WATER RESOURCES RESEARCH, VOL. 47, W11524, doi:10.1029/2011WR010714, 2011

Use of Gravity Recovery and Climate Experiment terrestrial water storage retrievals to evaluate model estimates by the Australian water resources assessment system

A. I. J. M. van Dijk,[1] L. J. Renzullo,[1] and M. Rodell[2]

Received 29 March 2011; revised 11 October 2011; accepted 12 October 2011; published 24 November 2011.

[1] Terrestrial water storage (TWS) estimates retrieved from the Gravity Recovery and Climate Experiment (GRACE) satellite mission were compared to TWS modeled by the Australian Water Resources Assessment (AWRA) system. The aim was to test whether differences could be attributed and used to identify model deficiencies. Data for 2003–2010 were decomposed into the seasonal cycle, linear trends and the remaining de-trended anomalies before comparing. AWRA tended to have smaller seasonal amplitude than GRACE. GRACE showed a strong (>15 mm yr^{-1}) drying trend in northwest Australia that was associated with a preceding period of unusually wet conditions, whereas weaker drying trends in the southern Murray Basin and southwest Western Australia were associated with relatively dry conditions. AWRA estimated trends were less negative for these regions, while a more positive trend was estimated for areas affected by cyclone Charlotte in 2009. For 2003–2009, a decrease of 7–8 mm yr^{-1} (50–60 km^3 yr^{-1}) was estimated from GRACE, enough to explain 6%–7% of the contemporary rate of global sea level rise. This trend was not reproduced by the model. Agreement between model and data suggested that the GRACE retrieval error estimates are biased high. A scaling coefficient applied to GRACE TWS to reduce the effect of signal leakage appeared to degrade quantitative agreement for some regions. Model aspects identified for improvement included a need for better estimation of rainfall in northwest Australia, and more sophisticated treatment of diffuse groundwater discharge processes and surface-groundwater connectivity for some regions.

Citation: 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.

1. Introduction

[2] The Gravity Recovery and Climate Experiment satellite mission (GRACE) [*Tapley et al.*, 2004] provides integrated estimates of variations in total terrestrial water storage (TWS) based on precise observations of Earth's time-variable gravity field. Although the effective "footprint" is large (>300 km) for hydrological applications, GRACE is unique in its ability to monitor integrated changes in surface, soil, and groundwater [e.g., *Leblanc et al.*, 2009; *Rodell et al.*, 2009; *Swenson et al.*, 2008]. Studies to date have tended to average GRACE observations over large basins to minimize error from the large footprint. Alternatively, one can spatially compare GRACE and model TWS over large areas. There is considerable uncertainty in the retrieval of water storage estimates from GRACE, and its errors are not spatially and temporally uniform. However, when used with caution GRACE data have proven valuable for evaluation of large-scale hydrological models [see reviews by *Güntner*, 2008; *Ramillien et al.*, 2008; *Syed et al.*, 2008]. Provided the model produces "GRACE-like" TWS estimates, such a comparison should help to identify any model deficiencies and suggest ways to improve the model structure or parameterization, although *Güntner* [2008] concluded that few studies have actually done so. If there is sufficient agreement and errors in GRACE TWS can be estimated, the data can also be used in data assimilation [*Zaitchik et al.*, 2008].

[3] In this paper, we aim to test whether a statistical comparison of GRACE and model TWS estimates provides sufficient information to spatially assess the performance of a continental water balance model for Australia. In particular, we wanted to test whether any divergence between GRACE and model estimates could be attributed with confidence and used to identify concrete opportunities for model improvement. We did this by spatially comparing GRACE TWS with model estimates, for the period 2003–2010. To evaluate different aspects of model behavior, we decomposed the TWS data into their seasonal cycle, 8 year linear trends, and de-trended anomalies.

2. Data

2.1. Modeled Terrestrial Water Storage

[4] The Australian Water Resources Assessment (AWRA) system [*Van Dijk and Renzullo*, 2011; *Van Dijk*,

[1]Black Mountain Laboratories, CSIRO Land and Water, Canberra, ACT, Australia.
[2]Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

0043-1397/11/2011WR010714

2010a] is a water balance monitoring system used by the Australian Bureau of Meteorology to support the production of water accounts and water resource assessments. The system combines a comprehensive spatial hydrological model with meteorological forcing data and remotely sensed land surface properties to produce estimates of water stored in the soil, surface water and groundwater. The AWRA system includes a grid-based spatial landscape water balance model, AWRA-L (version 0.5). Conceptual aspects of AWRA-L relevant here include: (1) shallow and deep soil layers are assumed to be explored by all vegetation and deep-rooted vegetation only, respectively; (2) a linear reservoir groundwater model has a drainage characteristic estimated from analysis of streamflow from several hundred small upland catchments [*Van Dijk*, 2009]; and (3) the only surface water storage considered is the stream network, which drains rapidly in response to reduced inflows. Meteorological forcing is derived by interpolation of station data on a regular 0.05° (~5 km) grid; model outputs have the same resolution. Full details on AWRA-L (version 0.5) can be found in the model technical document [*Van Dijk*, 2010a]. The AWRA-L model parameterization used in this analysis is the same as that evaluated by *Van Dijk and Warren* [2010]. AWRA-L ignores diffuse lateral water transport between grid cells. Additional AWRA system components describing deep groundwater systems and the lateral redistribution and subsequent evapotranspiration of surface water are being developed [*Van Dijk and Renzullo*, 2011] but were not yet implemented at the time of writing. The potential role of these processes, therefore, needs to be considered when interpreting the data.

[5] AWRA water balance estimates have received fairly extensive evaluation for Australian conditions, using streamflow and deep drainage observations from several hundred catchments and sites, respectively; evapotranspiration measurements at seven flux tower sites; radar and microwave remote sensing estimates of surface soil moisture content; and vegetation canopy cover and density estimated from optical remote sensing [*Liu et al.*, 2010; *Van Dijk and Warren*, 2010; *King et al.*, 2011; *Viney et al.*, 2011; L. Peeters, personal communication, 2011]. AWRA simulated TWS were also compared against GRACE-derived TWS aggregated for a few large Australian basins [*Van Dijk and Renzullo*, 2009]. Although only a preliminary analysis, this showed generally good agreement between the dynamic range and temporal patterns.

2.2. GRACE Estimated Terrestrial Water Storage

[6] Several TWS products derived from GRACE observations currently exist. For this study, we used the 1° resolution gridded estimates based on gravity solutions from the University of Texas Centre for Space Research (CSR), which were downloaded from the GRACE Tellus website (available at http://grace.jpl.nasa.gov/data/; data source version "ssv201008") [*Swenson and Wahr*, 2006]. Data for January 2003 through to December 2010 were downloaded; 1 month was missing. Due to the orbit of the satellite overpasses, GRACE TWS estimates do not represent all days of the month (see GRACE Tellus website for details). The raw solutions have been de-striped and smoothed; atmospheric mass variations estimated by weather models and mass effects of postglacial rebound (PGR) removed (see website for details). In addition, to correct for the smoothing due to filtering, it is recommended that the retrieved anomalies be multiplied with a spatially varying scaling coefficient also available from the GRACE Tellus Web site. This scaling grid was derived by applying the GRACE filtering process to output from the Community Land Surface model and, on a pixel basis, determining the multiplier that minimizes the root-mean-square difference (RMSD) between the filtered and scaled data and the original data, respectively. Because of the various ancillary measured and modeled data used in the retrieval the term "observation" is perhaps somewhat liberally applied, but we will use it here for textual clarity. The error in the GRACE TWS product is spatially correlated and will vary by location and month. A spatial grid of the estimated average error due to measurement and leakage (including scaling) are also available from the GRACE Tellus Web site. From these, prior estimates of total error were calculated. A mean monthly error of 62 mm can be calculated for Australia, of which 79% is contributed by leakage and concentrated around the coastlines. The PGR signal over Australia is comparatively small (<0.5 mm yr^{-1}) [*Paulson et al.*, 2007; *Peltier*, 2004] and therefore can be ignored as a source of uncertainty in interpretation.

