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

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

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

Abstract

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

1. Introduction

The changes of Terrestrial Water Storage (ΔTWS) derived from the Gravity Recovery And Climate Experiment (GRACE) data products have been used in the last decade to study global water resources, including groundwater depletion in India and Middle East (Rodell et al., 2009; Voss et al., 2013), water storage accumulation in Canada (Lambert et al., 2013), flood-influenced water storage fluctuation in Cambodia (Tangdamrongsub et al., 2016). The gravity data obtained from GRACE satellites are commonly processed and released in three different product levels (L) that increase in the amount of processing, L1B – satellite tracking data (e.g., Wu et al., 2006), L2 – global gravitational Stokes coefficients (e.g., Bettadpur,

43 2012), and L3 – global grids (e.g., Landerer and Swenson, 2012). The original (L1B)
44 GRACE information is inevitably altered or sheered due to data processing and successive
45 post-processing filterings, because the error covariance information is not propagated through
46 each post-processing step.

47 The GRACE-derived ΔTWS has been computed widely from the higher-level products (e.g.,
48 L2 and L3) on which various ad hoc post-processing filters were applied (e.g., Gaussian
49 smoothing filter (e.g., Jekeli, 1981), destripe filter (e.g., Swenson and Wahr, 2006)). ΔTWS
50 obtained from these filters lacks proper error covariance information and is attenuated and
51 distorted. To overcome the signal attenuation in GRACE high-level products, empirical
52 approaches have been developed, including the application of scale factors computed from
53 land surface models (Landerer and Swenson, 2012) to the GRACE L3 products. GRACE
54 uncertainty in high level product is usually unknown or assumed. For example, Zaitchik et al.
55 (2008) derived empirically a global average uncertainty that is variable depending on choices
56 of post-processing filters (Sakumura et al., 2014). Furthermore, GRACE error and sensitivity
57 is dependent on latitudes due to the orbit convergence toward poles (Wahr et al., 2006) and
58 any post-processing filters will alter the GRACE data and their error information. Rigorous
59 statistical error information is of equal importance to derivation of ΔTWS for data
60 assimilation and model calibration (Tangdamrongsub et al., 2017; Schumacher et al., 2016,
61 2018). ΔTWS and its uncertainty estimates should be formulated directly from L1B data
62 considering the complete statistical information.

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

81 This study aims to use the GRACE information of ΔTWS measurement directly from the raw
82 L1B data. The approach optimally combines the GRACE's least-squares normal equations
83 with the model simulation results from the Community Atmosphere Land Exchange
84 (CABLE, Decker, 2015) to improve ΔSM and ΔGWS estimates. The proposed approach
85 presents three main advantages. Firstly, one can exploit the full GRACE signal and error
86 information by using the normal equation data sets. Secondly, the approach is developed for
87 optimal least-squares combination (e.g., Ramillien et al., 2004), which maximizes the model

88 and observation strength while simultaneously supressing their weaknesses. Finally, the
89 method bypasses empirical, multiple-step post-processing filters.

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

98 This paper is outlined as follows: Firstly, the derivation of GC approach is presented in Sect.
99 2 while the description of GRACE data processing, including the use of GRACE normal
100 equation, is given in Sect. 3. Secondly, the CABLE modelling is outlined in Sect. 4. This
101 includes the derivation of model uncertainty based on the quality of precipitation data and the
102 model parameter inputs. The processing of validation data is also described in Sect. 4.
103 Thirdly, Sect. 5 presents the result of ΔSM and ΔGWS estimates and comparison to in-situ
104 data. The long-term trends in the Australian mass variation over the last 13 years is also
105 investigated in this section.
106

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

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

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

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

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

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

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

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

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

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

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

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

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

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

131 where $\mathbf{Y}$ converts ΔTWS to geopotential coefficients considering the load Love numbers
132 (e.g., Wahr et al., 1998) and $\mathbf{H}$ is the operational matrix converting ΔSM_{top}, ΔSM_{rz}, and
133 ΔGWS to ΔTWS. Eq. (5) is based on the assumption that the GRACE orbital perturbation is a
134 result of ΔTWS variation on the surface. If M is the number of model grid cells, N_{max} is the
135 maximum degree of the geopotential coefficients, and $L=(N_{max}+1)^2–4$ is the number of
136 geopotential coefficients from GRACE, the dimension of $\mathbf{Y}$, $\mathbf{H}$, and h are $L{\times}M$, $M{\times}3M$, and
137 $3M{\times}1$, respectively. Note that, Eq. (5) is defined with the global grid of h. For a regional
138 application, Eq. (5) can be modified as:

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

140 where the subscript R indicates the grid ΔTWS only in a region of interest, and o for the rest
141 of the globe. If the number of the model grid cells associated with R is J and that of the
142 outside cells is $M–J$. As such, the dimensions of $\mathbf{Y}_R$, $\mathbf{H}_R$, $\widehat{h}_R$, $\mathbf{Y}_o$, $\mathbf{H}_o$, $\widehat{h}_o$ are $L{\times}J$, $J{\times}3J$, $3J{\times}1$,
143 $L{\times}(M–J)$, $(M–J){\times}3(M–J)$, $3(M–J){\times}1$, respectively. The dimension of $\mathbf{N}$ and c remain
144 unchanged, since they are essentially from the normal equations of the original GRACE L1B
145 data (to be discussed in the following section).

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

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

149 or

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

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

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

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

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

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

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

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

3. GRACE data

3.1 GRACE least-squares normal equations

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

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

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

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

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

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

3.2 GRACE-derived ΔTWS products

Three monthly GRACE-derived ΔTWS products are also used, the ITSG-Grace2016 DDK5 solution (ITSG-DDK5 for short, http://icgem.gfz-potsdam.de/series/99_non-iso/ITSG-Grace2016), the CNES/GRGS Release 3 (RL3) (GRGS for short, Lemoine et al., 2015; http://grgs.obs-mip.fr/grace/variable-models-grace-lageos/grace-solutions-release-03) and the JPL RL05M mascon-CRI version 2 product (mascon for short, Watkins et al., 2015; Wiese et al., 2016; http://grace.jpl.nasa.gov/data/get-data/jpl_global_mascons). The ITSG-DDK5 product is the post-processed version of the ITSG L2 solution where the non-isotropic filter DDK5 (Kusche et al., 2009) is applied. The DDK5 solution is empirically selected here to be a good balance between the over-smoothed (e.g., DDK1) and noisy (e.g., DDK8) solutions. The GRGS solution provides ΔTWS at 1°×1° globally, derived from the Earth's geopotential coefficients up to the maximum degree and order 80, and no filter nor scale factor is applied (L2 data product). Mascon provides ΔTWS at equal-area 3° spherical cap grid globally. In

contrast to the ITSG-DDK5 and GRGS solutions, the mascon uses a gain factor derived from the land surface model (LSM) to restore mitigated signals and reduce leakage errors (L3 data products) (Watkins et al., 2015; Wiese et al., 2016). Additionally, mascon provides the ΔTWS uncertainty together with the solution. The uncertainty is computed based on several geophysical models (see Watkins et al. (2015) and Wiese et al. (2016) for more details). The uncertainty information is not available in the ITSG-DDK5 or GRGS product.

