Information pack on the habitat assessment framework underpinning Defra indicator B6 – Natural functions of water and wetland ecosystems

Natural England, May 2023

***Purpose**: This information pack is intended to support Defra's publication of an interim version of the B6 indicator, as part of the annual update of its Outcome Indicator Framework. It is divided into three parts: 1) explanatory notes; 2) assessment outputs; and 3) attribute information sheets. The primary storage location of this pack is presently the Document store on the Discovering priority habitats website.*

***Access to underlying data**: A supplementary Excel workbook (available on the same Document store) provides all of the aggregated naturalness scores on which this information pack is based. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets. The primary data used to generate naturalness scores for different attributes comes from multiple sources, explained in detail in this part of the information pack.*

PART 3 – ATTRIBUTE INFORMATION SHEETS

This part of the information pack provides information on individual attributes used within the assessment framework, including the rationale for inclusion, the definition of the attribute, the data used to populate the attribute, naturalness class boundaries, data limitations and other contextual information. For this interim version of the B6 indicator information is only provided on attributes used in the assessment of freshwater ecosystems, since these the are only ecosystems where interim assessments are being published.

Summary table of attributes for all ecosystems

	Running waters		Standing waters	
Component	**Rivers (i.e. excluding streams)**	**Headwater streams**	**Lakes**	**Ponds**
Hydrological	H1 River flows at Q95 H2 River flows at Q70 H3 River flows at Q50 H4 River flows at Q30 H5 Flooding regime H6 Groundwater inputs	H1 Flooding regime H2 Groundwater inputs H3 FBA hydrological assessment	H1 Naturalness of flow at Q95 H2 Naturalness of flow at Q70 H3 Naturalness of flow at Q50 H4 Naturalness of flow at Q30	H1 Artificial influences
Chemical	C1 Ammonia C2 Dissolved oxygen C3 Phosphorus C4 Nitrogen C5 pH C6 Specific pollutants C7 Macroinvertebrates C8 Macrophytes & diatoms	C1 Ammonia C2 Dissolved oxygen C3 Phosphorus C4 Nitrogen C5 pH C6 Specific pollutants C7 Macroinvertebrates C8 Macrophytes & diatoms	C1 Total Phosphorus C2 Total Nitrogen C3 ANC C4 Chlorophyll C5 Macrophytes and diatoms C6 Specific pollutants C7 Fish e-DNA	C1 Phosphorus C2 Nitrogen
Physical	P1 Fragmentation P2 Impoundment P3 Strategic connectivity P4 Streampower P5 Channelisation P6 Habitat Modification Score P7 Flow Habitat Mosaic P8 Riparian Trees P9 In-channel woody material P10 Riparian vegetation complexity P11 FBA physical assessment	P1 Fragmentation P2 Impoundment P3 Strategic connectivity P4 Streampower P5 Channelisation P6 Habitat Modification Score P7 Flow Habitat Mosaic P8 Riparian Trees P9 In-channel woody debris P10 Riparian vegetation complexity P11 FBA physical assessment	P1 Hydrological structures P2 Artificial shoreline P3 Sediment fluxes P4 Riparian habitat P5 Riparian trees P6 FBA physical assessment P7 Marginal fringe emergent vegetation	P1 Natural pond base P2 Natural shoreline P3 Semi-natural land-use P4 Shading P5 Grazing intensity
Biological	B1 Non-native species	B1 Non-native species	B1 Non-native species	B1 PSYM B2 Non-native species
Cross-cutting	CC1 Macroinvertebrate similarity index	CC1 Macroinvertebrate similarity index CC2 Catchment land cover		CC1 Landscape connectivity

Summary table of attributes for all ecosystems continued

	Wetlands		Estuaries & coasts	
Component	**Bogs**	**Fens**	**Estuaries**	**Coasts**
Hydrological	NA	NA	NA	NA
Chemical	NA	NA	NA	NA
Physical	NA	NA	NA	NA
Biological	NA	NA	NA	NA
Cross-cutting	NA	NA	NA	NA

NA – Not available at this time.

Attribute information sheet template

Principal habitat component(s) (rivers, streams, lakes, ponds, freshwater wetlands, estuaries, coasts): <>

Attribute title and code: <>

Rationale for inclusion: <>

Source data: <>

Outline description of dataset including spatial coverage, representativeness, limitations: <>

Data ownership and licensing restrictions (if any): <>

Data transfer arrangements: <>

Frequency of source data update/data transfer: <>

Form of attribute: <>

Data processing method for generating attribute output: <>

Naturalness class boundaries: <>

Attribute robustness: <>

Storage location for raw dataset and processed data: <>

Attribute information sheets for freshwater ecosystems

The information in these assessment sheets has been assembled by Chris Mainstone, Ruth Hall and Mel Fletcher of Natural England, working in close collaboration with Cedric Laize of UKCEH (who undertook all analyses of attributes for freshwater ecosystems) and a range of staff in the Environment Agency and other organisations (names noted under the individual sheets).

Attribute information sheet

Principal habitat components: Rivers

Flow regime, River codes H1-H4

Rationale for inclusion: The natural flow regime is the foundation of the natural functioning of river and stream ecosystems and their associated wetland habitats. All components of the flow regime are important in shaping these ecosystems and sustaining the habitat mosaics of characteristic biological assemblages. High flows perform most of the geomorphological work and generate natural inundation of the floodplain; natural low flow regimes sustain in-channel biota whilst providing seasonal exposed marginal and riparian areas for biota specialised in ephemeral habitats; mid-flows sustain the spatial extent and balance of the habitat mosaic. These attributes provide an assessment of the levels of deviation from the natural flow regime at different flow magnitudes.

Source data: Environment Agency Water Resources GIS system. Regular updates are supplied to Natural England for use in the B6 data framework. The dataset will be stored on Natural England systems.

Outline description of dataset including spatial coverage, representativeness, limitations:

The dataset provides actual flows and modelled naturalised flows at a number of naturalised flow (Qn) values. For the interim version of the B6 indicator, data on actual flows are based on abstraction returns from the period 2013-2018 for most catchments. A single set of values is provided for each Water Framework Directive waterbody.

Data field	Description
EA_WB_ID	Water body ID number.
WB_NAME	Waterbody name.
Type_IWB	Type of waterbody.
CATCHMENT	Name of catchment.
ScenRA%QN30	Recent actual scenario as a percentage of natural flows at Q30.
ScenRA%QN50	Recent actual scenario as a percentage of natural flows at Q50.
ScenRA%QN70	Recent actual scenario as a percentage of natural flows at Q70.
ScenRA%QN95	Recent actual scenario as a percentage of natural flows at Q95.

Data ownership and licensing restrictions (if any): The data owner is the Environment Agency. The data are available under open government licence.

Data transfer arrangements: A standard data specification (above) has been agreed with the Environment Agency.

Frequency of source data update/data transfer: Data on individual catchments are updated as and when possible. An update frequency of 3-5 years is considered appropriate.

Form of attribute: Calculated as the percentage deviation (negative or positive) of actual daily river flow from the modelled naturalised flow (flow in the absence of abstractions and discharges).

Data processing method for generating attribute output: Data are pre-processed by the Environment Agency within their Water Resources GIS system. These attributes only relate to the

'larger river' zone of waterbodies – no output is generated for the headwater zones of waterbodies since the assessment points for which data are available are towards the downstream end of waterbodies and cannot reflect the hydrological naturalness of headwater streams.

Naturalness class boundaries:

Qn	Class 1	Class 2	Class 3	Class 4	Class 5
95	<5	5-10	10-25	25-40	>40
70	<5	5-10	10-25	25-40	>40
50	<5	5-10	10-25	25-40	>40
30	<5	5-10	10-25	25-40	>40

Attribute robustness: The dataset is not able to characterise the effect of non-consumptive abstractions above the assessment points used in the EA WRGIS. Greater spatial discrimination in the dataset would therefore be preferable and is possible through additional modelling, e.g. via the EA's new hydroecology tool. The dataset is also based on comparison of frequency distributions of recent actual and naturalised flows, not a direct comparison of time series of data. Evaluation of data as time series would provide a more realistic picture of hydrological modification.

Storage location for source dataset and processed data: Source data are available from the Environment Agency's Water Resources GIS. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat components: Rivers, headwater streams

Flooding regime, Code H5

Rationale for inclusion: The natural flooding regime is critical to ecological interactions between the river/stream channel and the floodplain. Natural flooding provides seasonal connectivity between river/stream and floodplain habitats for native plants (seeds, propagules), invertebrates and fish and create conditions for floodplain habitats to thrive. Key ecological dependencies for this attribute are that: 1) floodplain land that is subject to a natural hydrological regime should be under suitable native vegetation, and 2) the water quality (chemical naturalness) of floodwaters is high so that floodplain vegetation and associated fauna are not adversely affected.

Source data: Ordnance Survey Mastermap dataset on 'Manmade slope landforms', Environment Agency AIMS dataset

Outline description of dataset including spatial coverage, representativeness, limitations:

Various approaches and datasets were investigated to identify all flood embankments along rivers and streams in England. Some relevant datasets are highly detailed and robust but only provide a partial picture; for instance, the Environment Agency's AIMS dataset which identifies formal flood defence assets but not the many other embankments created by historical and on-going channel dredging. LIDAR datasets provide a robust and high-resolution raw data source but require very resource-intensive analysis to identify all embankments on a national-scale.

The dataset chosen is a vector-based data layer formed from the high-resolution OS Mastermap dataset. It is not specific to flood embankments and includes road and rail embankments, but GIS processing has been used to eliminate as many irrelevant features as possible. This has been augmented by the AIMS dataset, as this provides more a more robust representation of flood embankments for the locations that the dataset covers. The combined dataset generates a **whole-inventory** attribute, providing characterisation of the whole river and stream habitat resource.

Data ownership and licensing restrictions (if any): The data owners are Ordnance Survey (Mastermap) and the Environment Agency (AIMS). The datasets are available under licence.

Data transfer arrangements: The Mastermap dataset required for the initial assessment is already available, but maintenance updates to the dataset are unlikely to be undertaken on an adequate scale and frequency to provide sufficient sensitivity to change for our purposes. Updates to the attribute will need to rely on new data capture processes that will allow changes to existing levels of embankments to be recorded, arising from river/stream restoration, abandonment (lack of maintenance) or new artificial modifications.

Frequency of source data update/data transfer: New data capture processes need to be developed that will aim to capture changes in the occurrence of flood embankments on an *ad hoc* basis. An update frequency of 3-5 years is appropriate.

Form of attribute: Calculated as the percentage of channel length in each waterbody (each divided into headwater and larger river zones) that is embanked on one or both sides.

Data processing method for generating attribute output: The dataset is filtered for relevance to rivers/streams by spatial overlay with the Environment Agency's GIS data on flood zones, selecting Zone 3 (relating to the natural functional floodplain). The length of embankment is expressed as a

percentage of river/stream channel length in each waterbody (split into headwater and larger river zones).

