

Thank you for giving me the opportunity to review this dataset. Especially due to its replication after 50 years it is a valuable and unique dataset that is worth to be published. However, I think that the dataset could be presented a bit more on point in the text. My review is split in two parts, one addressing the presented paper and the other one addressing the dataset itself.

PAPER:

General comments

The paper is well written and structured. However, it has a large focus on comparing the old and new data, whereas, the dataset itself consists of the newly sampled values. While the comparison of both datasets is definitely valuable, as a potential dataset user, I would appreciate more information on the dataset itself, including visualizations specifically describing the new dataset. Given that this is a data description paper, I would recommend shifting the focus more strongly toward the dataset itself, similar to the ESSD paper of Wood and Bunce (2016).

Furthermore, it only became clear to me at the very end of the paper, in the data availability statement, that the reference Wood and Bunce (2016) in the text is not only describing the study methods and the 1971 situation, but that it actually includes the dataset of the 1971 study. This should be clarified much earlier in the manuscript.

Specific comments

- Line 27 – Clicking the link of the dataset leads to a wrong URL
- Figure 1: Including site names or IDs (as used in the dataset and text) would improve clarity. Adding latitude and longitude axes would also help locate the sites
- Line 154 – The abbreviation SSSI is not defined
- Paragraph 2 – Data Collected
 - A table listing all plot coordinates, along with information on how precisely the original plots were relocated, would be helpful.
 - A figure illustrating the nested design and quadrats could improve understanding of the study design
- Line 225 – The soil moisture was determined, but it is not part of the dataset, is that correct? Please clarify why.
- Lines 255f – The text states: “willows (*Salix* spp.) and rowan (*Sorbus aucuparia*) have declined markedly across all sites, willows by a mean of 6 stems and rowan by a mean of 14 stems” – in Table S1, the mean for willows is -13.3 – please clarify this discrepancy

- Figure 3 – A logarithmic scale on the y-axis might help to visualize species with stems <250
- Table S2 – *Picea albies* should be corrected to *Picea abies*
- Line 296 - UK tree age calculator (*Felling UK, 2025*) – The link in the references is not accessible and the website says “felling.uk is for sale”. Having a look at another calculator (<https://virginiatreecare.com/tree-age-calculator/>), *Pinus sylvestris* of 20 cm dbh is said to be 38 to 60 years depending on growing conditions. As also that is only an estimate, the authors could give an age range here instead of a single number
- Figure 7 - A logarithmic scale on the y-axis might help to visualize species species with counts <40
- Line 347 – Please clarify whether the increase in wetness refers to a trend over time or a direct comparison between the years 1971 and 2022.
- Line 385 – To what extent could the change also be a result of different instruments used in the 1970’s and today?
- 397ff – This topic is quite isolated here. Additional context on Loss on Ignition (LOI), even if not directly comparable to the original survey, would be helpful.

DATASET:

General comments

The dataset is presented in CSV files, which facilitates the reuse, and organized in thematically structured files. The dataset is accompanied by useful metadata in the file `data_documentation_pinewoods_2023_eip.docx`.

However, several aspects of the data presentation and metadata could be improved to further enhance clarity, usability, and consistency.

Downloading “Supporting Documents” and “Download” (=all data) currently produces ZIP files with the same name but different contents. Would it be possible to assign distinct file names to these downloads? This would help differentiate them and avoid confusion.

The data are provided as CSV files, while the column description is given in as separate word-Table. It might be better to provide an additional CSV for each data-table in which the columns are described. This would make it easier to work with the data in scripting environments.

The naming of “Site ID” or “Site Number” should be consistent between all tables

Sampling dates are provided in a supporting document but should also be included directly in at least the file `scots_pine_2022_sites.csv`

Specific comments

Ground Flora data:

- The description for NEST_1_COV states Null = not present. How should empty cells be interpreted?
- Are BRC numbers searchable in an external reference? Some values (e.g., bare ground) are not listed in the field book tables.

Site Info data:

- Abbreviations in the comments should be written out (e.g., “Gld.3–12m mxd.”, “Std.dead <10cm”).
- What does an aspect value of –154 degrees represent for Site 7 (Glen Feshie), Plot 15?

Sites data

- Providing coordinates in WGS84 decimal degrees would improve usability.

Soil Data

- LOI values are reported with nine decimal places. Please confirm whether this level of precision is meaningful; if not, rounding would be appropriate.

Tree Data

- Some trees have IDs while others do not—please clarify why.
- Please clarify the reason for the different formats of ID, eg *A* vs *606b4232-43da-4550-8a93-1811b1fbc7b*
- Tree IDs do not appear to be unique (e.g., multiple “A” entries across sites). Please confirm whether this is intended.