nature water

News & Views | Published: 12 February 2024

Hydrogeodesy

Learning to downscale satellite gravimetry data through artificial intelligence

Alexander Sun

Nature Water **2**, 110–112 (2024)

Self-supervised learning offers a promising way of downscaling the total water storage anomaly data from the Gravity Recovery and Climate Experiment (GRACE) satellites, contributing to a better understanding of the impact of natural climate variability and human activities at basin scales.

Total water storage (TWS), encompassing all forms of water above and below the surface, is a holistic measure of the terrestrial water cycle, particularly in relation to perturbations caused by hydroclimatic extremes such as floods and droughts. Over the past 20 years, the Gravity Recovery and Climate Experiment (GRACE) and its follow-on (GRACE-FO) satellite missions have presented a unique opportunity to track global TWS anomalies[1], enabling the detection of emerging trends in freshwater availability[2] and the identification and quantification of hydroclimatic extremes at regional scales[3,4]. However, the coarse spatial resolution (about 300 kilometres) and long latency (45–60 days) of GRACE(-FO) data have largely constrained the practical utility of these data in water resources management.

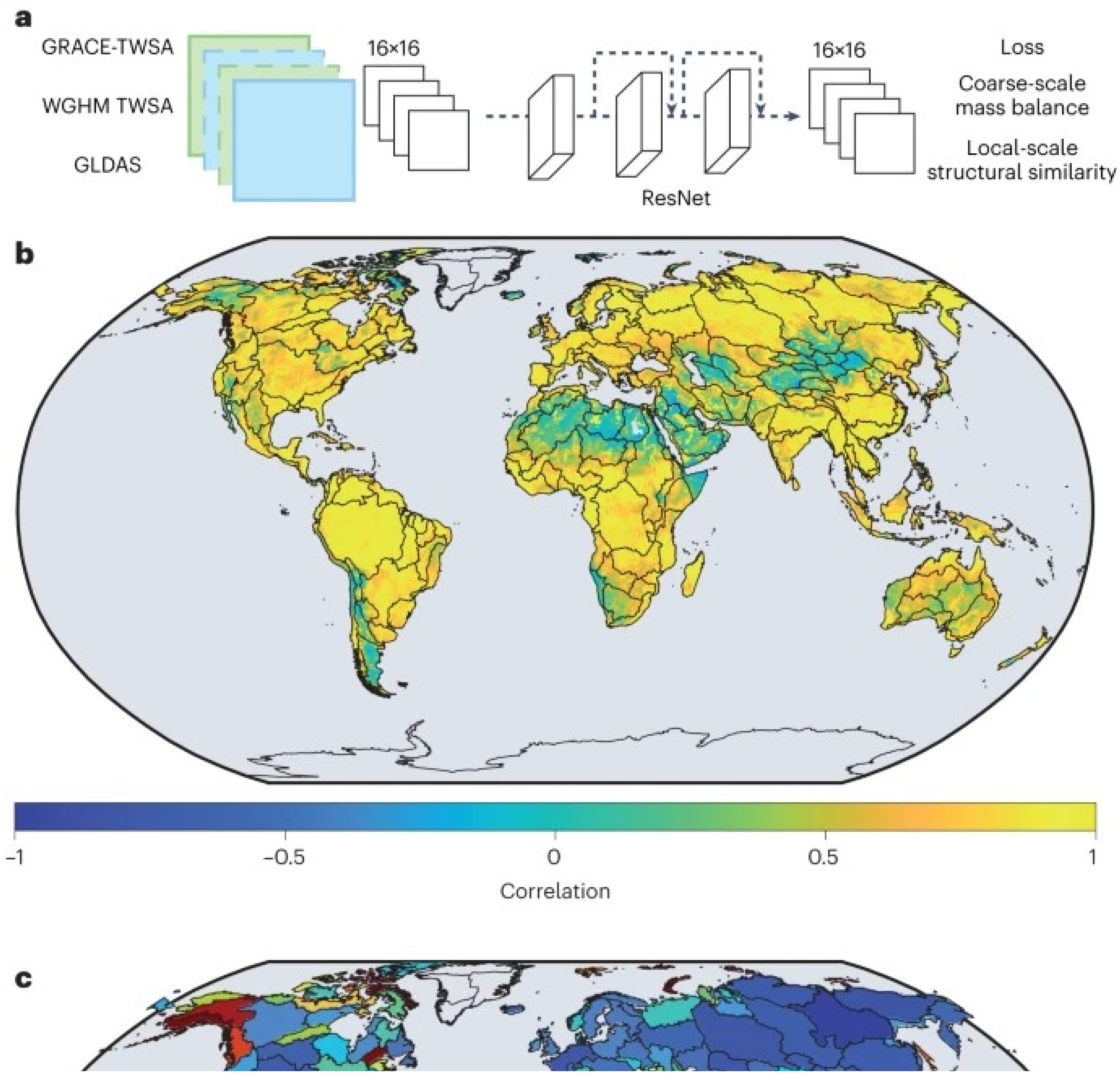
Spatial downscaling of TWS anomaly data would improve the applicability of satellite mass gravimetry. This goal has been pursued through dynamic and statistical downscaling, as well as data-driven machine learning, with the majority of studies so far focusing on regional applications[5]. In general, downscaling represents a type of inverse problem that aims to establish a mapping between low- and high-resolution data. The solutions are usually non-unique because of the inadequacy of high-resolution data to constrain the mapping. In the case of GRACE(-FO) TWS anomaly downscaling, in situ mass gravimetry measurements are essentially non-existent. Therefore, most previous studies have chosen to downscale a specific TWS component (such as groundwater) in regions with available in situ observations. Such a learning approach generally leads to region-specific models, with poor generalization and transferability to new study areas.