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

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

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

4. Hydrology model and validation data

4.1 Model setup

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

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

243 It is found that the water storage estimate is most sensitive to the replacement of precipitation
244 data, as expected, and relatively less sensitive to the change of other forcing components. We
245 use the GLDAS forcing data in this study and also further test 7 different precipitation data
246 products (see more details in Sect. 4.2). The forcing data are up/down sampled to a 0.5°×0.5°
247 spatial grid to reconcile with the CABLE spatial resolution.
248

4.2 Model uncertainty

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

263 From ensemble generations, total $K = 210$ sets of the ensemble water storage estimates (h_e)
264 are obtained:

$$\mathcal{H}_R = [h_e|_{k=1} \quad h_e|_{k=2} \quad h_e|_{k=3} \quad ... \quad h_e|_{k=K}] \qquad (15)$$

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

$$\tilde{h}_R = \frac{1}{K}\sum_{k=1}^{K} h_e|_k \qquad (16)$$

268 Note that due to the absence of GSWP3, Princeton, and MERRA data, the number of
269 ensembles reduces to $K = 180$ after December 2010, $K = 150$ after December 2012, and $K =$
270 120 after February 2016, respectively. The GC approach assumes that model errors are
271 normally distributed with zero mean. Any violation of this assumption will yield a bias in the
272 combined solutions. Therefore, the mean value is removed from each ensemble member,
273 $\mathcal{H}_R' = \mathcal{H}_R - \tilde{h}_R$, and the error covariance matrix of the model is empirically computed as:

$$C_R = \mathcal{H}_R'(\mathcal{H}_R')^T/(K-1) \qquad (17)$$

275 The $\tilde{h}_R$ (Eq. (16)) and C_R (Eq. (17)) terms can be directly used in Eq. (9). The distribution of
276 model errors is demonstrated in Fig. 2. The figure illustrates the histogram of model errors
277 ($\mathcal{H}_R'$) computed using 210 ensemble members of the model estimated ΔSM and ΔGWS in
278 Jan 2003. The histogram indicates that the model error may be approximately described by a
279 normal distribution as introduced in Eq. (1).

280 Furthermore, in practice, the sampling error caused by finite sample size might lead to
281 spurious correlations in the model covariance matrix (Hamill et al., 2001). The effect can be

282 reduced by applying an exponential decay with a particular spatial correlation length to $\mathbf{C_R}$. In
283 this study, the correlation length is determined based on the empirical covariance of model
284 estimated ΔTWS. The covariance function of ΔTWS is firstly assumed isotropic, and it is
285 computed empirically based on the method given in Tscherning and Rapp (1974). The
286 distance where the maximum value of the variance decreases to half is defined as the
287 correlation length. The obtained values vary month-to-month, and the mean value of 250 km
288 is used in this study.

289 It is emphasized that the model omission error caused by imperfect modeling of hydrological
290 process within the LSM is not taken into account in the above description. The omission error
291 may increase the model covariance and introduce a bias as well. We account for the omission
292 error by increasing 20% of the model covariance. (i.e., multiplying $\mathbf{C_R}$ by 1.2). We determine
293 such omission error based on trial-and-error such that it increases the model error (due to the
294 omission error) but not exceeds the model error value reported by Dumedah and Walker
295 (2014). We acknowledge that this is only a simple practical way of accounting for the
296 omission error into the total model error.
297

4.3 Validation data

4.3.1 Satellite soil moisture observation

300 The satellite observed surface soil moisture data is obtained from the Advanced Microwave
301 Scanning Radiometer-Earth Observing System (AMSR-E) using the Land Parameter
302 Retrieval Model (Njoku et al., 2003). The observation is used to validate our estimates of top
303 soil moisture changes (ΔSM_{top}). The AMSR-E product provides volumetric water content in
304 the top layer derived from a passive microwave data (from NASA EOS Aqua satellite) and
305 forward radiative transfer model. In this study, the level 3 product, available daily between
306 June 2002 and June 2011 at 0.25°×0.25° spatial resolution is used (Owe et al., 2008). The
307 measurements from ascending and descending overpasses are averaged for each frequency
308 band (C and X). Then, the monthly mean value is computed by averaging the daily data
309 within a month. To obtain the variation of the surface soil moisture, the long-term mean
310 between June 2002 and June 2011 is removed from the monthly data. Regarding the different
311 depth measured in CABLE and AMSR-E, the CDF-matching technique (Reichle and Koster,
312 2004) is used to reduce the bias between the top soil moisture model and the observation. The
313 CDF is built using the 2003-2004 data, and it is used for the entire period. There is no
314 satellite observed or ground measured root zone soil moisture data for meaningful
315 comparison with our results, particularly at continental scale. Validation of ΔSM_{rz} at regional
316 and continental scales is currently unachievable due to a complete lack of observations at this
317 spatial scale.
318

4.3.2 In-situ groundwater

320 The in-situ groundwater level from bore measurements are obtained from 2 different ground
321 observation networks (see Fig. 1). The data in Queensland are obtained from Department of
322 Natural Resources and Mines (DNRM) while the data in Victoria is from Department of
323 Environment and Primary Industries (DWPI). More than 10,000 measurements are available
324 from each network, but the data gap and outliers are present. Therefore, the bore

measurement is firstly filtered by removing the sites that present no data or data gap longer than 30 months during the study period.

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

5. Results

5.1 Model-only performance

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

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

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

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

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

374

375 **5.2 Impact of GRACE on storage estimates**

376 **5.2.1 Contribution of GRACE**

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

387 Figures 4 and 5 show the GC results of ΔTWS as well as ΔSM_{top}, ΔSM_{rz}, and ΔGWS in
388 different basins. The monthly time-series and the de-seasonalized time-series are shown. In
389 general, GRACE tends to increase ΔTWS when the model ΔTWS (MN) is predicted to be
390 underestimated (see e.g., LKE, MRD, NWP, SWP, TIM between 2011 and 2012) and by
391 decrease ΔTWS when determined to be overestimated (see all basins between 2008 and
392 2010). A clear example is seen over Gulf of Carpentaria (Fig. 4d), where CABLE
393 overestimates ΔTWS and produces phase delay between 2008 and 2010. The over estimated
394 amplitude and phase delay seen in CABLE ΔGWS during this above period (Fig. 4c) is
395 caused by an overestimation of soil and groundwater storage. The positively biased soil and
396 groundwater storage causes a phase delay by increasing the amount of time required for the
397 subsurface drainage (baseflow) to reduce to soil and groundwater stores. The overestimation
398 of water storage is the result of overestimated precipitation or underestimated
399 evapotranspiration. The amplitude and phase of the water storage estimate are adjusted
400 toward GRACE observation in the GC approach.

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

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

419

420 **5.2.2 Impact on long-term trend estimates**

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

430 GRACE shows significant changes in the ΔTWS trend estimates particularly over the
431 northern and western parts of the continent. The model estimates around the Gulf of
432 Carpentaria basin show a strong negative trend that is inconsistent from the GRACE data. It
433 is found that underestimated precipitation after 2012 is likely the cause of such an
434 incompatible negative trend (see Fig. 4d). Applying the GC approach clearly improves the
435 trend (Fig. 6d vs. 6e). The other example is seen over the western part of the continent (see
436 rectangular area in Fig. 6d, 6e) where the averaged long-term trend of ΔTWS was predicted
437 to be –0.4 cm/year but changed to be –1.2 cm/year (see also Sect. 5.4) by the GC approach.
438 The precipitation over the western Australia is understood to be overestimated after 2012,
439 evidently seen by that the model ΔTWS is always greater than the GC solution (see e.g., Fig.
440 4h, 5d, 5p). The GC approach reveals that the water loss over the western Australia is at least
441 twice greater than what has predicted by the CABLE model.

