Reviewer 1

We would like to acknowledge the insightful comments and suggestions provided by reviewer 1. We will consider the reviewer's suggestions in our revised manuscript. Followings are the responses (R) based on the comments:

(1) The description of the GC approach appears to be somewhat ambiguous: While Sect. 2 is claiming to use L1B KBRR data only, it becomes clear from Sect. 3 that in fact L2 monthly normal equations from ITSG2016 are applied. Those NEQ, however, include not only KBRR but all GRACE sensor information (KBRR, ACC, GPS, attitude) and a priori background models (AOD1B, earth, ocean, and atmospheric tides, third body effects).

R1: Reviewer is correct that the normal equation is built including measurements from all GRACE sensors, including KBR, and a priori background models. Therefore, lines **144 – 145** of the revised manuscript is revised as follows:

"… the observation vector containing various kinds of L1B data including the inter-satellite ranging"

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.

R2: The computation of ΔTWS 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 ΔTWS 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.

(2) 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.

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. For clarity, we modify lines **298 – 301** of the revised manuscript as follows:

"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, $\mathcal{H}_R{}' = \mathcal{H}_R - \tilde{h}_R$, and the error covariance matrix of the model is empirically computed as"

(3) line 284: It is optimistic to assume that the model omission error can be fully accounted by just increasing the model covariance by 20%, in particular since this assumes that omission errors do not contribute to biases (which is quite unlikely). More evidence need to be provided for the (approximate) validity of this assumption.

R4: Rigorous development of the statistical property of the land surface models (expectation and covariance) is crucial but a difficult task. The empirical way was taken in this study. As also noted by the reviewer, the omission will likely introduce a bias in the solutions as well.

We acknowledge this in the last paragraph of Section 4.2. Lines **313 – 319** of the revised manuscript now reads:

"It is emphasized that the model omission error caused by imperfect modelling 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 $\mathbf{C}_R$ 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) The statement of line 212ff is unclear (and apparently not picked up again in the remainder of the paper). Please elaborate.

R5: This is to introduce what other independent GRACE solutions (GRGS and JPL mascon) to be used to compare our GC results. The solutions are used in Sect. 5.2 and 5.5 in the revised manuscript.

(5) Line 289: What does "cooperating" mean in that case?

R6: Replaced with "using" in the revised manuscript.

(6) The specific yield for the Queensland and Victoria networks differ by a factor of 2: Is there any geologic evidence/argument available for those very different yield factors?

R7: We took the specific yield values simply from the literatures (Chen et al., 2016; Rassam et al., 2013, Welsh 2008).

(7) Sect. 6.2 appears to be rather an outlook to a future study. Since no actual results are presented, I am wondering if this section should not be better removed entirely?

R8: We thank for reviewer's suggestion. Section 6.2 is removed from the manuscript.as suggested.

(8) Major mining activities are currently taking place at the North West Plateau where GRACE picks up negative mass trends: What measures have been taken to reduce mass loss un-related to the terrestrial water cycle from the GRACE data in particular in that area?

R9: Reviewer made a good point. However, it is likely that the negative mass trends is mainly caused by the declining rainfall after 2000 (van Dijk et al., 2011), and less likely mining activity, as confirmed by the government WA Department of Water.

http://www.news.com.au/technology/environment/climate-change/nasa-study-says-the-canning-basin-in-wa-is-being-depleted-too-fast/news-story/9bf107b8299c19b57904ed719639a0ba

Reference

van Dijk, A. I. J. M., L. J. Renzullo, and M. Rodell (2011), Use of Gravity Recovery and Climate Experiment terrestrial water storage retrievals to evaluate model estimates by the Australian water resources assessment system, Water Resour. Res., 47, W11524, doi:10.1029/2011WR010714.

Reviewer 2

We would like to acknowledge the insightful comments and suggestions provided by reviewer 2. We will consider the reviewer's suggestions in our revised manuscript. Followings are the responses (R) based on the comments:

Line 152: Equation 9 is the most important equation in this study, but some of the information is provided in the later section 3. Also the model covariance matrix is provided in section 4.2. Authors might consider making the method section clearer and reduce some unnecessary equations.

R1: We organize the paper to introduce mathematical developments and followed by implementation. Section 2 describes the general methodology, while sections 3 and 4 describe the specifics needed to implement the algorithm with the ITSG data and CABLE model, respectively. We believe this is a logical presentation of our methods. This is clarified in the revised manuscript, lines **116 – 120** 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."

Line 170: Basically, the paper claimed "the use of intersatellite tracking data", but the data was the normal matrix N and vector c obtained from the ITSG-Grace 2016 as well as the gravity field coefficient from GOCO05s solution. No Level 1B data was actually used directly in this study, so I wonder whether the title is appropriate.

R2: As reviewer pointed out correctly, we started from the L1B normal equation data, not directly from L1B data. We revised the titles as follows: "On the use of GRACE normal equation of intersatellite tracking data for improved estimation of soil moisture and groundwater in Australia".

Line 195: The GRGS gridded TWS products were used in Equation 9 to work out the TWS values outside Australia. The L3 GRGS products derived from the Earth's geopotential coefficients up to degree and order 80, while ITSG data used in the study were up to 90. Why not using the ITSG TWS data? Can the ITSG normal equation represent the uncertainty in L3 GRGS products?

R3: Eq. (9) requires the knowledge of ΔTWS grid outside the study region. However, the ITSG does not provide such gridded data directly. GRGS solution is one of the best possible candidate suitable for correcting the outside effect since it is most compatible with normal equation dataset we use in terms of a resolution and it does not require the (subjective) post-processing of the solutions unlike ITSG and other L2 solutions. This is clarified in lines **240 – 241** of the revised manuscript.

Line 210: The gridded GRGS data was resampled to 0.5 degree spatially, but the normal equation only contains the information to degree 90. How did you deal with the different spatial scale in the error variance-covariance matrix?

R4: This study uses the least-square combination approach in spatial domain. It is possible to compute the ΔTWS from the normal equation at any spatial scale (here is 0.5 degree). However, it is noted that the 0.5 degree data are spatial correlated (because of 90 degree resolution), and such correlation in the 0.5 degree grid is already accounted in the GC approach described in Sect. 2.

Line 231: depth between 0.022 m not cm

R5: Reviewer is correct. We correct 0.022 cm to 0.022 m in the revised manuscript.

Line 256: The sensitivity study of the model parameters is an important process but not necessary to show in the paper. Author may consider removing table 2.

R6: We thank for reviewer's suggestion. We still believe that Table 2 is worth to present since it gives an insight about the model parameters in a direct relevance to the storage outputs. Therefore, we decide to keep Table 2 in the revised manuscript.

Line 298: Did you do the CDF matching for few years and validate the results for the rest of time period? Or did you match all the time series and validate the same time period? If so, your estimates and observations are not independent. The CDF matching may discard important signals of the observations. Since only correlation was calculated, CDF matching is not necessary.

R7: We built the CDF using 2003-2004 data and applied it to the rest. As the NS coefficient (not correlation coefficient) is used in this section, the remaining bias might result in poor NS values, the bias correction is then necessary. For clarity, we added the additional statement in lines **335 – 336** of the revised manuscript as follows:

"The CDF is built using the 2003-2004 data, and it is used for the entire period"

Line 301: The variability of soil moisture inside a basin is quite high. The average of basin and monthly soil moisture can smooth out lot of signals. Since your output is 0.5_x 0.5_ gridded products, why not validate at this scale instead of basin scale? Can you show some validation with in-situ measurements?

