 ± 0.04	–0.25 ± 0.05	0.10 ± 0.02	–0.16 ± 0.03

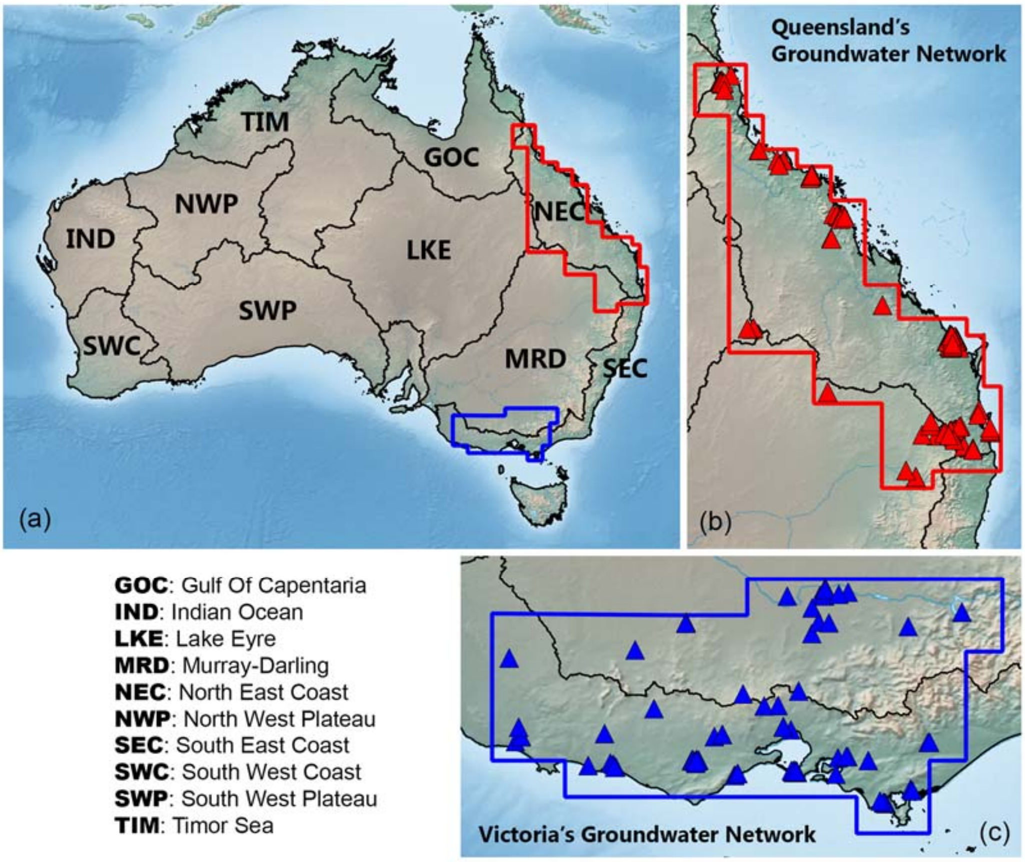

Figure 1. (a) Geographical location of 10 Australian river basins. Red and blue polygons indicate the boundaries of groundwater networks in Queensland (b) and Victoria (c), respectively. Triangles (in b and c) represent the selected bore locations used in this study.

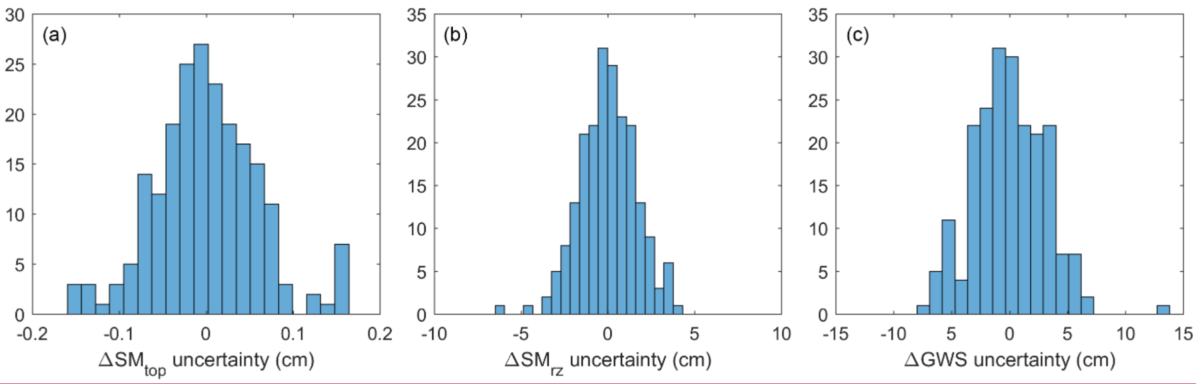

Figure 2: Histograms of the model errors computed from 210 ensemble members ($\mathcal{H}_R{}'$) without the mean. The basin averaged values (from all 10 Australian basins) of January 2003, for example, are shown.