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

447

448 **5.2.3 Reduction of uncertainty**

449 Influenced by climate pattern, the uncertainty of water storage estimates significantly varies
450 across Australia. The uncertainty of the model estimate is computed from the variability

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

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

465 It is worth mentioning that the model uncertainty is mainly influenced by the meteorological
466 forcing data. The uncertainty of precipitation derived from seven different precipitation
467 products is shown in Fig. 7e. The spatial pattern of the precipitation uncertainty is correlated
468 with the uncertainty of water storage estimates. The larger water storage uncertainty is
469 deduced from the larger precipitation uncertainty. The quality of precipitation forcing data is
470 found to be an important factor to determine the accuracy of water storage computation.
471

472 **5.3 Comparison with independent data**

473 **5.3.1 Soil moisture**

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

482 The GC approach leads to a small bit of improvement of the top soil estimate consistently
483 from C- and X-band measurements and from all basins. No degradation of the NS value is
484 observed in the GC solutions. The largest improvement is seen over LKE and NEC, where
485 NS increases by 10 – 15%. For other regions, the change in the NS coefficient may be
486 incremental.
487

488 **5.3.2 Groundwater**

489 The ΔGWS estimates from the model and the GC method are compared with the in situ data
490 obtained from 2 different ground networks in Queensland and Vitoria. For each network, all
491 ΔGWS data inside the groundwater network boundary (see polygons in Fig. 1) are used to
492 compute the average ΔGWS time series. From the comparison given in Fig. 8, it is found that

the GC solutions of ΔGWS follows the overall inter-annual pattern of CABLE but with a considerably larger amplitude. This results in a better agreement with the in situ ΔGWS data seen from both networks. The NS coefficient of ΔGWS between the estimates and the in situ data are given in Table 4. The CABLE ΔGWS performs significantly better in Queensland (NS = ~0.5) than Victoria (NS = ~0.3). Significant improvement is found from the GC solutions in both networks, where the NS value increases from 0.5 to 0.6 (~ 22 %) in Queensland and from 0.3 to 0.6 (~85 %) in Victoria. Even greater improvement is seen when the inter-annual patterns are compared. The NS value increase from 0.5 to 0.7 (~ 32 %), and 0.4 to 0.8 (~93 %) in Queensland and Victoria, respectively.

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

5.4 Assessment of mass variation in the past 13 years

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

During Big Dry (2003 – 2009), a significant loss of total storage (40 – 60 Gton over 7 years) is observed over LKE, MRD, NWP, and SWP basins. The largest groundwater loss of >20 Gton is found from LKE and MRD. No significant change is observed over the tropical climate regions (e.g., GOC, NEC). The mass loss mostly occurs in the root zone and

537 groundwater compartments where the sum of ΔSM_{rz} and ΔGWS explains more than 90% of
538 the ΔTWS value. The mass loss is also observed in ΔSM_{top} but >10 times smaller than
539 ΔSM_{rz} and ΔGWS.

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

546 The opposite behaviour is observed over the basins (such as NEC and GOC) that experienced
547 mass gain during Big Dry. The water storage gain is greater in ΔGWS compared to ΔSM_{rz}. In
548 NEC, ΔGWS gain is ~8 times larger than ΔSM_{rz} during Big Wet. The soil compartment may
549 be saturated during Big Dry and additional infiltration from the Big Wet precipitation leads to
550 an increased groundwater recharge. The ΔSM_{rz} loss observed over GOC is simply caused by
551 the timing selection of Big Wet period, which ends earlier (~2011) in GOC than in other
552 basins. The ΔSM_{rz} gain becomes ~26 Gton if the Big Wet period is defined as 2008 – 2011.
553 During the post-Big Wet period (2012 and afterwards), the decreasing trend of water storage
554 is observed from all basins (see Fig. 4, 5). This is mainly caused by the decrease in
555 precipitation after 2012 and by gradual water loss through evapotranspiration (Fasullo et al.,
556 2013).

557 The overall water storage change in the last 13 years demonstrates that the severe water loss
558 from most basins during Big Dry (the millennium drought) is balanced with the gain during
559 Big Wet (the La Niña). The negative ΔTWS estimated during Big Dry becomes positive in
560 LKE, MRD, and SEC and less negative in TIM, and the greatest gain is observed from NEC
561 by ~50 Gton during 13 year-period (see Fig. 9c). However, the water mass loss is still
562 detected over the western basins (e.g., IND, NWP, SWP, SWC), and their magnitudes are
563 even larger than the mass loss during Big Dry. For example, the greatest ΔTWS loss of ~79
564 Gton is observed over NWP, which is ~25 Gton greater than the loss during Big Dry (see Fig.
565 9a and 9c). The basin is less affected by the La Niña, and the rainfall during Big Wet is
566 clearly inadequate to support the water storage recovery in the basin. Rainfall deficiency also
567 reduces the groundwater recharge, resulting in even more decreasing of ΔGWS, compared to
568 the Millennium Drought period (see Fig. 9j and 9l). The continual decrease in water storage
569 over western basins is likely caused by the interaction of complex climate patterns like El
570 Niño Southern Oscillation, Indian Ocean Dipole, and Southern Annular Mode cycles
571 (Australian Bureau of Meteorology, 2012; Xie et al., 2016).
572

573 **5.5 Comparison of GC approach with alternatives**

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

$$\widehat{h}_R = \widetilde{h}_R + \mathbf{H}\mathbf{C}_R^T(\mathbf{H}\mathbf{R}\mathbf{H}^T + \mathbf{C}_R)^{-1}\big(b - \mathbf{H}\widetilde{h}_R\big) \qquad (18)$$

580

581 where $\widetilde{h}_R$, $\mathbf{H}$, $\mathbf{C}_R$ are described in Sect. 2, b is an observation vector containing GRACE-
582 derived ΔTWS, and $\mathbf{R}$ is an error variance-covariance matrix of the observation. The
583 GRACE-derived ΔTWS and its error information is obtained from the mascon solution. The
584 matrix $\mathbf{R}$ is a (diagonal) error variance matrix since no covariance information is given in the
585 mascon product. Note that the model uncertainty remains the same as in GC approach (Sect.
586 4.2). The different results from App1 and App2 are mainly attributed to the different estimates
587 of the uncertainty.

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

601

602 **6. Conclusion**

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

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

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

15

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

626 However, it is difficult to constrain different water storage components by only using total
627 storage observation like GRACE. In addition, it is challenging to improve surface soil
628 moisture varying rapidly in time, using a monthly mean GRACE observation. Tian et al.
629 (2017) utilized the satellite soil moisture observation from the Soil Moisture and Ocean
630 Salinity (SMOS, Kerr et al., 2001) in addition to GRACE data for their data assimilation and
631 showed a clear improvement in the top soil moisture estimate. The GC approach with
632 complementary observations at higher temporal resolution should be considered particularly
633 to enhance the surface soil moisture computation.

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

638

646

Appendix A: Least-squares normal equation of GRACE

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

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

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

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

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

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

It can be simplified as:

$$\mathbf{H}^T\mathbf{Y}^T\mathbf{N}\ \mathbf{Y}\mathbf{H}\ \widehat{\boldsymbol{h}} = \mathbf{H}^T\mathbf{Y}^T\boldsymbol{c}, \tag{A4}$$

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

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

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

$$NS = 1 - \frac{\sum_{i=1}^{N}(\boldsymbol{y_i} - \widehat{\boldsymbol{x_i}})^2}{\sum_{i=1}^{N}(\boldsymbol{y_i} - \overline{\boldsymbol{y}})^2} \tag{B1}$$

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:

$$\bar{Z} = \frac{\sum_{j=1}^{M} w_j \bar{z}_j}{\sum_{j=1}^{M} w_j} \tag{B2}$$

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

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

769 Owe, M.,de Jeu, R., Holmes, T.: Multisensor historical climatology of satellite-derived global
770 land surface moisture, J. Geophys. Res., 113, F01002, 17 pp., doi:10.1029/2007JF000769,
771 2008.