R8: We indeed conducted the analysis at 0.5x0.5 scale (for surface SM), and it provides very similar conclusions to the basin mean. Both results do not show the significant improvement for the surface soil moisture computation, which is supported by the recent publication (Tian, et al., 2017). We discussed why this is the case and the limitation of GRACE data in the original manuscript.

Line 307: The groundwater estimates were only validated for two states using the state average. It should be possible to validate all the states over Australia or at basin scale to be consistent with other results. Two states are not sufficient to support the improvement in groundwater storage estimates over Australia.

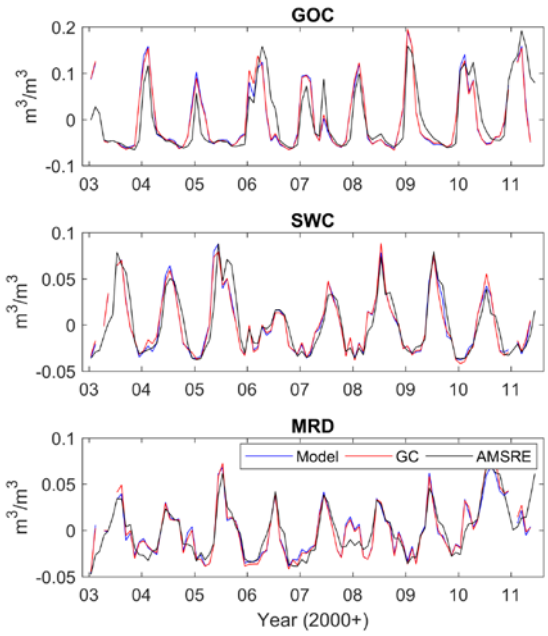
R9: We agree with reviewer. However, the primary purpose is to demonstrate a statistically rigorous way of using GRACE datasets (not high level processed data) for combining with the LSM models and eventually for data assimilation by emphasizing the importance of using error propagation and full covariance. The ground observation network of Australia is relatively sparse particularly at the state-wide level. Our study is initially based on what we could get.

Line 320: Is that only one value of specific yield per state was used to convert the groundwater level to storage? Will it be more appropriate to use different specific yields for different locations and calculate the average?

R10: We agree with reviewer that it is ideal to conduct the conversion at all grid cell independently. However, no such information is available at the grid cell level. We used the best possible knowledge of specific yield from the published literatures.

Line 468: The difference between model and GC approach for soil moisture is marginal here for basin monthly average. Can you show some time series examples of GC results and AMSR-E retrievals?

R11: Below are the examples of the time series between our estimates and AMSR-E (C-band):



The statistical value can be found in **Table 3** of the revised manuscript. As stated in the manuscript, no significant change is seen in GC solution, likely due to limitation of GRACE's temporal and spatial resolution. This same conclusion was reported by Tian et al. (2017). We provide a reference and recommendation how SM estimates can be improved in the conclusion section, lines **704 – 710** of the revised manuscript.

"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."

Line 482: In Table 4, the trend of GC approach is about one third of in-situ measurements for Queensland. What causes such big difference?

R12: As a matter of fact that the GC approach optimally combines the GRACE observation with the model results, the GC result is inevitably influenced by the model estimate. As seen in Fig. 7 and 9, the GC result moves toward the in situ measurement but it is still influenced by the ΔGWS estimation from the model.

Line 491: It will be interesting to see the groundwater storage change in Murray-Darling Basin after the GC approach compared with in-situ measurements, during the big drought and big wet period.

R13: The reviewer's suggestion is very interesting. However, we did not have the in situ groundwater data of the Murray-Darling Basin by the time of this study and so we conducted the analysis based on what we could obtain. We will consider the validation in the Murray-Darling Basin in our future study.

Line 502: This section investigates the mass variation in the past 13 years based on the GC approach. Figure 8 is a quite good demonstration of the mass variation at different layers of water storage. The top and root-zone soil moisture show quite different trends. The root-zone soil moisture has similar trends with TWS and groundwater for most of the basins. It will be better to have some validation of root-zone soil moisture estimates and more sufficient groundwater storage estimates to support the analysis in this section.

R14: We appreciate reviewer suggestion and we strongly agree. However, similar to what we stated above, we did not have an access to the in situ root-zone soil moisture of Australia, and the validation

of such a component was not possible. We would like to reiterate that this is the first demonstration of new technique of using more fundamental dataset from GRACE, applied to the Australian continent.

Line 579: 0.39 is App1 and 0.42 is App2? So the trend calculating from GRACE subtracting modeled soil moisture is the same with modeled groundwater trend (in Table 4). The NS value for App1 is 0.46, which is less than the CABLE model without GRACE? For Victoria, the NS value of App2 is 0.3 less than CABLE model without GRACE too. With the assimilation of GRACE in App2, the correlation is degraded. It seems model itself without GRACE is better compared to App1 or App2. Still, for only two states validation results, it's hard to demonstrate that GC approach works better due to the error information. It could be your model uncertainty is better estimated using the ensembles as explained in Section 4.2. When you do the App1 and App2, did you also used 7 precipitation dataset as the same as the GC approach? Please clarify.

R15: Reviewer is correct that App1 and App2 provide poorer results compared to CABLE (in some case) or GC approach. However, this section demonstrates the scenario when different ΔGWS computation approach are used (not GC). The model part remains the same (as in GC approach). App1 uses mascon solution with error free scenario while App2 uses mascon with its variance matrix, and the different outputs are mainly attributed to the different application of the uncertainty type. The poor results of App1 and App2 are mainly due to too simplified error information implemented. This has been clarified in the revised manuscript lines **630 – 632**:

"Note that the model uncertainty remains the same as in GC approach (Sect. 4.2). The different outputs between App1 and App2 are mainly attributed to the different application of the different estimates of the uncertainty."

Line 587: The future work in this section is interesting but no results were provided. Author may consider removing this section completely or providing the results in this paper together with the GC approach.

R16: We thank for reviewer's suggestion. Section 6.2 is removed from the manuscript.as suggested.

Reviewer 3

We would like to thank reviewer 3. Followings are the responses (R) based on the comments:

L15-L16–> This is not true that ”there is no covariance matrix for L2 products”. After filtering and conversion to TWS, this error can be propagated, which is addressed e.g., in DOI:10.1007/s10712-014-9309-8. L48 repeats the same claim!

R1: Review is correct that the covariance matrix is available for L2 product. However, it is not the case for the gridded product (Level 3) discussed here. This is clearly written in lines **13 – 16** of the submitted manuscript:

“… from the high level products (e.g., land grid). The gridded data products are subjected to several drawbacks such as signal attenuation and/or distortion caused by ad hoc posteriori filters, and a lack of error covariance information.”

For clarity, we include the additional information in line **16** of the revised manuscript as follows:

“… from the high level products (e.g., land grid from the Level 3 product). The gridded data products …”

L16-L17: The consequence is undesired alteration of ... data and its statistical property. It is not clear what this means. Are you suggesting that all other published papers are wrong!?

R2: The post-processed (high level grid) data are often used without a proper statistical information. It changes information GRACE data provides. For clarity, we revised the statement as follows:

“The post-processing process of GRACE data might lead to the undesired alteration of the signal and its statistical property.”

L21-L22–> This is not clear which approach has been used.