3. Methods

3.1. Processing of Model Estimates

[7] To generate model-based TWS estimates that could be directly compared with the GRACE data, for each grid cell the following computations were made: (1) TWS was estimated by summing water in all stores (in mm equivalent height of water), i.e., vegetation biomass water, soil layers, groundwater store and surface water store; (2) the 0.05° TWS estimates were averaged to the same geographic area and 1° resolution as the GRACE product; and (3) average TWS was calculated for each month.

3.2. Statistical Comparison

[8] The statistical comparison was performed by grid cell. The GRACE Tellus product has 1° grid cells but the actual signal originates from a larger area; therefore, interpretation for individual cells needs to be made with caution. Because GRACE TWS values are anomalies with respect to a reference period, a difference in mean temporal values has no useful meaning. Therefore, the bias between model and observations was removed, by subtracting the mean for those months the two data sets had in common. To help interpretation, the resulting data were decomposed into the seasonal cycle, the 8 year (2003–2010) linear trend, and the remaining de-trended anomalies. The seasonal cycle was calculated as the average anomaly pattern for all years. The linear trend was calculated by removing the seasonal cycle and then fitting a straight line to the de-trended values. Finally, the de-trended anomalies were calculated by subtracting both the seasonal cycle and linear trend from the original TWS anomaly time series. For each of the three components, the root-mean-square difference (RMSD) between AWRA and GRACE TWS was calculated, as well as the linear correlation coefficient (r) and, to allow comparison of the dynamic range, the amplitude of the seasonal cycle and the mean standard (i.e., root-mean-square) de-trended anomaly. Finally, the relative importance of the three components of

variability was calculated by dividing the sum of squared differences (SSD) for each component by the combined sum of SSD for all three components. To investigate the extent to which the relatively recently implemented GRACE TWS scaling coefficient affected the comparison, we also repeated the entire analysis without applying the scaling coefficient.

[9] Average net water balance estimates produced by *Guerschman et al.* [2009] were used in interpretation (Figure 1). These estimates were based on a combination of MODIS greenness and surface wetness indices and interpolated station-level climate data. Notwithstanding inevitable estimation errors, the resulting map clearly shows regions where rainfall exceeds evapotranspiration (blue colors in Figure 1) and areas where evapotranspiration exceeds rainfall, such as irrigation areas, floodplains and (salt) lakes (red colors). Region and state names are also shown for reference.

4. Results

4.1. Seasonal Cycle

[10] Maps comparing the seasonal cycle in observed and modeled TWS are shown in Figure 2 and the seasonal cycle for indicated grid cells are shown in Figure 3; selected statistics of the agreement are listed in Table 1. For most of the continent there is reasonable agreement in seasonal patterns, with median RMSD values of 16 mm and a median correlation coefficient (r) of 0.71 (Figures 2a and 2b, respectively). Some of the best agreement was found in northern Australia (e.g., area A; Figure 2 and Figure 3a). Large absolute disagreement between seasonal cycles (RMSD $>$ 50 mm) occurs along the Gulf of Carpentaria (area B, Figure 1). *Tregoning et al.* [2008] concluded that this is due to the influence of the large barotropic seasonal sea level variations of around 400 mm that occur in the nearby Gulf of Carpentaria, and which are not adequately accounted for in the GRACE retrievals. Differences of up to 50 mm occur in the wetter regions along the east Coast and southwest WA, whereas differences were $<$30 mm for 57% of the continent. For 79% the amplitude in the seasonal cycle (ASC) in modeled TWS was smaller than in the observed values. Indicators of the influence of the scaling coefficient are listed in Table 1. The scaling is temporally uniform and so does not affect the correlation coefficient. Not applying the scaling coefficient did reduce the difference between the median ASC of GRACE and AWRA and slightly reduced the RMSD and the number of cells where the ASC for GRACE was greater than for AWRA, or where the two differed by at least 30 mm.

4.2. Linear Trends

[11] Maps comparing the linear 8 year trend in observed and modeled TWS are shown in Figure 4 and time series for indicated grid cells are shown in Figure 5. The model did not reproduce negative TWS trends of $>$15 mm yr^{-1} in the Great Sandy Desert (Figure 5a; area A in Figure 4) and the southern Murray Basin (Figure 5d, area D), particularly for 2008–2009. The modeled positive trend along the central and north eastern coast was stronger than observed (Figure 5c, area C). In the Top End modeled trends varied regionally and more strongly than those observed (e.g., Figure 5b, area B). For the remaining areas neither GRACE nor AWRA suggested any strong trend (Figure 4c) and trends agreed within 10 mm yr^{-1} for 67% of the continent.

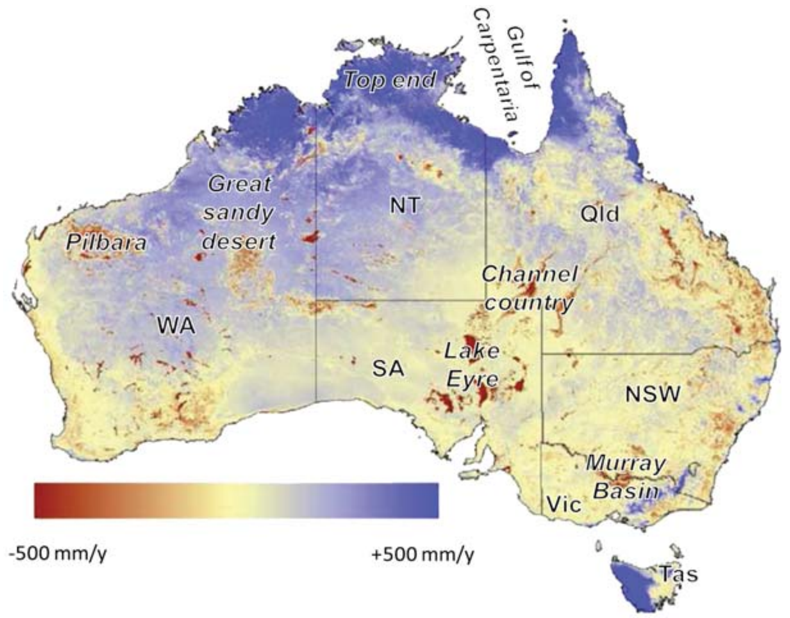

Figure 1. Map of the mean net water balance (2000–2006) estimated as the difference between precipitation and actual evapotranspiration [*Guerschman et al.*, 2009]. Lines represent state boundaries (WA, Western Australia; NT, Northern Territory; SA, South Australia; Qld, Queensland; NSW, New South Wales; Vic, Victoria; Tas, Tasmania). Also indicated are specific regions referred to in the text.