Naturalness class boundaries:

	Class 1	Class 2	Class 3	Class 4	Class 5
% of channel length with embankments	0	0-5	5-10	10-30	>30

Attribute robustness: Flood embankments are the most obvious artificial modification affecting natural flooding of the floodplain, but channel over-sizing (deepening, widening) and straightening also have major effects. These in-channel physical modifications are characterised by attributes within the physical component of naturalness and should be thought of as part of the consideration of natural flood regime. Within the dataset actually used for assessing this attribute there are man-made landforms that are not intrinsically linked to flooding, which within the functional floodplain may or may not have an effect on flood regime. This has to be considered to be part of the noise within the dataset, which cannot be removed without substantial detailed local analysis. Nevertheless, the attribute overall should be capable of providing a reasonable assessment of the habitat resources as a whole and is therefore considered to be fit for purpose.

Storage location for source dataset and processed data: Access to the source dataset is via Ordnance Survey, under licence. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat component: Rivers, headwater streams

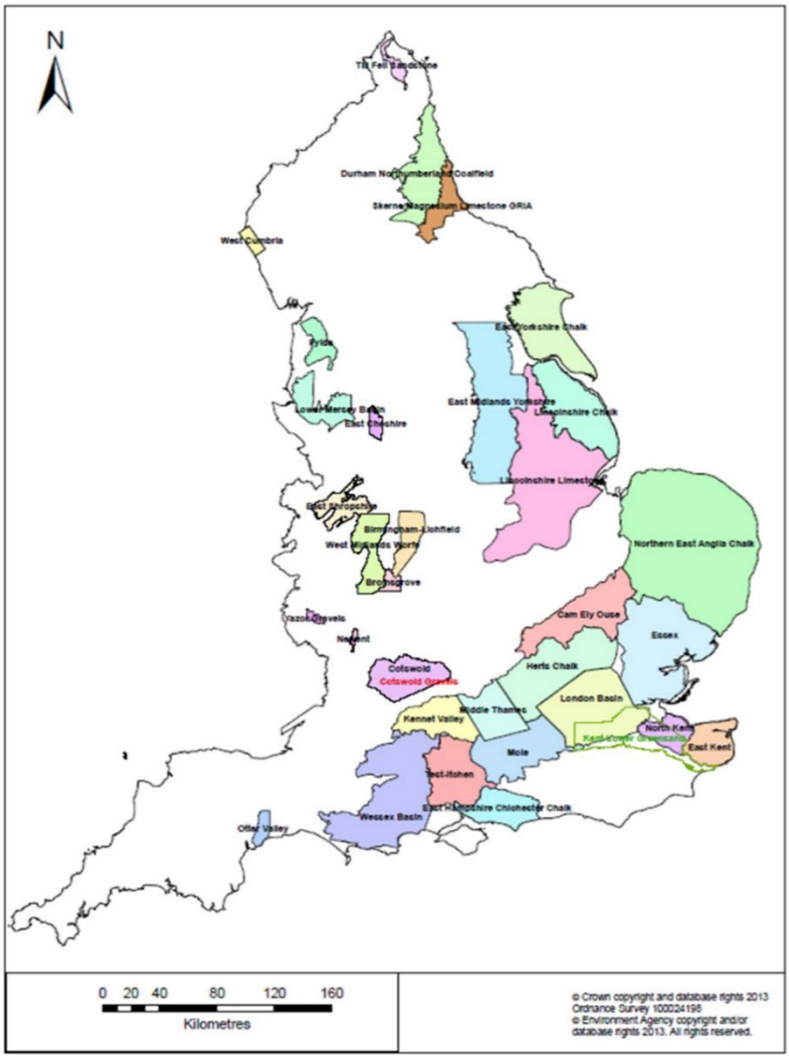
Groundwater inputs: H6 (rivers), H2 (streams)

Rationale for inclusion: Natural groundwater inputs are ecologically critical in sustaining the natural baseflow and water chemistry of groundwater-fed rivers and streams. They are particularly important for headwater streams, where groundwater baseflow is the dominant natural water supply, dictating the natural seasonality of flows in temporary stream sections within the headwater stream network. These attributes characterise levels of deviation from the natural regime of groundwater inputs to rivers and streams, indicating the level of impact from groundwater abstraction.

Source data: Environment Agency Hydroecology tool.

Outline description of dataset including spatial coverage, representativeness, limitations:

The Environment Agency Groundwater Hydro-ecology tool is a new tool developed from the outputs of a range of regional groundwater models. The models included in the tool cover the major exploited aquifers in England – see the map below.



There are some notable gaps in model coverage, particularly relating to complex geologies which are extremely difficult to model, but (with suitable funding) it is anticipated that model coverage will improve further in the future. These attributes are therefore considered to be '**whole-inventory'** (as opposed to representative).

The tool and the underlying models operate at a spatial resolution of 200 metres, providing data in a framework of 200 metre gridcells. The tool is capable of providing a range of outputs, but for the purposes of the B6 indicator the Environment Agency pre-processes the data for B6 by extracting relevant data fields on groundwater-to-surface water (GW-to-SW) flows for typical seasonal low flow conditions (Qn90%ile):

1) recent actual (abstraction scenario) – field name L90GSRA;
2) modelled naturalised (in the absence of abstractions and artificial recharges) – field name.

Data ownership and licensing restrictions (if any): The data owner is the Environment Agency. The data have been made available to Natural England under licence. There is Third Party IP as follows: Centre for Ecology & Hydrology (CEH); Met Office (MO); British Geological Survey (BGS); Ordnance Survey (OS); National Soil Resources Institute (NSRI).

Data transfer arrangements: A standard data specification (above) has been agreed with the Environment Agency.

Frequency of source data update/data transfer: Can be annual if deemed necessary but updates every 3-5 years would be most appropriate.

Form of attribute: Calculated as the percentage deviation (negative or positive) of recent actual groundwater inputs from the modelled naturalised groundwater input.

Data processing method for generating attribute output: For each 200m gridcell in a waterbody catchment (separately for headwaters zone and larger river zone):

1. L90GSRA values are summed
2. L90GSN values are summed
3. The ratio of Sum(GSRA) /Sum(GSN) is calculated
4. The ratio is transformed to an absolute % deviation from unity (e.g. if ratio is 0.8 or 1.2, % deviation is 20%).

Naturalness class boundaries:

	Class 1	Class 2	Class 3	Class 4	Class 5
% deviation	<5	5-10	10-25	25-40	>40

Attribute robustness: The available data are considered to be moderately robust for this purpose. There are gaps in data coverage associated with smaller or more complicated aquifers, but the dataset as it stands is considered to provide a reasonable reflection of the habitat resource of groundwater-fed rivers and streams. Refinements to existing models and the development of new models are needed to cover data gaps.

Storage location for source dataset and processed data: The Groundwater Hydro-ecology tool is maintained by the Environment Agency as the source dataset. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being

prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat components: Rivers, headwater streams

Chemical attributes Ammonia (C1), Dissolved Oxygen (C2), Phosphorus (C3), Nitrogen (C4), pH (C5), Macroinvertebrates (C7), Macrophytes & Diatoms (C8)

Rationale for inclusion: These attributes have been selected to provide a broad indication of chemical naturalness, using chemical determinands for the most basic elements of chemical naturalness and biological metrics to indicate other chemical issues (hazardous chemicals, episodic pollution). Hazardous chemicals are also given more specific consideration under attribute C6).

Source data: Environment Agency Catchment Data Explorer.

Outline description of dataset including spatial coverage, representativeness, limitations: The dataset provides the current ecological status class of each attribute at each monitoring point in the EA's Water Framework Directive monitoring programme, integrating data over a number of years to provide a complete picture for all monitoring locations. For the interim version of the indicator, the 2019 version of the dataset (the most recent available) has been used. No assessment of C4 (nitrogen) is currently possible since no data are currently provided by the monitoring programme. The current monitoring programme includes a reasonable range of headwater stream locations although only for larger streams. Environment Agency monitoring design is changing to a representative surveillance network – the implications of this for the future sourcing of data is currently unclear.

Data ownership and licensing restrictions (if any): The data owner is the Environment Agency. The data are available to the public via Catchment Data Explore under open government licence.

Data transfer arrangements: Data are downloaded directly from Catchment Data Explorer.

Frequency of source data update/data transfer: In future data will be sourced from the EA's new surveillance programme, which will monitor sites on a rolling basis and complete a full cycle every 5 years. It would therefore be sensible to update these attributes on a 5-year cycle.

Form of attribute: Ecological status class of each attribute, which is related to naturalness class directly.

Data processing method for generating attribute output: The data for each monitoring site are pre-processed into ecological status class allocations within Catchment Data Explorer. Once extracted from Explorer, monitoring sites are then resolved into the spatial framework of WFD waterbodies, each divided into larger river and headwater stream zones. Where more than one site occurs within a waterbody zone the worst performing site is selected to represent the zone. For this interim version of the indicator, data for rivers (but not streams) were used that were pre-processed into WFD waterbodies, so that a more complete set of attributes could be portrayed (one attribute was not included in the site-based dataset available on Catchment Data Explorer).

Naturalness class boundaries:

	Class 1	Class 2	Class 3	Class 4	Class 5
Ecological status	High	Good	Moderate	Poor	Bad

Attribute robustness: The robustness of these attributes will be dependent on the shape of future Environment Agency surveillance programmes. The density of monitoring sites, the adequacy of coverage of detailed river/stream types, and the coverage of individual determinands are all key factors. In the future the intention is that data will need to be sourced from Environment Agency national river and headwater stream surveillance programmes. These programmes are similarly designed to be representative of the river and stream network, and discussions are on-going about ensuring they are also representative of detailed habitat types assessed within the assessment framework used in B6.

Storage location for source dataset and processed data: Source dataset currently stored in Catchment Data Explorer. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat component: Rivers, headwater streams, lakes

Specific pollutants (Code C6)

This information sheet was prepared in collaboration with Lindsey Sturdy, Ben Bradford, Helen Wilkinson and Isobel Bain (Environment Agency).

Rationale for inclusion: Whilst biological metrics (particularly Attribute C7 on macroinvertebrates) are good at detecting impacts from a range of chemicals, the inclusion of an explicit and comprehensive attribute on hazardous substances provides a more robust assessment of chemical naturalness.

Source data: Environment Agency Catchment Data Explorer. As these data are stored on Environment Agency systems there is no need for Natural England to store them separately.

Outline description of dataset including spatial coverage, representativeness, limitations: The dataset is derived from the Environment Agency's chemicals monitoring programmes. The pollutants currently included in this attribute are based on Specific pollutants as defined by The Water Framework Directive (Standards and Classification) Directions (England and Wales) 2015 (legislation.gov.uk) (Schedule 3, Part 2, Table 1). A list is provided below.