R3: The approach used in this study is the least-squares combination. This is clearly mentioned in the earlier sentence (lines **19 – 22** of the submitted manuscript):

“The approach combines the GRACE’s least-squares normal equation (full error variance-covariance information) of L1B data with the results from the Community Atmosphere Land Exchange (CABLE) model to improve soil moisture and groundwater estimates.”

L61-L64–> Inversion techniques for signal separation have been applied, which consider errors in GRACE and complementary data used for signal separation. DOI:10.1016/j.jog.2012.03.001; DOI:10.1016/j.jog.2011.02.003; DOI:10.1007/s10712-016-9403-1

R4: Based on the suggested literatures, the text was revised as follows.

“Several techniques have been developed to separate different signals considering the errors in GRACE and other data (Rietbroek et al., 2012; Schmeer et al., 2012; Forootan et al., 2017). However, the GRACE uncertainty is often derived empirically, not necessarily reflecting the actual GRACE error characteristics. The empirical GRACE errors have been also used in the data assimilation (e.g., Zaitchik et al., 2008; Tangdamrongsub et al., 2015; Tian et al., 2017).”

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.

R5: As written in the original manuscript, we use the normal equation data from the ITSG-2016. The description of the data can be found in the data webpage provided in the manuscript

(https://www.tugraz.at/institute/ifg/downloads/gravity-field-models/itsg-grace2016). As the derivation of the normal equation is not the focus of this study, we do not discuss it further but refer to the description in the data webpage and references therein (this is stated clearly in **Sect. 3.1** of the original manuscript). The element of the matrix A is mainly the partial derivative of the variational equation respect to the orbital information and gravitational coefficients. The variational equation include both orbital information and various kinds of L1B data including KBR data.

The accuracy of recovery has not been justified, which is essential for any scientific appliction to show that the accuracy of software is comparable with official products. Please include comparisons with the official ITSG2016 monthly solutions.

R6: The objective of this study is to combine GRACE normal equation with land surface model result to improve the estimated SM and GWS, and do not independently resolve the GRACE solution. We already presented the validation of SM and GWS computations. This is stated clearly in the introduction lines **80 – 82** of the submitted manuscript:

"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).

Results of the inversion might be compared with those that assimilate GRACE into hydrological models to improve the surface/sub-surface storage compartments. Recent studies over Australia include: DOI:10.1002/2016WR019641; DOI:10.1016/j.advwatres.2017.07.001

R7: We thank for reviewer's suggestion. We already cited Tian et al. (2017) and approached the same conclusion about the benefit of GRACE on the groundwater estimate, and its limitation of the surface soil moisture.

New references added

Forootan, E., Safari, A., Mostafaie, A., Schumacher, M., Delavar, M., and Awange, J. L.: Large-Scale Total Water Storage and Water Flux Changes over the Arid and Semiarid Parts of the Middle East from GRACE and Reanalysis Products, Surv. Geophys., 38:591-615, doi:10.1007/s10712-016-9403-1, 2017.

Rietbroek, R., Fritsche, M., Brunnabend, S.-E., Daras, I., Kusche, J., Schröter, J., Flechtner, F., and Dietrich, R.: Global surface mass from a new combination of GRACE, modelled OBP and reprocessed GPS data, J. Geodyn., 59 - 60:64 - 71, doi:10.1016/j.jog.2011.02.003, 2012.

Schmeer, M., Schmidt, M., Boschb, W., and Seitz, F.: Separation of mass signals within GRACE monthly gravity field models by means of empirical orthogonal functions, J. Geodyn., 59 - 60:124 - 132, doi:10.1016/j.jog.2012.03.001, 2012.

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., 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 ad hoc 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. To exploit the GRACE information rigorously and negate these limitations, this study uses the fundamental GRACE satellite tracking Level 1B (L1B) data, not the post-processed ΔTWS grid data. Our approach combines the GRACE's least-squares normal equation (full error variance-covariance information) of L1B data 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 raw GRACE data. The GRACE-combine (GC) approach is developed for optimal least-squares combination maximizing the strength of the model and observations while suppressing the weaknesses. The approach is applied to estimate the soil moisture and groundwater over 10 Australian river basins and 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 likely due to limitation of GRACE's temporal and spatial resolution.

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

Deleted: intersatellite tracking data

Deleted: process

Deleted: The consequence is undesired alteration of ΔTWS data and its statistical property

Deleted: The

different product levels (L) that increase in the amount of processing, L1B – satellite tracking data (Wu et al., 2006), L2 – global gravitational Stokes coefficients (Bettadpur, 2012), and L3 – global grids (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 (Jekeli, 1981), destripe filter (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). 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 GRACE and of the land surface model outputs are neglected in the combination of two noisy data. Several techniques have been developed to separate different signals considering the errors in GRACE and other data (Rietbroek et al., 2012; Schmeer et al., 2012; Forootan et al., 2017). However, the GRACE uncertainty is often derived empirically, not necessarily reflecting the actual GRACE error characteristics. The empirical GRACE errors have been also used in the data assimilation (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 (Eicker et al., 2014, Schumacher et al., 2016; Tangdamrongsub et al., 2017), however, the GRACE signal used were still affected by the post-processing filters.

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

Deleted: signal separation

Deleted: , which considered

Deleted: complementary

Deleted: in the signal separation process

Deleted: commonly

Deleted: true

Deleted: Similar issue is seen

Deleted: application

Deleted: In data assimilation application, albeit its importance, the GRACE uncertainty is commonly derived empirically not necessarily reflecting the true GRACE error characteristics (e.g., Zaitchik et al., 2008; Tangdamrongsub et al., 2015; Tian et al., 2017).

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, 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). 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.

Deleted: In Sect. 6, the purposed approach is discussed in terms of effectiveness, and data assimilation implementation.

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

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

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

where $\boldsymbol{h}$ is the "truth" (unknown) model state vector while $\widetilde{\boldsymbol{h}}$ 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., $\boldsymbol{h_R}$, $\mathbf{C_R}$) 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:

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

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:

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

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 linearized GRACE satellite-tracking observation equation is formulated as:

$$\boldsymbol{y} = \mathbf{A}\boldsymbol{x} + \boldsymbol{e}; \boldsymbol{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \quad (4)$$

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. (4) can be modified explicitly in terms of soil moisture and groundwater storage variations as:

Deleted: the

$$\boldsymbol{y} = \mathbf{A}\mathbf{S}\bar{\mathbf{Y}}\mathbf{H}\boldsymbol{h} + \boldsymbol{e}; \boldsymbol{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \tag{5}$$

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, and $\mathbf{H}$ is the operational matrix converting $\Delta \boldsymbol{SM}_{top}$, $\Delta \boldsymbol{SM}_{rz}$, and $\Delta \boldsymbol{GWS}$ to ΔTWS. This model is based on the assumption that the GRACE orbital perturbation is a result of ΔTWS variation on the surface, which is very particular in Australia. For convenience, the term $\mathbf{Y} = \mathbf{S}\bar{\mathbf{Y}}$ is used in the further derivation. 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, the dimension of $\mathbf{Y}$, $\mathbf{H}$, and $\boldsymbol{h}$ are $L\times M$, $M\times 3M$, and $3M\times 1$, respectively.