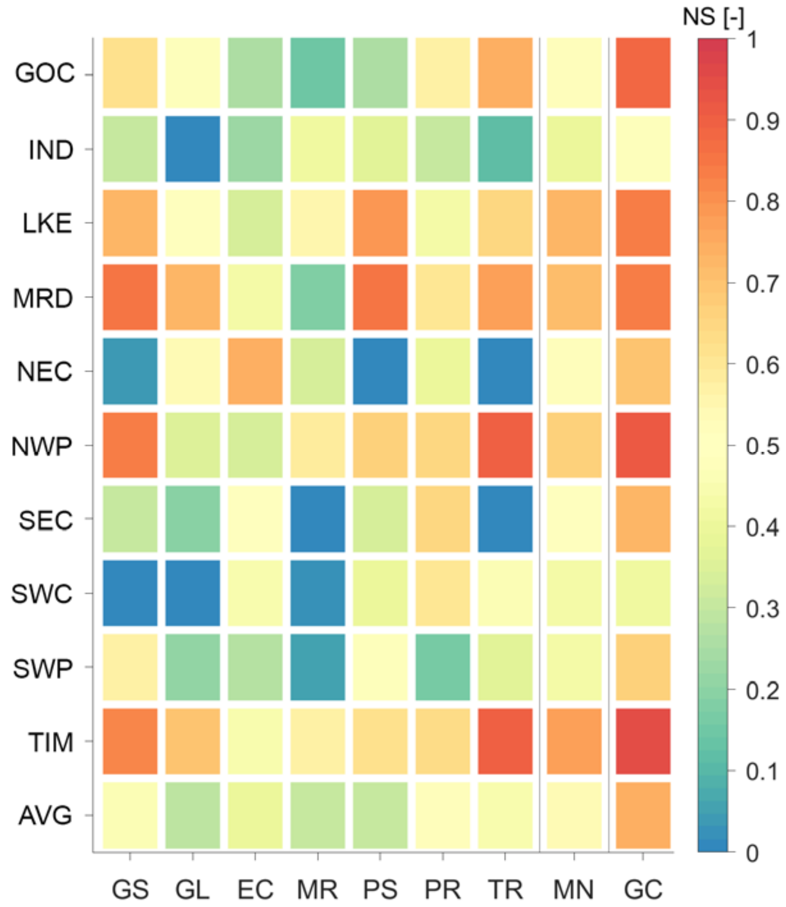

Figure 3. NS coefficients between the model and GRACE-mascon ΔTWS over 10 Australian basins (in ordinate). The NS values were computed based on CABLE ΔTWS computed with 7 different precipitation data (in abscissa), GSWP3 (GS), GLDAS (GL), ECMWF (EC), MERRA (MR), PERSIANN (PR), TRMM (TR). The NS value of the mean ΔTWS estimates (the average of 7 variants) is also shown (MN). The area-weighted average NS value over all basins is also shown (AVG). The NS value of ΔTWS from the GRACE-combined (GC) approach is shown in the last column. The full name of the basins can be found in Fig. 1.