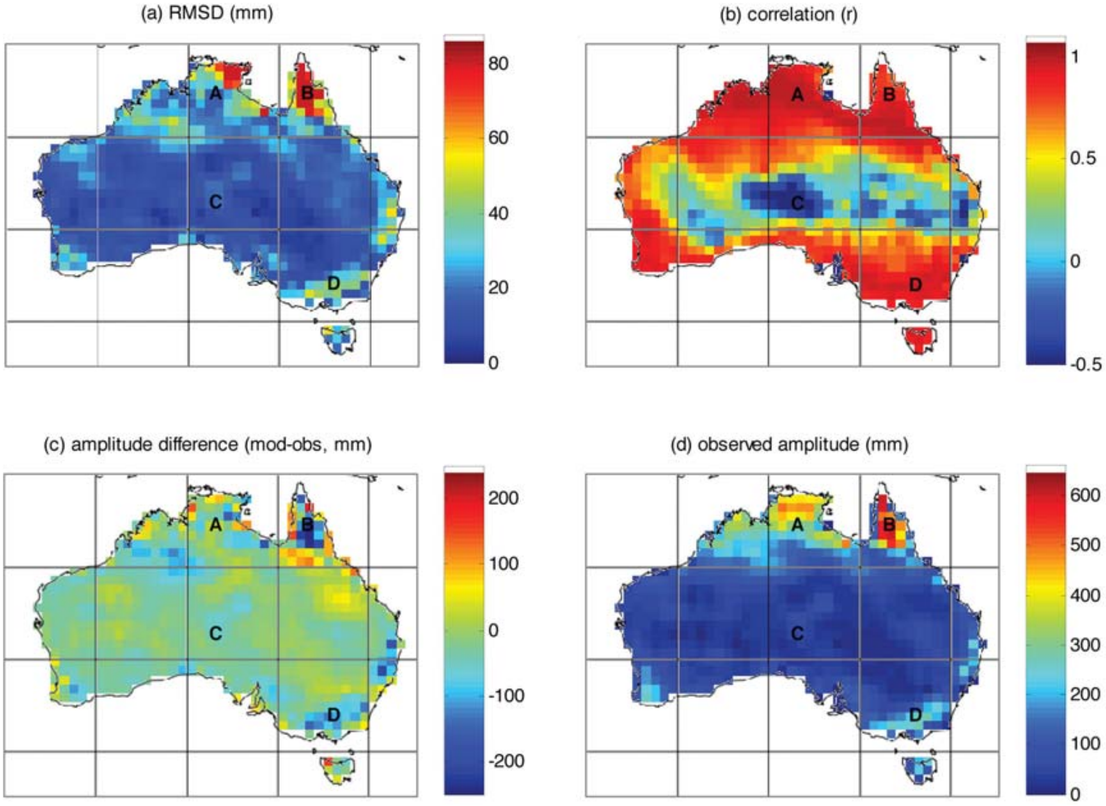

Figure 2. Comparison of seasonal cycle in GRACE TWS and AWRA TWS. Letters A–D indicate location of data shown in Figure 3 and discussed in the text.

The mean continental observed trend over this period was -5.6 mm yr^{-1}, compared with a modeled trend of $+0.8$ mm yr^{-1}. The scaling coefficient accounted for a small part of the disagreement between GRACE and AWRA trends: not applying the scaling coefficient reduced the continental trend to –4.5 mm yr^{-1} and reduced the area with strong trends and area with large differences between GRACE and AWRA trends (Table 1).

[12] Temporal patterns in the average observed and modeled continental water storage are shown in Figure 6a, along with values for the Nino3.4 index as an indicator of the phase of the El Niño–Southern Oscillation (ENSO). Most of the divergence in trends appears to occur between 2006 and 2009 (Figure 6b). The monthly and seasonal variability superimposed on this longer-term drying trend shows qualitative agreement with the temporal pattern in ENSO.

4.3. De-trended Anomalies

[13] Maps comparing the agreement in de-trended anomalies are shown in Figure 7, and trends for indicated grid cells are shown in Figure 8. The RMSD in de-trended monthly anomalies showed the same pattern spatial pattern as the

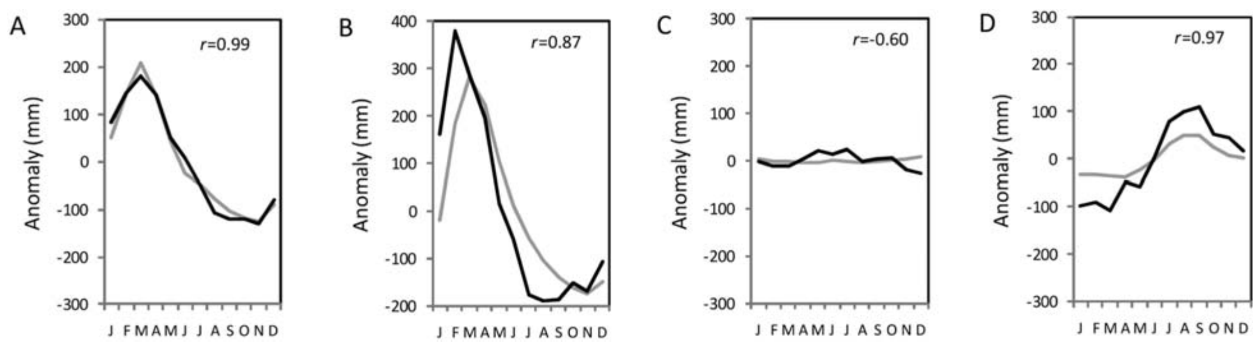

Figure 3. Seasonal cycle in GRACE TWS (black) and AWRA TWS (shaded) for selected grid cells (refer to Figure 2 for locations).

Table 1. Statistics of the Absolute Agreement Between AWRA and GRACE TWS, Showing the Effect of Not Applying the Scaling Coefficient to the GRACE Product[a]

	GRACE, Original	GRACE, Scaling Removed	AWRA
	Seasonal Cycle (SC)		
Median $RMSD_{SC}$ (mm)	15.9	14.3	n/a
Median ASC (mm)	58.8	49.5	38.7
$ASC_{GRACE} > ASC_{AWRA}$ (% cells)	79%	66%	n/a
$\lvert ASC_{GRACE} - ASC_{AWRA} \rvert > 30$ mm (% cells)	43%	38%	n/a
	Linear Trends (LT)		
Mean LT (mm yr^{-1})	−5.6	−4.5	0.8
$LT < -15$ mm yr^{-1} (% cells)	19%	10%	4%
$(LT_{GRACE} - LT_{AWRA}) < -10$ mm yr^{-1} (% cells)	29%	19%	n/a
	De-trended Anomalies (DA)		
Median SDA (mm)	36.1	31.3	19.7
Median $RMSD_{DA}$ (mm)	33.7	28.0	n/a
$RMSD_{DA} > 50$ mm (% cells)	25%	16%	n/a
$SDA_{GRACE} > SDA_{AWRA}$ (% cells)	75%	68%	n/a
	Total Anomalies (TA)		
Median $RMSD_{TA}$ (mm)	48.3	37.7	n/a
$RMSD_{TA} > 50$ mm (% cells)	48%	28%	n/a

[a]ASC, amplitude of seasonal cycle; LT, linear trend; SDA, standard de-trended anomaly; TA, total anomaly.

observed de-trended anomalies; RMSD was greatest in the wetter coastal regions (>80 mm, Figures 8a, 8b, and 8c), decreasing to <15 mm in the arid interior (Figure 8d). The correlation between anomalies was greatest in northern and eastern Australia (0.40–0.85), and lowest (<0.10) in the arid interior. Correlation was also lower in southwest WA and along the Queensland coast. The model estimated lesser TWS anomalies than retrieved from GRACE along the southeast coast and in southwest WA, but higher anomalies in parts of northern Queensland; differing by >50 mm in some cases. Not applying the scaling coefficient improved indicators of absolute agreement between GRACE and AWRA (Table 1). In particular, the area where the RMSD of anomalies exceeds 50 mm is reduced from 25% to 16% of grid cells.

4.4. Overall Agreement

[14] Maps comparing the overall agreement in total (i.e., not de-trended) observed and modeled TWS are shown in Figures 9a and 9c; Figure 9b shows the prior error estimates based on the analysis of the retrieval process as available from the GRACE Tellus Web site; and Figure 9d shows the fraction in total differences explained by the three components. Large differences (RMSD > 100 mm) occur in north Queensland and along the southeast coast and are mostly because of diverging linear trends (Figure 9a and 9d). Large differences were also found along the Gulf of Carpentaria (attributed to unaccounted oceanic influences) and southwestern Australia (not associated with any eight-year trend). Smaller differences (50–90 mm) occur in the Great Sandy Desert and are primarily due to different linear trends. The remaining areas show better agreement between observed and modeled TWS (difference <50 mm); in the areas with the best agreement (<15 mm), differences in de-trended anomalies generally dominate the other two components. Correlations are highest in southern and northern Australia and lowest in the arid interior, reflecting differences in variability rather than being a measure of absolute agreement (Figure 9c). The spatial pattern in prior GRACE retrieval error estimates agrees rather well with the calculated pattern in RMSD between AWRA and GRACE (Figure 9a). A direct comparison of pixel values suggests (Figure 10) suggests that the prior GRACE error estimates may be conservative estimates, considering that there must also be error in the model (see section 5). The GRACE data providers caution against the possible degradation in trend information due the scaling applied. Table 1 provides some evidence for this; indeed all metrics of agreement are improved if the scaling coefficients are not applied. It is reiterated that a marked difference in continental mean trend remains when scaling is not applied (Figure 6b).