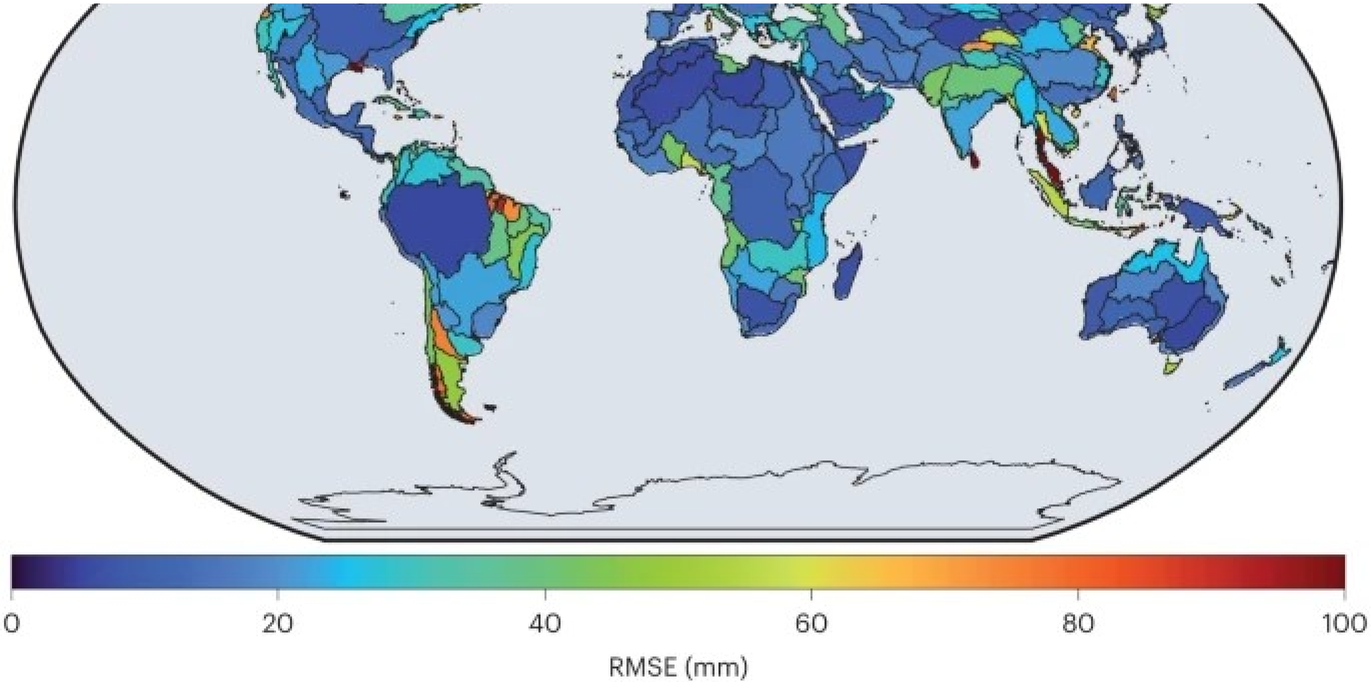
Now, writing in *Nature Water*, Junyang Gou and Benedikt Soja[6] present a hybrid deep learning model for downscaling GRACE(-FO) TWS anomaly data at the global scale by using outputs from the WaterGAP Hydrology Model (WGHM), a higher-resolution (0.5°) global hydrological model. Recently, hybrid methods that combine physics-based process modelling with machine learning have been developed[7]. Unique to this work is that the authors apply a self-supervised learning (SSL) technique to TWS anomaly downscaling.

SSL is an evolving machine learning training technique that reduces dependence on labelled data. In computer vision, labelled data refer to data annotated with context or class information. In time series regression, labelled data may be in situ observations of the target variable itself. SSL is directly responsible for the recent breakthrough of the large language models[8] and artificial intelligence global weather-forecasting models that have demonstrated superior performance over numerical weather models[9,10]. The general philosophy of SSL is to design prediction tasks (also known as pretext tasks) from the data itself without the need for paired annotations. This is particularly important because the amount of unlabelled data is generally much greater than the amount of labelled data. With well designed pretext tasks, a large amount of data from diverse sources and domains, and access to high-end computing platforms, models with billions of parameters can be trained to adapt to many different downstream tasks, possibly in never-before-seen settings. Models trained in this way are known as artificial intelligence foundation models[11].