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

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

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

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

783 Rienecker, M. M., Suarez, M. J., Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich MG,
784 Schubert SD, Takacs L, Kim G-K, Bloom S, Chen J, Collins D, Conaty A, da Silva A, Gu W,
785 Joiner J, Koster RD, Lucchesi R, Molod A, Owens T, Pawson S, Pegion P, Redder CR,
786 Reichle R, Robertson FR, Ruddick AG, Sienkiewicz M, Woollen J. 2011. MERRA—NASA's
787 Modern-Era Retrospective Analysis for Research and Applications. J. Climate, DOI:
788 10.1175/JCLI-D-11-00015.1.

789 Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault,
790 K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D.,
791 and Toll, D.: The global land data assimilation system, Bull. Amer. Meteor. Soc., 85(3), 381–
792 394, 2004.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

838　Ukkola, A. M., Pitman, A. J., Decker, M., De Kauwe, M. G., Abramowitz, G., Kala, J., and
839　Wang, Y.-P.: Modelling evapotranspiration during precipitation deficits: identifying critical
840　processes in a land surface model. Hydrol. Earth Syst. Sci., 20, 2403–2419, doi:10.5194/hess-
841　20-2403-2016, 2016.

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

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

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

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

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

859　Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., and Landerer, F. W.: Improved
860　methods for observing Earth's time variable mass distribution with GRACE using spherical
861　cap mascons, J. Geophys. Res. Solid Earth, 120, 2648–2671, doi:10.1002/2014JB011547,
862　2015.

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

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

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

871　Wu, S. C., Kruizinga, G., and Bertiger, W.: Algorithm theoretical basis document for
872　GRACE Level-1B data processing V1.2, JPL D-27672, Jet Propul. Lab., Pasadena, Calif,
873　2006.

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

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

881

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

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

898　**Table 3**. NS coefficients between top soil moisture estimates and the satellite soil moisture
899　observations from AMSR-E products over 10 different Australian basins. The area-weighted
900　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

901

902

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

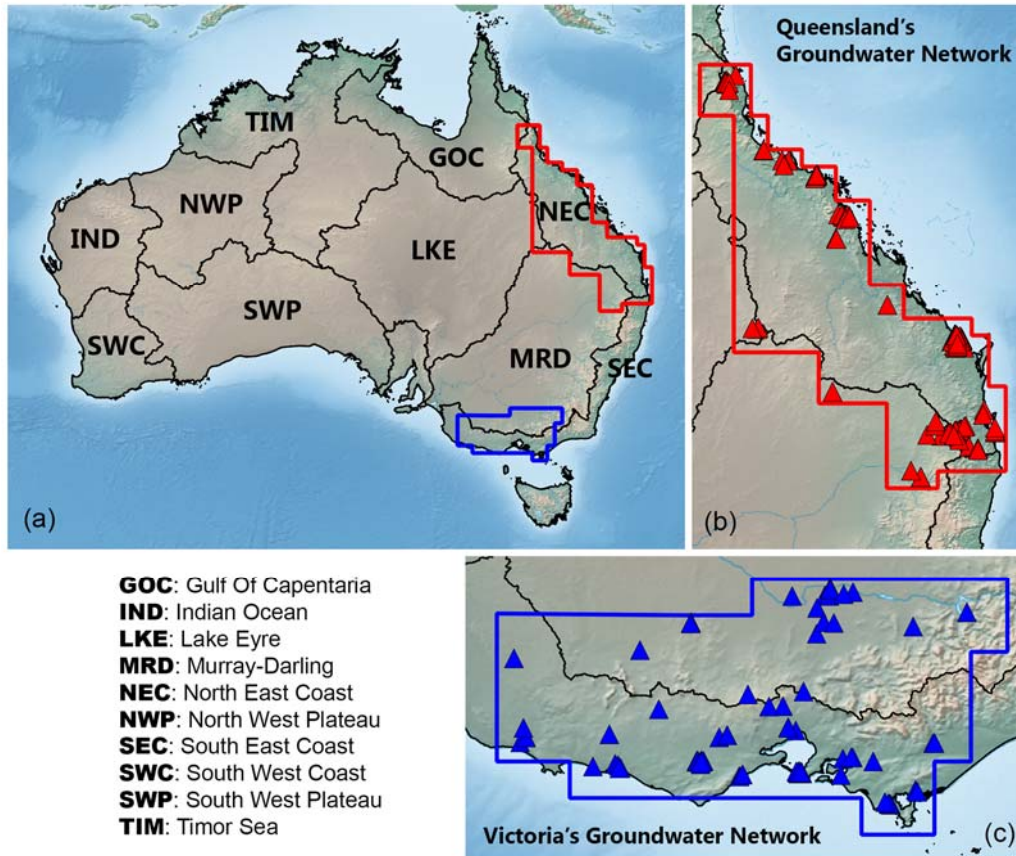|	Queensland			Victoria		
	In-situ	CABLE	GC	In-situ	CABLE	GC
Original time-series						
NS [-]	-	0.49	0.60	-	0.34	0.63
Trend [cm/year]	1.60 ± 0.05	0.39 ± 0.02	0.63 ± 0.05	-0.27 ± 0.05	0.10 ± 002	-0.18 ± 0.03
De-seasonalized time-series						
NS [-]	-	0.50	0.66	-	0.43	0.83
Trend [cm/year]	1.60 ± 0.05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