Unionised ammonia (as nitrogen)	Diazinon	Methiocarb
Arsenic	2,4 Dichlorophenol	Pendimethalin
Benzyl butyl phthalate	2, 4 Dichlorophenoxyacetic acid	Permethrin
Carbendazim	3, 4 Dichloroaniline	Phenol
Chlorine	Dimethoate	Tetrachloroethane
Chlorothalonil	Glyphosate	Triclosan
Chromium III	Iron	Toluene
Chromium VI	Linuron	Zinc
Copper	Manganese	
Cyanide	Mecoprop	

NB Cypermethrin was removed from this list in 2013 when it became a Priority Substance

The dataset provides the WFD ecological status class for Specific pollutants as a group at each monitoring site. The only ecological status classes recorded within the dataset are High or Moderate, due to the nature of the assessment undertaken by the Environment Agency to generate the dataset. The assessment is based on a pass/fail compliance test of each pollutant against their respective Environment Quality Standard, nominally assigning High status if the standard is passed and Moderate Status if the standard is failed. A status class is then assigned to Specific pollutants as a group, adopting the worst class recorded of any of the listed pollutants at each monitoring site (i.e. if any one pollutant fails its standard, the group is assigned an ecological status class of Moderate).

The coverage of the habitat resource is relatively sparse compared to other attributes used for assessing chemical naturalness, with a large number of waterbodies without data.

The dataset used for this exercise has not been updated on Catchment Data Explorer since before 2019. Data for 2019 exist but have not yet been uploaded.

Data ownership and licensing restrictions (if any): The data owner is the Environment Agency. The data are available to the public via Catchment Data Explorer under open government licence.

Data transfer arrangements: Data are downloaded directly from Catchment Data Explorer.

Frequency of source data update/data transfer: The source dataset on Catchment Data Explorer is updated every 3 years, which provided a suitable update frequency for the attribute.

Form of attribute: Ecological status class of Specific pollutants as a group.

Data processing method for generating attribute output: The data for each monitoring site are pre-processed into an aggregated ecological status class allocation within Catchment Data Explorer. Once data are extracted from Catchment Data Explorer, monitoring sites are then resolved into the spatial framework of WFD waterbodies, each divided into larger river and headwater stream zones. Where more than one site occurs within a waterbody zone, the worst performing site is selected to represent the waterbody zone. For this interim version of the indicator, data for rivers (but not streams) were used that were pre-processed into WFD waterbodies, because data of Specific Pollutants was not included in the site-based dataset available on Catchment Data Explorer).

Naturalness class boundaries:

Owing to the nature of data pre-processing by the Environment Agency, only two naturalness classes can be defined.

	Class 1	Class 2	Class 3	Class 4	Class 5
Ecological status	High	-	Moderate	-	-

Attribute robustness: The future robustness of this attribute will be dependent on the shape of future Environment Agency chemical surveillance programmes; the density of monitoring sites, the adequacy of coverage of detailed river/stream types, and the coverage of individual determinands are all key factors. In particular, the current attribute only covers a small suite of chemicals so is not currently a comprehensive indicator of hazardous chemicals. Future refinement of the attribute to address this issue is not a simple matter and will need to be considered as part of broader deliberations about the design of the Environment Agency's chemicals monitoring programmes. Any refined attribute for B6 would sensibly continue to feed off future Environment Agency assessment and reporting arrangements for ecological status rather than creating a separate grouping of chemicals and separate data processing. Related Environment Agency assessment and reporting of 'chemical status' is complicated by its use of human health-related standards rather than ecologically-based standards, which affects the ecological relevance of the datasets generated.

Storage location for source dataset and processed data: Source dataset is stored in Catchment Data Explorer and is likely to be stored in the same location in the future. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat components: Rivers, headwater streams

Attributes relating to artificial in-channel structures: P1 Fragmentation, P2 Impoundment, P3 Strategic connectivity

Rationale for inclusion: These attributes cover the key aspects of habitat modification arising from artificial in-channel structures. They fragment natural processes associated with coarse and fine sediment delivery and the movement of organisms (addressed by P1). They permanently impound water, removing natural variations in current velocities, water levels and depths and generating excessive fine sediment deposition (addressed by P2). They eliminate connectivity between the sea and the river network which has particular effects on natural saline transition zones and the movement and life cycles of long-distance migratory species such as eels, shads, lampreys and salmon (addressed by P3).

Source data: Environment Agency.

Outline description of dataset including spatial coverage, representativeness, limitations: The dataset is intended to provide a full inventory of all in-channel structures in the English river network. Information is stored on the type of structure, location and the drop in water level (or head). There are known gaps in the dataset that the EA and CaBA initiative (Catchment-Based Approach) are progressively rectifying. There is a particular lack of data on small structures in the headwater stream network, but also other gaps in the coverage of larger structures including some major dams. Data on head of structures is both patchily recorded and of relatively low confidence.

Data ownership and licensing restrictions (if any): The dataset is licensed by the Environment Agency under open government licence.

Data transfer arrangements: Data can be downloaded directly from the CaBA Data Hub whenever required for B6 updates.

Frequency of source data update/data transfer: The Environment Agency and partners have developed an 'app' to record River Obstacles and discussions are on-going about including additions that ensure the app captures the data required for B6. A particular B6 requirement is the ability to record removals of structures, or loss by degradation. A similar app is being developed at European level, which provides an additional vehicle for data capture. The source dataset is updated as and when new information is available so updates to the attributes can be annual if deemed necessary. However, 3-5 year intervals between updates are probably more appropriate.

Form of attributes:

- ***P1 Fragmentation*** – Number of in-channel structures in each waterbody zone (either larger rivers or headwater streams)
- ***P2 Impoundment*** – Percentage of total channel length that is impounded by in-channel structures in each waterbody zone.
- ***P3 Strategic connectivity*** – Percentage of free-flowing channel length to the sea per river basin.

Data processing method for generating attribute output: Data for relevant in-channel structures are extracted from the national dataset (data for the current analysis was downloaded on 15 June 2022). Only those generating (or likely to generate) an observable head are used. Since data on the

head of structures is not reliable, filtering currently has to be undertaken based on structure type. Those included in the assessment are: weir, dam, sluice, ford, culvert, bridge footing, pipe bridge, bridge apron

- ***P1 Fragmentation –*** The locations of structures are resolved into waterbody zones and the number of structures per waterbody zone is calculated.
- ***P2 Impoundment*** – The head of each structure is multiplied by channel gradient (in metres /per kilometre) at its location, using the UKCEH Digital Rivers Network (DRN), to generate an estimate impounded length (in kilometres). These lengths are then summed for each waterbody zone, and the total impounded length is then expressed as a percentage of total channel length in each waterbody zone.
- ***P3 Strategic connectivity*** - The digital layer of 'Integrated Hydrological Units (IHA) of the United Kingdom' has been used to delineate river basins because it is consistent with the UKCEH intelligent river network used in this and other parts of the assessment framework used for B6. It is very similar to the river basin delineation used for the Water Framework Directive. Very small basins with no river/stream channels evident at 1:50,000 map scale are excluded from the analysis. A GIS algorithm is used to identify the most downstream structures on the UKCEH DRN and calculate, for each basin, the total channel length downstream of those structures. This channel length is then apportioned between headwater streams and larger rivers and expressed as the percentage of total headwater stream and larger river length within each basin.

Naturalness class boundaries:

Attribute	Class 1	Class 2	Class 3	Class 4	Class 5
P1	0	1-5	6-10	11-30	>30
P2	0	>0-5	>5-10	>10-30	>30
P3	100	90-100	50-90	10-50	<10

Attribute robustness: The attribute is currently not robust because of gaps in the source dataset. This should be progressively resolved as the app is used and structures are added, but in analytical terms these additions to the database over time may spuriously manifest as a loss of naturalness unless some form of correction is included in the B6 data analysis. New records of structures will need to be flagged as 'historical' or 'new' to allow relevant structures to be retrospectively added to the baseline scoring of attributes rather than to subsequent updates to attribute scores. Specifically in relation to Attribute P3, the skewed distribution of basin areas towards small basins gives greater prominence to small basins dominated by headwater streams that are less likely to have tidal structures and more likely to have missing records of in-channel structures in the River Obstacles dataset. Effectively, the River Thames basin generates only one value within the histogram, having no greater weight in the analysis of larger rivers than a small coastal basin with a short section of larger river.

Storage location for source dataset and processed data: Source dataset available from the CaBA Data Hub at this location. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat components: Rivers, headwater streams

Streampower, Code P4

Rationale for inclusion: Streampower relates to the capacity of a river or stream to do geomorphological work, moving bed sediments and eroding banks in ways that generate natural physical form and dynamism. It is therefore a fundamental determinant of the natural character of the physical habitat mosaic, which in turn determines the natural character of biological assemblages. Artificial modifications to the natural streampower of a river or stream generate impacts on its ability to provide characteristic habitat mosaics and support characteristic biological assemblages. Both artificial reductions and increases from natural streampower values generate changes in the character and dynamism of the physical habitat mosaic and hence biological assemblages and therefore constitute reductions in naturalness.

Source data: Natural England/UKCEH dataset on observed-to-reference (putative natural) streampower values.

Outline description of dataset including spatial coverage, representativeness, limitations:

The dataset provides modelled existing and reference streampower values across the English river and stream network, on a 50-metre grid that follows the blue line of UKCEH's Intelligent River Network. It is therefore a **whole-inventory** dataset rather than a representative dataset. Existing and reference values of streampower have been calculated as described in the box below.

Specific stream power (SSP) was derived using the equations described in O'Hare et al (2011):

(1) TSP = SWW x Q x S

Where:

TSP is Total Stream Power (W.m-1)
SWW is Specific Weight of Water (N.m-3)
Q is Discharge (m3.s-1)
S is Slope (m.m-1)

TSP is then standardised by dividing it by wetted width W (m) to give SSP(W.m-2), which allows comparing rivers of different sizes:

(2) SSP = TSP / W

For Q the 'median annual maximum flood peak' or QMED (i.e. the 1 in 2 year flood) was used. For W the bankfull width was used as this matches QMED.

We used a value of 9807 for SWM as per OHare et al. (2011):

(3) SSP = (9807 x QMED x S) / W

QMED is one of the underlying datasets of the Flood Estimation Handbook (FEH). QMED was re-calculated from FEH catchment descriptors to derive natural values, using the following formula (Kjeldsen. 2010):

(4) QMED = (8.3062 * AREA^0.851)*(0.1536^(1000/SAAR))*(FARL^3.4451)*(0.0460^(bfihost^2))

Where

QMED is the median annual flow rate; the 1:2 year event.
AREA is the area of the catchment in km^2.
SAAR is the standard average annual rainfall for the period 1961 to 1990 in mm.
FARL is a reservoir attenuation function (predominantly natural ie lakes and ponds, but also a few artificial)
BFIHOST is the base flow index derived using the HOST classification.

