

# Dataset Documentation

## Freshwater river macrophyte surveys (Biosys)

17/06/2022

This document will help you to understand and use the Freshwater river macrophyte survey dataset held in Biosys

### Dataset description

This is for Approval for Access product Afa307.

Biosys is a database for storing, manipulating and reporting data from freshwater and marine biological surveys at any taxonomic level. These samples are collected and analysed by the Environment Agency and by third parties.

All the relevant data associated with Biosys freshwater macrophyte surveys, such as site information, sample information and analysis information can be obtained from the [Ecology and Fish data explorer](#). The data can be accessed using the interactive map to produce small downloadable files or they can be downloaded in bulk as a set of relational CSVs that can be linked via unique codes in each file.

The data downloads are split into three different files within the Ecology and Fish data explorer: Sites, metrics and taxa. You will see them labelled in the downloads as follows:

- MACP\_OPEN\_DATA\_SITE (Contains freshwater macrophyte site information)
- MACP\_OPEN\_DATA\_METRICS (Contains freshwater macrophyte sample and metric information)
- MACP\_OPEN\_DATA\_TAXA (Contains freshwater macrophyte analysis and taxa information)

If you are bulk downloading the full dataset there is also a fourth file available which contains additional taxonomic information for all of the taxa listed in our database:

- OPEN\_DATA\_TAXON\_INFO (Contains taxonomic information for all taxon names available in Biosys and is also referred to as Biosys Taxon Info)

The tables can be related to each other using the joins below:

- MACP\_OPEN\_DATA\_SITE.SITE\_ID = MACP\_OPEN\_DATA\_METRICS.SITE\_ID
- MACP\_OPEN\_DATA\_METRICS.ANALYSIS\_ID = MACP\_OPEN\_DATA\_TAXA.ANALYSIS\_ID
- OPEN\_DATA\_TAXON\_INFO.TAXON\_LIST\_ITEM\_KEY = MACP\_OPEN\_DATA\_TAXA.TAXON\_LIST\_ITEM\_KEY

Please refer to the dataset content table, at the end of this document, for an explanation of the fields/columns in each file.

### API

dataset can also be accessed using an API (Application Programming Interface): [Ecology and Fish Data API](#). For more information regarding the service please visit the dedicated Frequently Asked Questions section of the Defra Data Services Portal (DSP) Forum: [Ecology and Fish Data FAQs](#).

## Update frequency

This dataset is extracted monthly on the first Sunday of each month. It can then take a week or so for the data to appear on the explorer. Data entered onto Biosys immediately prior to extraction may not have been quality assured and so may not reflect the final version of the data. You can use SAMPLE\_DATE in this dataset to identify recent data.

## Related datasets

Sites in this dataset may also be listed in the Freshwater river macroinvertebrate (Biosys) or Freshwater river diatom (Biosys) dataset.

## Common questions & known issues

### Known issues:

#### Reporting Area and Agency Area names:

The Reporting Area field was added to the Explorer as part of the 2022 update. The Reporting Area identifies the team who report on that data. It is more closely aligned to the [Environment Agency public face Area names](#) than the older, Agency Area field. The table below shows the comparison between all the fields.

Please be aware that the geographical boundaries for the old Southern and Thames Agency Areas do not match the geographical boundaries for the current Areas replacing them. Therefore, a small proportion of sites located in the old Area may not be located in the matched new Area:

