nature portfolio

Corresponding author(s): Mukesh Kumar

Last updated by author(s): Mar 24, 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	No software was used to collect the data.
Data analysis	Data analyses were performed using Python (version 3.10.9), R (version 4.2.1), and MATLAB (version R2022a). Major libraries used in Python include Pandas, Numpy, Scipy, os, glob, Matplotlib, and Tensorflow (2.11.0). Major libraries used in R were dplyr, lubridate, tidyr, and ggplot2. MATLAB was used to run the empirical model.

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 datasets generated and/or analyzed during the current study and related scripts are available in the Zenodo repository at https://zenodo.org/record/7768479. Scripts also can be found at https://github.com/praghav444/Modeling_Stomatal_Conductance_for_ET_Partitioning.

Human research participants

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

Reporting on sex and gender	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

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	Using machine learning and eddy covariance flux tower data of 642 years, the study shows that structural constraints in both empirical and plant hydraulics-based models of stomatal conductance limit their effectiveness for predicting evapotranspiration and its components, i.e., evaporation and transpiration.
Research sample	Dataset used in this study were obtained from multiple sources. PLUMBER2 fluxnet data was obtained from http://doi.org/10.25914/5fdb0902607e1. Soil property data was obtained from http://globalchange.bnu.edu.cn/research/soil5d.jsp.
Sampling strategy	Fluxnet sites with data record of less than one year were excluded from the study. Also, sites with no measurements of soil moisture were not included in the study. Further only good quality data (based on multiple criterion outlined in previous scientific literature, which have been referenced in the manuscript) were used in the analysis.
Data collection	Data required for the study was obtained from online available datasets, which have been referenced in the manuscript.
Timing and spatial scale	Data was site-specific with each site having different length of data record. For temporal scale, the original resolution of Flux data (i.e., 30 minutes or 1 hour) was used. Sites were chosen to cover a wide range of biomes across the globe.
Data exclusions	Fluxnet sites with data record of less than one year were excluded from the study. Also, sites with no measurements of soil moisture were not included in the study. Further only good quality data (based on multiple criterion outlined in previous scientific literature, which have been referenced in the manuscript) were used in the analysis.
Reproducibility	The results of the study can be reproduced using the input data and codes that have been uploaded online at https://zenodo.org/record/7768479.
Randomization	Data was shuffled randomly for training and validation of the models.
Blinding	Blinding was not used during the analysis as the study results were obtained by the development of different models.

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