This represents the naturalised QMED (i.e. 'as rural'). In order to capture modified catchments, a correction factor for urbanised area was calculated and then applied to QMED to derive a second set of QMED values ('QMED Urban'). The correction factor, called Urban Adjustment Factor (UAF) is based on the formula described in Kjeldsen (2010):

(5) PRUAF = 1 + 0.47 * URBEXT2000 * (BFIHOST / (1 - BFIHOST))

(6) UAF = ((1 + URBEXT2000)^0.37) * (PRUAF^2.16)

(7) QMED Urban = QMED * UAF

Where PRUAF is the percentage runoff urban adjustment factor, ie an estimate of the increase in run-off volume that occurs as a consequence of urbanisation and is a function of urban extent and catchment type, and URBEXT2000 is a composite index of urban and suburban extent.

Slope (S) was deived from UKCEH digital terrain model by taking elevation differences 500 meters upstream and downstream of any given site.

To derive naturalised **bankfull width**, we used the equation described by Soar and Thorne (2001) for typical UK rivers:

(8) W = 2.48 * QMED^0.5

Actual observed bankfull widths were extracted from OS data.

In order to generate estimates of observed SSP with which to generate a ratio of observed to natural SSP values, two SSP series were derived:

- Naturalised SSP using equations (3), (4) and (8), ie QMED and modelled W
- Observed SSP using QMED Urban instead of QMED and observed widths from OS data instead of modelled widths.

Data ownership and licensing restrictions (if any): The data owner is Natural England. The data will be made available under open government licence.

Data transfer arrangements: The dataset required for the initial assessment is already available and represents a one-off modelling exercise that will not be repeated. Updates to the attribute will rely on new data capture processes that will allow changes to existing streampower values based on changes in the level of physical channel modification, arising from channel restoration or new artificial modifications.

Frequency of source data update/data transfer: New data capture processes need to be developed that will aim to capture changes in the levels of physical channel modification on an *ad hoc* basis. An update frequency of 3-5 years is appropriate.

Form of attribute: Calculated as the percentage deviation of modelled existing-to-reference streampower across the river/stream network.

Data processing method for generating attribute output: The existing dataset is pre-processed into modelled existing and reference streampower values on a 50-metre grid. Values within each waterbody and waterbody zone (larger rivers and headwater streams) are aggregated into a single value by averaging.

Naturalness class boundaries:

	Class 1	Class 2	Class 3	Class 4	Class 5
% deviation of modelled existing from modelled reference	<10	10-25	25-40	40-75	>75

Attribute robustness: The modelling process provides only a coarse representation of alterations to natural streampower. The data that need to be captured to provide updates to the attribute, to refine existing streampower values within the dataset, are derived from a completely different data model and the updating process must be considered to be a crude representation. Nevertheless, the attribute overall should be capable of provide a reasonable assessment of the habitat resources as a whole and is therefore considered to be fit for purpose. There are some concerns that streampower might not provide an adequate representation of stream energy in small streams with high channel gradients, due to the lack of 'weight of water' in the channel to do geomorphological work. This may limit the robustness of the attribute in these instances and is worthy of further consideration.

Storage location for source dataset and processed data: Source dataset is available from Natural England or UKCEH. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

References:

Kjeldsen, T. (2010) Modelling the impact of urbanization on flood frequency relationships in the UK. Hydrology Research 41 (5): 391–405. https://doi.org/10.2166/nh.2010.056

O'Hare, J.M., O'Hare, M.T., Gurnell, A.M., Dunbar, M.J., Scarlett, P.M., Laize, C. (2011) Physical constraints on the distribution of macrophytes linked with flow and sediment dynamics in British rivers. *River Research and Applications* 27/6, 671-683.

Soar, P.J., Thorne, C. (2001) Channel Restoration Design for Meandering Rivers. US Army Corps of Engineers. ERDC/CHL CR-01-1. Washington.

Attribute information sheet

Principal habitat components: Rivers, headwater streams

Channelisation, Code P5

Rationale for inclusion: Natural habitat mosaics in river and stream ecosystems are complex and dynamic, shaped and constantly reformed by natural hydrological and geomorphological processes. Channelisation of river and stream channels results in the wholesale elimination of this dynamic complexity in channel and riparian habitats. The availability of niches for characteristic species is reduced to a set of uniform environmental conditions supporting impoverished biological assemblages. Channel dredging constantly resets the system to an artificially uniform state.

Source data: GIS layer of predicted channel and bank re-sectioning across England, provided by Mark Naura of the River Restoration Centre.

Outline description of dataset including spatial coverage, representativeness, limitations:

The dataset provides a modelled index of channel resectioning, based on predictive GIS modelling using a range of environmental variables from RHS survey sites. Predicted index values are provided for 500-metre channel sections on the 1:50,000 scale UKCEH river network. It is therefore a **whole-inventory** dataset rather than a representative dataset. Further details on the predictive modelling are provided by Naura (2017).

Data ownership and licensing restrictions (if any): The data owner is Marc Naura of the River Restoration Centre. The dataset is only available with the permission of Marc Naura.

Data transfer arrangements: The dataset required for the initial assessment is already available and represents a one-off modelling exercise that will not be repeated. Updates to the attribute will rely on new data capture processes that will allow changes to existing levels of channelisation to be recorded, arising from channel restoration or new artificial modifications.

Frequency of source data update/data transfer: New data capture processes need to be developed that will aim to capture changes in the levels of channelisation on an *ad hoc* basis. An update frequency of 3-5 years is appropriate.

Form of attribute: Calculated as the mean Channel Resectioning Index within each waterbody, each divided into headwater and larger river zones.

Data processing method for generating attribute output: The dataset is pre-processed into an index of resectioning (channelisation), using the Environment Agency's digital rivers network. The index ranges from 0-30 and is classified into 5 classes by Naura, which have been adopted as naturalness classes. Values within each waterbody and waterbody zone (larger rivers and headwater streams) are aggregated into a single value by averaging.

Naturalness class boundaries:

	Class 1	Class 2	Class 3	Class 4	Class 5
Mean channel resectioning index	0-1	2	3-6	7-16	17-30

Attribute robustness: The modelling process provides only a coarse representation of channel resectioning. The data that need to be captured to provide updates to the attribute are derived from a completely different data model and the updating process must be considered to be a crude

representation. Nevertheless, the attribute overall should be capable of provide a reasonable assessment of the habitat resources as a whole and is therefore considered to be fit for purpose.

Storage location for source dataset and processed data: Access to the source dataset is by request to Marc Naura at the River Restoration Centre. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

References

Naura, M. (2017) Mapping channel re-sectioning in England and Wales. Riverdene Consultancy.

Attribute information sheet

Principal habitat components: Rivers, headwater streams

Attributes derived from River Habitat Survey: P6 Habitat Modification Score, P7 Flow Habitat Mosaic, P8 Riparian Trees, P9 In-channel woody debris, P10 Riparian vegetation complexity

Rationale for inclusion: The ability of river and stream ecosystems to provide dynamic, complex and resilient habitat mosaics for characteristic biological assemblages is dependent on natural geomorphological processes and their intimate interaction with in-channel and riparian vegetation. This family of attributes covers key aspects of artificial physical habitat modification relating to impacts these natural processes.

- ***P6 Habitat Modification Score*** – This well-established index provides an aggregate view of range of artificial channel modifications such as resectioning and straightening.
- ***P7 Flow Habitat Mosaic*** – This attribute aggregates information on the meso-scale habitats occurring in the channel, providing an indication of habitat complexity which is broadly inversely associated with levels of physical modification.
- ***P8 Riparian Trees*** – This attribute characterises the extent of (and by extension the level of artificial absence of) riparian trees and their interaction with the river/stream channel, providing an indication of the level of naturalness of their ecological influence.
- ***P9 In-channel woody material*** – This attribute characterises the extent to which fallen tree trunks, boughs and small branches are left in the channel to play their critical natural role in shaping in-channel and riparian habitat mosaics.
- ***P10 Riparian vegetation complexity*** – This attribute characterises the extent to which riparian vegetation is allowed to express its natural complexity, in terms of variation in physical character.

Source data: Environment Agency River Habitat Survey database, part of the Agency's internal EcoSys tool.

Outline description of dataset including spatial coverage, representativeness, limitations: The source dataset is River Habitat Survey data fields, which provide information on a wide range of aspects of physical habitat and modification. For this interim version of the B6 indicator data have been sourced from previous RHS baseline surveys (1995/96 and 2007/08) and Countryside Surveys (2000 and 2007), which is assumed to provide a spatially representative picture of the English river and stream network as a whole. In the future the intention is that data will need to be sourced from Environment Agency national river and headwater stream surveillance programmes. These programmes are similarly designed to be representative of the river and stream network, and discussions are on-going about ensuring they are also representative of detailed habitat types assessed within the assessment framework used in B6.

Data ownership and licensing restrictions (if any): The RHS database owner is the Environment Agency. There is no public access to the EcoSys tool but some specific RHS data fields are open data.

Data transfer arrangements: To be agreed but the most likely route is via specific request to the Environment Agency to extract relevant surveillance data from EcoSys. It may be possible to organise pre-processing of B6 attributes prior to data transfer.

Frequency of source data update/data transfer: The Environment Agency undertake surveillance monitoring on a rolling basis so new data on a proportion of sites will become available each year. Updates to these B6 attributes would sensibly be at 3-5-year intervals.

Form of attributes:

- ***P6 Habitat Modification Score*** – Aggregate score of physical modifications present.
- ***P7 Flow Habitat Mosaic*** – Aggregate score of meso-scale habitats present.
- ***P8 Riparian Trees*** – Aggregate score of presence and habitat influence.
- ***P9 In-channel woody material*** - Aggregate score of presence and habitat influence.
- ***P10 Riparian vegetation complexity*** – Aggregate score of complexity values.

Data processing method for generating attribute output: RHS sites are resolved into WFD waterbodies, each divided into larger river and headwater stream zones. For each RHS site data are processed as below.