Gou and Soja[6] designed a variant of a frequently used encoding–decoding machine learning task to learn to project the input data (consisting of GRACE and WGHM TWS anomalies, and outputs from the Global Land Data Assimilation System, GLDAS) into a low-dimensional latent space, to extract and fuse the patterns from the multi-scale data, and finally, to generate the downscaled TWS anomalies. To this end, they devised an innovative loss function to maximize similarity with high-resolution patterns simulated by the WGHM while keeping mass conservation at the basin scale to be consistent with the GRACE data (Fig. 1a). The underlying assumption is that the WGHM simulates the spatial-variation patterns of TWS anomalies with reasonable fidelity, despite its known biases in capturing basin-scale, long-term trends[12]. The resulting TWS anomalies product covers 2002–2019.

Fig. 1: Machine learning model architecture and results of the GRACE data downscaling.





a, Design of the artificial intelligence/machine learning total water storage anomaly (TWSA) downscaling framework used by Gou and Soja[6]. The global grid is divided into 16 × 16 patches and trained using a residual network (ResNet). The final result is obtained by stitching all patches together. GRACE, Gravity Recovery and Climate Experiment; WGHM, WaterGAP Hydrology Model; GLDAS, Global Land Data Assimilation System. **b**, Correlation with the WGHM. **c**, Basin-average root-mean-square error (RMSE).

Gou and Soja[6] show that the overall Pearson's correlation between the downscaled product and the WGHM is high, with a median value of 0.80 (Fig. 1b). The root-mean-square errors of the downscaled product are lower than 30 millimetres in most of the land areas, resulting in a global average root-mean-square error of 21.9 millimetres, a 56%

improvement over the WGHM (Fig. 1c). Over 160 global basins with a drainage area greater than 200,000 square kilometres, the downscaled product is shown to capture the long-term trend of TWS anomalies well, reaching a correlation of 0.94 with GRACE, versus 0.47 by WGHM. Correlation on the annual (and semi-annual) amplitudes is 0.97 (0.95), which is also higher than those obtained by the WGHM: 0.83 (0.83). The results are promising, suggesting that the SSL-based training is capable of assimilating information from multi-scale data sources. The authors further demonstrate the merit of the downscaled product through two downstream tasks: drought monitoring and flood potential prediction.

The WGHM does not fully resolve the impact of human intervention in some basins (for example, the Yangtze river basin in China), which affects the quality of the downscaled product. Although not considered in this work, in situ data related to river stage, groundwater, snowpack and soil moisture could also be incorporated to fine-tune the pre-trained model, further improving the resolution and fidelity of the downscaled TWS anomalies product. To be truly useful for operational use, the temporal resolution of the downscaled product must also be improved, and so must the product latency. Nevertheless, Gou and Soja’s work is timely, representing a meaningful first step towards training an artificial intelligence foundation model for hydrogeodesy, a burgeoning field encompassing the fusion of satellite mass gravimetry, global navigation satellite system, interferometric synthetic aperture radar and satellite altimetry data, for the continuous global monitoring of TWS anomalies. With increased accessibility to high-end computing infrastructures, we should soon be able to consume, process and potentially produce new content (such as downscaled TWS anomaly maps) from high-volume, multimodality Earth observation data.

References

1. Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F. & Watkins, M. M. *Science* **305**, 503–505 (2004).

2. Rodell, M. et al. *Nature* **557**, 651–659 (2018).

3. Rodell, M. & Li, B. *Nat. Water* **1**, 241–248 (2023).

4. Pokhrel, Y. et al. *Nat. Clim. Change* **11**, 226–233 (2021).

5. Miro, M. E. & Famiglietti, J. S. *Remote Sensing* **10**, 143 (2018).

6. Gou, J. & Soja, B. *Nat. Water* https://doi.org/10.1038/s44221-024-00194-w (2024).

7. Sun, A. Y. et al. *Water Resour. Res.* **55**, 1179–1195 (2019).

8. Brown, T. et al. *Adv. Neur. Inform. Process. Syst.* **33**, 1877–1901 (2020).

9. Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X. & Tian, Q. *Nature* **619**, 533–538 (2023).

10. Lam, R. et al. *Science* **382**, 1416–1421 (2023).

11. Bommasani, R. et al. Preprint at https://arxiv.org/abs/2108.07258 (2021).

12. Scanlon, B. R. et al. *Proc. Natl Acad. Sci.* **115**, 1080–1089 (2018).

Author information

Authors and Affiliations

Bureau of Economic Geology, The University of Texas at Austin, Austin, TX, USA

Alexander Sun

Corresponding author

Correspondence to Alexander Sun.

Ethics declarations

Competing interests

The author declares no competing interests.

Rights and permissions

About this article

Cite this article

Sun, A. Learning to downscale satellite gravimetry data through artificial intelligence. *Nat Water* **2**, 110–112 (2024). https://doi.org/10.1038/s44221-024-00199-5

Published
12 February 2024

Issue Date
February 2024

DOI
https://doi.org/10.1038/s44221-024-00199-5

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Provided by the Springer Nature SharedIt content-sharing initiative

Subjects

- Hydrology
- Natural hazards

Nature Water (*Nat Water*) | ISSN 2731-6084 (online)