906

Figure 1. (a) Geographical location of 10 Australian river basins. Red and blue polygons indicate the boundaries of groundwater networks in Queensland (b) and Victoria (c), respectively. Triangles (in b and c) represent the selected bore locations used in this study.
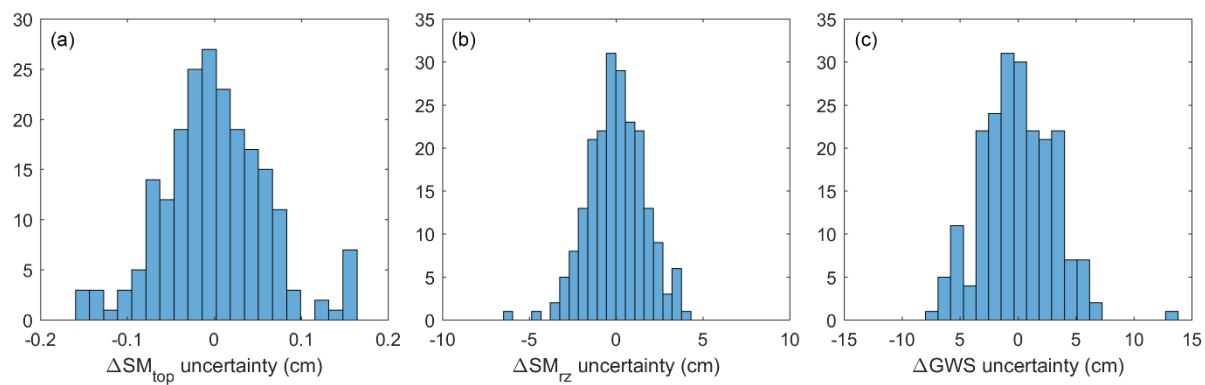
Figure 2: Histograms of the model errors computed from 210 ensemble members ($\mathcal{H}_R{}'$) without the mean. The basin averaged values (from all 10 Australian basins) of January 2003, for example, are shown.
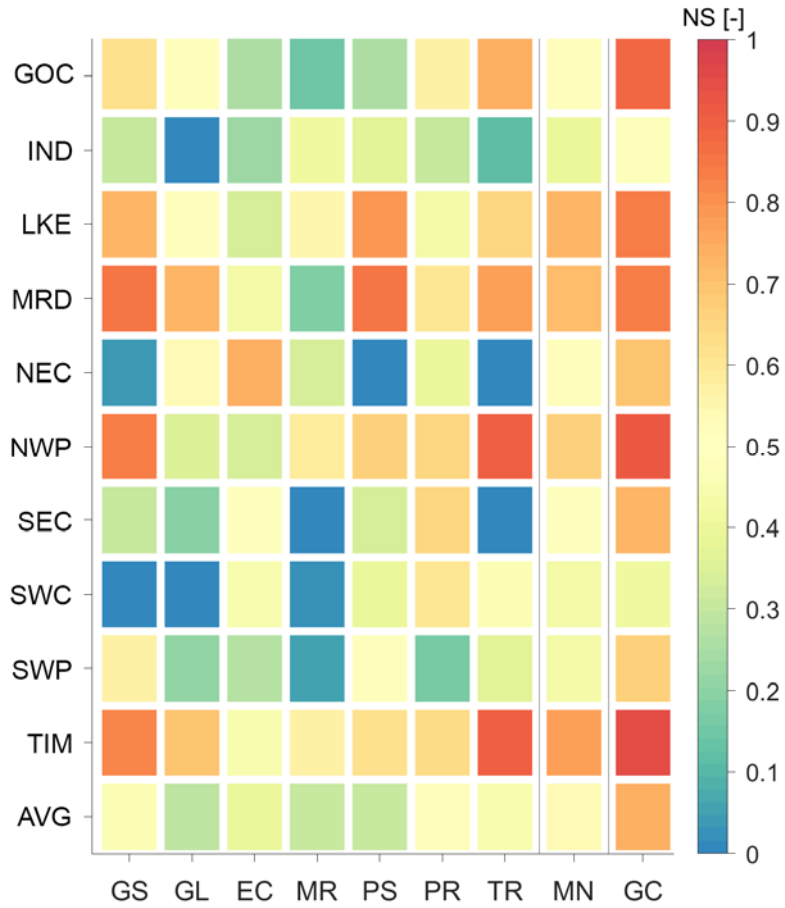
Figure 3. NS coefficients between the model and GRACE-mascon ΔTWS over 10 Australian basins (in ordinate). The NS values were computed based on CABLE ΔTWS computed with 7 different precipitation data (in abscissa), GSWP3 (GS), GLDAS (GL), ECMWF (EC), MERRA (MR), PERSIANN (PR), TRMM (TR). The NS value of the mean ΔTWS estimates (the average of 7 variants) is also shown (MN). The area-weighted average NS value over all basins is also shown (AVG). The NS value of ΔTWS from the GRACE-combined (GC) approach is shown in the last column. The full name of the basins can be found in Fig. 1.
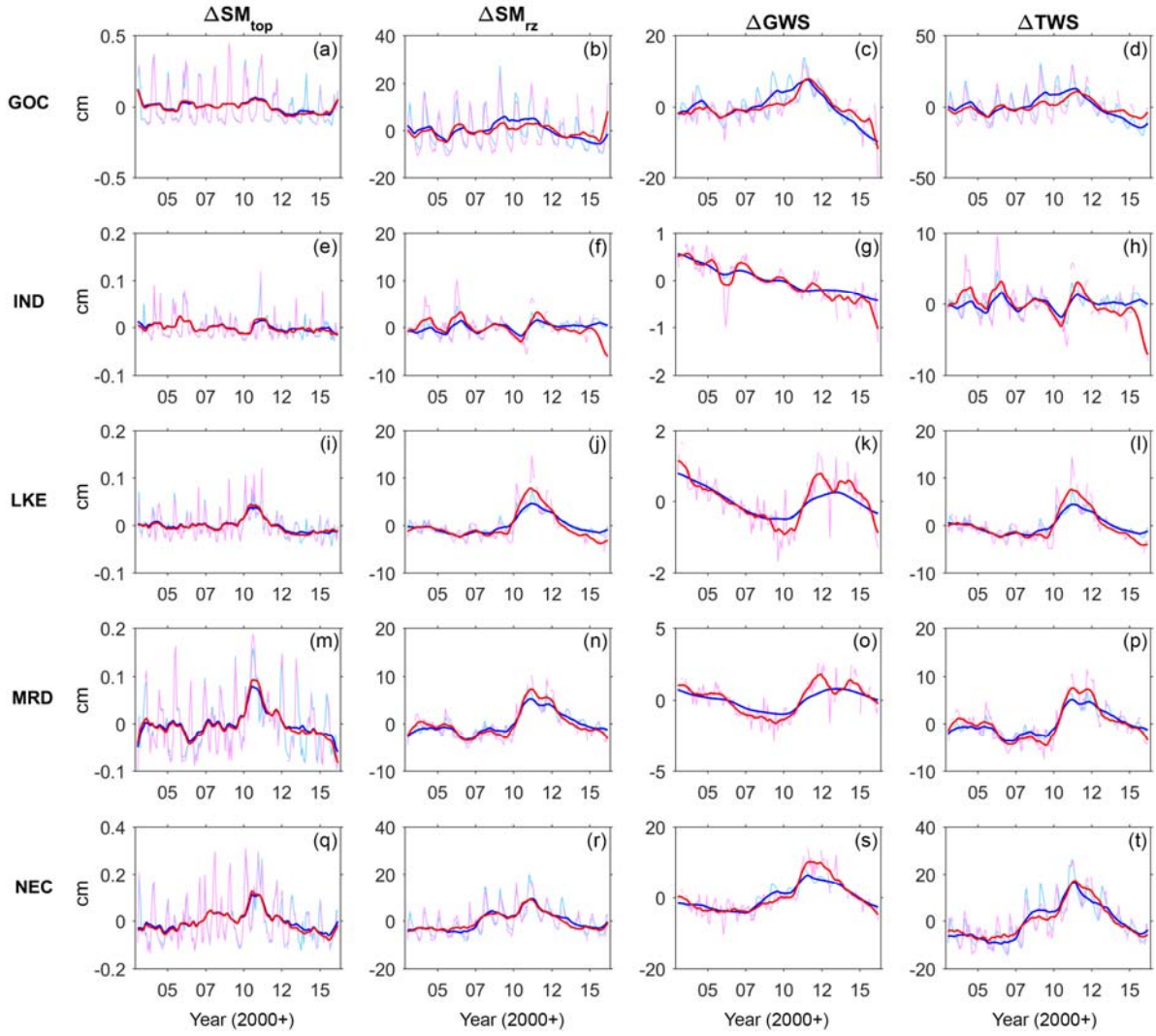
Figure 4. The monthly time series of ΔSM_{top}, ΔSM_{rz}, ΔGWS, and ΔTWS estimated from model (blue) and GC (red) solutions over Gulf of Carpentaria (GOC), Indian Ocean (IND), Lake Eyre (LKE), Murray-Darling (MRD), and North East Coast (NEC). The de-seasonalized time series is also shown.
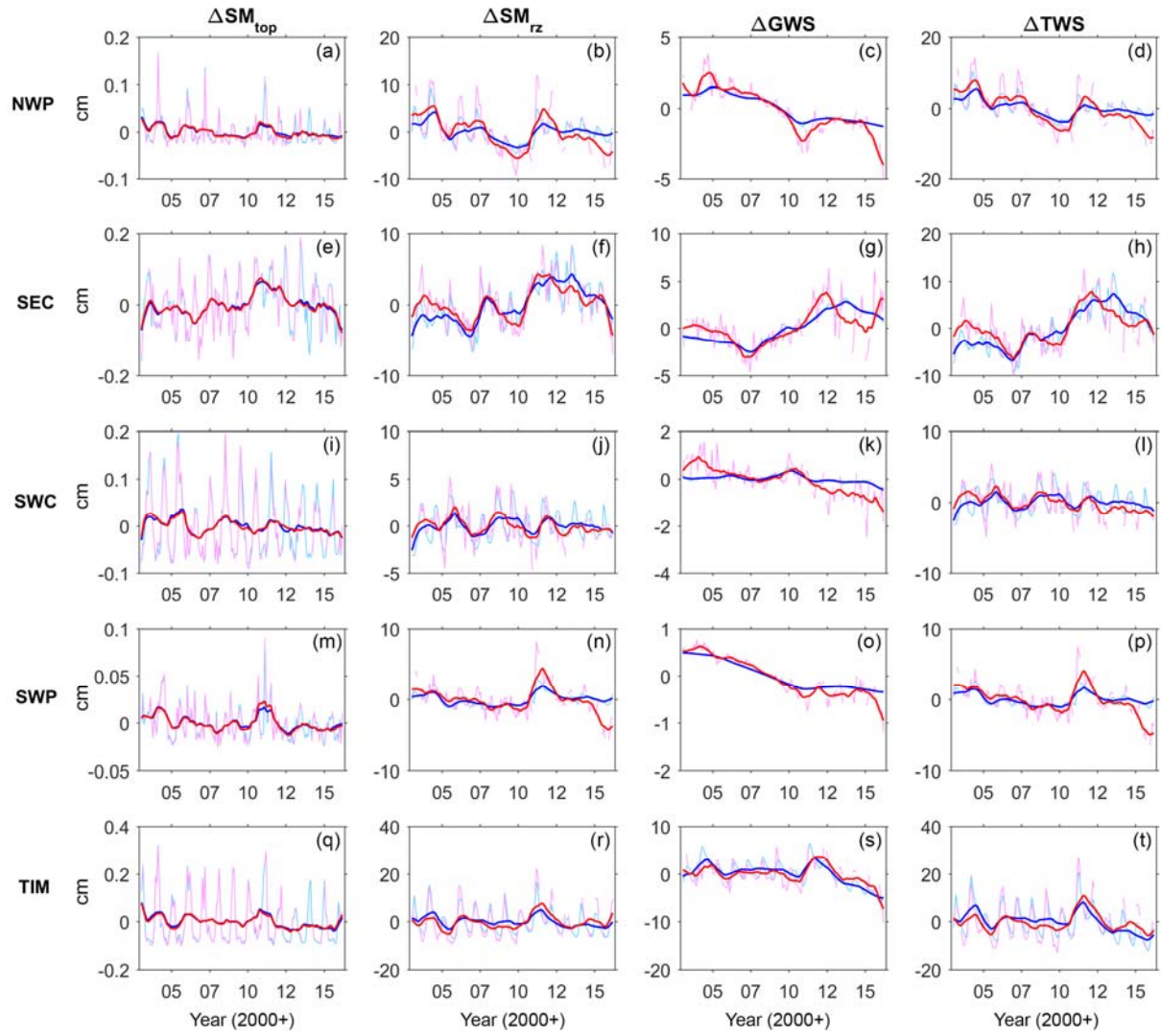
932

