IMPROVING ACCURACY OF GROUNDWATER STORAGE IN THAILAND USING GRACE DATA ASSIMILATION TECHNIQUE

Natthchet Tangdamrongsub

Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, USA
Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

ABSTRACT
Groundwater storage (GWS) is a fundamental component of the terrestrial hydrology and climate system. Accurate GWS measurements are required for a reliable assessment of regional water resource availability and climate variation but obtaining GWS is very challenging due to the sparsity of the groundwater measurement network. To date, the only measurements of large-scale GWS available are gravity data from the Gravity Recovery And Climate Experiment (GRACE) and GRACE follow-on missions. The coarse spatial resolution of GRACE (e.g., 100,000 km^2), on the other hand, limits its application to groundwater research in large river basins, whereas effective communication with the private sector or public stakeholders would necessitate a much higher spatial resolution. Therefore, this study employs data assimilation (DA) techniques to statistically combine the strengths of satellite data and model estimates to improve the spatiotemporal resolution and accuracy of GWS. The analysis is carried out in Thailand's northern Ping River basin, where in situ groundwater data are available to evaluate DA performance. The results reveal that GRACE DA significantly improves GWS estimates by increasing correlation coefficients (relative to the ground truth) by up to 0.53, highlighting the GRACE DA approach as an effective tool for a groundwater monitoring system in Thailand. The proposed algorithm is entirely based on publicly available data, and the approach is easily adaptable to other regions of Thailand.

Keywords— Data assimilation, GRACE, Groundwater, Thailand

1 INTRODUCTION

Groundwater is critical for supporting ecosystems and facilitating human adaptation to climate change, which may exacerbate water accessibility and food security issues (Taylor et al., 2013). Accurate groundwater data is thus essential for improving the reliability of the country's strategic plan. However, obtaining groundwater data from ground measurements is difficult due to the scarcity of sampling sites, despite providing accurate groundwater information. In some provinces, for instance, there may be only a few available measurement sites. As a result, groundwater information is frequently obtained through a hydrology or land surface model. A strong point of models is their ability to generate spatially distributed estimates and comprehensive environmental variables. The model's limitation is the significant uncertainty surrounding, for example, inaccurate parameter calibrations and underrepresented model physics, and relying solely on model simulation may result in faulty or invalid predictions.

Groundwater data can also be estimated by measuring the Earth's regional gravity fields over time. These have been observed by the Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On (GRACE-FO) satellite missions since 2002 (Tapley et al., 2004). GRACE-FO is a continuous gravity mission to its successor (GRACE), and this study refers to both missions as GRACE. GRACE is a twin satellite tracking system that flies in a near-polar orbit 500 km above the Earth's surface. The K-band ranging system is used to calculate the range (or range rate) deviation between the twin satellites as a result of changes in Earth's gravity. This range measurement is used to derive variations in terrestrial water storage (TWS), which is the sum of soil moisture, groundwater, canopy, snow, and surface water storage. Satellite gravimetry is distinct from other remote sensing techniques in that it can detect total column of mass variations (including groundwater), whereas other techniques are only sensitive to a few centimeters of depth. The disadvantage of GRACE is its limited spatiotemporal resolution, which only provides monthly-average data at a footprint of approximately 100,000 km^2. This limits GRACE applications to a large area, while direct use of GRACE data is unlikely to benefit local farmland or urban planning.

This study employs the data assimilation (DA) technique to combine the strengths of model simulation and GRACE data. In DA, the model states are statistically adjusted using satellite observations, taking into account the uncertainties in the model states and observations (e.g., Tangdamrongsub et al., 2021). Despite the success of GRACE DA techniques, they have never been used in Thailand. The purpose of this paper is to demonstrate the GRACE DA application in improving groundwater information in Thailand. The development is demonstrated in the northern part of the Ping River Basin, where in situ groundwater measurements are available to evaluate the results (Fig. 1). GRACE data are assimilated into the Community Atmosphere–Biosphere Land Exchange (CABLE) land surface model (LSM) to improve the groundwater storage (GWS) at a resolution of approximately 5x5 km^2. The analysis from this study highlights the relevance of the GRACE DA technique in assisting a groundwater monitoring system in Thailand.

