

# Metadata for the Peatland Code and Woodland Carbon Code biodiversity reference database

## **Background:**

The Peatland Code and Woodland Carbon Code are exploring the use of the Wallacea Trust method for measuring biodiversity change in peatland restoration and woodland creation sites in the UK. Once quantified, this biodiversity benefit could be incorporated into the carbon credit, potentially generating greater private finance to fund further habitat restoration. The Wallacea Trust method depends on the use of reference site (a near-natural or previously restored site) that sets the benchmark for biodiversity measurements. However, conducting biodiversity surveys on both the project site and a reference is likely to be prohibitively costly. A main outcome of this project was to aggregate available biodiversity data for peatlands and woodlands in the UK that are in relatively intact condition. This database could then be used to identify appropriate reference data for projects. This removes the need to pay for a reference site to be surveyed, especially where comparable reference sites are not available in vicinity of the project site.

## **Method:**

A comprehensive search of the published literature was carried out using the search terms included in Appendix A, as well as a comprehensive search for publicly available databases (including searching the NBN Atlas for full datasets). In a previous phase of the project, suitable taxonomic groups for measuring peatland restoration and woodland creation success were identified through expert opinion. Thus, we limited our search to the target taxa of vegetation, arthropods, birds, as well as bats for woodland. Any articles where data collection did not occur in the UK were removed. We did not include data from meta-analyses to avoid duplication. For datasets to be included as suitable, there needed to be a reproducible sampling protocol, using methods that are widely used for the taxonomic group of interest (eg. quadrat surveys for vegetation, pitfall traps for beetles etc.). Data where no location was given were excluded, although we kept data where information on approximate

location was provided. We included both abundance and occurrence (i.e. presence/absence) datasets.

We then filtered the data to reference sites that were in good condition. Because of the lack of UK-wide information on peatland and woodland condition, we used a variety of approaches to filter reference sites. Woodland sites in Scotland were filtered to good condition using the Native Woodland Survey condition score of 3 and above (Forestry Commission, n.d.). For sites in England and Wales, no information on woodland condition was available in publicly available datasets or GIS layers (we ruled out the use of SSSI status since assessments are rarely up-to-date). Studies have found that woodland structural diversity, as measured by the standard deviation of tree diameter at breast height (DBH), is closely linked to successful establishments of woodland plants (Waddel et al, 2024) and the diversity of other species groups (Wagenaar et al., 2025). Thus, for datasets where tree DBH was measured alongside biodiversity data, we calculated standard deviation of DBH and used this to filter out sites with low structural diversity, including only sites within the top quartile of structural diversity. For peatlands, there are no reliable datasets showing peatland condition across the UK. Thus, we gathered expert opinion on the condition of peatlands within our dataset, and removed anything that was deemed by the expert to be in poor condition.

Missing data is identified with 'NA'

X: Row number

Data\_code: Identifies all data that came from the same dataset

SiteID: Site name or identifier where the data was collected.

Species: Scientific name of species following the binomial nomenclature (Genus species). In some cases species are aggregated into a group ('agg'). Species that are only identified to Genus level are named by their genus followed by 'sp' or 'spp'.

Order: Order of species

**Date:** Date that the data was collected. In many cases this is the year only. However, more information on collection dates can often be found by following the links available in the 'Further.method.information' column.

**Latitude:** Coordinate. In some cases this is an approximate coordinate

**Longitude:** Coordinate. In some cases this is an approximate coordinate

**Abundance:** The total number of individual species

**Taxon:** Common taxonomic name enabling a broad search for a specific taxonomic group (eg. birds, invertebrates etc)

**Abundance\_occurrence:** Identifies if abundance ('AB'), presence ('OC') or frequency ('FR') data was collected.

**Method:** Identifies the broad method used for surveys – further information can often be found using the links in the 'Further.method.information' or 'Citation' columns.

**Site\_area:** Total area of the site in hectares. For woodlands, this is the size of the whole woodland, if known. For peatlands, this is often the size of the Nature reserve and may be surrounded by peatland areas outside of the boundary of varying quality.

**Habitat\_type:** Identifies the peatland or woodland habitat type in more detail (eg. Native pinewood, Raised bog)

**Licence:** Licence of the original dataset

**Citation:** Citation of the original dataset

**Further.method.information:** Column with links or names of documents where more detailed information can be found relating to survey methods used

Survey\_area: Total area surveyed if known. Where this is NA, more information may be available via links in 'Further.method.information'.

Plot\_no: Some sites were divided into plots for surveys. This designates the plot number where this is the case

Time\_since\_habitat\_creation: number of years since woodland was planted or peatland was created. Where woodland age is not known but has been identified as mature using the Native woodland survey of Scotland, this is identified as 'Mature'. For ancient woodland or peatland this is designated as 'ancient', 'near-natural' or '250' years.

Distance\_nearest\_woodland: Only applicable for woodland data. Number of meters between the surveyed woodland and the nearest native woodland.

Proportion\_woodland\_3000m: Only applicable for woodland data. Proportion of woodland within 3000m of the survey area. Where this is NA, this information isn't available with the original dataset but can be obtained using habitat maps.

Former\_land\_use: Where this is known, land use before woodland creation or peatland restoration.

Visit\_no: where multiple visits occurred, this is the visit number

Sample\_no: This identifies the sample number that the data was collected from (eg. the quadrat number, pitfall trap, or transect for example).

Source: link to the original source of the dataset

Further\_license\_info: Any further information relating to the data license

Unique\_site\_idenfifer: Uniquely identifies each site that the data was collected from

Broad\_habitat: Broad habitat type, either 'woodland' or 'peatland'