A least-squares solution of Eq. (5) is given as:

$$(\mathbf{H}^T\mathbf{Y}^T\mathbf{A}^T\boldsymbol{\Sigma}^{-1}\mathbf{A}\mathbf{Y}\mathbf{H})\hat{\boldsymbol{h}} = \mathbf{H}^T\mathbf{Y}^T\mathbf{A}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y}. \tag{6}$$

It can be simplified as:

$$\mathbf{H}^T\mathbf{Y}^T\mathbf{N}\,\mathbf{Y}\mathbf{H}\,\hat{\boldsymbol{h}} = \mathbf{H}^T\mathbf{Y}^T\boldsymbol{c}, \tag{7}$$

where $\mathbf{N} = \mathbf{A}^T\boldsymbol{\Sigma}^{-1}\mathbf{A}$ and $\boldsymbol{c} = \mathbf{A}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y}$. (The rationales of introducing $\mathbf{N}$ and $\boldsymbol{c}$ are explained in the following section). Note that, the above derivations (Eq. (5) – Eq. (7)) are defined with the global grid of $\boldsymbol{h}$. For a regional application, Eq. (7) can be modified as:

$$\left[\mathbf{H}_R^T\mathbf{Y}_R^T \mid \mathbf{H}_o^T\mathbf{Y}_o^T\right]\mathbf{N}\begin{bmatrix}\mathbf{Y}_R\mathbf{H}_R\\ \mathbf{Y}_o\mathbf{H}_o\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{h}}_R\\ \hat{\boldsymbol{h}}_o\end{bmatrix} = \left[\mathbf{H}_R^T\mathbf{Y}_R^T \mid \mathbf{H}_o^T\mathbf{Y}_o^T\right]\boldsymbol{c}, \tag{8}$$

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 $\mathbf{Y}_R$, $\mathbf{H}_R$, $\hat{\boldsymbol{h}}_R$, $\mathbf{Y}_o$, $\mathbf{H}_o$, $\hat{\boldsymbol{h}}_o$ are $L\times J$, $J\times 3J$, $3J\times 1$, $L\times$ (M–J), (M–J)×3(M–J), 3(M–J)×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. (8), the normal equations associated with ΔTWS in the region of interest can then be written as

$$\mathbf{H}_R^T\mathbf{Y}_R^T\mathbf{N}\mathbf{Y}_R\mathbf{H}_R\hat{\boldsymbol{h}}_R = \mathbf{H}_R^T\mathbf{Y}_R^T\boldsymbol{c} - \mathbf{H}_R^T\mathbf{Y}_R^T\mathbf{N}\mathbf{Y}_o\mathbf{H}_o\hat{\boldsymbol{h}}_o \tag{9}$$

or

$$\mathbf{N}_R\hat{\boldsymbol{h}}_R = \boldsymbol{c}_R \tag{10}$$

where $\mathbf{N}_R = \mathbf{H}_R^T\mathbf{Y}_R^T\mathbf{N}\mathbf{Y}_R\mathbf{H}_R$ and $\boldsymbol{c}_R = \mathbf{H}_R^T\mathbf{Y}_R^T\boldsymbol{c} - \mathbf{H}_R^T\mathbf{Y}_R^T\mathbf{N}\mathbf{Y}_o\mathbf{H}_o\hat{\boldsymbol{h}}_o$. As seen, Eq. (9) is the regional representation of Eq. (7) 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. (10), the optimal combined solution of $\widehat{\boldsymbol{h}}_{\boldsymbol{R}}$ can be resolved as follows:

$$\widehat{\boldsymbol{h}}_{\boldsymbol{R}} = \left(\mathbf{C}_{\boldsymbol{R}}^{-\mathbf{1}} + \mathbf{N}_{\boldsymbol{R}}\right)^{-1}\left(\mathbf{C}_{\boldsymbol{R}}^{-\mathbf{1}}\widetilde{\boldsymbol{h}}_{\boldsymbol{R}} + \boldsymbol{c}_{\boldsymbol{R}}\right) \quad (11)$$

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

$$\widehat{\boldsymbol{\Sigma}} = (\mathbf{C}_{\boldsymbol{R}}^{-1} + \mathbf{N}_{\boldsymbol{R}})^{-1}, \quad (12)$$

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

$$\boldsymbol{\sigma}_{\widehat{\boldsymbol{h}}} = \sqrt{diag(\widehat{\boldsymbol{\Sigma}})}, \quad (13)$$

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 (Mayer-Gürr et al, 2016; https://www.tugraz.at/institute/ifg/downloads/gravity-field-models/itsg-grace2016) 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. (7)) 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:

$$\mathbf{N}\,\Delta\boldsymbol{x} = \boldsymbol{d} \quad (14)$$

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. (14), and the mean gravity field ($\overline{\boldsymbol{x}}_{\boldsymbol{0}}$) computed from all $\boldsymbol{x_0}$ between January 2003 and March 2016 is then removed as follows:

$$\mathbf{N}\,(\Delta\boldsymbol{x} + \boldsymbol{x_0} - \overline{\boldsymbol{x}}_{\boldsymbol{0}}) = \boldsymbol{d} + \boldsymbol{N}(\boldsymbol{x_0} - \overline{\boldsymbol{x}}_{\boldsymbol{0}}) \quad (15)$$

$$\mathbf{N}\,\boldsymbol{x} = \boldsymbol{d} + \boldsymbol{N}(\boldsymbol{x_0} - \overline{\boldsymbol{x}}_{\boldsymbol{0}}) \quad (16)$$

Therefore, in Sect. 2 (e.g., Eq. (7) – (11)), the matrix $\mathbf{N}$ remains unchanged while the vector $\boldsymbol{c}$ can be simply replaced by $\boldsymbol{c} = \boldsymbol{d} + \boldsymbol{N}(\boldsymbol{x_0} - \overline{\boldsymbol{x}}_{\boldsymbol{0}})$.

3.2 GRACE-derived $\Delta\boldsymbol{TWS}$ products

Two monthly GRACE-derived ΔTWS products are also used, the CNES/GRGS Release 3 (RL3) (GRGS for short, Lemoine et al., 2015) and the JPL RL05M mascon-CRI version 2 product (mascon for short, Watkins et al., 2015; Wiese et al., 2016). 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 GRGS solution, 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 GRGS product.

The ΔTWS products are obtained between January 2003 and March 2016. The GRGS solution is retrieved from http://grgs.obs-mip.fr/grace/variable-models-grace-lageos/grace-solutions-release-03 while the mascon is from http://grace.jpl.nasa.gov/data/get-data/jpl_global_mascons. 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 spatial resolution of two datasets are resampled to 0.5°×0.5° using the nearest grid values.

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

1. To obtain the ΔTWS values outside Australia. As shown in Eq. (9), the $\widehat{\boldsymbol{h}}_{\boldsymbol{o}}$ vector needs to be known, which can be from the GRACE-derived ΔTWS solution. We use the GRGS solutions as the GRGS solution provides ΔTWS at a spatial resolution comparable to the normal equation data.
2. To compare with the ΔTWS estimates from our approaches. Both GRGS and JPL mascon 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 for this study. CABLE 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 modelled to be 20 m uniformly. Recalling Eq. (3), $\Delta \boldsymbol{SM}_{top}$ is defined as the soil moisture storage variation at the top 0.022 m thick layer, while $\Delta \boldsymbol{SM}_{rz}$ 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-

Deleted: c

Deleted: (Dirmeyer et al., 2011; http://hydro.iis.u-tokyo.ac.jp/GSWP3)

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:

$$\boldsymbol{\mathcal{H}_R} = [\boldsymbol{h_e}|_{k=1} \quad \boldsymbol{h_e}|_{k=2} \quad \boldsymbol{h_e}|_{k=3} \quad \dots \quad \boldsymbol{h_e}|_{k=K}] \tag{17}$$

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

$$\tilde{\boldsymbol{h}}_{\boldsymbol{R}} = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{h_e}|_k \tag{18}$$

Note that due to the absence of GSWP3, Princeton, and MERRA data, the number of ensembles reduce 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:

$$\mathbf{C}_{\boldsymbol{R}} = \boldsymbol{\mathcal{H}_R}'(\boldsymbol{\mathcal{H}_R}')^T/(K-1) \tag{19}$$

The $\widetilde{\boldsymbol{h}}_{\boldsymbol{R}}$ (Eq. (18)) and $\mathbf{C}_{\boldsymbol{R}}$ (Eq. (19)) terms can be directly used in Eq. (11).