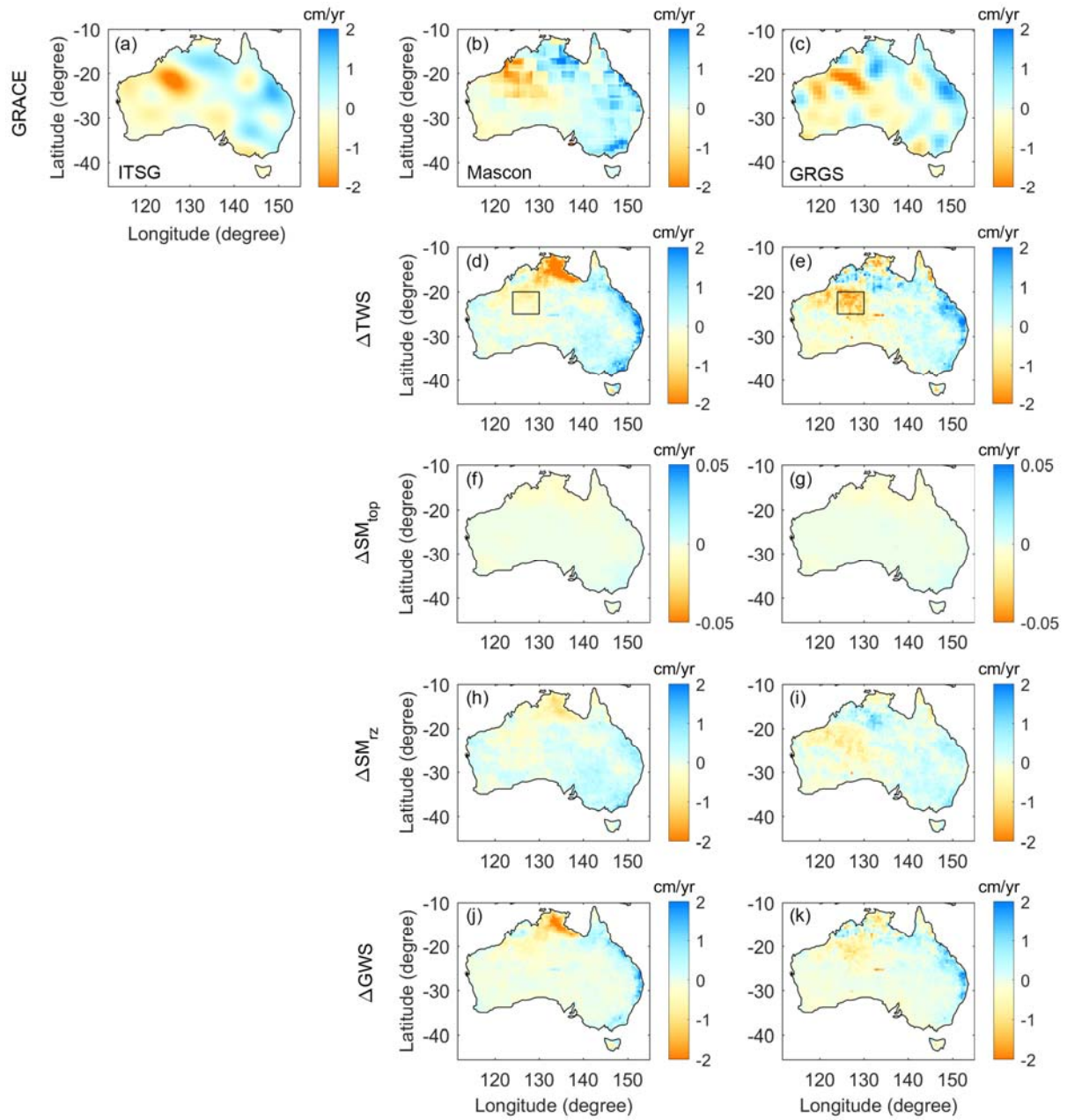**Figure 5.** Similar to Fig. 3, but estimated over North West Plateau (NWP), South East Coast (SEC), South West Coast (SWC), South West Plateau (SWP), and Timor Sea (TIM).