5. Discussion

5.1. Uncertainties in GRACE Terrestrial Water Storage Estimates

[15] The scaling applied to the GRACE TWS retrievals [*Swenson and Wahr*, 2006] appeared to contribute to differences between GRACE and AWRA. Without this scaling the absolute agreement is improved, particularly in terms of the unexplained trend and the RMSD of de-trended anomalies (Table 1). Geographically, the greatest improvements are not in coastal areas per se (where signal leakage to the ocean would be expected), but apparently mainly in areas with high rainfall gradients, such as southwestern Australia, Top End, and the Murray uplands (Figure 11, cf. Figure 1). These are also regions were the scaling factors are largest (>2); therefore, it seems plausible that the retrievals have been "over-corrected." This limits the confidence with which absolute differences in the amplitude of the seasonal cycle, the magnitude of 8 year trends, and the RMSD in (de-trended or total) anomalies between GRACE and AWRA can be attributed. The calculated correlation coefficients are not affected, however.

[16] There was encouraging spatial agreement between the prior GRACE TWS error estimates and estimates obtained through the model-data comparison. The median

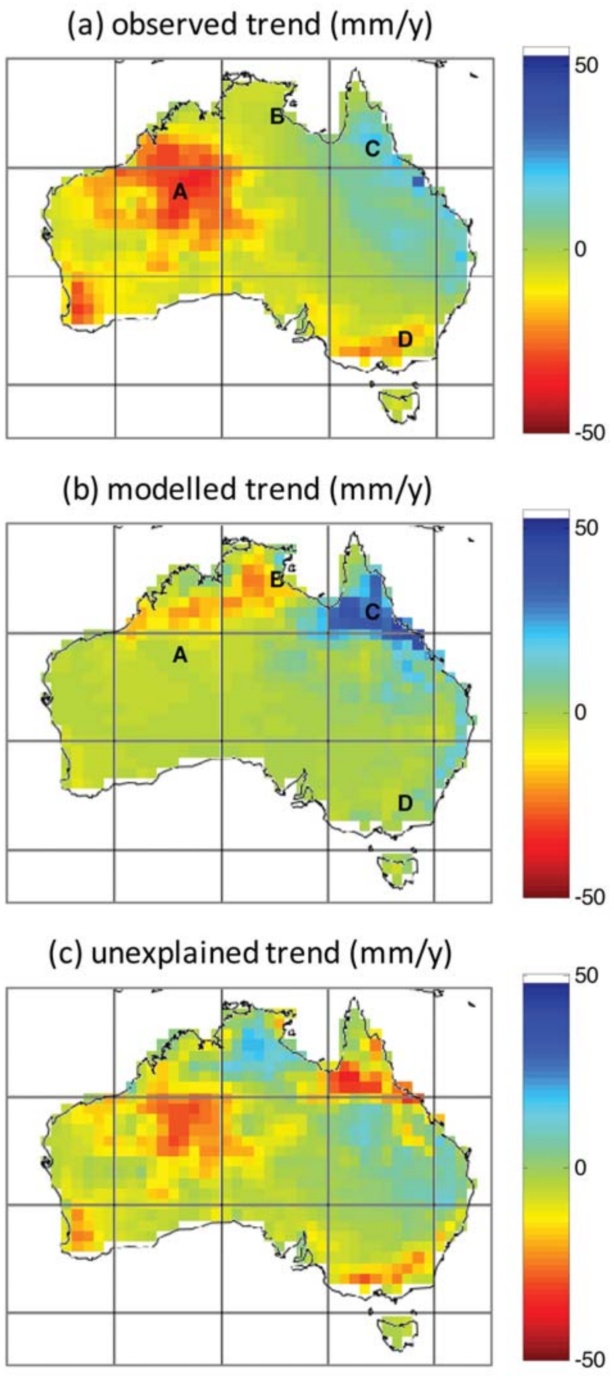

Figure 4. Comparison of linear trends in GRACE TWS and AWRA TWS. Letters A–D indicate location of data shown in Figure 5 and discussed in the text (unexplained trend was calculated as observed minus modeled trend).

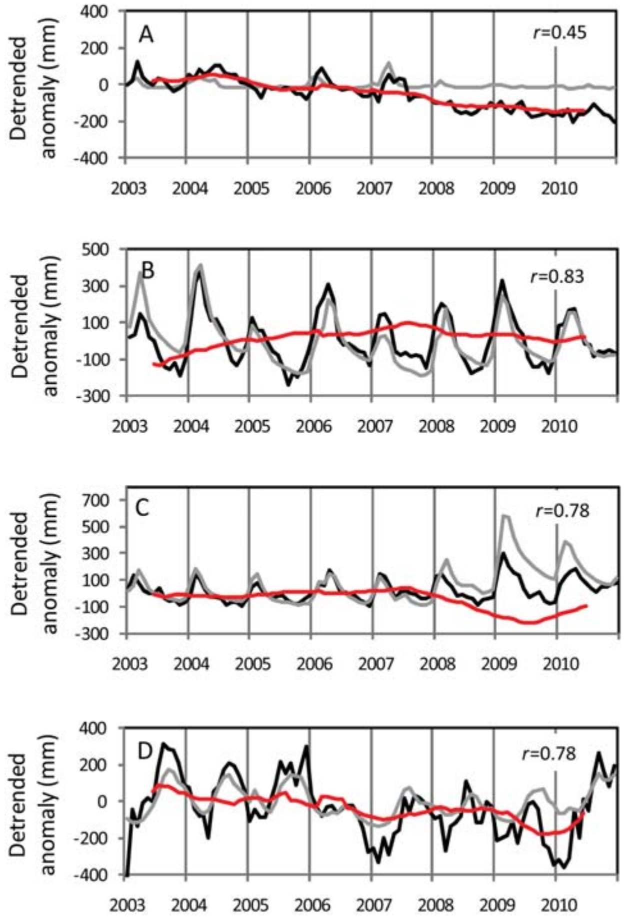

Figure 5. Time series of GRACE TWS (black) and AWRA TWS (shaded) along with 12 month running averages of the difference between the two (red) (refer to Figure 4 for locations).

prior estimate of 47 mm (mean 58 mm) was close to the calculated standard difference of 48 mm when scaling was applied and 38 mm when it was not. Since the model estimates must also have error, these numbers suggest that the actual error in GRACE is considerably less than estimated from the retrieval. For example, part of total disagreement between model and data is because of the divergence in trends which presumably is mainly attributable to model error (Figure 9d).

[17] In addition to uncertainty about the scaling, the coarse resolution of the GRACE hydrology products also makes them more challenging to use in spatial parameter calibration and assimilation, although some early progress toward such formal model-data fusion approaches has been made [*Zaitchik et al.*, 2008]. The dominance of systematic trend-related differences between observation and model also suggest that model structure and parameter improvements (i.e., calibration) should be made before attempting assimilation of GRACE TWS retrievals. Combined with the expected 2 year or longer observation gap between the current set of GRACE satellites and their successors, the observations are currently not considered for assimilation in the operational AWRA system.

5.2. Limitations of the Model

[18] Despite uncertainties in the accuracy of the GRACE TWS estimates, several valuable inferences could be made. The model appeared to underestimate the seasonal TWS amplitude, suggesting a tendency of the modeled soil, groundwater and/or surface water systems to drain too fast.