| Agency Area (Old EA Region and Area Name) | Reporting Area                             | EA Public Face Area Name                   |
|---|--|--|
| Anglian - Central                         | East Anglia - West                         | East Anglia                                |
| Anglian - Eastern                         | East Anglia - East                         | East Anglia                                |
| Anglian - Northern                        | Lincolnshire and Northamptonshire          | Lincolnshire and Northamptonshire          |
| North West - North                        | Cumbria and Lancashire                     | Cumbria and Lancashire                     |
| North West - South                        | Greater Manchester Merseyside and Cheshire | Greater Manchester Merseyside and Cheshire |
| South West - Devon                        | Devon and Cornwall                         | Devon and Cornwall                         |
| South West - Cornwall                     | Devon and Cornwall                         | Devon and Cornwall                         |
| South West - North Wessex                 | Wessex                                     | Wessex                                     |
| South West - South Wessex                 | Wessex                                     | Wessex                                     |
| Midlands - Central                        | West Midlands                              | West Midlands                              |
| Midlands - East                           | East Midlands                              | East Midlands                              |
| Midlands - West                           | West Midlands                              | West Midlands                              |
| North East - Yorkshire                    | Yorkshire                                  | Yorkshire                                  |

|                                 |                                   |                                   |
|---------------------------------|-----------------------------------|-----------------------------------|
| North East - North East         | North East                        | North East                        |
| Thames - North East             | Hertfordshire and North London    | Hertfordshire and North London    |
| Thames - South East             | Kent South London and East Sussex | Kent South London and East Sussex |
| Thames - West                   | Thames                            | Thames                            |
| Southern - Kent & E. Sussex     | Kent South London and East Sussex | Kent South London and East Sussex |
| Southern - Solent & South Downs | Solent and South Downs            | Solent and South Downs            |

### WFD Waterbody ID:

Care should be taken when using this field to search for data. The WFD Waterbody ID field is populated on Biosys with the Water Framework Directive waterbody ID that the site *sits* in. This may be different to the Waterbody that the site is used to produce a WFD status assessment for. This field is often null for historic sites and may also be null where site sits in a waterbody that is too small to have a WFD ID. The ID may also relate to the WFD cycle 1 set of waterbodies or cycle 2 waterbody IDs. If you are interested in a particular waterbody ID then it is also worth using this field in conjunction with a search on the interface map and/or an extract of the data using the Waterbody name.

The [catchment data explorer](#) can also be used to give further high level information on a particular waterbody and could be used to compliment/support interpretation of this dataset.

## Common questions:

### What data are included in this dataset?

All freshwater (river) macrophyte samples that have been collected using our standard sampling methods are included in these datasets. The standard sampling methods are:

- MACROPHYTE: Standard method for macrophytes (WFD compliant)
- MACROPHYTE: ECN Survey

Please note that we do not include pollution sample data or protected species where releasing details of their location could have a detrimental effect e.g. freshwater pearl mussels.

Site name is excluded from this dataset. This is because it is a free-text field and may contain information or data that cannot be released under an Open Government Licence.

The sample reason field in this dataset displays a generic primary sample reason such as 'National monitoring' or 'Local investigative monitoring'. It does not include detail on all of the reasons for which the sample has been collected.

### Can I use this data to run through LEAFPACS2 (River Macrophyte classification tool) to produce a Water Framework Directive (WFD) classification?

Yes, you can get the data required by LEAFPACS2, from these datasets where the data has been collected. For additional information on running a LEAFPACS2 classification please visit the [WFD UK Tag website](#). You can also download the River LEAFPACS2 Class calculator from there.

Not all sites have environmental data listed. If this is the case it will not be possible to run LEAFPACS2 unless you have your own additional data for the site.

For some fields, additional data can be gathered from other sources e.g. source altitude from Ordnance Survey maps. Alkalinity data are also available on the [Water Quality Archive](#) from our water quality monitoring points. Alkalinity does not vary that much from site to site or year to year, so it's possible to interpolate a value based on monitoring sites nearby, particularly on the same geology and/or watercourse. However please be aware that alkalinity should be at reference status to produce an accurate classification. Therefore measured values of alkalinity, especially from waterbodies that are heavily impacted by anthropogenic discharges may not be suitable for use in producing classifications and should be used with caution.

**Have sample & analysis methods changed over time?**

Yes, when analysing long-term trends bear in mind that from 2006 onwards we moved from the standard MTR (Mean Trophic Rank) methodology to the WFD survey method.