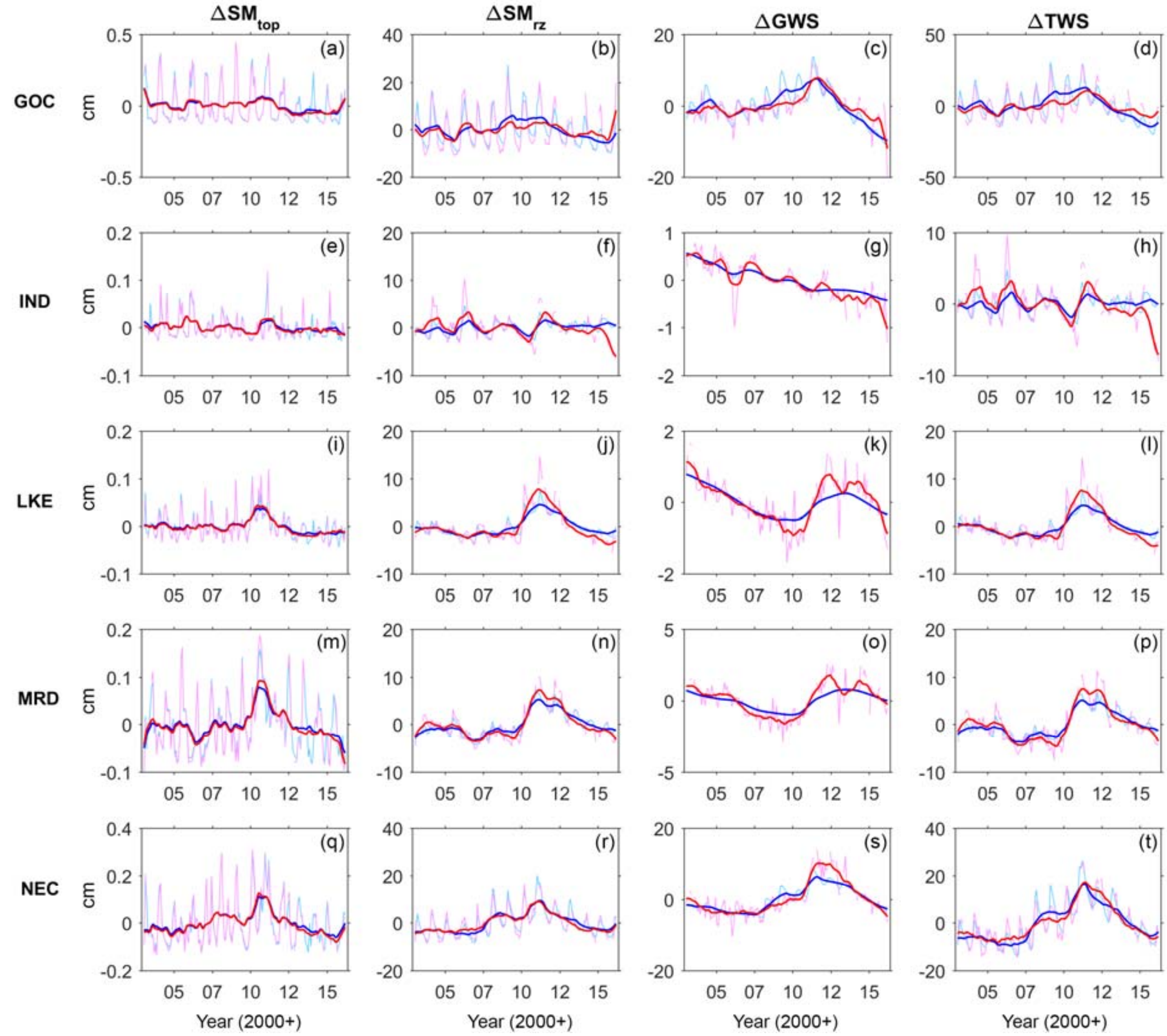

Figure 4. The monthly time series of ΔSM_{top}, ΔSM_{rz}, ΔGWS, and ΔTWS estimated from model (blue) and GC (red) solutions over Gulf of Carpentaria (GOC), Indian Ocean (IND), Lake Eyre (LKE), Murray-Darling (MRD), and North East Coast (NEC). The de-seasonalized time series is also shown.