933
934
935

Figure 6. Long-term trends of ΔTWS (d, e), ΔSM_{top} (f, g), ΔSM_{rz} (h, i), and ΔGWS (j, k) estimated from the model-only (second column) and the GC solutions (third column). Results of GRACE ΔTWS independently from ITSG-DDK5 (a), mascon (b), and GRGS solution (c) are also shown. The eastern part of North West Plateau basin is shown as a rectangle polygon in (d) and (e).
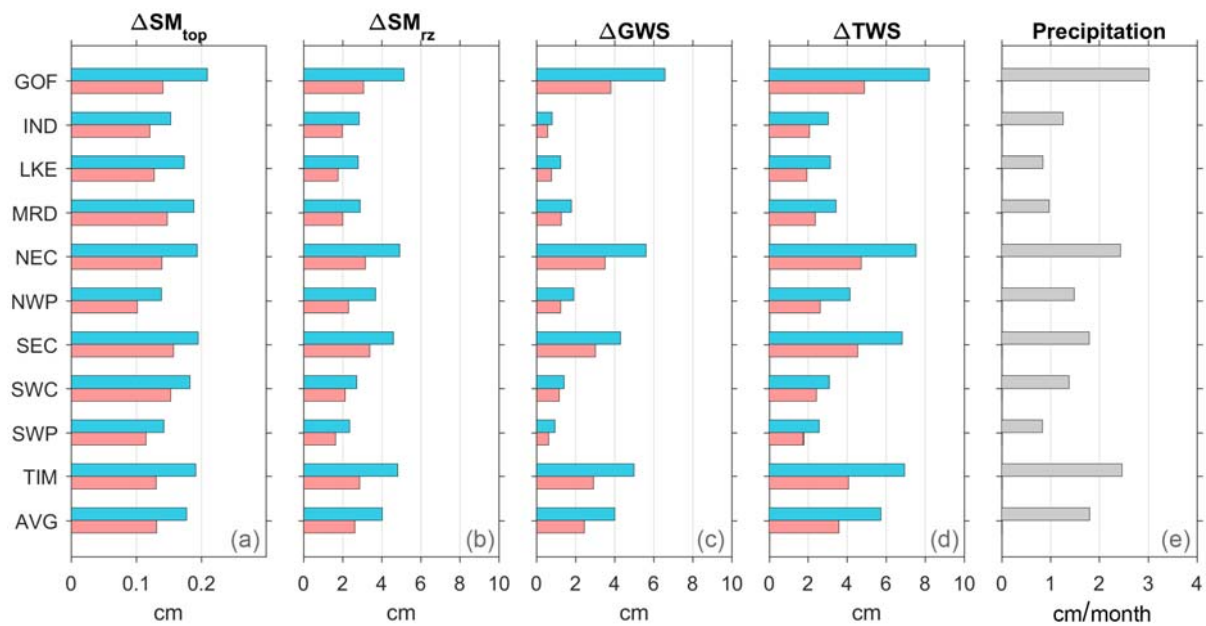
Figure 7. Uncertainties of ΔSM_{top}, ΔSM_{rz}, ΔGWS, and ΔTWS estimated from the model (blue) and the GC solutions (red) in 10 different Australian basins. The uncertainty of the precipitation is shown in (e). The area-weighted average value (AVG) is also shown.
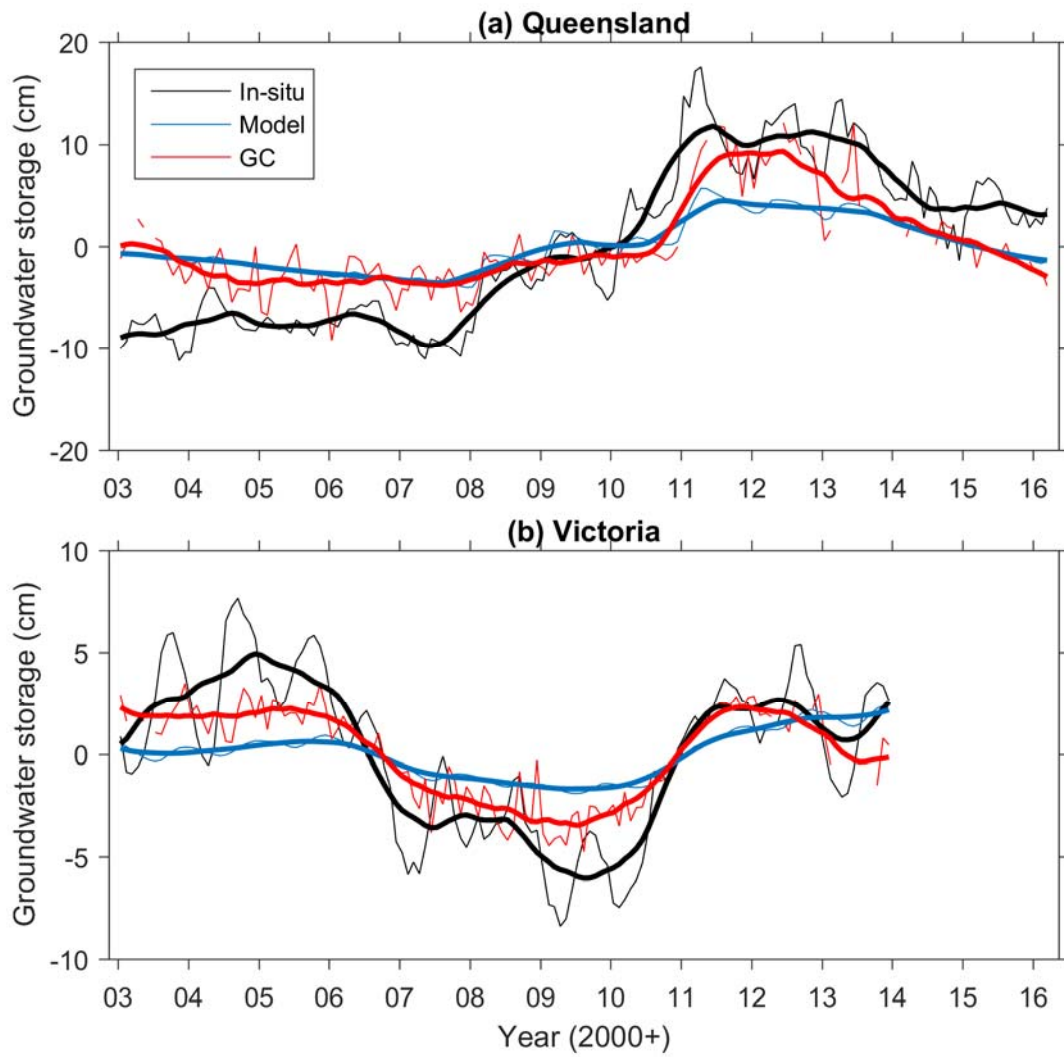
Figure 8. The monthly time series of ΔGWS estimated from the model, GC solutions, and measured from the in situ groundwater network in Queensland (a) and Victoria (b). Deseasonalized time series are shown in thick lines.
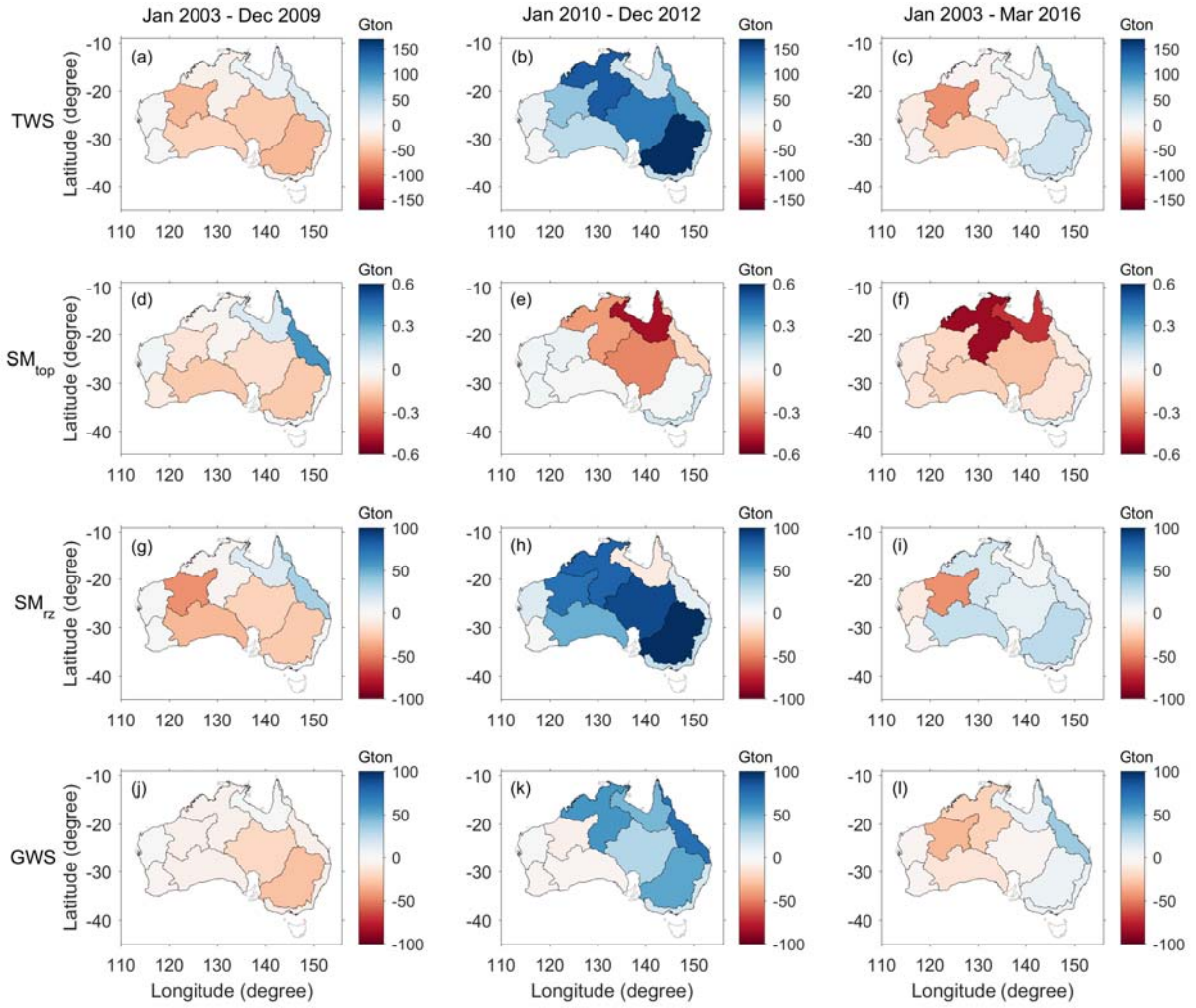
953