Note that 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 modelling 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

Deleted: assume

Deleted: for such model error

Deleted: assume

Deleted: the

Deleted: this way

Deleted: t

Deleted: o be a good compromise between

Deleted: ing of

Deleted: and

Deleted: ing

Deleted: upper limit of the

Deleted: suggested

Deleted: However, it should be noted that t

Deleted: assumption has a limitation of representing the error characteristic and more investigation is needed in order to acquire a more realistic omission error

Deleted: cooperating

Deleted: Since t

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. (A1)) (Fig. 2).

Figure 2 demonstrates CABLE ΔTWS varies noticeably by precipitation as well as locations. The area-weighted average values (see Eq. (A2)) 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 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 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. 2.

Deleted: 7

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 2 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.

Figure 3 shows 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. 3d), 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. 3c) 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. 5). For comparison, the long-term trends of ΔTWS derived from the mascon and GRGS solutions are also shown (Fig. 5a, 5b). From Fig. 5d, 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. 5e) and the GC approach does not change (Fig. 5f). The largest adjustment is seen in ΔSM_{rz} and ΔGWS components, to be consistent with the GRACE data in most basins (Fig. 5h, 5j).

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. 3d). Applying the GC approach clearly improves the trend (Fig. 5c vs. 5d). The other example is seen over the western part of the continent (see rectangular area in Fig. 5c, 5d) 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. 3h, 4d, 4p). 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. 5) 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 is 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. (13). 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. 6). 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 6 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. 6e. 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.

Deleted: 7

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. 7, 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. 7a). 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. 8. 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. 3, 4). 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. 8c). 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. 8a and 8c). 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. 8j and 8l). 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

Deleted: 6. Further development of GC approach¶
6.1

of model outputs and GRACE data, the water storage states are updated through a Kalman filter:

$$\widehat{\boldsymbol{h}}_{\boldsymbol{R}} = \widetilde{\boldsymbol{h}}_{\boldsymbol{R}} + \mathbf{H}\mathbf{C}_{\boldsymbol{R}}^{\boldsymbol{T}}(\mathbf{H}\mathbf{R}\mathbf{H}^{\boldsymbol{T}} + \mathbf{C}_{\boldsymbol{R}})^{-\mathbf{1}}(\boldsymbol{b} - \mathbf{H}\widetilde{\boldsymbol{h}}_{\boldsymbol{R}}) \quad (20)$$

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 $\mathbf{R}$ 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 $\mathbf{R}$ 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. 9. 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 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 occur in groundwater layer. Rainfall inadequacy is attributed to the

Deleted: outputs

Deleted: between

Deleted: 6.2 GC data assimilation approach¶
We so far discussed the GC approach to update the water storage estimates independently every month. The approach can be easily expanded to sequentially update the model initial states whenever the GRACE observation is available (for example, every day) as in data assimilation (DA) like ensemble Kalman filter (Evensen, 2003) and particle filter (Weerts and El Serafy, 2006). We briefly describe a way of modifying the GC approach suitable for DA. The ensemble of simulated monthly water storage estimates is predicted based on the set of ensemble forcing data and model parameters. This is simply running CABLE for *K* (number of ensemble) times. When GRACE observation is available, the updated state is computed: ¶
$\widehat{\boldsymbol{h}}_{\boldsymbol{Re}} = (\mathbf{C}_{\boldsymbol{R}} + \mathbf{N}_{\boldsymbol{R}})^{-\mathbf{1}}(\mathbf{C}_{\boldsymbol{R}}\widetilde{\boldsymbol{h}}_{\boldsymbol{Re}} + \boldsymbol{c}_{\boldsymbol{Re}})$ . . . . (21)¶
where the subscript *e* represents the ensemble or perturbed version of the original vector or matrix (see e.g., Eq. (11)). The dimension of $\widehat{\boldsymbol{h}}_{\boldsymbol{Re}}$, $\widetilde{\boldsymbol{h}}_{\boldsymbol{Re}}$, $\boldsymbol{c}_{\boldsymbol{Re}}$ is $3J\times K$. The estimated $\widehat{\boldsymbol{h}}_{\boldsymbol{Re}}$ can be directly used as in the initial state for the next time step for CABLE run (Eicker et al., 2014; Tangdamrongsub et al., 2015; Tian et al., 2017), or used in the repeated run to avoid any spurious jump of the water storage estimates between the each step (Forman et al., 2012; Tangdamrongsub et al., 2017). This sequential update process can be carried out as long as desired. The feasibility of GRACE DA has been demonstrated with "devised" uncertainty (covariance) information. As a future work, we will develop new DA approach on the basis of full error information of GRACE data by using the least-squares normal equation and thus carrying the error information from the fundamental (satellite tracking) data level. ¶
¶
7

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.

Finally, 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 was funded by NASA projects on GRACE and GRACE Follow-On missions and partly by Australian Research Council (DP160104095). 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.

Deleted: is

Deleted: University of Newcastle to support

Deleted: 's

Deleted: .

Deleted: ark

Deleted: ecker

Deleted: yungjun

Deleted: im

Deleted:

Appendix: Nash-Sutcliff coefficient and area-weighted average

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

$$NS = 1 - \frac{\sum_{i=1}^{N}(\boldsymbol{y_i} - \widehat{\boldsymbol{x}_i})^2}{\sum_{i=1}^{N}(\boldsymbol{y_i} - \overline{\boldsymbol{y}})^2} \quad \text{(A1)}$$

where $\boldsymbol{y}$ is an observation vector, $\overline{\boldsymbol{y}}$ is the mean of the observation, $\widehat{\boldsymbol{x}}$ 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:

$$\bar{Z} = \frac{\sum_{j=1}^{M} w_j \bar{z}_j}{\sum_{j=1}^{M} w_j} \quad \text{(A2)}$$

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.

Deleted: Dirmeyer, P. A.: A history and review of the Global Soil Wetness Project (GSWP), J. Hydrometeorol., 12, 729–749, doi:10.1175/JHM-D-10-05010.1, 2011.¶

Forootan, E., Safari, A., Mostafaie, A., Schumacher, M., Delavar, M., and Awange, J. L.: Large-Scale Total Water Storage and Water Flux Changes over the Arid and Semiarid Parts of the Middle East from GRACE and Reanalysis Products, Surv. Geophys., 38:591-615, doi:10.1007/s10712-016-9403-1, 2017.

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.

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.

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.

Rietbroek, R., Fritsche, M., Brunnabend, S.-E., Daras, I., Kusche, J., Schröter, J., Flechtner, F., and Dietrich, R.: Global surface mass from a new combination of GRACE, modelled OBP and reprocessed GPS data, J. Geodyn., 59 - 60:64 - 71, doi:10.1016/j.jog.2011.02.003, 2012.