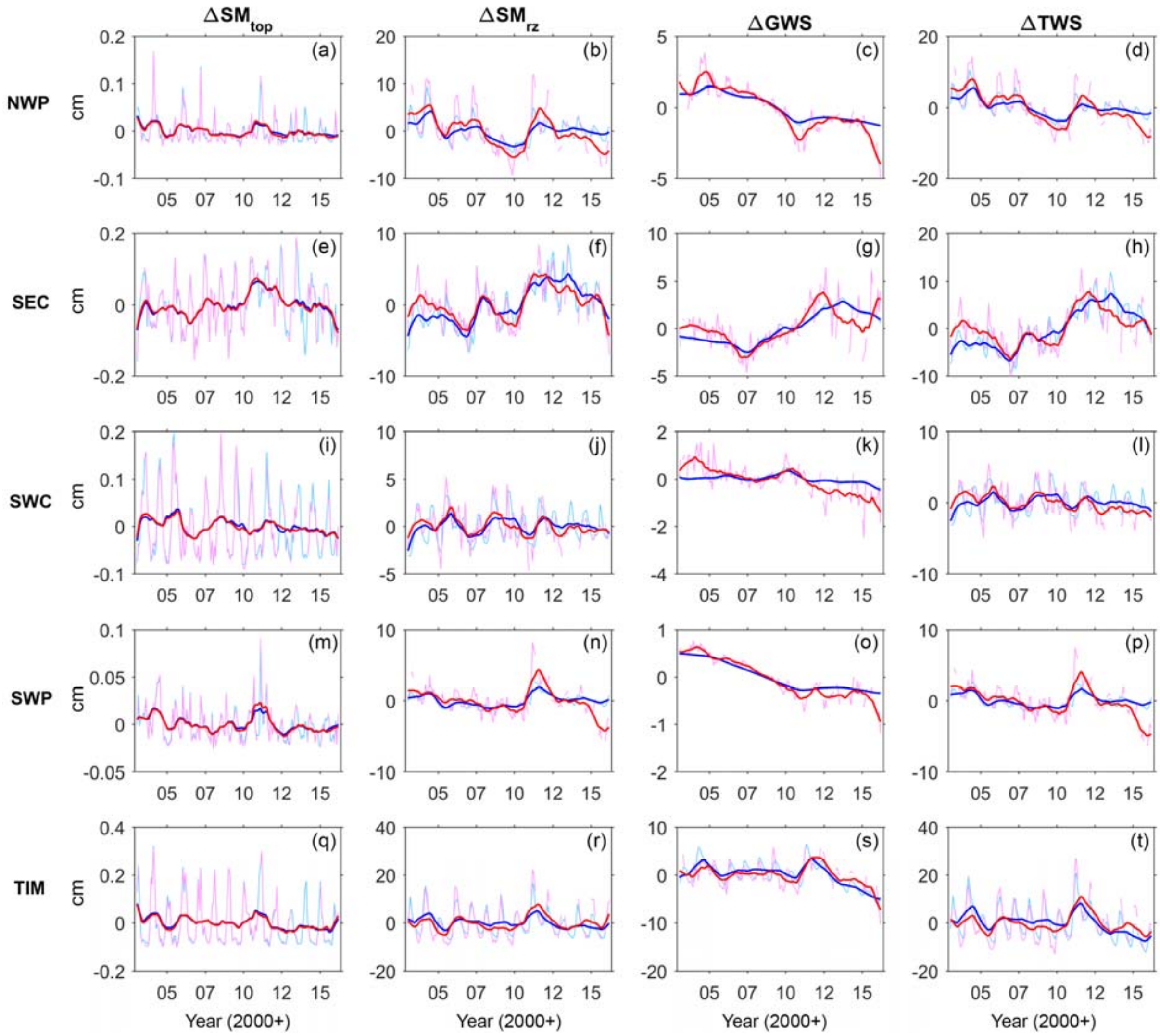

Figure 5. Similar to Fig. 3, but estimated over North West Plateau (NWP), South East Coast (SEC), South West Coast (SWC), South West Plateau (SWP), and Timor Sea (TIM).

