nature portfolio

Corresponding author(s):   Christopher R. Hakkenberg

Last updated by author(s):   Oct 15, 2024

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
Data analysis	All data analysis was done using custom software in the in the R language version 4.3.1 as well QGIS version 3.38.3. In addition to base functions, raster data was processed using the terra package. Generalized Linear Mixed Models were prototyped using the lme4, MASS, and nlme packages, and final models were produced using name. Maps were produced in QGIS and charts employed the ggplot2, ggpubr, and ggrepel packages.

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

All datasets used are publicly available. The GEDI L2 dataset is available from the LP DAAC (https://lpdaac.usgs.gov/products/gedi02_av002/, 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	We employed GEDI spaceborne lidar to consistently assess how pre-fire forest fuel structure affects wildfire severity levels across 42 California wildfires between 2019-2021. Inference into underlying relationships between fuels, topography, weather and severity was based on univariate and multivariate hierarchical spatial Generalized Linear Mixed Models (spGLMMs). Both univariate and multivariate spGLMMs were parameterized with random intercepts and random slopes, with random effects determined at the level of the wildfire.
Research sample	Model data consist of 830,709 samples centered on the estimated centroid location of all GEDI footprints. Samples consist of 45m diameter circular footprints within the bounds (90m) of all 42 fire perimeters. Within all samples, data on pre-fire structure derived from GEDI waveforms as well as spatially coincident data on topography, fire weather, climate and dNBR-based burn severity was extracted.
Sampling strategy	All data consist of 45m diameter circular samples centered on the estimated centroid location of each GEDI footprint. The 45m diameter extent of each sample was determined from the 25m diameter circular GEDI footprint, buffered outwards by 10m to account for potential geolocational error, estimated as ±10m (1 standard deviation). All samples were extracted from within the perimeter of all large (>2000 ha) wildfire in the study domain buffered inwards by 90m to minimize edge effects and geolocation imprecision
Data collection	No data were collected.
Timing and spatial scale	The study domain was defined as the North Coast, Central Coast, and Sierras regions of California (Fig. 1), for fires occurring during years of temporally-coincident GEDI observations; namely, 2019-2021
Data exclusions	Excluded data were those within our study domain that did not satisfy the following conditions: (1) possessing a minimum canopy height (RH100) of 5m; (2) having been acquired during leaf-on dates (corresponding to day of year 105 to 319); (3) occurring on slopes less than 25 degrees owing to challenges with ground-finding in steep terrain; (4) possessing footprint elevations <150m divergent from a TanDEM-X digital elevation model to prevent the erroneous inclusion of low-lying clouds; (5) having minimal positional degradation as indicated by the degrade flag; (6) possessing surface flag equal to 1 and the stale return flag equal to 0; and (7) having beam sensitivity >0.9 for the default ground finding algorithm and sensitivity of >0.95 for ground finding algorithm 2.
Reproducibility	All data are publicly available (See Data Availability statement). Code is psoted on the corresponding author's GitHub accpount (See Code Availability).
Randomization	GEDI footprint locations exhibit a non-random, linear, lattice-like spatial structure whose spatial dependence could violate assumptions of independence underlying our parametric modeling framework and potentially underestimate model prediction error. We therefore implemented a spatial blocking routine to remove autocorrelated GEDI shots while providing a framework for permutation-based inference.To do this, we first developed empirical semivariograms for canopy height (RH98) for each fire. RH98

was used for this assessment of spatial autocorrelation because tests have shown it to be the most accurate among our selected canopy structural attributes based on independent validation with a global data set of field sites and airborne lidar. We then applied a pairwise distance filter initiated from the semivariogram range so that only points beyond that nominal Euclidean distance were retained for a secondary Moran's I test. When significant (i.e. autocorrelated), we iterated larger blocking distances until remaining points produced a nonsignificant Moran's I. This process was repeated ten times (with replacement) for each wildfire, each initiated with a different random seed, to produce ten distinct GEDI sample combinations per fire, whereby all shots in each of the ten sample sets were sufficiently uncorrelated as determined from re-evaluated Moran's I values. All subsequent models were iterated over each of the ten sample sets to produce mean ensemble coefficient estimates (including intercepts, slopes, and standard errors) based on the resulting distributions.

Blinding | Blinding was not relevant to our study. We used all available data (i.e. GEDI footprints fulfilling the aforementioned criteria)

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