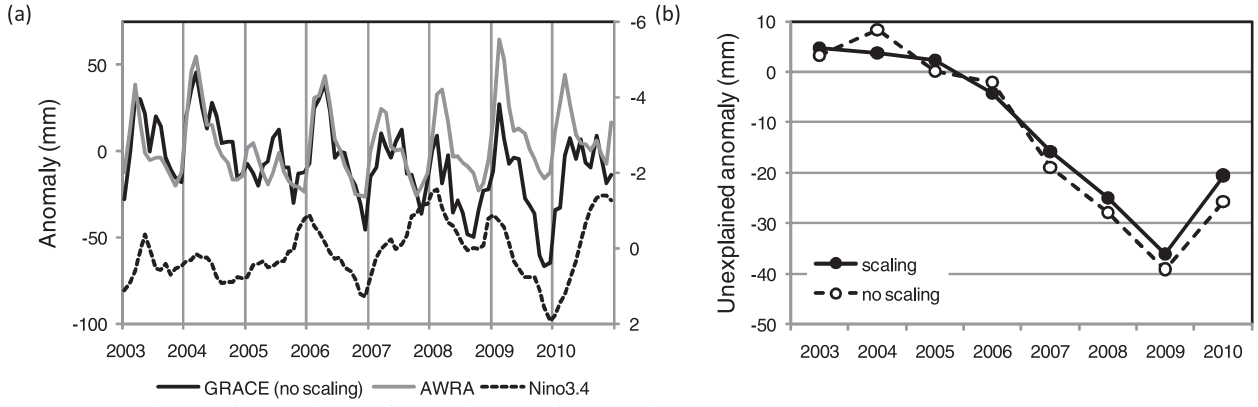

Figure 6. (a) Time series of average Australian GRACE (black line) and AWRA (shaded) TWS anomalies, (b) difference between mean annual AWRA TWS and GRACE TWS for Australia.

The AWRA model structure and parameterization were developed using concepts and streamflow observations that were probably biased toward small, well-defined upland catchments, often with medium to high rainfall [cf. *Van Dijk*, 2009; *Van Dijk*, 2010b]. By contrast, most of Australia is covered by extensive plains with often poorly developed drainage networks that drain internally into aquifers, wetlands and (salt) lakes (several are visible in Figure 1).

[19] Across most of Australia, AWRA showed lesser decreases—or greater increases—in TWS than GRACE over the 8 year period. This divergence was not explained by the scaling of GRACE TWS estimates (Figures 11a and 11b). Unaccounted changes in water stored in public reservoirs are unlikely to explain the difference either: total storage in public reservoirs across Australia at the start of 2003 (data available at http://www.water.gov.au/WaterAvailability/) and

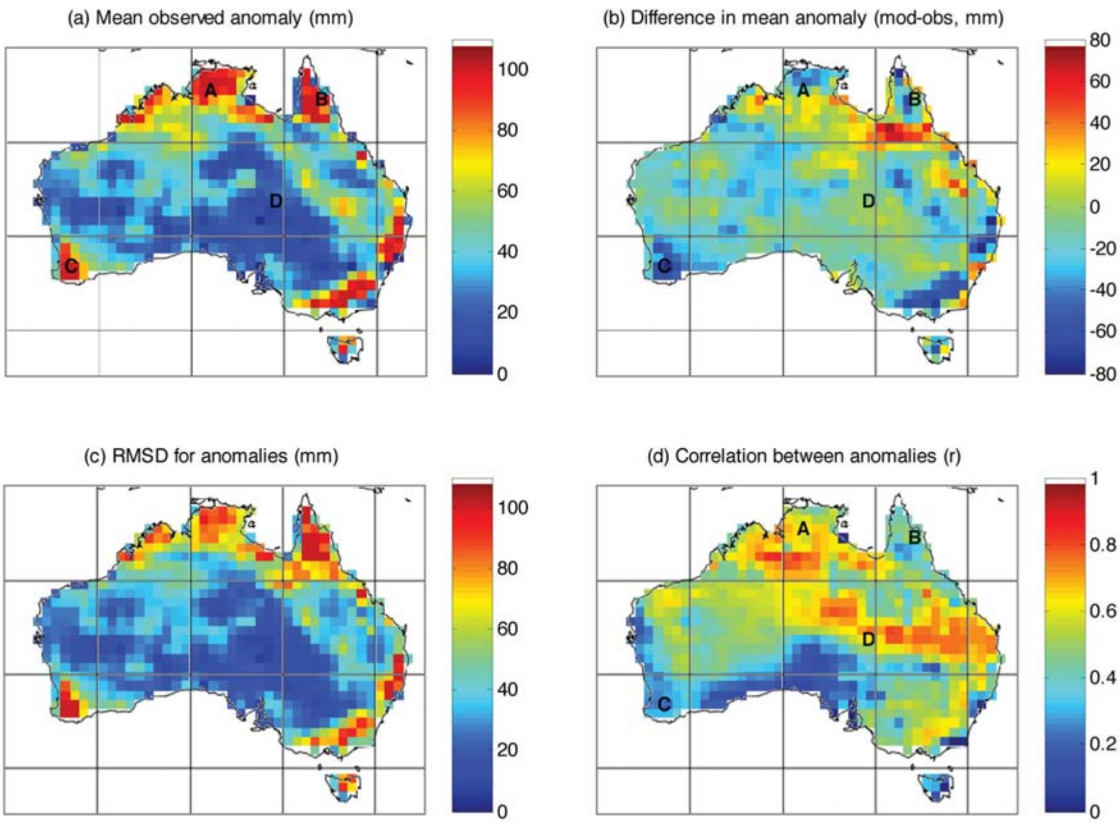

Figure 7. (a and b) Comparison of the mean GRACE and AWRA TWS de-trended anomaly (calculated as the root-mean-square of anomalies), and (c and d) indicators of agreement between de-trended anomalies. Letters A–D indicate location of data shown in Figure 8 and discussed in the text.

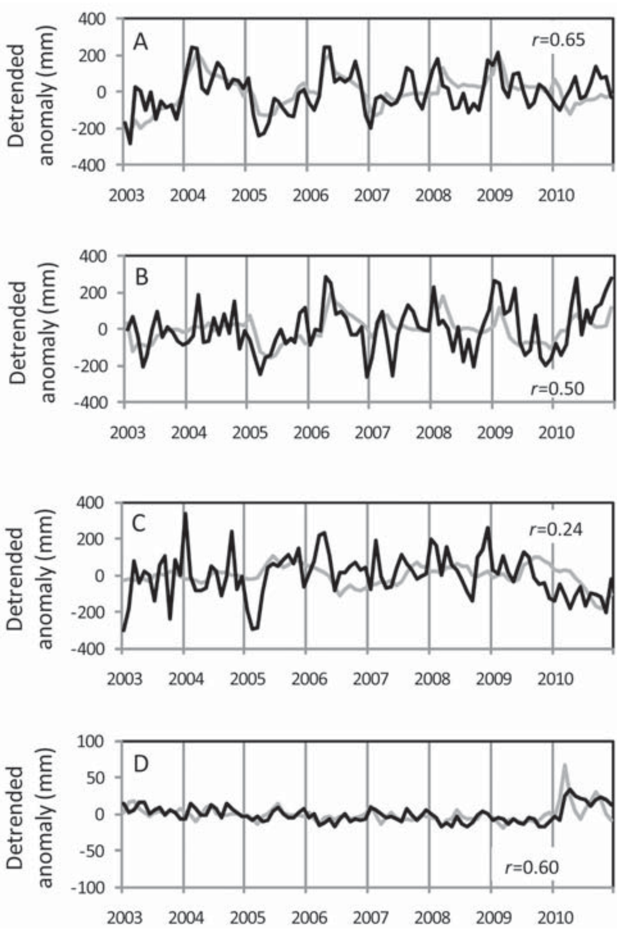

Figure 8. Time series of de-trended anomalies in GRACE TWS (black) and AWRA TWS (shaded) for selected grid cells (refer to Figure 7 for locations). Note the different vertical scale of Figure 8d.

the end of 2009 (data available at http://water.bom.gov.au/waterstorage/awris/) were almost identical (39 km^3, equivalent to 5 mm) and increased to $\sim$60 km^3 by the end of 2010. They may account for a small part of unexplained trends in southwest and southeast Australia; for example, using the same data sources, the average contribution to the linear trend for the state of Victoria was estimated at -1.6 mm yr^{-1} during 2003–2009. In the other regions with unexplained GRACE trends there are much fewer or no public storages, while the volume of private storages is negligible.