- ***Habitat Modification Score*** – This is a pre-processed index automatically generated within the Environment Agency RHS database. It is also pre-processed into five Habitat Modification Classes which are adopted as naturalness classes in the B6 attribute.
- ***P7 Flow Habitat Mosaic*** - This attribute uses data from a combination of the spot check and sweep up stages of the River Habitat Survey methodology. The flow types are recorded at each of the ten spot checks (free fall, chute flow, broken standing wave, rippled flow, upwelling, smooth flow, no flow, dry channel, not visible). Each flow type scores 1 if recorded in the reach, 2 if recorded at 2 or 3 spot checks, 3 if 4 or more spot checks. Dry river beds and 'not visible' occurrences score 0. Then, at the sweep up stage, 1 is added to the score for each flow type recorded that was not recorded in the spot checks, and another 1 is added for the occurrence of marginal dead-water. The maximum possible value of the score (indicating the highest diversity in current velocities and therefore habitat provision) is 14 (maximum from spot checks is 10, max from sweep up is 4).
- ***P8 Riparian trees*** - Data for this attribute are derived from the sweep-up stage of the River Habitat Survey method, specifically the 'trees' section. It is based on the presence and extent of 4 elements: shading of the channel, boughs overhanging the channel, bankside roots and submerged roots. Each is recorded as absent, present or extensive. RHS sites are classified into one of five naturalness classes according whether the 4 tree-related elements are present or not at the site, and if present, how many are extensive. Class 5 sites show none of the elements whilst in Class 1 at least 3 of the 4 elements are extensive.
- ***P9 In-channel woody material*** - Data for this attribute are derived from the sweep-up stage of RHS, specifically the 'trees' and 'special features' section. The attribute is based on the presence and extent of 3 elements: fallen trees, large woody material and debris dams. Each is recorded as absent, present or extensive. The attribute classifies sites into five naturalness classes according whether the 3 elements are present or not at the site, and if present, how many are extensive. Class 5 sites show none of the elements whilst in class 1 all 3 elements are extensive.
- ***Riparian vegetation complexity*** - Data for this attribute are derived from the spot check stage of the standard River Habitat Survey method. At each spot check the vegetation structure of both bank tops is assessed as bare (scores 0), uniform (scores 1), simple (scores 2) or complex (scores 3). The maximum possible value of the score is 60 (equivalent to complex vegetation on both banks at all 10 spot checks). Scores are banded equally into 5 naturalness classes, with Class 1 representing the highest bank vegetation complexity and Class 5 the lowest.

Naturalness class boundaries:

Attribute	Class 1	Class 2	Class 3	Class 4	Class 5
P6 Habitat Modification Score	<17	17-199	200-499	500-1399	>1400
P7 Flow Habitat Mosaic	12-14	9-11	6-8	3-5	0-2
P8 Riparian Trees	3 or 4 'extensive'	2 'extensive'	1 'extensive'	>= 1 'present'	All 'absent'
P9 In-channel woody material	3 'extensive'	2 'extensive'	1 'extensive'	>= 1 'present'	All 'absent'
P10 Riparian vegetation complexity	48-60	36-48	24-36	12-24	0-12

Attribute robustness: The dataset used for this version of the B6 indicator is deemed to be representative of the river and stream habitat resource as a whole but was not designed to be representative of individual river and stream habitat types included in the assessment framework used for B6 (such as chalk streams/rivers or active shingle rivers). It is intended that Environment Agency surveillance regimes will be refined to ensure that they provide a representative picture at the level of this detailed types.

Storage location for source dataset and processed data: Source data are stored in the Environment Agency's internal EcoSys tool. Summarised processed data on naturalness are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat components: Rivers, headwater streams, lakes

Attributes relating to FBA Naturalness assessment - Codes P11 (rivers and headwater streams) H3 (Headwater streams only), P6 (Lakes)

Rationale for inclusion: Citizen science is an increasingly important element of evidence gathering and strategic efforts need to be made to incorporate citizen science data into formal evaluations of the habitat resource. It is a challenging work area due to the increasing diversity of citizen science methods and issues associated with data quality assurance and securing a representative picture of the habitat resource. Simple methods for evaluating the naturalness of rivers, streams and lakes have been designed as part of a partnership project between Natural England and the Freshwater Biological Association (FBA) and a range of key partners including the Environment Agency. These methods are intended to provide both entry-level assessment of habitat naturalness for citizen scientists and a framework for resolving data from other citizen science and professional survey methods into the five-class classification used in the B6 indicator. Over time the database of naturalness assessments will grow as surveys are added directly and data from other survey methods are incorporated by 'read-across' processes.

Source data: The Data Portal of the FBA Discovering priority habitats website, although within this data are stored by Cartographer which provides the engine for the Data Portal. Arrangements are being made for the national dataset to be hosted by Natural England on its ARCGIS open data repository. Regular updates to the dataset would be made.

Outline description of dataset including spatial coverage, representativeness, limitations: The dataset provides naturalness class allocations for a range of river, stream and lakes sites across England. The naturalness framework is structured by the same key components as the framework used for the B6 indicator, i.e. hydrological, chemical, physical and biological. Only certain components of the naturalness database are currently included in the B6 assessment framework, selected because they are deemed to add most value:

- **Hydrological assessment of streams** – Included because available data on the naturalness of flow regimes are geared towards larger rivers. This component of the FBA naturalness assessment allows the identification of local water diversions and abstractions.
- **Physical assessment of rivers, streams and lakes** – Included because this component of FBA naturalness assessments is considered to be the most robust (simplest to evaluate) and can augment other B6 physical attributes, which are mainly representative in nature and rely on (inevitably) limited surveillance programmes.

Other components (particularly chemical naturalness) could be added in due course depending on how the FBA naturalness datasets grow over time and the success of securing read-across data transfers from related citizen science initiatives such as the River Fly partnership.

Data ownership and licensing restrictions (if any): The data owner is Natural England. The data are available for the public to view via the Display Data facility of the Discovering priority habitats website and in the future will be available to download under open government licence from Natural England.

Data transfer arrangements: Data can be downloaded directly from Cartographer as and when B6 assessments are required. No data download has yet been made because the database has not been

sufficiently populated to make an assessment. The database is now growing rapidly and it is anticipated that an assessment of these attributes could be made in Spring 2024.

Frequency of source data update/data transfer: The source dataset on the Data Portal is updated in real-time, so updates to the attribute can be annual if deemed necessary. However, an update frequency of 3-5 years would be most appropriate.

Form of attribute: Naturalness class.

Data processing method for generating attribute output: When the database is sufficiently populated and naturalness data on the relevant components has been extracted, the data will need to be filtered to extract a representative subset. This avoids any bias in the assessment due to variations in the spatial intensity of assessments. Data will be aggregated to waterbody-level (for river waterbodies data will be aggregated into larger river and headwater stream zones).

Naturalness class boundaries:

	Class 1	Class 2	Class 3	Class 4	Class 5
FBA naturalness class	1	2	3	4	5

Attribute robustness: The simplicity of the FBA naturalness assessment and the intended incorporation of data from other methods suggests relatively low data confidence at the level of individual survey sites. However, the robustness of the assessment at the habitat resource level should be considered to be considerably higher, in terms of providing a general picture across the habitat resource. Robustness is also critically affected by the spatial intensity of data – if data accumulate rapidly and relatively evenly across England then the representativeness of the assessment becomes more robust.

Storage location for source dataset and processed data: The storage location for source data should be considered to be Natural England's ARCGIS open data facility, but the dataset is not expected to be available there until Spring 2024. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat component(s): Rivers, headwater streams, lakes, ponds

Non-native species (Rivers/streams/lakes code B1, Ponds code B2)

This information sheet was prepared in collaboration with Gavin Measures (Natural England), Alice Hiley (Environment Agency) and Helen Roy and Steph Rorke (UKCEH).

Rationale for inclusion: Non-native species constitute a critical biological source of modification to ecosystem naturalness. Even if hydrological, chemical and physical components of naturalness are at high levels, non-native species can have highly damaging direct effects on native biological assemblages as well as effects on broader ecosystem integrity.

Source data: National Biodiversity Network, including most data available from UK Biological Records Centre.

Outline description of dataset including spatial coverage, representativeness, limitations: Point-based records (all available records, mostly post-1990 but with no fixed time cut-off) of all the freshwater and riparian non-native species on the impact lists maintained by the Water Framework Directive UK Technical Advisory Group (WFD UKTAG). Analysis of species records at different levels of spatial resolution (point, 1km, 2km and 10km grid cells) revealed that point records are the recording norm and excluding grid cell-based records does not affect the spatial coverage of the dataset. The UKTAG impact lists are used to screen species for relevance to freshwater habitats, using impact score as a surrogate for likely population abundance (extent of modification).

It had originally been hoped that we could use the species records dataset and associated data processing used in the production of Defra OIF indicator H2 (Distribution and spread of non-native species and diseases), since this would provide us with a pre-standardised and regularly updated dataset to plug into the B6 data framework. Indicator H2 is based on a long-standing UK biodiversity indicator on non-native species spread (Coded RB6), which is reported by the JNCC. Unfortunately the dataset used for RB6/H2 is too coarse for B6 purposes.

Data ownership and licensing restrictions (if any): NBN data are open to the public.

Data transfer arrangements: Records for relevant species can be extracted from the NBN at any time. All point records for freshwater and riparian species on the most recent version of the UKTAG impact lists will need to be extracted.

Frequency of source data update/data transfer: The source databases are updated as records are added so updates to the attributes can be annual if deemed necessary. However, an update frequency of 3-5 years would be more appropriate, sensibly preceded by a proactive trawl for new records from key organisations. Ideally older records would be filtered out of the dataset as new data are added, but this can only be done in a robust way if new records (of both species absence and presence) replace old records of presence, or if there is some other valid reason to assume that a species has disappeared from a location. This will have to be considered at the update stage but the default position is that old records are retained in the dataset, on the basis that methods for eradication are highly limited.

Form of attribute: Aggregate presence of non-native species in each waterbody or (for rivers/streams) waterbody zone, weighted by the UKTAG impact score for each species.

Data processing method for generating attribute output: Records for each species are overlain on the digital river, stream and lake networks. A 20-metre spatial buffer is then used to allocate point records to each digital network (buffers of various sizes were analysed but 20 metres provided the best fit). Records within the spatial buffers for rivers, streams or lakes are assigned to those ecosystem types. Records outside of these buffers are nominally assigned to ponds (NB it is recognised that many of these latter records will originate from ditches, but a distinction between ponds and ditches is not practical for the purposes of this attribute).

Relevant records are then aggregated to the following spatial units and a naturalness class is allocated to each unit: 1) river and stream records aggregated to river/stream waterbody zones; 2) lake records aggregated to lake waterbody (smaller lakes are aggregated to waterbody catchment where appropriate); 3) putative pond records are aggregated to 10km grid cells.

Naturalness class boundaries:

Class 1	Class 2	Class 3	Class 4	Class 5
No species on UK TAG lists	Only low impact species	Only low or moderate impact species	No more than 1 high impact species	More than 1 high impact species

Attribute robustness: The robustness of the non-native species attributes is highly influenced by the intensity and distribution of species recording. Most records for freshwater non-native species (particularly invertebrates) come from the Environment Agency, added to the NBN in batches at intervals. There is potential to provide a more representative picture by using records gathered as part of the Environment Agency's future environmental surveillance programme, since this programme is being designed to provide a representative picture of the freshwater habitat resource. However, this will need further consideration in terms of the coverage of species and spatial coverage of the habitat resource. A hybrid data model would still be necessary, involving EA surveillance data to provide a robust core and *ad hoc* recording data to enhance the dataset, increasingly making use of citizen science and eDNA techniques. The treatment of ponds in the GIS analysis of records is a weakness, in that there is no discrimination between ponds and ditches. This could be rectified through the use of a spatial layer of ponds.