From 2010 onwards all Environmental Change Network (ECN) sites have been surveyed using the WFD standard method for macrophytes. Prior to this the ECN methodology was used and data entered separately in sections of 10m. You should not compare ECN survey results with WFD survey results. For further information on the ECN please access the [ECN website](#).

Since the inception of the Water Framework Directive, guidance, training and quality assurance of macrophyte survey methods and identification has improved; bringing greater consistency, precision and accuracy within our datasets.

**What are the abundance categories?**

For the standard WFD macrophyte survey method, rather than recording the actual percentage cover of each species, it is recorded as a cover band as below:

| Cover Band | Percentage cover of the macrophyte species in survey section |
|------------|--|
| C1         | <0.1%  |
| C2         | 0.1 to 1%  |
| C3         | 1 to 2.5%  |
| C4         | 2.5 to 5%  |
| C5         | 5 to 10%   |
| C6         | 10 to 25%  |
| C7         | 25 to 50%  |
| C8         | 50 to 75%  |
| C9         | >75%   |

For example: a cover value of 3 means that the plant occupied between 1 and 2.5% of the survey area.

Taxa without a % cover band are not included in the index calculations. These are likely to be taxa noted as present in the general area but not within the confines of the survey section. Historic surveys may have the actual % cover recorded but displayed in a separate field.

**What macrophyte indices are supplied in this dataset?**

The indices provided in this output (listed below) are calculated using information on macrophyte species and groups of such species. These taxa are given individual taxon values and those scores are used for a survey to generate the indices.

**Please note** that Biosys will calculate indices for all macrophyte samples as long as there is at least 1 'scoring' taxon present. Care should be taken when interpreting indices, and the values produced, to ensure that there is sufficient taxonomic data present to have confidence in the output.

Indices **should not** be used for surveys carried out using ECN methodology. We will adapt the data extract in the coming months to remove indices for that sample method.

All surveys carried out prior to 2005 used the MTR methodology and as such the indices in this extract should not be used to make a WFD status assessment. They are included for information as they may be applicable for undertaking temporal comparisons of different data collected from the same site. Note, however, that the indices are calculated from a much reduced (pre-LEAFPCAS) taxa list. Indices calculated on surveys between 2005-2007 should also be treated with caution as this was a transition and training period when surveys were being moved over to the WFD method and data accuracy cannot be guaranteed.

A brief overview of the indices are given below. Further information on Water Framework Directive (WFD) classifications, and the indices can be found on the UK Tag pages [here](#).

**RMNI** - The RMNI (River Macrophyte Nutrient Index) can be used to assess the impacts of eutrophication on macrophyte communities. Individual macrophyte taxa are scored according to nutrient preference (0, low - 10,high). The final RMNI score is the average RMNI, weighted by percentage cover band (PCT). Higher RMNI indices indicate macrophyte communities with a preference for higher nutrient levels. A high RMNI index does not automatically indicate eutrophication, instead it must be interpreted in context by comparing it with the expected RMNI score for the site based on reference conditions.

**RN\_A\_TAXA** - This assesses the richness of aquatic taxa and is the number of LEAFPCAS 2 aquatic macrophyte taxa (hydrophytes) recorded in the field survey.

**RFG** - This assesses whole community richness and is the number of functional groups (determined by morphology) that the aquatic macrophyte taxa are grouped in. There are 23 possible groups.

**RFA\_PC** - This is the percentage cover of green filamentous algae. Not all filamentous algae taxa contribute to this index. For information on those that do refer to the LEAFPCAS 2 methodology document on UKTAG.

**RMHI** - RMHI (River Macrophyte Hydraulic Index). This is an historic index that is no longer maintained within our database in terms of taxonomy or taxon values. It is displayed for information and was used for assessing the impacts of flow on macrophyte communities in the first version of the LEFAPCS tool (for Cycle 1 of the WFD).