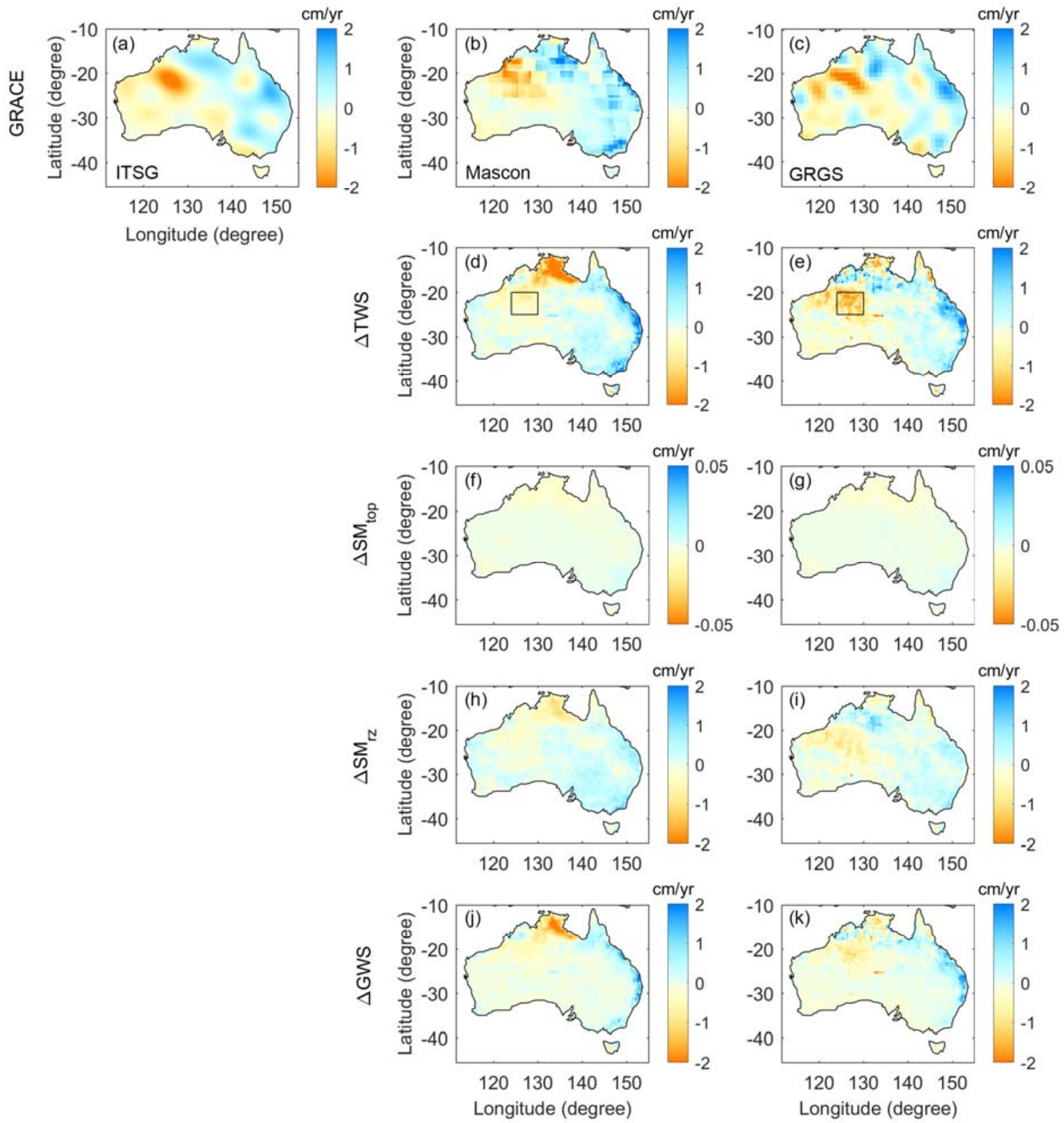

Figure 6. Long-term trends of ΔTWS (d, e), ΔSM_{top} (f, g), ΔSM_{rz} (h, i), and ΔGWS (j, k) estimated from the model-only (second column) and the GC solutions (third column). Results of GRACE ΔTWS independently from ITSG-DDK5 (a), mascon (b), and GRGS solution (c) are also shown. The eastern part of North West Plateau basin is shown as a rectangle polygon in (d) and (e).