Storage location for source dataset and processed data: Source data are stored in the NBN and Biological Records Centre databases. Summarised processed data are stored in a structured spreadsheet of naturalness scores attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat components: Rivers, headwater streams

Macroinvertebrate similarity index (CC1)

This information sheet was prepared in collaboration with John Murray-Bligh and Judy England (Environment Agency) and Andrew Johnson, Lee Brown and Megan Klaar (University of Leeds)

Rationale for inclusion: A number of macroinvertebrate metrics are currently used as part of assessing ecological status under the Water Framework Directive However, the headline metrics used for reporting purposes (for which the most comprehensive datasets are available) are geared towards detection of organic pollution, and more generally water quality (chemical naturalness) rather than holistic evaluation of the naturalness of the species assemblage. Other macroinvertebrate metrics have been developed to indicate other pressures (such as changes in current velocities and fine sediment deposition); data on these are more restricted and again none of them assesses changes to the biological assemblage directly.

Attribute CC1 is intended to provide direct characterisation of the level of deviation of the observed macroinvertebrate species assemblage from the assemblage expected under unimpacted reference conditions. It therefore provides a more biodiversity-relevant assessment, which is not only valuable for biodiversity reporting purposes but also better reflects the broad intentions of the Water Framework Directive. Since this attribute should be capable of reflecting impacts on any of the key components of naturalness (hydrological, chemical, physical or biological), it has been placed in the cross-cutting category of attributes. However, its ability to do this depends not only on the nature of the attribute but also the nature of the monitoring regime that generates the data.

Source data: Data on observed macroinvertebrate assemblages come from Environment Agency biological surveillance programmes – data need to be at species-level as far as this is practical (a small number of taxonomic groups are difficult to identify to species-level). These observed data need to be compared with parallel data on the reference (unimpacted) assemblage, which come from model predictions based on the natural environmental characteristics of each monitoring site. Data from Environment Agency surveillance programmes are available on BIOSYS. Data on reference communities can be generated by the predictive model RIVPACS (River Invertebrate Prediction and Classification System). Data tailored for use in this attribute can be generated by a new tool developed by the University of Leeds in association with the Environment Agency (Johnson *et al*., in prep). This tool represents an expansion of the current RIVPACS model and can produce reference assemblage predictions for sites in any region of England. An assessment of national trends in macroinvertebrate similarity has been undertaken for a 10-year period between 2010-2019 as part of Leeds' research programme.

Outline description of dataset including spatial coverage, representativeness, limitations: The Environment Agency has a national surveillance network designed to provide representative data suitable for evaluating the river and stream network as a whole. Historically such monitoring has been biased towards sizeable rivers at lower altitudes within intensive agricultural land uses, but reforms to the surveillance network are providing the basis for a more representative picture suited to the needs of the assessment framework used for the B6 indicator. This said, it is important to understand that monitoring and therefore assessment is restricted to in-channel and marginal assemblages – it does not extend to riparian or other ephemeral habitats which contain assemblages of high biodiversity importance, such as wetland vegetation and seasonally exposed shingle shoals.

The environmental data needed to generate predictions of the reference (in-channel) assemblage are recorded for each site within the Environment Agency's surveillance programme.

There are some issues to resolve that require some further work before the attribute can be populated with data and the limitations of the assessment can be properly characterised. These include the way in which data on predicted reference assemblages are generated and the ability of survey methods to characterise impacts relating to the overall spatial extent river and stream habitats and detailed character of meso-scale habitat mosaics.

The tool generated by the University of Leeds employs Monte Carlo simulation to compare the observed assemblage to the range of possible reference assemblages that might be present at a site, usefully accounting for the dynamic nature of natural fluvial environmental conditions and the macroinvertebrate assemblages they support. The tool js capable of using any similarity calculation or community standardisation technique, which can assign different importance to rare or common taxa. An understanding of what aspects of freshwater macroinvertebrate assemblages are most important for assessing river health, and how similarity changes as rivers improve in health, is required. More fundamentally, Environment Agency surveillance methods are designed with water quality assessment in mind, providing an integrated but undifferentiated assessment of the assemblages utilising the in-channel habitat mosaics provided by rivers and streams. These methods are not designed to detect impacts on the natural spatial extent and character of the habitat mosaic, of the type generated by hydrological and physical modifications (abstraction, water diversion, channel engineering etc.).

Data ownership and licensing restrictions (if any): The owner of data on observed assemblages is the Environment Agency. Data for the attribute are available to the public on the Ecology and Fish Data Explorer https://environment.data.gov.uk/ecology/explorer/. Data on predicted reference assemblages need to be generated and the intention is that the tool generated by the University of Leeds is used for this purpose. The creator of the tool (Andrew Johnson) and the university are making the tool available for this and other purposes.

Data transfer arrangements: Data on observed assemblages will be extracted from the Environment Agency's internal BIOSYS database. No arrangements have yet been made for generating predicted reference assemblages – further work on the precise form of the attribute and the approach to prediction will be required before this is possible. The Leeds tool could potentially be incorporated into an on-line tool called RICT (River Invertebrate Classification Tool), which is used for calculating macroinvertebrate metrics for classifying ecological status.

Frequency of source data update/data transfer: BIOSYS is updated frequently so updates to the attribute can be annual if deemed necessary, although data on individual sites are updated on a rolling basis so only a proportion of the results for monitoring sites are updated each year. An update frequency of 3-5 years is considered appropriate.

Form of attribute: Similarity of observed macroinvertebrate assemblage to the predicted reference (unimpacted) assemblage.

Data processing method for generating attribute output: Comparison of the observed and predicted reference assemblages at each surveillance site will be made using one or more standard similarity indices. The Bray-Curtis index is most likely to be employed but further thought needs to be given to data transformation techniques that place different levels of emphasis on rare and characteristic species. For each surveillance site in the national dataset, one or more indices will need to be generated for the time period being assessed (using data from the most recent 3-5 years

depending on the attribute update frequency). Monitoring sites would then be resolved into the spatial framework of WFD waterbodies, each divided into river and headwater stream zones. Where more than one site occurs within a waterbody zone the site exhibiting most deviation from the predicted reference community would be selected to represent the zone.

Naturalness class boundaries:

	Class 1	Class 2	Class 3	Class 4	Class 5
Similarity Index value	TBA	TBA	TBA	TBA	TBA

Attribute robustness: The robustness of this attribute will ultimately be dependent on the shape of future Environment Agency surveillance programmes. The density of macroinvertebrate monitoring sites and the adequacy of coverage of detailed river/stream types are key factors. It remains to be seen if there is scope to modify site-level surveying methods to allow better characterisation of impacts on hydrological and physical naturalness. Without changes to survey methods the attribute is likely to be limited in its sensitivity, placing heavy reliance on direct indicators of hydrological and physical naturalness to detect these impacts. The necessary incorporation of all possible dynamic expressions of the natural assemblage within the prediction of reference assemblages will make it more difficult to detect human modifications over and above natural variation, unless the characterisation of the assemblage can be made more nuanced.

Storage location for source dataset and processed data: Source data on observed assemblages is available from the Environment Agency's BIOSYS database. In the future, data on predicted reference assemblages will hopefully be available via RICT. Summarised processed data of similarity index values will eventually be added to the structured spreadsheet attached to the B6 information pack. As part of the publishing Leeds research, computer code (R environment) for calculating similarity indices will be made openly available for use with future biological monitoring datasets. However, the code could potentially be developed into a user-friendly interface to allow wider use by non-programmers. It is possible that this tool will be integrated into the RICTs implementation of the RIVPACS model as a functional tool accessible to non-programmers.

References

Johnson *et al*. (in prep) Assessing the use of taxonomic targets as a candidate for freshwater management.

Attribute information sheet

Principal habitat components: Headwater streams

Catchment land cover (Code CC2)

Rationale for inclusion: The land use of the catchment is often used as a surrogate for a wide range of pressures and impacts on freshwater ecosystems, particularly in international assessments where more detailed datasets are lacking. For this assessment framework, the information available on rivers is considered to provide a more nuanced evaluation that possible with catchment land use information. However, we have less information available to evaluate headwater streams, as well as biodiversity ambitions to protected or restore high levels of natural or semi-nature land cover in a significant proportion of the headwater catchment resource. It therefore makes sense to include an attribute for headwater streams that measures this directly.

Source data: The Living England map has been used as the source data layer, since this will be a key vehicle reporting change in the countryside in the future and will be updated on suitable timescales to enable this. The dataset provides coverage of the whole of England at 50-metre spatial resolution, which is considered adequate for this purpose.

Data ownership and licensing restrictions (if any): The dataset is licensed by Natural England under open government licence.

Data transfer arrangements: An up-to-date version of the dataset is available to Natural England and can be used at any time.

Frequency of source data update/data transfer: An update frequency of 3-5 years is considered appropriate but the precise frequency will need to coincide with updates to the Living England layer.

Form of attribute: Naturalness of land cover in headwater catchments

Data processing method for generating attribute output: The following land cover types within the Living England map have been allocated to an aggregate natural/semi-natural land cover category:

- Acid, Calcareous, Neutral Grassland
- Bare Ground
- Bare Sand
- Bog
- Bracken
- Broadleaved, Mixed and Yew Woodland
- Coastal Saltmarsh
- Coastal Sand Dunes
- Dwarf Shrub Heath
- Fen, Marsh and Swamp
- Scrub
- Water

Data on this aggregated category, and on the urban land cover category, are resolved into the spatial framework of headwater zones of waterbody catchments widely used in this assessment framework. For the headwater zone of each waterbody catchment, the total area under natural/semi-natural land cover is calculated and divided by the total area of the headwater zone. The same calculation is made for the urban land cover category. Each waterbody is then allocated to naturalness classes

using the class boundaries below. The overall value for the attribute for each waterbody catchment is taken as the worst class of % natural/semi-natural and % urban.

Naturalness class boundaries:

Sub-attributes	Class 1	Class 2	Class 3	Class 4	Class 5
1. % natural/semi-natural vegetation	100	80-100	50-80	25-50	<25
2. % urban	0	0-5	5-10	10-25	>25

Attribute robustness: This whole-inventory attribute is considered to provide a robust assessment of land cover within headwater catchments across the habitat resource. The updated frequency of the Living England layer should be adequate to provide updates of the attribute at suitable intervals.

Storage location for source dataset and processed data: Source data on land cover are available from the Living England layer. Summarised processed data of similarity index values will eventually be added to the structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat component: Lakes

Hydrological attributes: Naturalness of flow at lake outflow, Lake codes H1-H4

Rationale for inclusion: Natural hydrological regimes are fundamental to healthy lake ecosystems. Both extreme fluctuations and loss of fluctuations can potentially cause the loss of species. Residence times and flushing rates also influence water quality. These attributes provide an assessment of the levels of deviation from the natural flow regime in the outflow of the lake, at different flow magnitudes.