### How do I use the taxonomic information?

The Natural History Museum maintain the UK Species Inventory (UKSI) and we use that to update our 'taxon dictionary' in Biosys. They maintain a bespoke checklist for us that limits the taxon names we can record against and also ensures we have the taxon hierarchy and names required by our calculations. This does mean that sometimes we have bespoke names or names that are slightly out-of-date. You can use our dataset to understand the latest name that we use (as this may be different in older samples), the type of taxon (e.g. macrophyte) but also the recommended NBN Taxon\_Version\_Key. If you wish to link to other datasets then it is the taxon\_version\_key (TVK) that you need.

## Disclaimers:

Note: We do our best to avoid quality problems but this dataset reflects the data we hold. Our datasets may contain errors.

### External websites:

We have provided links to external websites as a convenience and to provide additional information. We are not responsible for the reliability or content of those external websites. We are not liable for any loss or damage that may come from using the external website links in this document.

customer service line  
03708 506 506

incident hotline  
0800 80 70 60

floodline  
0345 988 1188

[www.gov.uk/environment-agency](http://www.gov.uk/environment-agency)

## Sort code:

We list the Sort code in Biosys which is derived from the Maitland list, owned and maintained by [CEH](#). Our sort code list contains omissions, additions and errors and should *not* be used as a definitive list. It can be used as an aid to ordering taxa in a taxonomic hierarchy.

## Dataset content

### Description of column headings

| Field name                        | Description   | Freshwater<br>Macrophyte Site | Freshwater<br>Macrophyte Metrics | Freshwater<br>Macrophyte Taxa | Taxon information |
|-----------------------------------|---|-------------------------------|----------------------------------|-------------------------------|-------------------|
| <b>AGENCY_AREA</b>                | EA Area name as given in Biosys for site (e.g. NORTH WEST - NORTH)  | Yes                           |                                  |                               |                   |
| <b>REPORTING_AREA</b>             | EA Area name based on more-up-to-date Areas e.g. Cumbria and Lancashire.  | Yes                           |                                  | Yes                           |                   |
| <b>SITE_ID</b>                    | Unique numeric Site identifier generated when the site was created in Biosys (e.g. 66155)                       | Yes                           | Yes                              |                               |                   |
| <b>WATER_BODY</b>                 | Water body in Biosys, free-text field (e.g. LUNE (NW CENTRAL))  | Yes                           |                                  |                               |                   |
| <b>WATERBODY_TYPE_DESCRIPTION</b> | Waterbody type recorded in Biosys from a pick-list (e.g. RIVER: Natural/semi-natural flowing fresh watercourse) | Yes                           |                                  |                               |                   |
| <b>NGR_PREFIX</b>                 | Two letter prefix of National Grid Reference (e.g. SS)  | Yes                           |                                  |                               |                   |
| <b>EASTING</b>                    | Five-digit Easting of NGR (e.g. 58450)  | Yes                           |                                  |                               |                   |

|                         |   |     |  |  |  |
|-------------------------|---|-----|--|--|--|
| <b>NORTHING</b>         | Five-digit Northing of NGR (e.g. 03430)   | Yes |  |  |  |
| <b>NGR_10_FIG</b>       | Ten-digit National Grid Reference (e.g. SS5845003430)   | Yes |  |  |  |
| <b>FULL_EASTING</b>     | Full Easting of the site (e.g. 258450)  | Yes |  |  |  |
| <b>FULL_NORTHING</b>    | Full Northing of the site (e.g. 103430)   | Yes |  |  |  |
| <b>WFD_WATERBODY_ID</b> | WFD ID code given in Biosys (e.g. GB112072065980).  | Yes |  |  |  |
| <b>SLOPE</b>            | Slope of the site (m/km)  | Yes |  |  |  |
| <b>DIST_FROM_SOURCE</b> | Distance from Source (km)   | Yes |  |  |  |
| <b>SOURCE_ALTITUDE</b>  | Altitude of the source of the river that the site sits in (m)   | Yes |  |  |  |
| <b>ALKALINITY</b>       | Alkalinity (mg/l)   | Yes |  |  |  |
| <b>BASE_DATA_DATE</b>   | Date the site Base Data (e.g. Alkalinity etc.) was entered against the site in Biosys (e.g. 01/01/1995) | Yes |  |  |  |
| <b>MIN_SAMPLE_DATE</b>  | The first Macrophyte sample date at the site.   | Yes |  |  |  |
| <b>MAX_SAMPLE_DATE</b>  | The latest Macrophyte sample date at the site.  | Yes |  |  |  |
| <b>COUNT_OF_SAMPLES</b> | The number of Macrophyte sampled taken at the site.   | Yes |  |  |  |
| <b>ECN_SITE_MPHYTE</b>  | This field identifies if a site is sampled for Macrophytes as part of the Environmental Change Network. | Yes |  |  |  |