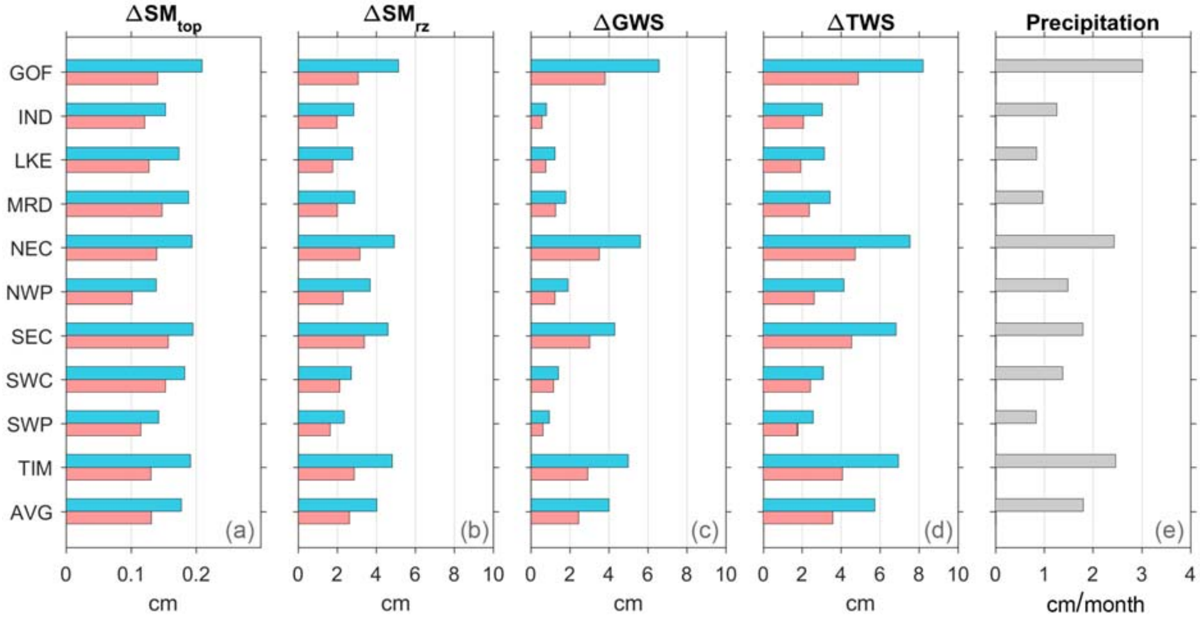

Figure 7. Uncertainties of ΔSM_{top}, ΔSM_{rz}, ΔGWS, and ΔTWS estimated from the model (blue) and the GC solutions (red) in 10 different Australian basins. The uncertainty of the precipitation is shown in (e). The area-weighted average value (AVG) is also shown.