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.

Schmeer, M., Schmidt, M., Boschb, W., and Seitz, F.: Separation of mass signals within GRACE monthly gravity field models by means of empirical orthogonal functions, J. Geodyn., 59 - 60:124 - 132, doi:10.1016/j.jog.2012.03.001, 2012.

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.

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 are 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

961 **Table 3**. NS coefficients between top soil moisture estimates and the satellite soil moisture
962 observations from AMSR-E products over 10 different Australian basins. The area-weighted
963 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

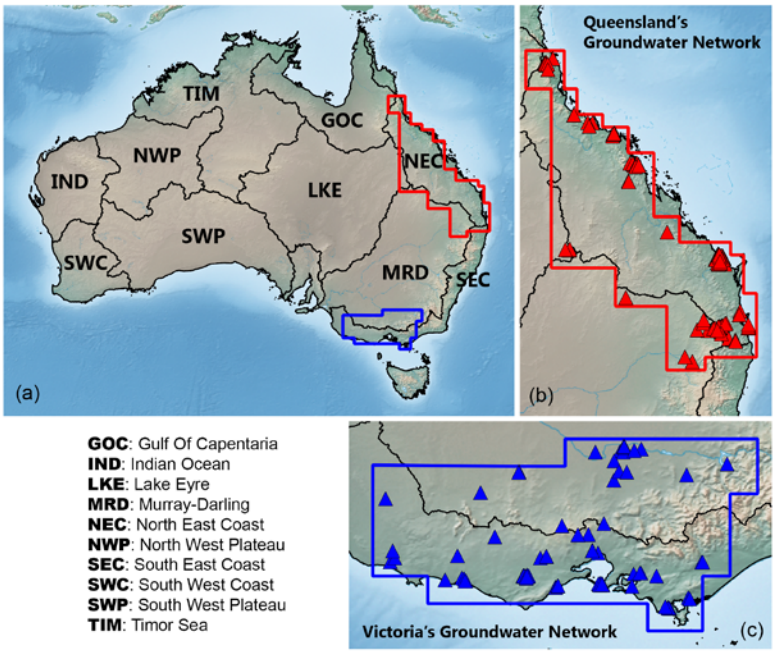
964

965

966 **Table 4**. NS coefficient and long-term trend of ΔGWS estimated from the model-only and
967 GC solutions in Queensland and Victoria groundwater network. The long-term trend of the
968 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

969



970

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

974

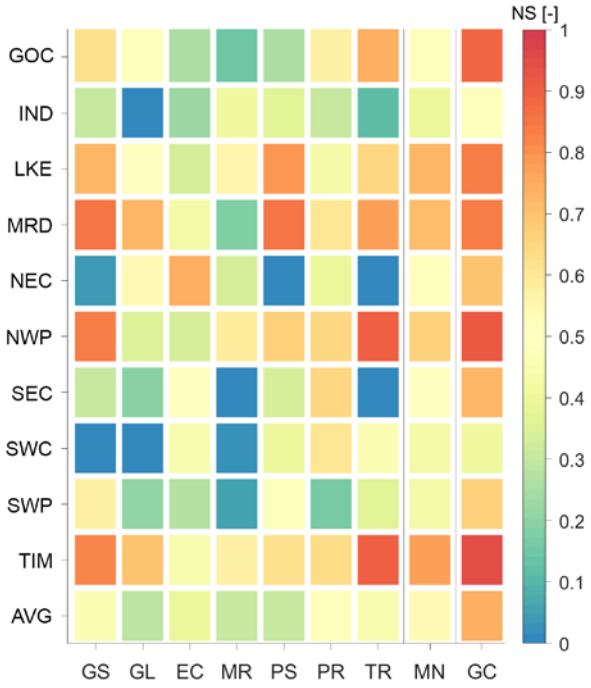


Figure 2. NS coefficients between the model and GRACE-mascon ΔTWS over 10 Australian basins (in ordinate). The NS values were computed based on CABLE ΔTWS computed with 7 different precipitation data (in abscissa), GSWP3 (GS), GLDAS (GL), ECMWF (EC), MERRA (MR), PERSIANN (PR), TRMM (TR). The NS value 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.

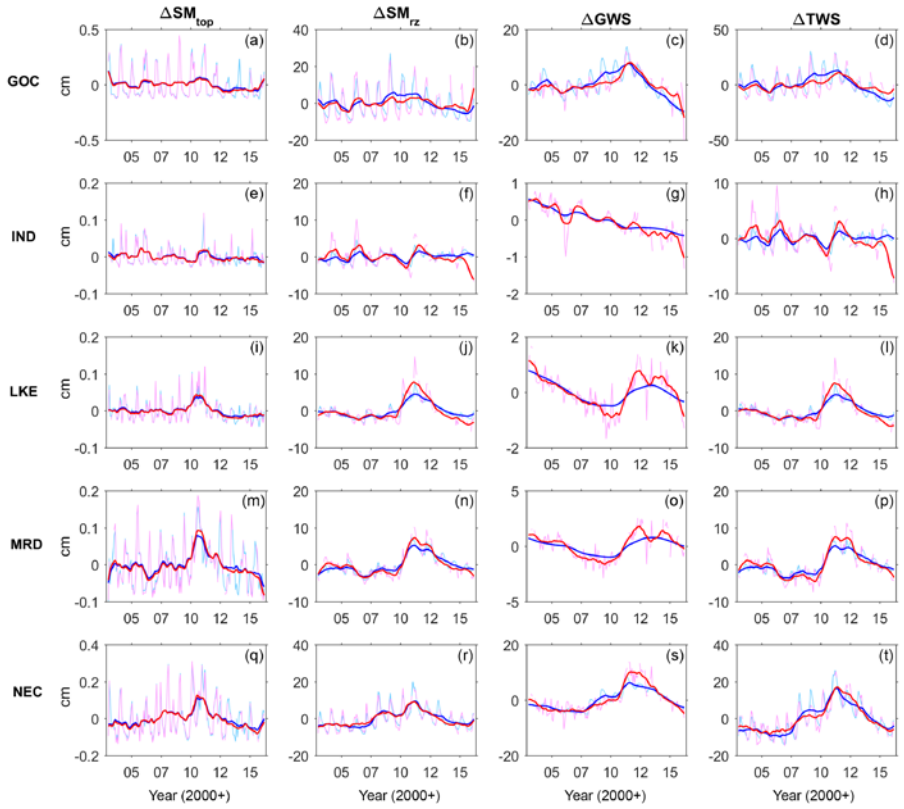
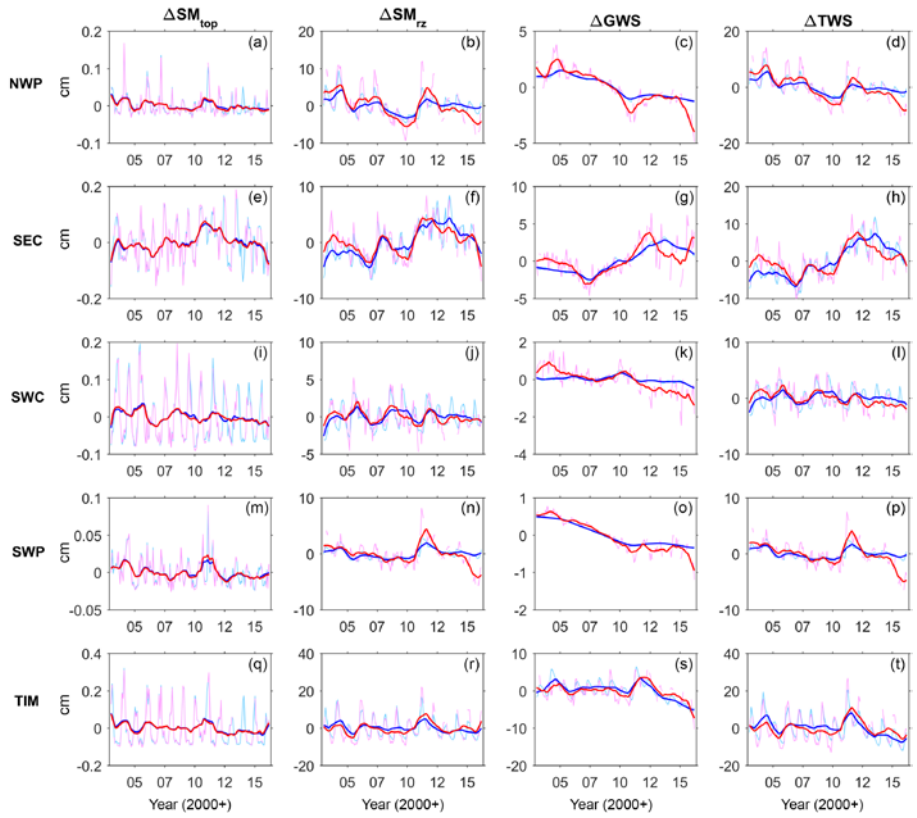