customer service line  
03708 506 506

incident hotline  
0800 80 70 60

floodline  
0345 988 1188

[www.gov.uk/environment-agency](http://www.gov.uk/environment-agency)



|                                  |   |     |     |     |  |
|----------------------------------|---|-----|-----|-----|--|
| <b>SAMPLE_ID</b>                 | Sample ID generated when the sample is added to Biosys (e.g. 282922)  |     | Yes |     |  |
| <b>SAMPLE_VERSION</b>            | Version number of sample in Biosys. Together with the SAMPLE_ID this denotes a unique record.                                     |     | Yes |     |  |
| <b>REPLICATE_CODE</b>            | A free-text field used to differentiate between replicate samples taken at the same site, date and time.                          |     | Yes |     |  |
| <b>SAMPLE_DATE</b>               | Date the sample was taken (e.g. 05/05/1994)   |     | Yes |     |  |
| <b>SAMPLE_TYPE_DESCRIPTION</b>   | The type of sample collected (FRESHWATER: MACROPHYTE<br>All macrophyte samples (including WFD))                                   |     | Yes |     |  |
| <b>SAMPLE_METHOD_DESCRIPTION</b> | The type of sample method used (MACROPHYTE: Standard method for macrophytes (WFD compliant))                                      | Yes | Yes |     |  |
| <b>SAMPLE_REASON</b>             | Displays a generic sample reason e.g. National/Investigative/Local purposes etc.  |     | Yes |     |  |
| <b>ANALYSIS_ID</b>               | Analysis ID created when analysis has been added against the sample (e.g. 264143)   |     | Yes | Yes |  |
| <b>DATE_OF_ANALYSIS</b>          | Date the sample was analysed. For macrophyte surveys this is the same as the sample date as analysis is carried out in the field. |     | Yes |     |  |

customer service line  
03708 506 506

incident hotline  
0800 80 70 60

floodline  
0345 988 1188

[www.gov.uk/environment-agency](http://www.gov.uk/environment-agency)

|                                    |   |  |     |     |     |
|------------------------------------|---|--|-----|-----|-----|
| <b>ANALYSIS_TYPE_DESCRIPTION</b>   | Analysis type description as recorded in Biosys e.g. FIELD ANALYSIS   |  | Yes |     |     |
| <b>ANALYSIS_METHOD_DESCRIPTION</b> | Analysis method description as recorded in Biosys e.g. MACROPHYTE: Cover class (MTR, ECN)   |  | Yes |     |     |
| <b>RMHI</b>                        | River Macrophyte Hydraulic I index  |  | Yes |     |     |
| <b>RMNI</b>                        | River Macrophyte Nutrient Index   |  | Yes |     |     |
| <b>RN_A_TAXA</b>                   | Number of aquatic taxa contributing to the RMNI index   |  | Yes |     |     |
| <b>N_RFG</b>                       | Number of river macrophyte functional groups.   |  | Yes |     |     |
| <b>RFA_PC</b>                      | The percentage cover of Filamentous algae   |  | Yes |     |     |
| <b>IS_THIRD_PARTY_DATA</b>         | Lists whether the data is from a third-party organisation. <b>Note:</b> a 'No' entry may still mean data has been provided by a third-party (particularly for historic data) as this is a relatively new field in Biosys. |  | Yes |     |     |
| <b>TAXON_LIST_ITEM_KEY</b>         | This is a unique ID, assigned by the NHM, to each name in the EA checklist. This code is required by Biosys.  |  |     | Yes | Yes |
| <b>TAXON_NAME</b>                  | Name of the taxon unit e.g. <i>Potamogeton pectinatus</i>   |  |     |     | Yes |
| <b>PERCENTAGE_COVER_BAND</b>       | This is the % cover band category for a taxon and will range from 1-9   |  |     | Yes |     |