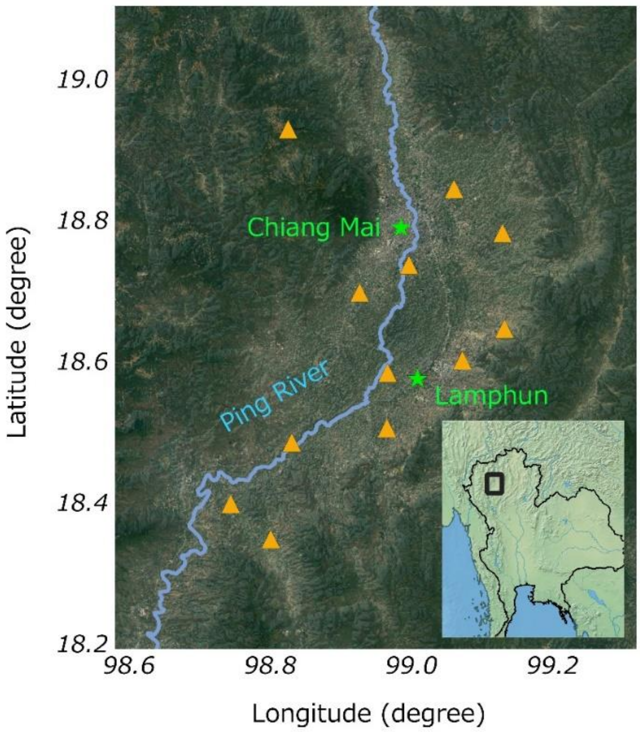

Fig. 1. The northern part of Ping River Basin in Thailand and geolocations of in situ groundwater sites. The inset shows the geolocation of the study area (black rectangle).

2 MODEL SETUP AND DATA

2.1 Model configurations

TWS and GWS estimates are obtained from CABLE LSM simulations, providing comprehensive hydrologic variables at 0.05° (~5 km) resolution (Tangdamrongsub et al., 2021). The model was developed by the Commonwealth Scientific and Industrial Research Organization in Australia and has been widely used for global terrestrial hydrology analyses. CABLE TWS includes soil moisture storage (SMS), GWS, snow water equivalent (SWE), canopy storage (CNP), of which SWE and CNP are neglectable in Thailand. Forcing data required for CABLE simulations are precipitation, air temperature, wind speed, humidity, surface pressure, and radiation. Precipitation is obtained from the Climate Hazards Group Infrared Precipitation with Station data (CHIRPS; Funk et al., 2015), while other forcing variables are derived from the Modern-Era Retrospective Analysis for Research and Applications, Version 2 products (MERRA-2; Gelaro et al., 2017). To overlay with model spatial resolution, all forcing variables are spatially resampled to 0.05° model grid space using nearest-neighbor interpolation.

2.2 GRACE observations

This study uses the GRACE and GRACE-FO mascon solutions (RL06M.MSCNv02) of the Jet Propulsion Laboratory (JPL), California Institute of Technology, downloaded from http://grace.jpl.nasa.gov. The mascon approach parameterizes Earth's mass variation using mass concentration function, yielding a more accurate TWS estimate than the spherical harmonic basis. From April 2002, the JPL mascon product provides monthly TWS variations and uncertainties at a spatial resolution of about 3° (~300 km). The solutions from April 2002 to December 2020 are used in this study. To obtain TWS variations relative to this study period, the long-term mean computed from all data in time series is removed from each monthly data at each mascon cell.

2.3 In situ groundwater measurements

In situ groundwater is obtained from the Department of Groundwater Resources in Thailand (DRG; http://www.dgr.go.th/th/home). DGR provides groundwater level (H) data on a map-based basis. The data at each site are manually downloaded. Only sites with more than three years of data that do not contain significant data gaps or outliers are used in this study. Because model estimates (GWS) and in situ data (H) are in different domains and cannot be compared directly, the evaluation is only performed in terms of temporal correlation between the two, as specific yield for the conversion from H to GWS is not available.