[20] Alternative explanations include errors in model forcing and model physics. The greatest trend deviations ($>$15 mm yr^{-1}; with or without scaling) occurred in North Queensland, the Great Sandy Desert, and the southern Murray Basin (Figure 11a). The difference in trends for North Queensland appeared mainly associated with cyclone Charlotte in 2009 (Figure 5c). The rainfall gauge interpolation procedure for this event may have led to rainfall overestimation, but probably more likely is that runoff to the ocean occurred faster and more effectively than estimated by the model. Rainfall in the Great Sandy Desert is basically ungauged (Figure 12a). Satellite rainfall estimates derived from the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) product [*Huffman et al.*, 2007] show a stronger negative rainfall trend for 2003–2010 (-38 mm yr^{-1}; Figure 12b) than the interpolated gauge data (-25 mm yr^{-1}). It follows that errors in the gauge interpolation may explain some of the difference in GRACE and AWRA TWS trends for this area (see Figure 5a), although "true" rainfall remains unknown. A longer record from a station nearby shows that the observed rainfall decrease is part of a return to drier conditions after a sequence of unusually wet years (Figure 12b). In addition to possible rainfall forcing errors, Figure 5a shows that the model stores are effectively depleted after 2007, whereas the observations shows continuing water loss.

[21] The disagreement found for the Great Sandy Desert and also for the Murray Basin suggests a model deficiency that leads to underestimation of the rate of groundwater discharge, recharge, or both. *Leblanc et al.* [2009] compared GRACE TWS estimates over the Murray-Darling Basin with soil moisture estimates produced by the Global Land Data Assimilation System (GLDAS) driving the Noah land surface model [*Rodell et al.*, 2004] and found a similar negative trend in GRACE TWS that was only partially explained by modeled soil moisture and could attribute the remaining trend to widespread groundwater lowering as observed in monitoring wells. Greater than estimated groundwater discharge to the surface water network is unlikely to explain the slow trends observed. Diffuse groundwater discharge may well have been underestimated, however. This process is described in the model but is assumed negligible once groundwater level reaches the base of the surface drainage network. In reality, groundwater discharge can continue after connectivity with the surface water network has been lost, through deep root water uptake [e.g., *Petrone et al.*, 2010] and capillary rise [*Costelloe et al.*, 2008]. Even capillary rise can occur at sufficiently high rates (e.g., up to 30 mm yr^{-1}, [*Costelloe et al.*, 2008]) to explain the modest observed 8 year trends as a delayed response to a sequence of below-average rainfall years. In the Great Sandy Desert, permeable sand deposits overlay and obstruct the surface drainage network. This would be expected to enhance groundwater recharge and reduce fast groundwater discharge, leaving the stored water available for slower discharge through diffuse processes.

5.3. Australia's Contribution to Sea Level Change and Oceanic Influences on Water Storage

[22] Although not part of our main objective, it is of interest to consider the contribution of the continental TWS trend to global sea level rise between 2003 and 2009 (we did not have access to more recent updates). The surface area of Australia and the world's oceans is 7.7 and 335 million of km^2, respectively. Therefore, the observed continental loss of water for 2003–2009 (6.7–8.0 mm yr^{-1} without and with scaling, respectively) equals 50–60 km^3 yr^{-1} or an equivalent $+0.15$–0.18 mm yr^{-1} sea mass change. This is of the same magnitude as the net sea level lowering of 0.22 $\pm$ 0.05 mm yr^{-1} (2002–2009) estimated from GRACE for 33 of the world's largest rivers basins combined [*Llovel et al.*, 2010a]. This depletion was not reproduced by the hydrological model. For the period considered, the average rate of sea level rise estimated from satellite altimetry was

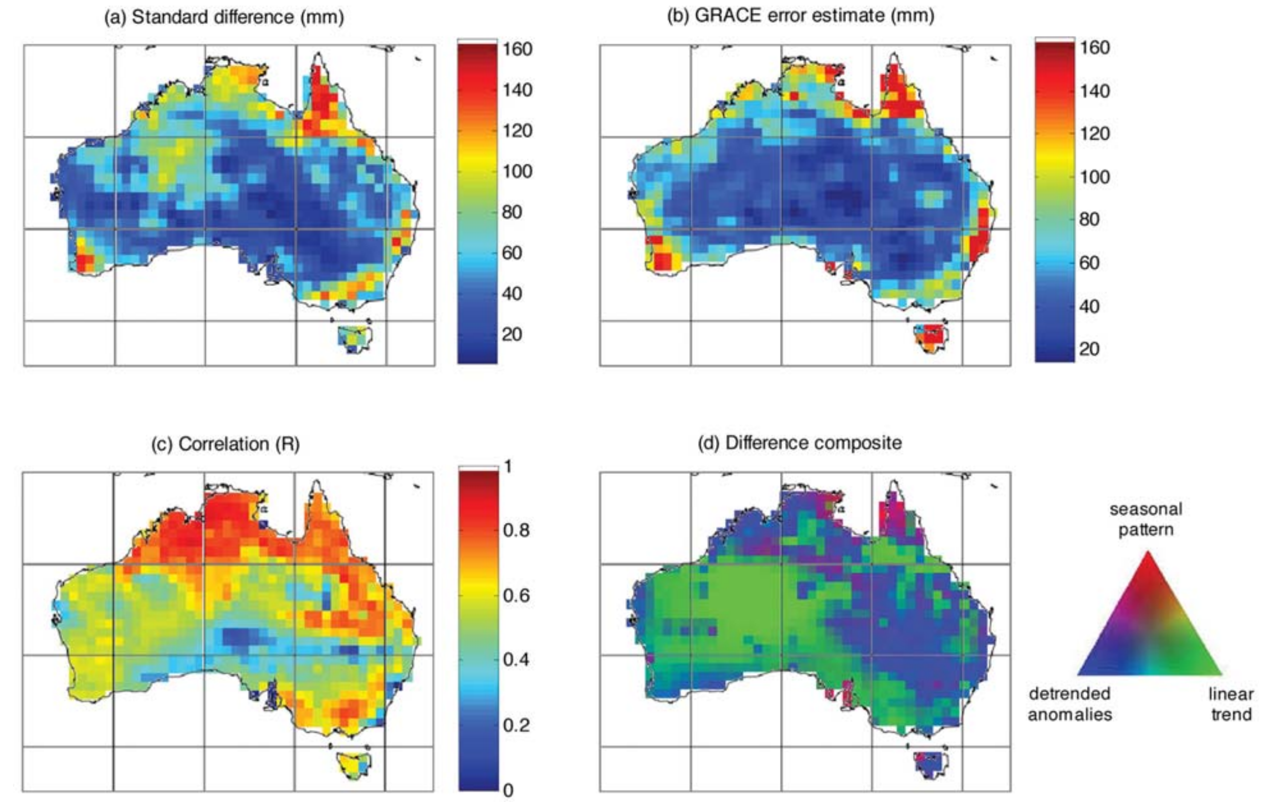

Figure 9. (a to c) Indicators of overall agreement between original GRACE and AWRA TWS anomalies. (d) Color composite showing the relative contribution of the three signal components (seasonal cycle, eight-year trends, de-trended anomalies) to the overall disagreement between GRACE and AWRA TWS.

2.6 mm yr^{-1} [*Ablain et al.*, 2009]. GRACE based estimates of ocean mass increase (i.e., not including thermal and salinity influences on sea level) show a large range of 0.3–2.1 mm yr^{-1} [*Cazenave et al.*, 2009; *Leuliette and Miller*, 2009; *Llovel et al.*, 2010b]. It follows that Australia's TWS changes account for anywhere between 7% and 60% of

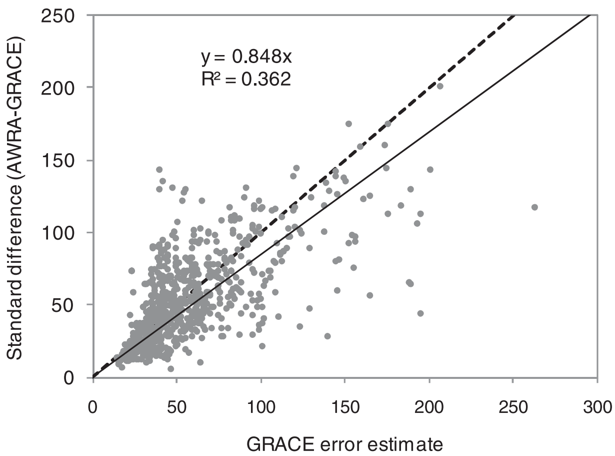

Figure 10. Relationship between GRACE retrieval error estimates and standard difference estimates derived from this analysis. Shown are the 1:1 line (dashed) and a linear regression equation passing through the origin (solid line).

global sea mass increase and 6%–7% of sea level rise over this period. Understanding Australia's groundwater dynamics therefore appears relevant to understanding sea level change.