customer service line  
03708 506 506

[www.gov.uk/environment-agency](http://www.gov.uk/environment-agency)

incident hotline  
0800 80 70 60

floodline  
0345 988 1188

|                             |   |  |  |     |     |
|-----------------------------|---|--|--|-----|-----|
| <b>PERCENTAGE_COVER</b>     | The % cover of a taxon.   |  |  | Yes |     |
| <b>TAXON_VERSION_KEY</b>    | This is a unique ID, assigned by the NHM to different versions of a taxon name e.g. where authority or context (e.g. <i>sensu stricto</i> ) is different. This is the NBN code that can be used to link to data with other organisations & the National Biodiversity Network. |  |  |     | Yes |
| <b>AUTHORITY</b>            | Naming authority e.g. (L.) Gaertn.; Linnaeus, 1758.   |  |  |     | Yes |
| <b>SORT_CODE</b>            | This is a code (also known as the maitland code) defined by CEH to enable us to sort taxa in a hierarchy. NOTE: our list is not the definitive list and contains omissions/errors. The latest list can be downloaded from the CEH website.                                    |  |  |     | Yes |
| <b>TAXON_NAME_CURRENT</b>   | Yes/No field to denote where the taxon is the most current name in use on our system.   |  |  |     | Yes |
| <b>PREFERRED_TAXON_NAME</b> | This is the most current and preferred name for the taxon on Biosys. If the TAXON_NAME_CURRENT field is 'No' then you will see the latest name listed here.   |  |  |     | Yes |
| <b>PREFERRED_NAME_TVK</b>   | The Taxon Version Key for the Biosys preferred taxon name. This is the NBN code that can be used to link to data with other   |  |  |     | Yes |

|                            |   |  |  |  |     |
|----------------------------|---|--|--|--|-----|
|                            | organisations & the National Biodiversity Network.  |  |  |  |     |
| <b>NBN_RECOMMENDED_TVK</b> | This is the NHM recommended Taxon Version Key (taken from their Nameserver table). This is the NBN code that can be used to link to data with other organisations & the National Biodiversity Network.                          |  |  |  | Yes |
| <b>PARENT_TLIK</b>         | The parent Taxon_List_Item_Key for the parent taxon in Biosys   |  |  |  | Yes |
| <b>PARENT_TAXON_NAME</b>   | The Parent taxon name.  |  |  |  | Yes |
| <b>TAXON_RANK</b>          | Taxon rank e.g. Family, species etc.  |  |  |  | Yes |
| <b>TAXON_TYPE</b>          | Lists the taxon type in Biosys e.g. Algae; Other Macroinvertebrate; Macrophyte.   |  |  |  | Yes |
| <b>TAXON_GROUP_NAME</b>    | Lists the taxon group as defined by the NHM e.g. alga; crustacean.  |  |  |  | Yes |
| <b>NON_NATIVE_SP</b>       | Lists whether the taxon is a non-native species to England.   |  |  |  | Yes |
| <b>PROTECTED_TAXA</b>      | Lists whether the taxon is a protected species. Note we use JNCC information to help us list protected species but we only flag taxa that have the potential to be impacted by activities we permit or activities we carry out. |  |  |  | Yes |