Source data: Environment Agency Water Resources GIS system.

Outline description of dataset including spatial coverage, representativeness, limitations:

The dataset provides actual flows and modelled naturalised flows at a number of naturalised flow (Qn) values. For the interim version of the indicator, data on actual flows are based on abstraction returns from the period 2013-2018 for most catchments. A single set of values is provided for each Water Framework Directive waterbody.

Data field	Description
EA_WB_ID	Water body ID number.
WB_NAME	Waterbody name.
Type_IWB	Type of waterbody.
CATCHMENT	Name of catchment.
ScenRA%QN30	Recent actual scenario as a percentage of natural flows at Q30.
ScenRA%QN50	Recent actual scenario as a percentage of natural flows at Q50.
ScenRA%QN70	Recent actual scenario as a percentage of natural flows at Q70.
ScenRA%QN95	Recent actual scenario as a percentage of natural flows at Q95.

For these lake attributes, the dataset is filtered using waterbody type, selecting only lake waterbodies (583 in all). Since the dataset is based on a network of river gauging stations, it can only provide information on on-line lakes where the hydrology of the lake is directly influenced by the flow regime of the river network. Generally the dataset provides more information of large lakes, and many of these are artificial in nature (man-made reservoirs including artificially impounded rivers). This means that the dataset does not provide good representation of the many small lakes which form the large majority of the habitat resource by number. The coverage of non-WFD lakes will vary across Environment Agency regions, according to abstraction management in that area.

Data ownership and licensing restrictions (if any): The data owner is the Environment Agency. The data are available under open government licence.

Data transfer arrangements: A standard data specification (above) has been agreed with the Environment Agency.

Frequency of source data update/data transfer: Data on individual catchments are updated as and when possible. An update frequency of 3-5 years is considered appropriate.

Form of attribute: Calculated as the percentage deviation (negative or positive) of actual daily river flow from the modelled naturalised flow (flow in the absence of abstractions and discharges).

Data processing method for generating attribute output: Data are pre-processed by the Environment Agency within their Water Resources GIS system.

Naturalness class boundaries:

Qn	Class 1	Class 2	Class 3	Class 4	Class 5
95	<5	5-10	10-25	25-40	>40
70	<5	5-10	10-25	25-40	>40
50	<5	5-10	10-25	25-40	>40
30	<5	5-10	10-25	25-40	>40

Attribute robustness: The dataset is not able to characterise the effect of non-consumptive abstractions above the assessment points used in the EA WRGIS. Greater spatial discrimination in the dataset would therefore be preferable and is possible through additional modelling, e.g. via the EA's new Groundwater Hydro-ecology tool. The dataset is also based on comparison of frequency distributions of recent actual and naturalised flows, not a direct comparison of time series of data. Evaluation of data as time series would provide a more realistic picture of hydrological modification. As these data are derived from river locations adjacent to on-line lakes, they are only currently possible to produce for on-line lakes, i.e. they only represent part of the overall lake habitat resource.

Storage location for source dataset and processed data: To be arranged. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack.

Attribute information sheet

Principal habitat component: Lakes

Chemical attributes - Total phosphorus (C1); Total Nitrogen (C2); Acid Neutralising Capacity (ANC); Chlorophyll (C4); Macrophytes and diatoms (C5)

Rationale for inclusion: High water quality is a critical requirement for protecting and restoring characteristic biological communities including priority species. Nutrient status is a key factor, and nutrient enrichment is implicated in a range of ecosystem effects. Other water quality issues include acidification, and toxic pollution. These attributes have been selected to provide a broad indication of chemical naturalness, using chemical determinands for the most basic elements of chemical naturalness and biological metrics to indicate other chemical issues (hazardous chemicals, episodic pollution). Hazardous chemicals are also given more specific consideration under attribute C6).

Source data: Environment Agency Catchment Data Explorer.

Outline description of dataset including spatial coverage, representativeness, limitations: The dataset provides the current ecological status class of each attribute for each lake waterbody in the EA's Water Framework Directive monitoring programme.. For the interim version of the indicator, the 2019 version of the dataset (the most recent available) has been used. WFD lake sampling is focused primarily on lakes over 50 ha and therefore provides a non-representative sample of the lake habitat resource as a whole. It includes many SSSIs and reservoirs, so potentially over-emphasising those sites where measures are implemented to improve water quality, but it is the most comprehensive data set available on water quality. The biological metric included in this information sheet does not measure the naturalness of the biological assemblage per se, and instead is geared towards detecting pollution stress, which is why it is included here. Outside of protected sites and WFD monitoring there is little other data to draw on, making assessment difficult without further monitoring. Environment Agency monitoring design is changing to a representative surveillance network which improve the picture..

Data ownership and licensing restrictions (if any): The data owner is the Environment Agency. The data are available to the public via Catchment Data Explore under open government licence.

Data transfer arrangements: Data are downloaded directly from Catchment Data Explorer.

Frequency of source data update/data transfer: In future data will be sourced from the EA's new representative surveillance programme, which will monitor sites on a rolling basis and complete a full cycle every 5 years. It would therefore be sensible to update these attributes on a 5-year cycle.

Form of attribute: Ecological status class of each attribute, which is related to naturalness class directly.

Data processing method for generating attribute output: The data for each lake waterbody are pre-processed into ecological status class allocations within Catchment Data Explorer.

Naturalness class boundaries:

	Class 1	Class 2	Class 3	Class 4	Class 5
Ecological status	High	Good	Moderate	Poor	Bad/Fail

Attribute robustness: The robustness of these attributes will be dependent on the shape of future Environment Agency surveillance programmes. The density of monitoring sites, the adequacy of coverage of detailed lake types, and the coverage of individual determinands are all key factors. In the future the intention is that data will need to be sourced from Environment Agency national lake surveillance programmes. These programmes are similarly designed to be representative of the lake habitat resource, and discussions are on-going about ensuring they are also representative of detailed habitat types assessed within the assessment framework used in B6.

Storage location for source dataset and processed data: Source dataset currently stored in Catchment Data Explorer. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat component: Lakes

Chemical attribute Fish e-DNA, Code C7

Rationale for inclusion: The fish assemblage has a major bearing on the natural functioning of lake ecosystems, particularly in shallow water lakes where their natural balance is critical to the maintenance of a clear-water state that allows submerged plant communities and their associated fauna to thrive. Owing to uncertainties around the ability of this metric to reflect alterations to natural assemblages, it has been placed in the chemical naturalness component as an indicator of water quality.

Source data: This dataset is in development. It has been provided by the Environment Agency.

Outline description of dataset including spatial coverage, representativeness, limitations: The dataset currently includes data from 49 lakes. It provides an ecological status class for the fish assemblage for each lake, based on comparison of e-DNA samples with an expected e-DNA profile under reference conditions.

Data ownership and licensing restrictions (if any): The Environment Agency, available under licence as a dataset in development.

Data transfer arrangements: Future transfer arrangement to be arranged once data from a surveillance programme are generated.

Frequency of source data update/data transfer: To be arranged but probably every 5 years as part of future Environment Agency environmental surveillance programmes.

Form of attribute: Ecological status class.

Data processing method for generating attribute output: The data for each monitoring site are pre-processed into ecological status class allocations, which are converted into naturalness classes as outlined below.

Naturalness class boundaries:

	Class 1	Class 2	Class 3	Class 4	Class 5
Ecological status	High	Good	Moderate	Poor/Bad	-I

Currently no discrimination is provided of poor and bad ecological status. For the interim indicator these status classes have nominally been assigned to naturalness Class 4. After further development of the e-DNA metric it is anticipated that these classes will be discriminated and naturalness classes will be assigned as for other attributes depending on ecological status class (i.e. Poor – Class 4 and Bad = Class 5).

Attribute robustness: The current robustness of the attribute is judged to be low as the e-DNA metric is still in development and the dataset covers only a small set of lakes.

Storage location for raw dataset and processed data: The raw dataset should be requested from the Environment Agency. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat component: Lakes

Physical attributes derived from Lake Habitat Survey (LHS) P1 Hydrological structures; P2 Artificial shoreline; P3 Sediment fluxes; P4 Riparian habitat; P7 Marginal fringe emergent vegetation

Rationale for inclusion: The ability of lake ecosystems to provide dynamic, complex and resilient habitat mosaics for characteristic biological assemblages is dependent on natural geomorphological processes and their intimate interaction with in-lake and marginal and riparian vegetation. This family of attributes covers key aspects of artificial physical habitat modification relating to impacts on these natural processes.

The littoral zone substrate is essential for fish spawning, invertebrate diversity and abundance and macrophyte anchorage and nutrition. It is impacted by sedimentation and reduction in substrate heterogeneity due to water level manipulations, as well as the introduction of artificial substrates for various reasons.

Riparian habitat has intrinsic conservation value as part of the lake habitat, supporting a range of characteristic species. Riparian vegetation has been lost and reduced through drainage of riparian land and alternative land use in land adjacent to lakes.

Source data: Environment Agency Lake Habitat Survey database. The interim version of the indicator uses data from 2012 as this is the most recent data available.

Outline description of dataset including spatial coverage, representativeness, limitations: The Lake Habitat Survey method (Rowan *et al*., 2006, 2008) is broadly equivalent to River Habitat Survey but has not yet been operationalised in the same way. It provides data on a range of variables, some of which are extracted for the purposes of B6. The sites included in the available database were not selected as a representative sample of English lakes and so do not reflect the condition of the habitat resource as a whole. Some sites had been chosen specifically to trial the LHS method, representing the full range of naturalness from artificial to natural lakes including the extremes. Unlike RHS there has been no accreditation for LHS so inter-surveyor variation may be expected to be higher. There are also no reference sites, and consequently no method of assessing what could naturally be expected at a site. Despite these drawbacks it is the most comprehensive dataset and most worked-up method for assessing lake physical habitat. The assessment of this attribute in the interim version of the B6 indicator should therefore be seen as illustrative, which will be improved on over time as LHS is refined and operationalised.

Data ownership and licensing restrictions (if any): The Environment Agency. Some LHS data are available under open government licence from the Agency's data portal.

Data transfer arrangements: Future transfer arrangement to be arranged once data from a surveillance programme can be generated.

Frequency of source data update/data transfer: To be arranged but probably every 5 years as part of future Environment Agency environmental surveillance programmes.