3 METHODOLOGIES

3.1 Data assimilation

The GRACE-derived TWS variations are assimilated into the CABLE model using the 3-dimensional ensemble Kalman smoother (EnKS 3D). The EnKS 3D accounts for spatial correlations in model and observation errors, given that the latter are highly correlated at neighboring 0.05°×0.05° grid cells (e.g., within 3°×3° observation space). This section provides a high-level overview of the EnKS 3D concept, while more in-depth information can be found in Tangdamrongsub et al. (2021). The EnKS 3D comprises forecast, analysis, and update distribution steps (Fig. 2). The forecast step propagates the ensemble model states forward in time. The analysis step calculates the monthly model state correction using GRACE observations (and uncertainties). The final step reinitializes the initial states and repeats the forecast step with the correction distributed across the month.

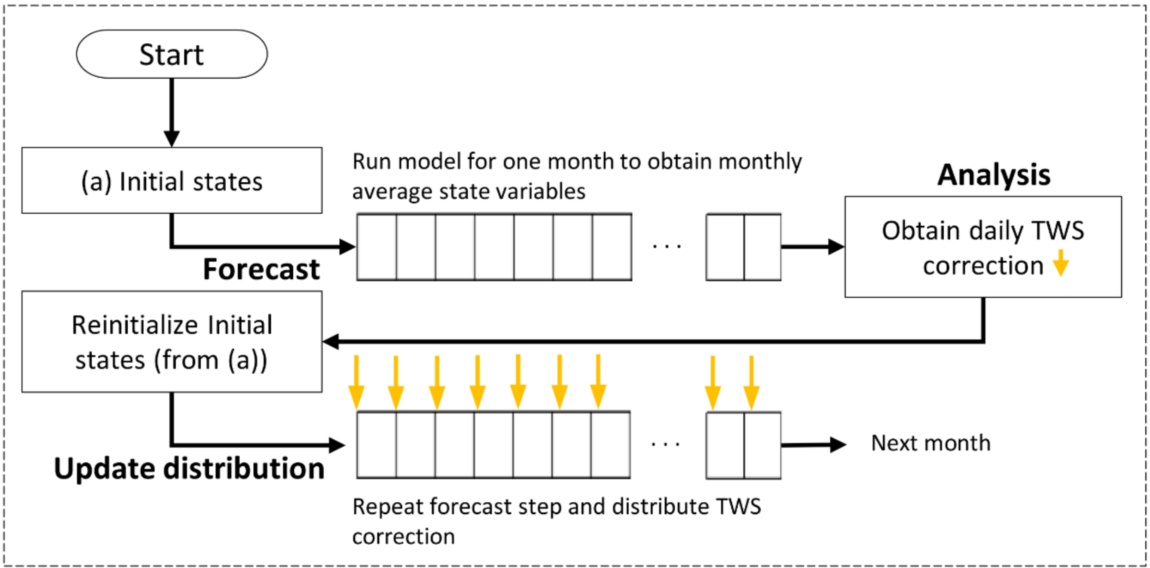

Fig. 2. Data processing diagram of GRACE data assimilation

The formulation of EnKS is as follows. Firstly, the forecast step performs the ensemble model run without data assimilation:

$$x_t^i = \mathcal{F}\left(x_{t-1}^i, f_t^i, \alpha^i\right), \tag{1}$$

where $\boldsymbol{\mathcal{F}}$ is the model operator used to propagate the model states from $t-1$ to t, $\boldsymbol{x}$ is the model state vector (containing SMS and GWS), $\boldsymbol{\alpha}$ represents the model parameters, and $i = 1,2,3,\dots,N$ denotes the index of ensemble member. Note that the ensemble member is generated by adding random noises to the nominal values. The forecast step is run for one month to obtain the monthly average state variables consistent with GRACE temporal resolution. Secondly, the analysis step computes the updated state vector ($\hat{\mathbf{x}}_T^{i}$) as follows:

$$\hat{\mathbf{x}}_T^i = \mathbf{x}_T^i + \mathbf{K}\left(\mathbf{y}_T^i - \mathbf{H}\mathbf{x}_T^i\right) = \mathbf{x}_T^i + \Delta\mathbf{x}_T^i, \tag{2}$$

where the subscript T indicates monthly average, $\Delta\mathbf{x}_T^{i}$ denotes the monthly correction, $\mathbf{y}_T^{i}$ is the observation vector containing GRACE monthly TWS data, $\mathbf{K}$ is the Kalman gain matrix, and $\mathbf{H}$ is the measurement operator relating model state vector to observation. After obtaining the monthly correction $\Delta\mathbf{x}_T^{i}$, the daily correction is computed by dividing $\Delta\mathbf{x}_T^{i}$ by the number of days in that month. Finally, the update distribution step is carried out by repeating the forecast step but with the daily correction applied to the initial state at the beginning of each day.