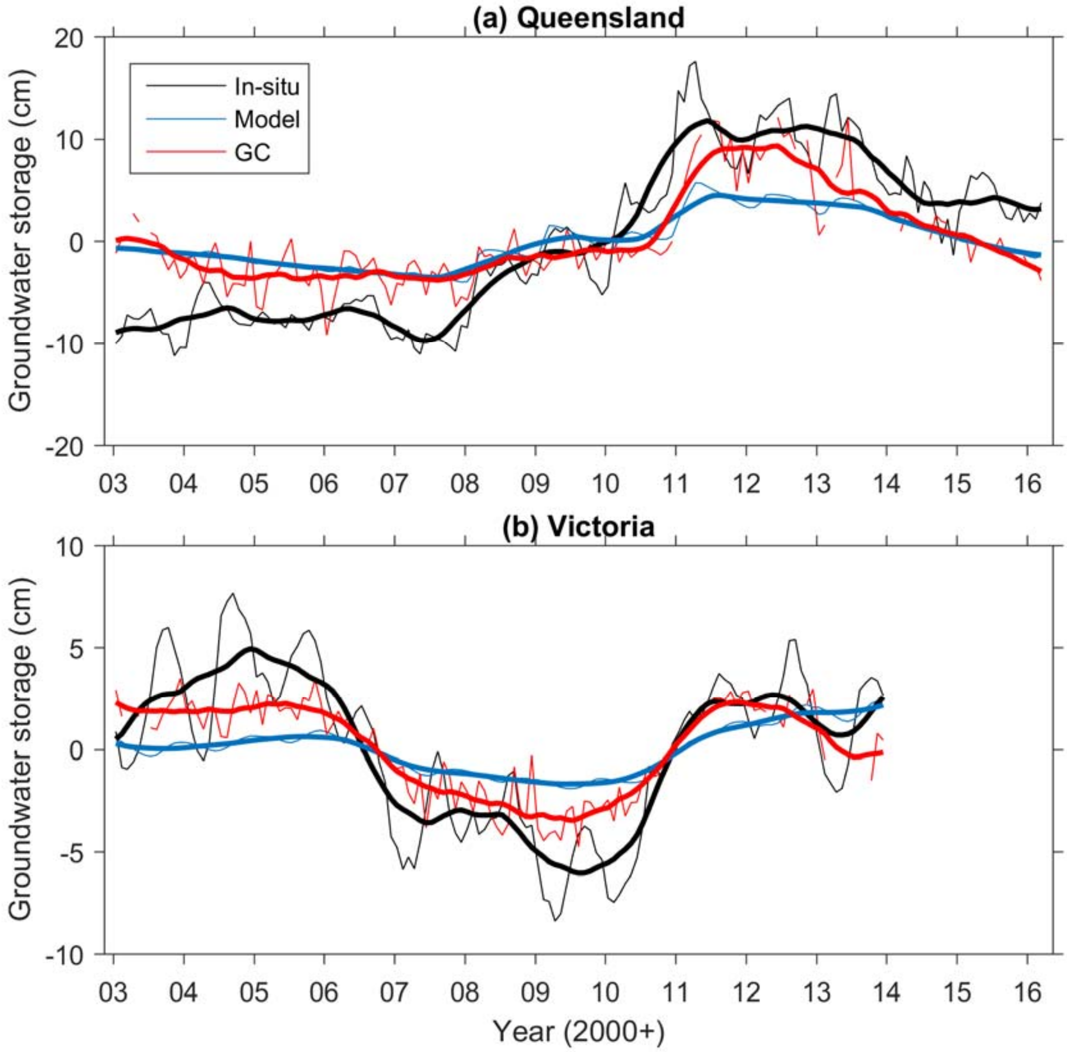

Figure 8. The monthly time series of ΔGWS estimated from the model, GC solutions, and measured from the in situ groundwater network in Queensland (a) and Victoria (b). De-seasonalized time series are shown in thick lines.

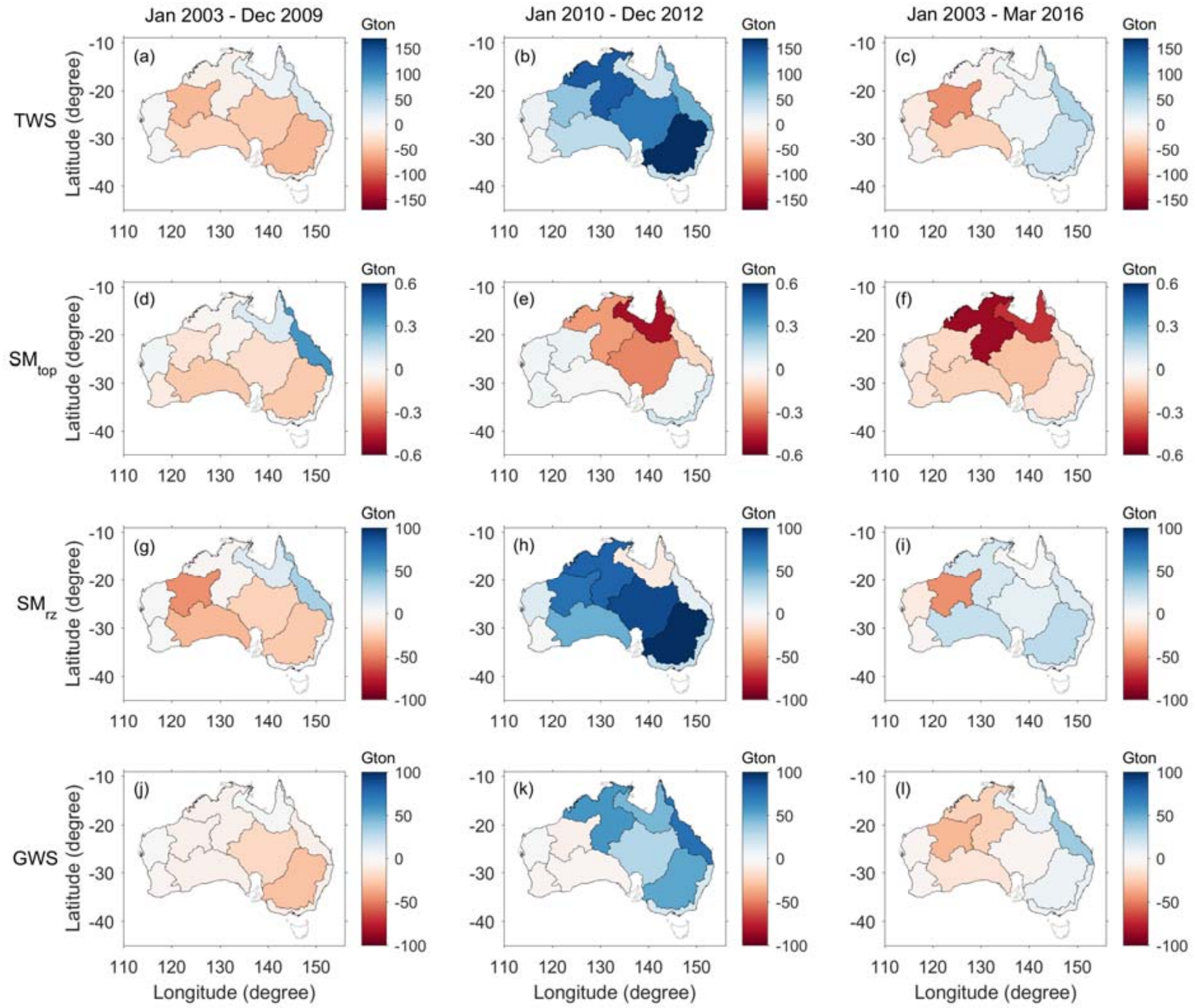

Figure 9. Mass changes (Gton, Giga tonne) of ΔTWS, ΔSM_{top}, ΔSM_{rz}, and ΔGWS estimated from GC solutions over 10 Australian basins in 3 different periods, Big Dry (January 2003 – December 2009), Big Wet (January 2010 – December 2012), and entire period (January 2003 – March 2016).