[23] Conversely, both model and GRACE data suggest that total water storage across the Australian continent is strongly influenced by ENSO conditions (Figure 6a). The influence of ENSO and other ocean phenomena (e.g., the Indian Ocean Dipole) on Australian rainfall are well published. Following the strong influence detected in remotely sensed shallow soil moisture [*Liu et al.*, 2007], our analysis shows that a similarly strong signal is propagated into total water storage but with stronger and lingering accumulative effects.

6. Conclusions

[24] Our objective was to test whether a comparison between GRACE and AWRA TWS would allow us to identify concrete needs for model improvement. The comparison benefited from decomposition of the data into their seasonal cycle, linear trends, and de-trended anomalies. We found evidence that the scaling applied to GRACE artificially degraded the comparison for some regions. The generally good agreement between GRACE and model TWS estimates suggests that the GRACE TWS are likely to be more accurate than estimated as part of the retrieval process.

[25] AWRA tended to have smaller seasonal amplitude than GRACE. GRACE showed a strong (>15 mm yr^{-1})

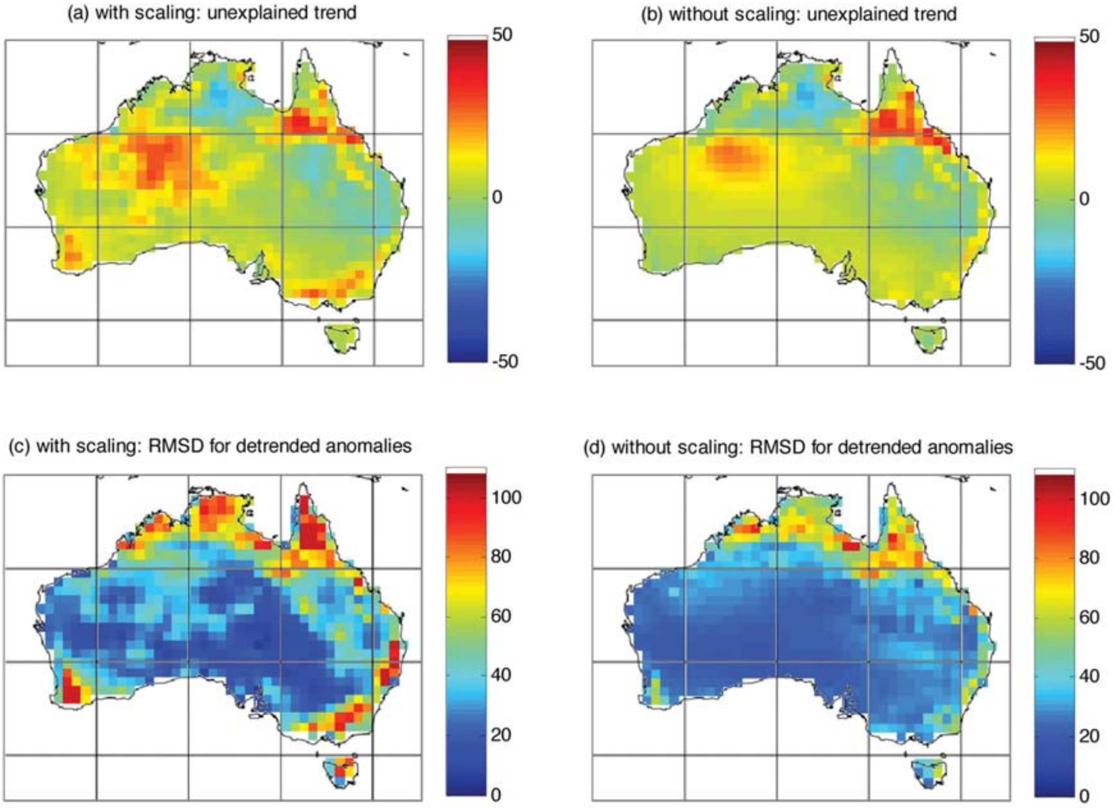

Figure 11. Effect of the GRACE scaling on agreement with AWRA TWS: (a and b) unexplained trend (mm yr^{-1}); (c and d) RMSD in detrended anomalies (mm).

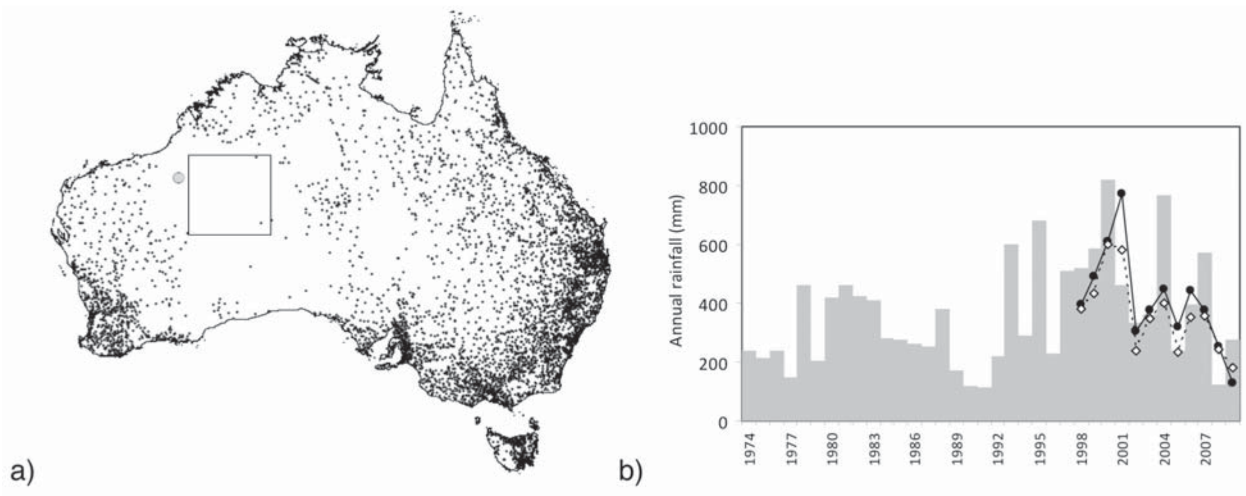

Figure 12. (a) Distribution of rain gauges (points) used in generating AWRA rainfall forcing data. Also shown are the location of the Telfer aerodrome gauge (21.71°S, 122.23°E, shaded dot) and region of interest (rectangle), (b) Annual rainfall at Telfer aerodrome (shaded bars) and region average rainfall derived from AWRA model forcing (open diamonds) and the TMPA satellite product [*Huffman et al.*, 2007] (solid dots).

drying trend in northwest Australia that was associated with a preceding period of unusually wet conditions, whereas weaker drying trends in the southern Murray Basin and southwest Western Australia were associated with relatively dry conditions. AWRA estimated trends were less negative for these regions, while a more positive trend was estimated for areas affected by cyclone Charlotte in 2009. For 2003–2009, a decrease of 7–8 mm yr^{-1} (50–60 km^3 yr^{-1}) was estimated from GRACE, enough to explain 6%–7% of the contemporary rate of global sea level rise. This trend was not reproduced by the model.

[26] The analysis provided several insights into model performance that could not have been obtained with any other source of observations, and indicated two priorities to improve the AWRA system. First, precipitation forcing in poorly gauged areas is one such opportunity and is being addressed through blending of gauged and remotely sensed precipitation [*Van Dijk and Renzullo*, 2011; *Renzullo et al.*, 2011]. Second, a tendency for water storage to decline for longer and more gradually than estimated by AWRA was observed and indicates that greater sophistication in the model assumptions about diffuse groundwater discharge and surface-groundwater connectivity is needed.