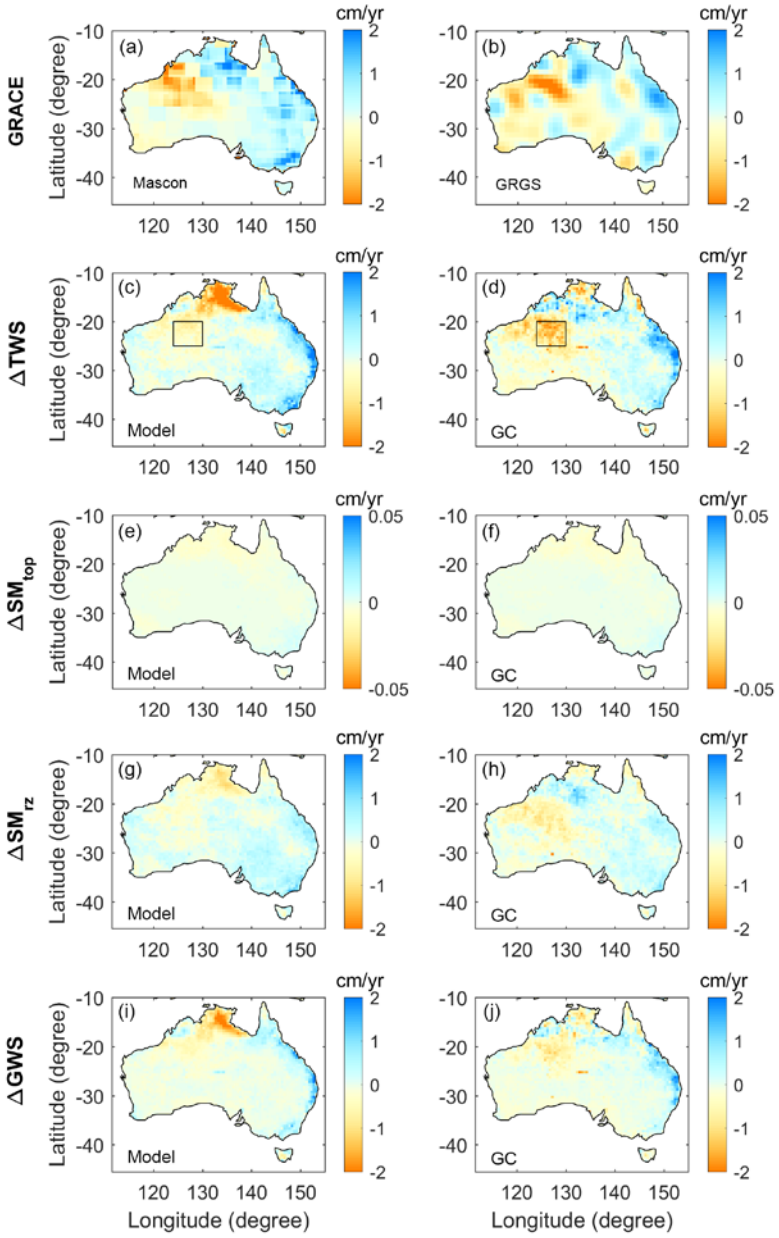
Figure 3. The monthly time series of ΔSM_{top}, ΔSM_{rz}, ΔGWS, and ΔTWS estimated from model (blue) and GC (red) solutions over Gulf of Carpentaria (GOC), Indian Ocean (IND), Lake Eyre (LKE), Murray-Darling (MRD), and North East Coast (NEC). The de-seasonalized time series is also shown.



990

991 **Figure 4.** Similar to Fig. 3, but estimated over North West Plateau (NWP), South East Coast
992 (SEC), South West Coast (SWC), South West Plateau (SWP), and Timor Sea (TIM).

993



994

995 **Figure 5.** Long-term trends of ΔTWS (c, d), ΔSM_{top} (e, f), ΔSM_{rz} (g, h), and ΔGWS (i, j)
996 estimated from the model-only (left) and the GC solutions (right). Results of GRACE ΔTWS
997 independently from mascon (a) and GRGS solution (b) are also shown. The eastern part of
998 North West Plateau basin is shown as a rectangle polygon in (c) and (d).

999

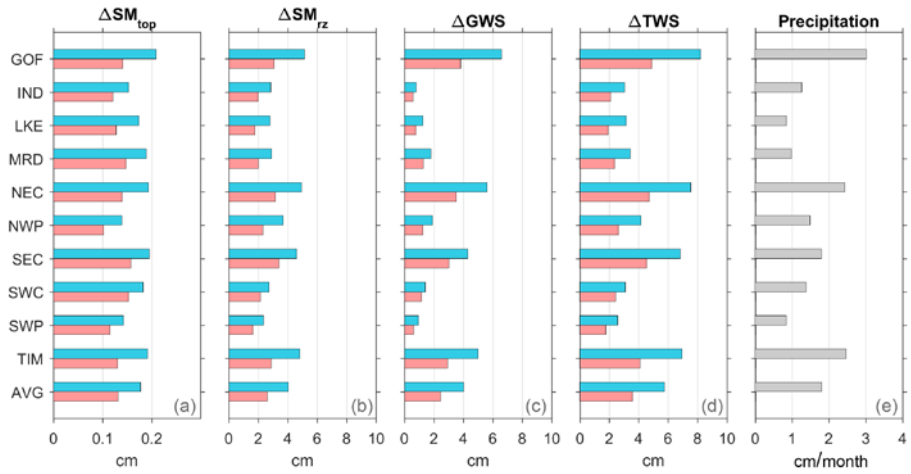


Figure 6. Uncertainties of ΔSM_{top}, ΔSM_{rz}, ΔGWS, and ΔTWS estimated from the model (blue) and the GC solutions (red) in 10 different Australian basins. The uncertainty of the precipitation is shown in (e). The area-weighted average value (AVG) is also shown.

