

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	All the used raw data in this study are publicly available. The links to the data are listed in the Data availability section. No specific software was used.
Data analysis	The deep learning part of this study is conducted using Tensorflow V2.6.0. The data preprocessing and analysis were conducted using Matlab R2022a.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The downscaled TWSA product generated in this study is publicly available at <https://doi.org/10.3929/ethz-b-000648738>. The raw data used in this study are publicly available. The JPL GRACE mason product is available at <https://grace.jpl.nasa.gov/data/get-data/>. The WGHM TWSA simulations are available a <https://>

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	N.A.
Population characteristics	N.A.
Recruitment	N.A.
Ethics oversight	N.A.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	In this study, we proposed a self-supervised deep learning algorithm to assimilate total water storage anomalies (TWSAs) from satellite measurements and hydrological simulations. The achieved model has remarkable generalizability to provide accurate high-resolution TWSAs of 0.5 degrees globally. With the evaluation against other TWSA products and discussion about downstream applications, we proved the quality of our products and the significance of the proposed approach.
Research sample	We focus on the intersecting period of the GRACE mission and WGHM simulations (2002 to 2019) containing 180 months data in total. All the pixels on land are considered as the central pixels to generate a 16 degrees times 16 degrees patch around each central pixel.
Sampling strategy	We use all available data from 2002 to 2019
Data collection	The raw data used in this study are publicly available. The JPL GRACE mason product is available at https://grace.jpl.nasa.gov/data/get-data/ . The WGHM TWSA simulations are available at https://doi.org/10.1594/PANGAEA.918447 . The GLDAS-Noah products are available at https://disc.gsfc.nasa.gov/ . The HydroBasin data are available at https://www.hydrosheds.org/products/hydrobasins . The datasets and models generated and/or analysed during the current study are provided to the editor(s) and reviewers during the review process. The trained model and the generated downscaled TWSAs will be made publicly available upon publication.
Timing and spatial scale	The temporal resolution is approximately monthly, affected by some missing values of GRACE measurements. The spatial resolution is 0.5 degrees times 0.5 degrees.
Data exclusions	The data over Greenland were excluded due to the deficiency of hydrological models.
Reproducibility	The trained model and weights are provided to the editor(s) and reviewers during the review process and will be made publicly available upon publication.
Randomization	Multiple models were trained with different random initialisation to prove the robustness. Moreover, the uncertainties of the ensemble results are reported.
Blinding	Blinding is not relevant for our study as none of such experiments was involved.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a

Involvement in the study

☒

☐

Antibodies

☒

☐

Eukaryotic cell lines

☒

☐

Palaeontology and archaeology

☒

☐

Animals and other organisms

☒

☐

Clinical data

☒

☐

Dual use research of concern

Methods

n/a

Involvement in the study

☒

☐

ChIP-seq

☒

☐

Flow cytometry

☒

☐

MRI-based neuroimaging