3.2 Evaluation metrics

The agreement between the estimated variable and the in situ data is assessed using the Pearson correlation coefficient (*R*) calculated as:

$$R = E[(\boldsymbol{p}-\overline{\boldsymbol{p}})(\boldsymbol{q}-\overline{\boldsymbol{q}})]/(\sigma_{\boldsymbol{p}}\sigma_q), \tag{3}$$

Where the $\boldsymbol{p}$ vector contains the model estimates, the $\boldsymbol{q}$ vector contains the validation data (e.g., in situ measurements), $E[\cdot]$ is the expectation operator, and $(\overline{\boldsymbol{p}},\overline{\boldsymbol{q}})$ and $(\sigma_{\boldsymbol{p}},\sigma_q)$ are the mean and standard derivations of $\boldsymbol{p}$ and $\boldsymbol{q}$, respectively.

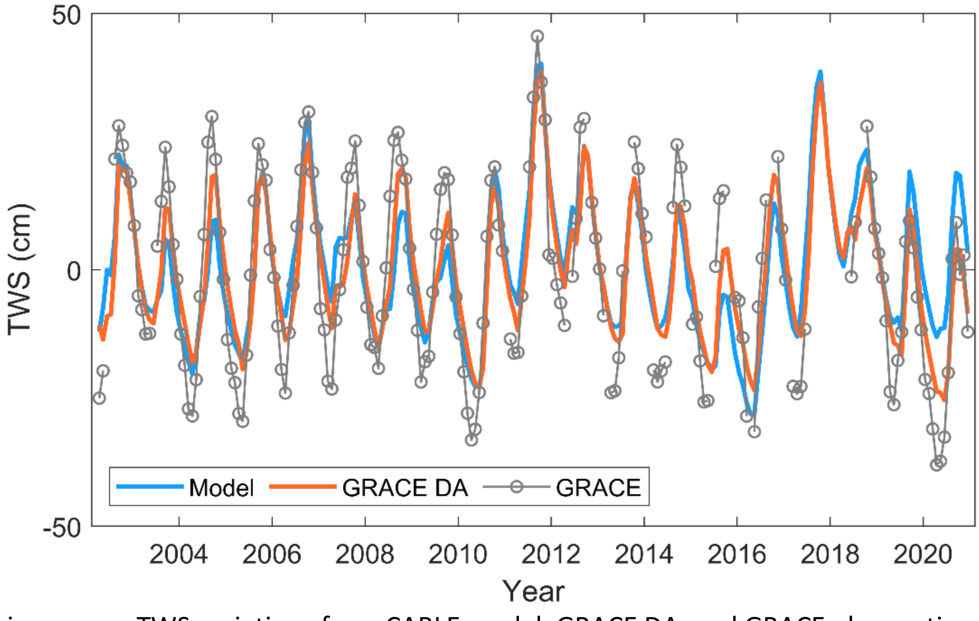

Fig. 3. Basin average TWS variations from CABLE model, GRACE DA, and GRACE observations between 2002 and 2020.

Table 1: Statistical results (correlation and RMSE) calculated from the model and GRACE DA estimates with respect to GRACE data. The long-term trend of all datasets is also shown.

	Model	GRACE DA	GRACE
Correlation	0.82	0.91	-
RMSE (cm)	10.46	8.45	-
Long-term trend (cm/year)	0.13 ± 0.06	-0.13 ± 0.01	-0.49 ± 0.06

