nature portfolio

Corresponding author(s):   John Volk

Last updated by author(s):   Nov 27, 2023

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 (n) 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 *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	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. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	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 d, Pearson's r), 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	The in situ measured evapotranspiration data analysed during the current study were processed using the "flux-data-qaqc" Python package, version 0.1.6 (https://github.com/Open-ET/flux-data-qaqc).
Data analysis	The "flux-data-footprint" Python package was used to generate temporally dynamic flux footprints for sampling of of daily and monthly model ET data (https://github.com/Open-ET/flux-data-footprint). The Numpy (version 1.17.2) and statsmodels (version 0.12.1) Python packages were used for data analyses during the current study.

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 in situ measured evapotranspiration data analysed during the current study are available in the Zenodo repository, with identifier http://dx.doi.org/10.5281/

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	The current study did not involve human participants, their data, or their biological material.
Population characteristics	The current study did not involve human participants, their data, or their biological material.
Recruitment	The current study did not involve human participants, their data, or their biological material.
Ethics oversight	The current study did not involve human participants, their data, or their biological material.

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	The current study is focused on one-to-one comparisons between evapotranspiration data as modeled by remote-sensing methods and as measured on the ground; well known goodness-of-fit metrics were used to evaluate model data against measured data. Modeled and measured data were paired based on temporally overlapping records at multiple measurement stations. Accuracy metric results were grouped by different shared characteristics of the stations, using weighted averaging. Spatial long-term model data were mapped for individual models and differenced from the model ensemble average.
Research sample	The measured daily and monthly evapotranspiration data in the current study came from a public dataset that is available here: http://dx.doi.org/10.5281/zenodo.7636781. The measurements used were collected between 1995-2021. OpenET model data were generated and paired to the daily and monthly measurements, however model data was not available prior to 2001. The total number of stations with paired data was 152, and among them there were 16,444 days and 4,107 months of paired data.
Sampling strategy	As a conservative measure, we required a minimum of 3 months of paired data per measurement station to be included in weighted average accuracy metrics. In order to avoid skewing grouped average metrics, we weighted each station by the square root of the number of paired data.
Data collection	Measured ET data was previously curated and publicly archived on Zenodo (http://dx.doi.org/10.5281/zenodo.7636781). Model ET data was generated for the study using the current OpenET version.
Timing and spatial scale	Measured data was collected primarily from eddy covariance systems and post-processed to daily and monthly aggregated periods. Model ET was sampled at daily (the dates of satellite overpass) and monthly intervals at the measurement stations using pixel footprints which were determined by either long-term wind direction or from a physically based flux footprint prediction model. Flux footprints rarely exceeded the size of a 7x7 (30m resolution) grid. Gaps in measured data exist due to lapses in sensor operation or faulty data, and gaps exist in model data due to cloud coverage.
Data exclusions	No data available to us were excluded from the analyses during the current study.
Reproducibility	Data processing steps for both modeled and measured data used in the current study were made reproducible by our use of well-documented open source software. We developed Python code for eddy covariance data processing and footprint development. Similarly, the OpenET modeled data was generated using the operational models which are open source and their data are also publicly available through various mechanisms including the OpenET API and the Google Earth Engine Data Catalog. The statistical methods used in the study were limited to simple techniques, and the results were independently tested by multiple members of the OpenET group using different statistical packages such as Python and Microsoft Excel.
Randomization	Randomization of data into groups was not applicable as groups were defined by biophysical characteristics such as climate and land cover type.
Blinding	This study did not employ blinding of data due to the limited amount of high quality measured evapotranspiration data available to us. However, additional data that has not yet been compared against OpenET models was held out of this study for a future blind intercomparison and accuracy assessment of OpenET.

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	Involved in the study
☒ ☐	Antibodies
☒ ☐	Eukaryotic cell lines
☒ ☐	Palaeontology and archaeology
☒ ☐	Animals and other organisms
☒ ☐	Clinical data
☒ ☐	Dual use research of concern

Methods

n/a	Involved in the study
☒ ☐	ChIP-seq
☒ ☐	Flow cytometry
☒ ☐	MRI-based neuroimaging