Open Research Awards 2026

Environment, Science and Economy disciplines prize

Enimhien Akhabue (Geography)

African plant trait–to–Plant Functional Type dataset for improving land-surface model representation

Introduction to the case study

We created and published an open dataset that systematically maps African plant species (from the TRY database) to Plant Functional Types (PFTs) consistent with the Joint UK Land Environment Simulator land-surface model (LSM). The resource addresses a key barrier to reproducible model evaluation and parameterisation in Africa: fragmented, inconsistent, and hard-to-reuse trait-to-PFT classifications. Our dataset can be used to integrate plant trait data into PFT parameterization in LSMs. By making the classification outputs and supporting documentation openly available alongside a peer-reviewed data descriptor, the work enables transparent reuse and extension by other researchers working on African ecosystems and global modelling.

Detail the practices you have used to support openness/reproducibility

• Open publication route: Published an open access peer-reviewed data descriptor so the workflow, rationale, technical validation, usage notes and dataset structure are transparently documented and citable.

• FAIR-by-design outputs: Released an open access and structured dataset (lookup table) linking species to JULES-aligned PFTs, with clear metadata, variable descriptions, and reproducible identifiers (species names, PFT labels, and supporting trait context). All the analysis code freely available on GitHub and Zenodo which are public repositories to enable others replicate, validate and use our dataset.

• Reproducible processing narrative: Documented the classification criteria and decision rules used to move from raw records to harmonised PFT assignments (e.g., handling synonyms/ambiguous records and ensuring consistent PFT definitions).

• Reusability for others: Designed the dataset for plug-and-play use in model evaluation and parameterisation (e.g., enabling community comparison of “what PFT a species belongs to” under consistent rules).

• Attribution and traceability: Provided citation guidance so downstream users can properly acknowledge the dataset and underlying sources.

Outline any barriers or challenges you encountered, and how they were handled

A major challenge was assembling the classification parameters needed to assign species to PFTs. The classification required up to four attributes for some species (growth form, leaf type, phenology, photosynthetic pathway, and climate zone), but these were rarely available from a single source. As a result, parameters often had to be compiled and validated across multiple databases and botanical references while resolving inconsistencies in species names and taxonomy. This was particularly demanding because the workflow involved thousands of species, and some had to be excluded when insufficient information was available. I addressed this by systematically cross-checking sources and documenting decision rules to ensure consistent and defensible classifications.

What benefits have you identified from your case?

The dataset improves reproducibility (shared, consistent mappings), reusability (others can directly use the lookup for model parameterisation/evaluation), and transparency (decision rules are documented and citable). It also supports better representation of African ecosystems and, by extension, global ecosystem modelling by making trait-based PFT work easier to replicate, compare, and extend thereby reducing duplicated effort and accelerating community progress. Early uptake has already been encouraging as several researchers have contacted me about using the dataset, and the paper has received over 1,000 accesses within two months according to Springer Nature Insights.

How do you plan to continue this case moving forward?

I plan to maintain the dataset as a versioned community resource by expanding plant species and trait coverage, incorporating feedback from users, and linking updates to open repositories (e.g., Zenodo/GitHub) so changes remain transparent and citable. I also intend to reuse the dataset in my PhD modelling work where I am optimizing LSMs by reparameterization and evaluation across multiple African ecosystems and to share the workflow with peers through publications, presentations and short training-style documentation to encourage uptake and reproducible reuse.

Conclusions

This case study demonstrates how publishing an open, well-documented dataset can make trait-enabled land-surface modelling for Africa more reproducible, transparent,

and reusable. It turns a common “hidden” modelling step (PFT assignment) into a citable, shareable research output.

Links to any materials/resources relevant to the case

• Data descriptor paper: https://www.nature.com/articles/s41597-026-06728-z

• Zenodo dataset DOI link: https://doi.org/10.5281/zenodo.16533069

• GitHub repository (code/workflow and dataset): https://github.com/TESS-Laboratory/Akhabue_PFT_Classification/tree/v1.4.1