[27] **Acknowledgments.** This work is part of the water information research and development alliance between the Bureau of Meteorology and CSIRO's Water for a Healthy Country Flagship. GRACE land data were processed by Sean Swenson, supported by the NASA MEASURES Program, and are available at http://grace.jpl.nasa.gov. Helpful suggestions from Paul Tregoning, Sean Swenson, Richard Cresswell, Glenn Harrington, Brian Smerdon, Don McFarlane, Tom van Niel and John Church are gratefully acknowledged.

References

Ablain, M., A. Cazenave, G. Valladeau, and S. Guinehut (2009), A new assessment of the error budget of global mean sea level rate estimated by satellite altimetry over 1993–2008, *Ocean Sci.*, *5*(2), 193–201.

Cazenave, A., K. Dominh, S. Guinehut, E. Berthier, W. Llovel, G. Ramillien, M. Ablain, and G. Larnicol (2009), Sea level budget over 2003–2008: A reevaluation from GRACE space gravimetry, satellite altimetry and Argo, *Glob. Plan. Change*, *65*(1–2), 83–88.

Costelloe, J., E. Irvine, A. Western, V. Matic, J. Walker, and M. Tyler (2008), Quantifying near-surface, diffuse groundwater discharge along the south-west margin of the Great Artesian Basin, in *Proceedings of Water Down Under 2008*, pp. 831, Engineers Australia, Modbury, Australia.

Guerschman, J. P., A. Van Dijk, G. Mattersdorf, J. Beringer, L. B. Hutley, R. Leuning, R. C. Pipunic, and B. S. Sherman (2009), Scaling of potential evapotranspiration with MODIS data reproduces flux observations and catchment water balance observations across Australia, *J. Hydrol.*, *369*(1–2), 107–119.

Güntner, A. (2008), Improvement of global hydrological models using GRACE data, *Surv. Geophys.*, *29*(4), 375–397.

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

King, E. A., et al. (2011), An operational actual evapotranspiration product for Australia, paper presented at WIRADA Symposium, Water Inf. Res. and Dev. Alliance, Melbourne, Victoria, Australia.

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

Leuliette, E. W., and L. Miller (2009), Closing the sea level rise budget with altimetry, Argo, and GRACE, *Geophys. Res. Lett.*, *36*, L04608, doi:10.1029/2008GL036010.

Liu, Y., R. A. M. De Jeu, A. I. J. M. Van Dijk, and M. Owe (2007), TRMM-TMI satellite observed soil moisture and vegetation density (1998–2005) show strong connection with El Niño in eastern Australia, *Geophys. Res. Lett.*, *34*, L15401, doi:10.1029/2007GL030311.

Liu, Y., J. Evans, M. McCabe, A. I. J. M. Van Dijk, and R. A. M. De Jeu (2010), Soil moisture estimates over Murray Darling Basin in Australia (1992–2008), paper presented at 4th International Workshop on Catchment-Scale Hydrological Modeling and Data Assimilation, Natl. Nat. Sci. Found. China, Lhasa, China, 21–23 July 2010.

Llovel, W., M. Becker, A. Cazenave, J. F. Cretaux, and G. Ramillien (2010a), Global land water storage change from GRACE over 2002–2009: Inference on sea level, *C. R. Geosci.*, *342*(3), 179–188, doi:10.1016/j.crte.2009.12.004.

Llovel, W., S. Guinehut, and A. Cazenave (2010b), Regional and interannual variability in sea level over 2002–2009 based on satellite altimetry, Argo float data and GRACE ocean mass, *Ocean Dyn.*, *60*(5), 1193–1204.

Paulson, A., S. Zhong, and J. Wahr (2007), Inference of mantle viscosity from GRACE and relative sea level data, *Geophys. J. Int.*, *171*(2), 497–508.

Peltier, W. R. (2004), Global glacial isostasy and the surface of the ice-age Earth: The ICE-5G (VM2) Model and GRACE, *Annu. Rev. Earth Planet. Sci.*, *32*, 111–149,

Petrone, K. C., J. D. Hughes, T. G. Van Niel, and R. P. Silberstein (2010), Streamflow decline in southwestern Australia, 1950–2008, *Geophys. Res. Lett.*, *37*, L11401, doi:10.1029/2010GL043102.

Ramillien, G., J. S. Famiglietti, and J. Wahr (2008), Detection of continental hydrology and glaciology signals from GRACE: A review, *Surv. Geophys.*, *29*(4), 361–374.

Renzullo, L. J., A. Chappell, T. Raupach, P. Dyce, M. Li, and Q. Shao (2011), An assessment of statistically blended satellite-gauge precipitation data for daily rainfall analysis in Australia, in Proceedings of the 34th ISRSE symposium, Sydney, New South Wales, Australia.

Rodell, M., P. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, and M. Bosilovich (2004), The global land data assimilation system, *Bull. Am. Meteorol. Soc.*, *85*(3), 381–394.

Rodell, M., I. Velicogna, and J. S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in India, *Nature*, *460*(7258), 999–1002.

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

Swenson, S., J. Famiglietti, J. Basara, and J. Wahr (2008), Estimating profile soil moisture and groundwater variations using GRACE and Oklahoma Mesonet soil moisture data, *Water Resour. Res.*, *44*, W01413, doi:10.1029/2007WR006057.

Syed, T. H., J. S. Famiglietti, M. Rodell, J. Chen, and C. R. Wilson (2008), Analysis of terrestrial water storage changes from GRACE and GLDAS, *Water Resour. Res.*, *44*, W02433, doi:10.1029/2006WR005779.

Tapley, B. D., S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins (2004), GRACE measurements of mass variability in the Earth System, *Science*, *305*(5683), 503–505.

Tregoning, P., K. Lambeck, and G. Ramillien (2008), GRACE estimates of sea surface height anomalies in the Gulf of Carpentaria, Australia, *Earth Planet. Sci. Lett.*, *271*(1–4), 241–244.

Van Dijk, A. I. J. M. (2009), Climate and terrain factors explaining streamflow response and recession in Australian catchments, *Hydrol. Earth Syst. Sci.*, *14*, 159–169.

Van Dijk, A. I. J. M. (2010a), Landscape Model (version 0.5) technical description, *AWRA Tech. Rep. 3*, WIRADA/CSIRO Water for a Healthy Country Flagship, Canberra, ACT, Australia.

Van Dijk, A. I. J. M. (2010b), Selection of an appropriately simple storm runoff model, *Hydrol. Earth Syst. Sci.*, *14*(3), 447–458.

Van Dijk, A. I. J. M., and L. J. Renzullo (2009), The Australian water resources assessment system: Blending water cycle observations and models at local and continental scale, paper presented at 6th International Scientific Conference on the Global Energy and Water Cycle and 2nd Integrated Land Ecosystem—Atmosphere Processes Study Science Conference, Monash Univ., Melbourne, Victoria, Australia, 24–28 Aug.

Van Dijk, A. I. J. M., and L. J. Renzullo (2011), Water resource monitoring systems and the role of satellite observations, *Hydrol. Earth Syst. Sci.*, *15*, 39–55.

Van Dijk, A. I. J. M., and G. A. Warren (2010), Evaluation against observations, *AWRA Tech. Rep. 4*, WIRADA/CSIRO Water for a Healthy Country Flagship, Canberra, ACT, Australia.

Viney, N. R., et al. (2011), Comparison of models and methods for estimating spatial patterns of streamflow across a continental domain, paper presented at WIRADA Symposium, Water Inf. Res. and Dev. Alliance, Melbourne, Victoria, Australia.

Zaitchik, B. F., M. Rodell, and R. H. Reichle (2008), Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River Basin, *J. Hydrometeorol.*, *9*(3), 535–548.

L. J. Renzullo and A. I. J. M. van Dijk, Black Mountain Laboratories, CSIRO Land and Water, Canberra, ACT 2601, Australia. (albert.vandijk@csiro.au)

M. Rodell, Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.