4 RESULTS AND DISCUSSIONS

4.1 Impact of GRACE DA on TWS simulations

The impact of GRACE DA can be apparently seen from the TWS estimates (Fig. 3). To begin with, the CABLE model performs particularly well in northern Thailand, where model TWS estimates show similar TWS variations to GRACE with a slightly underestimated annual amplitude. The Spearman correlation coefficient of 0.8 confirms this significant agreement (Table 1). Implementing GRACE DA brings TWS estimates closer to GRACE observations, improving correlation and root mean square error (RMSE) by 13% (from 0.8 to 0.9) and 24% (from 10.5 cm to 8.5 cm), respectively. It is also seen that GRACE DA provides continuous TWS estimates even when satellite observations are not available. In addition, GRACE DA provides a trend estimate closer to GRACE observations and informs a decreased water storage in the study area by 0.13 cm/year. GRACE DA also improves the model's ability to capture the drought signature in 2019 – 2020. The model, on the other hand, shows increased water storage and underestimates the drought feature. Using such incorrect information (from the model) may result in faulty or invalid water resource assessments.

4.2 Accuracy assessment of groundwater estimates

The evaluation is carried out by comparing the correlation (R) of GRACE DA GWS results (with respect to in situ data) with the correlation value of the model estimates. Specifically, the difference is computed by $R_{\text{GRACE DA}} - R_{\text{model}}$, where positive and negative values reflect GRACE DA improving and degrading GWS estimates, respectively. Figure 4 shows that GRACE DA provides a significant improvement in GWS estimations, with higher correlation values of up to 0.53, or by 0.3 on average (Fig. 4a). It is of particular note that this improvement is substantial in light of the fact that the model estimate is already quite accurate. The positive impact of GRACE DA on GWS is consistent with the most of GRACE DA research (e.g., Li et al., 2019; Yin et al., 2020). The 2019 – 2020 drought characteristic is found in both GRACE DA GWS estimations and in situ groundwater level (Fig. 4b), similar to TWS (Fig. 3), indicating the sensitivity of the groundwater component to droughts. Model simulations are unable to depict the declining GWS.

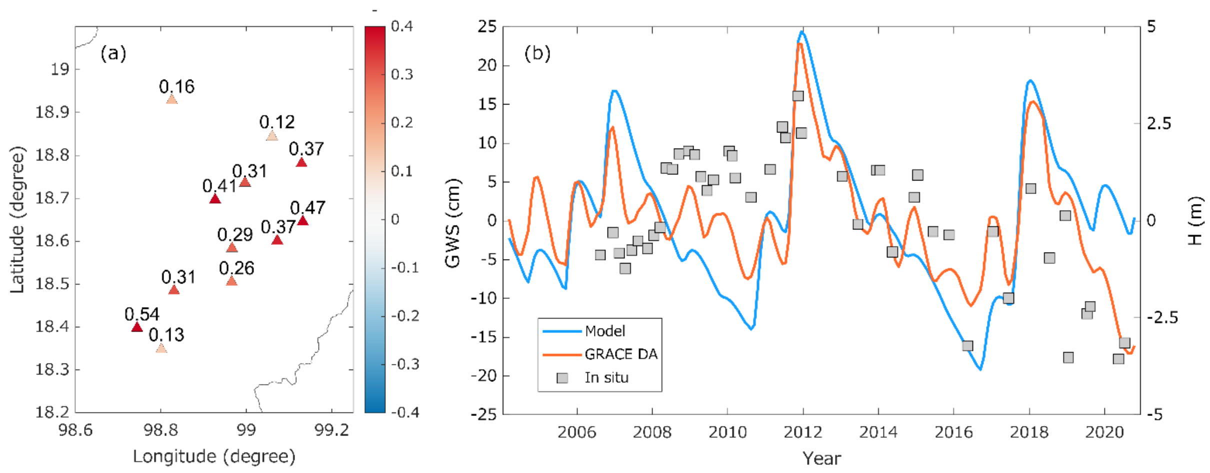

Fig. 4. (a) Differences in correlation between GRACE DA and model estimations (DA minus model). Positive (red) and negative (blue) readings imply that GRACE DA improves or degrades GWS estimates, respectively. The given values are the correlation differences at the measurement sites. (b) The average GWS estimates from model and GRACE DA between 2004 and 2020. The time series from all sites are used in the averaging. The average in situ groundwater level (H) is also shown.