Figure 9. Mass changes (Gton, Giga tonne) of ΔTWS, ΔSM_{top}, ΔSM_{rz}, and ΔGWS estimated from GC solutions over 10 Australian basins in 3 different periods, Big Dry (January 2003 – December 2009), Big Wet (January 2010 – December 2012), and entire period (January 2003 – March 2016).
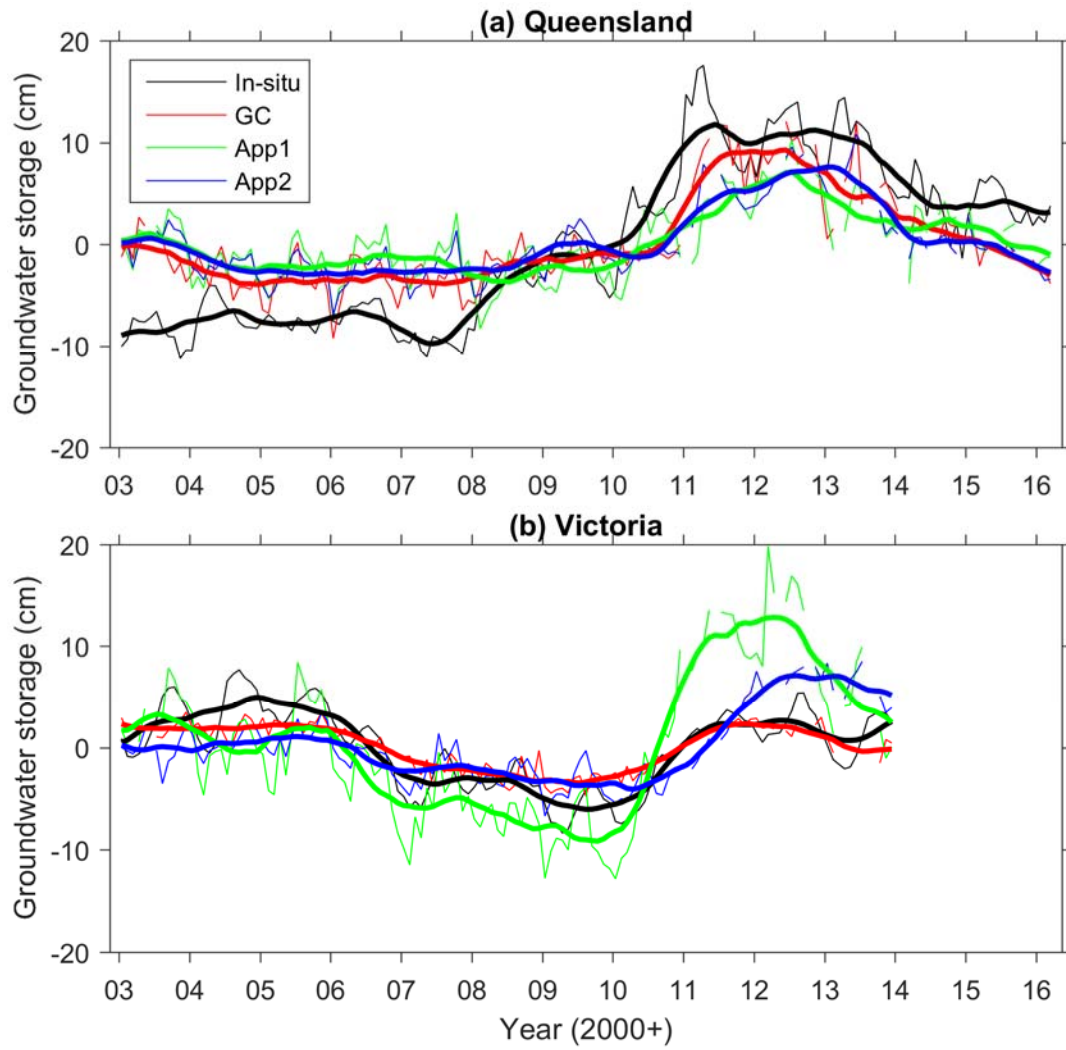
958

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