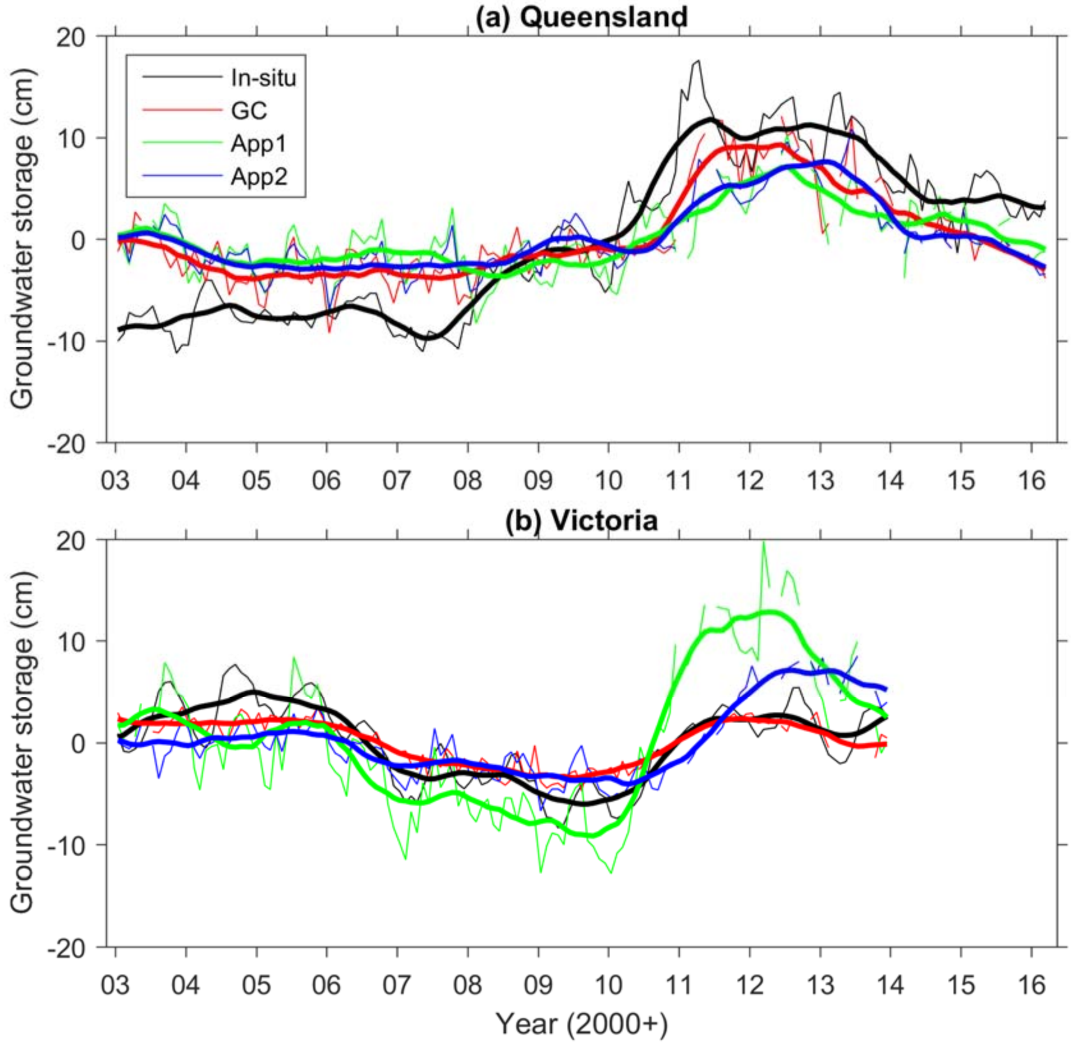

Figure 10. ΔGWS estimated from Approach 1 (App1) and Approach 2 (App2) in Queensland (a) and Victoria (b). The in-situ groundwater network data and the GC solutions are also shown. De-seasonalized time series are shown in thick lines.