4.3 Future development for improved hydrology analyses in Thailand

The development from this study can be applied to any area in Thailand, given that all data are globally and publicly available. Because the fundamental concept of EnKS remains unchanged, the DA approach described in Sect. 3 is also applicable with other satellite data such as the Soil Moisture Active Passive (SMAP; https://smap.jpl.nasa.gov) and surface reflectance products from the Moderate Resolution Imaging Spectroradiometer (MODIS; e.g., https://modis.gsfc.nasa.gov). This paper demonstrates the success of univariate DA, in which only one or a few state variables are updated with a single satellite data set. A multivariate DA could be used in a future study to incorporate multiple satellite data sets and improve a suite of hydrologic variables simultaneously (e.g., Tangdamrongsub et al., 2020). The GWS (and other hydrologic variables) can also be developed at ultra-resolution (i.e., near real-time 1 km or higher) over entire Thailand. Despite the potential utility of data collection across interdisciplinary sectors, the development is still in its early stages. It will necessitate substantial resources and financial support to maintain a public archive for reanalysis and operation purposes.

5 SUMMARIES

The success of using satellite data assimilation to enhance the accuracy of regional groundwater storage in Thailand is demonstrated in this study. The method improves hydrologic variables at the very high spatiotemporal resolution, and spatially/temporally continuous fields can always be produced even where/when satellite observations are absent. These accurate groundwater products may improve the robustness of water resources and climate-related decision-making, including (but not limited to) agriculture and urban planning. Future work aims to improve product resolution and incorporate multiple satellite data into a multivariate data assimilation framework to produce a suite of accurate hydrologic and climate variables for interdisciplinary studies.

REFERENCES

(1) Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. and Michaelsen, J.: The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes, Sci. Data, 2(1), 1–21, doi:10.1038/sdata.2015.66, 2015.

(2) Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M. and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), J. Climate, 30(14), 5419–5454, doi:10.1175/JCLI-D-16-0758.1, 2017.

(3) Li, B., Rodell, M., Kumar, S., Beaudoing, H. K., Getirana, A., Zaitchik, B. F., Goncalves, L. G. de, Cossetin, C., Bhanja, S., Mukherjee, A., Tian, S., Tangdamrongsub, N., Long, D., Nanteza, J., Lee, J., Policelli, F., Goni, I. B., Daira, D., Bila, M., Lannoy, G. de, Mocko, D., Steele-Dunne, S. C., Save, H., and Bettadpur, S.: Global GRACE Data Assimilation for Groundwater and Drought Monitoring: Advances and Challenges, 55, 7564–7586, https://doi.org/10.1029/2018WR024618, 2019.

(4) Tangdamrongsub, N., Han, S.-C., Yeo, I.-Y., Dong, J., Steele-Dunne, S. C., Willgoose, G., and Walker, J. P.: Multivariate data assimilation of GRACE, SMOS, SMAP measurements for improved regional soil moisture and groundwater storage estimates, Advances in Water Resources, 135, 103477, https://doi.org/10.1016/j.advwatres.2019.103477, 2020.

(5) Tangdamrongsub, N., Jasinski, M. F., and Shellito, P. J.: Development and evaluation of 0.05° terrestrial water storage estimates using Community Atmosphere Biosphere Land Exchange (CABLE) land surface model and assimilation of GRACE data, 25, 4185–4208, https://doi.org/10.5194/hess-25-4185-2021, 2021.

(6) Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., and Watkins, M. M.: GRACE Measurements of Mass Variability in the Earth System, 305, 503–505, https://doi.org/10.1126/science.1099192, 2004.

(7) Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J. S., Edmunds, M., Konikow, L., Green, T. R., Chen, J., Taniguchi, M., Bierkens, M. F. P., MacDonald, A., Fan, Y., Maxwell, R. M., Yechieli, Y., Gurdak, J. J., Allen, D. M., Shamsudduha, M., Hiscock, K., Yeh, P. J.-F., Holman, I., and Treidel, H.: Ground water and climate change, Nature Clim Change, 3, 322–329, https://doi.org/10.1038/nclimate1744, 2013.

(8) Yin, W., Li, T., Zheng, W., Hu, L., Han, S.-C., Tangdamrongsub, N., Šprlák, M., and Huang, Z.: Improving regional groundwater storage estimates from GRACE and global hydrological models over Tasmania, Australia, Hydrogeol J, https://doi.org/10.1007/s10040-020-02157-3, 2020.