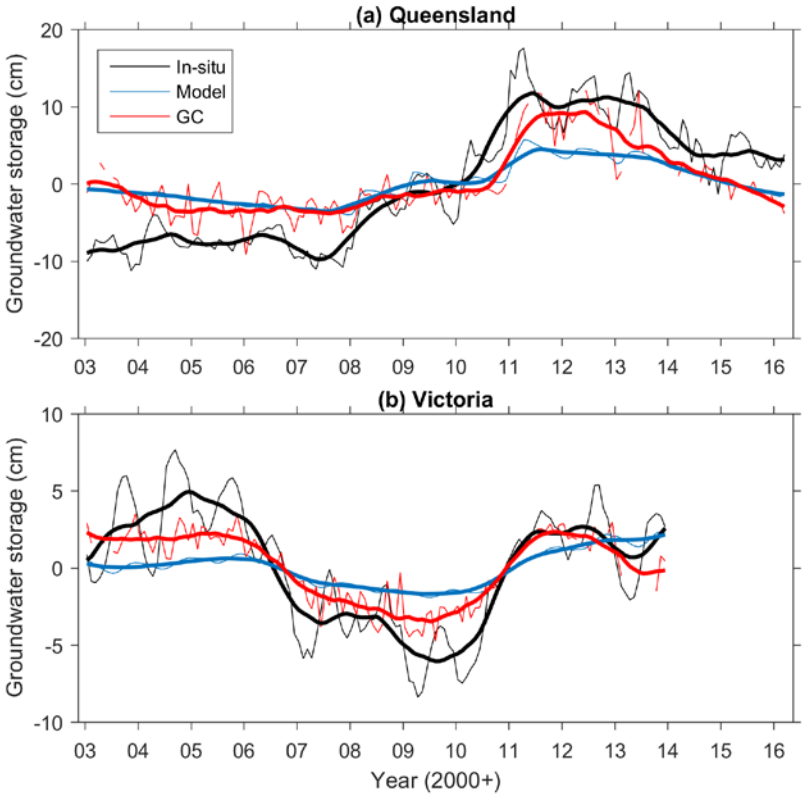


Figure 7. The monthly time series of ΔGWS estimated from the model, GC solutions, and measured from the in situ groundwater network in Queensland (a) and Victoria (b). De-seasonalized time series are shown in thick lines.

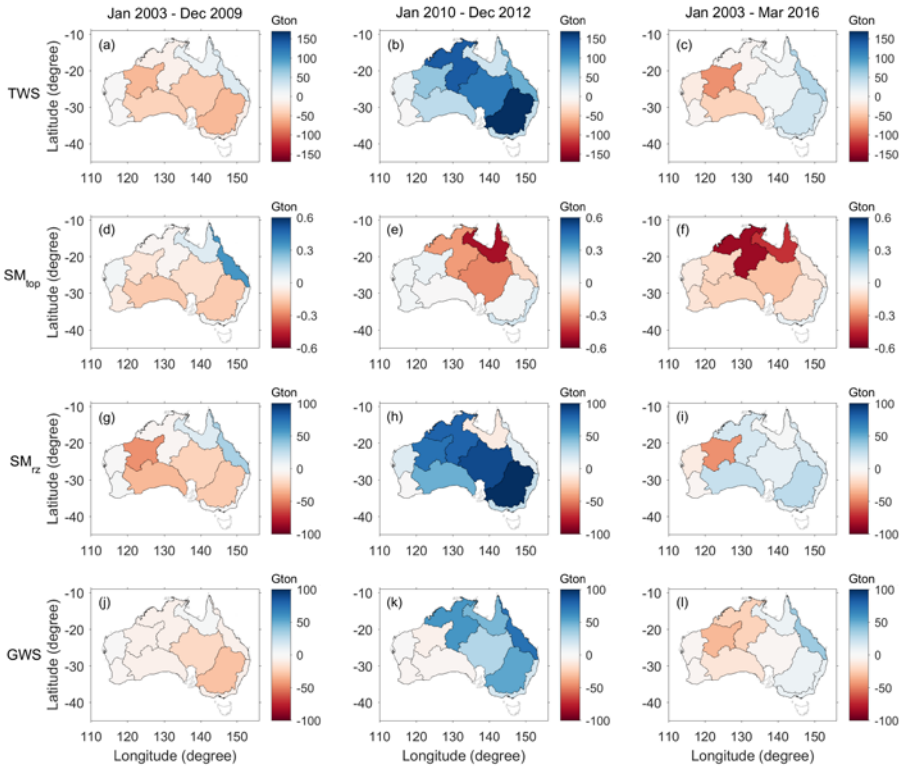


Figure 8. Mass changes (Gton, Giga tonne) of ΔTWS, ΔSM_{top}, ΔSM_{rz}, and ΔGWS estimated from GC solutions over 10 Australian basins in 3 different periods, Big Dry (January 2003 – December 2009), Big Wet (January 2010 – December 2012), and entire period (January 2003 – March 2016).

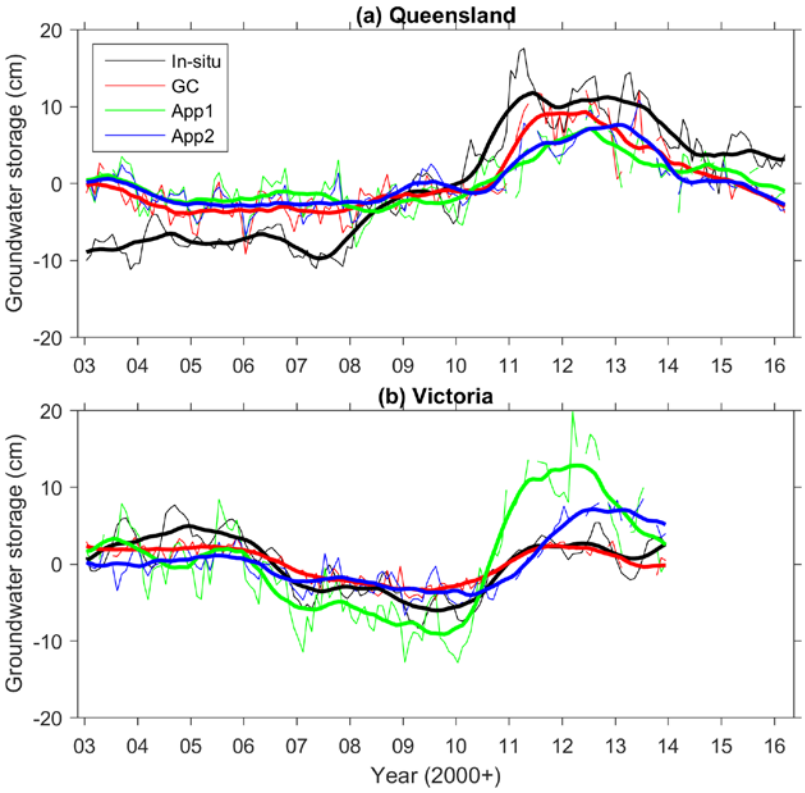


Figure 9. ΔGWS estimated from Approach 1 (App1) and Approach 2 (App2) in Queensland (a) and Victoria (b). The in-situ groundwater network data and the GC solutions are also shown. De-seasonalized time series are shown in thick lines.