Form of attribute:

- ***P1 Hydrological structures*** - Within LHS the presence of a range of structures are recorded. An estimate of the maximum height from the bed to top of principal retaining structure is also recorded. The attribute is assessed by combining these two pieces of information as described in the naturalness class boundaries table below.
- ***P2 Artificial shoreline*** - LHS records the extent of artificial bank construction across multiple perimeter sections that cover at least seventy-five percent of the shoreline. The extent of artificial bank construction along the shoreline is summed and weighted based on the length of the shoreline section surveyed. Naturalness class values are assigned to percentage values of artificial shoreline as described in the naturalness class boundaries table below.
- ***P3 Sediment fluxes*** - The presence of signs of sedimentation or depositional imbalances are recorded in LHS and are used to assess this attribute as described in the naturalness class boundaries table below.
- ***P4 Riparian habitat*** - LHS includes data on riparian land cover 15 metres from the lake edge, which is used to assess naturalness as described in the naturalness class boundaries table below.
- ***P7 Marginal fringe emergent vegetation*** - The ambition for this metric is to record the extent of emergent vegetation as a percentage of the lake perimeter, but this is not recorded at present as part of LHS. The presence of certain types of emergent vegetation are currently recorded in LHS, particularly reed beds and floating vegetation mats, but this excludes other emergent vegetation particularly those found in more nutrient-poor water bodies (such as *Eleocharis* and *Equisetum*). Consequently, at present only eutrophic and mesotrophic lakes are assessable using the data available from LHS. Discussions are needed to incorporate measurement of all marginal emergent vegetation types (as a percentage of lake perimeter) within the NCEA lake survey method.

Data processing method for generating attribute output:

P1 Hydrological structures –For the interim indicator the following fields were extracted and processed from the LHS 2012 Access database: "Max.height.of.retaining.structure", "Dams.no.fish.pass", "Dams.with.fish.pass","Barrage","Sluice","Lock","Weir". Sites were allocated to naturalness classes as follows:

- Class 1: No structures ie sum of all counts in columns "Dams.no.fish.pass","Dams.with.fish.pass", "Barrage","Sluice","Lock",and "Weir" is ZERO and max height is zero
- Class 2: Small structures <50cm ie Sum of all columns >0 AND "Max.height.of.retaining.structure" within 0-0.5m
- Class 3: Structure 50cm-1m ie Sum of all columns >0 AND "Max.height.of.retaining.structure" within 0.5-1m
- Class 4: Structure 1m+ ie "Dams.no.fish.pass"=0 AND Sum of other columns >0 AND "Max.height.of.retaining.structure" >1m
- Class 5: Water level control structure (no fish pass) 1m+ ie "Dams.no.fish.pass" >0 AND "Max.height.of.retaining.structure" >1m.

For some of the 2012 LHS surveys the height of a structure was recorded but not its presence, and for others there was presence but no height. In order to populate the wheel diagram such sites were assigned to naturalness class 3, but in the future it is important to ensure these data are collected systematically through the Environment Agency's lake surveillance programme.

P2 Artificial shoreline - Based on table “3 Shoreline survey information” from LHS 2012 database. For each lake, shorelines were surveyed sometimes partially (e.g. 80% of shoreline only), or per chunk (e.g. 4 distinct surveys to survey 100% of lake shoreline). All shoreline survey entries were extracted for each lake and discarded lakes the shoreline of which was surveyed at less than 75%. Then fields were extracted that specifically characterise artificial shoreline: impoundments, hard or soft engineering, docks, marinas, jetties, all within 15 m of shoreline. For each individual survey of a partial shoreline, and each item (e.g. docks), the database provides a range of the total impact on the partial shoreline (e.g. 10-40% of partial shoreline is impacted by docks). We converted the ranges into an actual figure by selecting the mid-point of each range. We then assigned Class 1-5 as described in the table of class boundaries below.

P3 Sediment fluxes – This used a combination of data from tables “2 Hab-Plot information” and "3 Shoreline survey information”. From the shoreline surveys, the process was identical to that for P2, P4 and P7 but using the field erosion within 15 m of shoreline. This was re-classified as described in table of class boundaries below, as the first criteria “% of shore affected by erosion”. From the Hab Plot table, information was extracted for each habitat plot for each lake about the geomorphological imbalance (ie erosional or depositional imbalance, whether there is active loss eroding or active gain deposition) and about visible sediment deposition over substrate (any type of sediment). These three criteria were combined to assign classes as described in table of class boundaries below.

P4 Riparian habitat - Like for P2, the analysis for the interim indicator used table "3 Shoreline survey information” and the partial shoreline surveys for each lake. The following fields characterising riparian habitat were used: wet woodland, bog, fen/marsh, broadleaf mixed woodland, scrub/shrubs, moorland heath, open water, rough grassland, tall herb, rock/scree/dunes, all within 15 m of shoreline. Information was then processed as for P2. The resulting data were then classified as described in the table below.

P7 Marginal fringe emergent vegetation - P7 was derived like P2 and P4 but using fields relevant to marginal vegetation: emergent reed bed within 15 m of shoreline and floating vegetation mats within 15 m of shoreline. It is not expected that lakes would necessarily have emergent vegetation around their entire perimeter, as this can naturally be limited by factors such as exposure, substrate and lake profile. To take this into account lakes with more than 40% of their perimeter supporting emergent vegetation are assigned to Class 1 (see the naturalness class boundaries table below). Whilst some lakes will support a significantly more expansive emergent fringe than this, the class boundaries ensure lake-wide pressures are not causing significant declines. To allow this attribute to be included in the wheel diagram for lakes, oligotrophic lakes have been removed from the current LHS data and assessment has been undertaken on mesotrophic and eutrophic lakes only. This creates a known bias in the data but the wheel diagram is only for illustrative purposes at this stage.

Naturalness class boundaries:

Attribute	Class 1	Class 2	Class 3	Class 4	Class 5
P1 Hydrological structures (Structure type/ height)	No structures	Small structure <50cm	Structure 50cm-1m	Structure 1m+	Water level control structure (no fish pass) 1m+
P2 Artificial shoreline (% artificial shoreline)	0%	≥0 ≤5 %	>5 ≤33.3 %	>33.3 ≤ 66.7 %	>66.7 ≤ 100 %
P3 Sediment fluxes (Evidence of artificial sediment flux)	<5 % of shore affected by erosion AND signs of sedimentation or depositional imbalance recorded at < 2 Hab-plots AND sedimentation over natural substrate recorded at < 2 Hab-plots	≥ 5% < 20 % of shore affected by erosion OR signs of sedimentation or depositional imbalance recorded at 2 Hab-plots OR sedimentation over natural substrate recorded at 2 Hab-plots	≥ 20% < 40 % of shore affected by erosion signs of sedimentation or depositional imbalance recorded at 3-4 Hab-plots OR sedimentation over natural substrate recorded at 3-4 Hab-plots	≥ 40% < 60 % of shore affected by erosion OR signs of sedimentation or depositional imbalance recorded at 4-6 Hab-plots OR sedimentation over natural substrate recorded at 4-6 Hab-plots	≥ 60 % of shore affected by erosion OR signs of sedimentation or depositional imbalance recorded at >6 Hab-plots OR sedimentation over natural substrate recorded at >6 Hab-plots
P4 Riparian habitat (Riparian land use 15m from lake edge)	Riparian land is all semi-natural.	Riparian land is predominantly semi-natural (90%).	Riparian land semi-natural for at least 2/3 of its extent	Riparian land semi-natural for at least 1/3 of its extent	Riparian land semi-natural for less than 1/3 of its extent
P7 Marginal fringe emergent vegetation (% of perimeter supporting emergent marginal fringing vegetation)	40% +	>30<40%	>20-<30%	>10<20%	<10%

Attribute robustness: The robustness of these attributes will be dependent on the shape of future Environment Agency surveillance programmes. The density of monitoring sites, the adequacy of coverage of detailed lake types, and the coverage of individual determinands are all key factors. In the future the intention is that data will need to be sourced from Environment Agency national lake surveillance programmes. These programmes are designed to be representative of the lake habitat resource, and discussions are on-going about ensuring they are also representative of detailed habitat types assessed within the assessment framework used in B6.

Storage location for raw dataset and processed data: The raw dataset should be requested from the Environment Agency. Summarised processed data are stored in a structured spreadsheet attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

References:

Rowan, J.S., Soutar, I., Bragg, O.M., Carwardine, J & Cutler, M.E.J. (2006) Lake Habitat Survey in the United Kingdom: field survey guidance manual, version 3.1, May 2006. SNIFFER project WFD42. Scotland and Northern Ireland Forum for Environmental Research, Edinburgh.

Rowan, J.S., Soutar, I., Bragg, O.M., Carwardine, J & Cutler, M.E.J. (2008) Lake Habitat Survey Field Survey Form, version 4, December 2008. Scotland and Northern Ireland Forum for Environmental Research, Edinburgh.

Attribute information sheet

Principal habitat component: Lakes

Physical attribute P5 Riparian Trees

Rationale for inclusion: Riparian trees have an important role to play in providing habitat and food source to in-lake assemblages, but heavy shading across substantial parts of the lake habitat resource would be detrimental to biodiversity. This attribute seeks to ensure that there is a suitable ecological balance across the habitat resource as would be expected within naturally functioning ecosystems.

Source data: Not currently collected

Outline description of dataset including spatial coverage, representativeness, limitations: Lake Habitat Survey (LHS - Rowan *et al*., 2006, 2008) collects data on the presence of trees from approximately 10 'hab-plots' per lake, which are 15m wide so can represent only a small fraction of the shoreline. LHS does record the presence of woodland in the riparian zone but this does not include the occurrence of riparian trees outside woodlands. LiDAR could potentially be used to identify where trees surround lakes, but regular updates of LiDAR datasets across the habitat resource are not a practical proposition. Consequently, there are no data currently available to assess this attribute. An assessment of the percentage of perimeter of the lake with riparian trees could relatively easily be incorporated in LHS, which records other attributes for the lake perimeter.

Data ownership and licensing restrictions (if any): Not applicable.

Data transfer arrangements: To be arranged when data sources are in place.

Frequency of source data update/data transfer: To be arranged but probably every 5 years as part of future Environment Agency environmental surveillance programmes.

Form of attribute: Percentage of lake perimeter with riparian trees.

Data processing method for generating attribute output: To be arranged.

Naturalness class boundaries: To be arranged.

Attribute robustness: The robustness of this attribute will be dependent on the shape of future Environment Agency surveillance programmes. The density of monitoring sites, the adequacy of coverage of detailed lake types, and the coverage of individual determinands are all key factors. In the future the intention is that data will need to be sourced from Environment Agency national lake surveillance programmes. These programmes are similarly designed to be representative of the lake habitat resource, and discussions are on-going about ensuring they are also representative of detailed habitat types assessed within the assessment framework used in B6. Possibilities for using remote sensing in the future.

Storage location for raw dataset and processed data: To be arranged.

References:

Rowan, J.S., Soutar, I., Bragg, O.M., Carwardine, J & Cutler, M.E.J. (2006) Lake Habitat Survey in the United Kingdom: field survey guidance manual, version 3.1, May 2006. SNIFFER project WFD42. Scotland and Northern Ireland Forum for Environmental Research, Edinburgh.

Rowan, J.S., Soutar, I., Bragg, O.M., Carwardine, J & Cutler, M.E.J. (2008) Lake Habitat Survey Field Survey Form, version 4, December 2008. Scotland and Northern Ireland Forum for Environmental Research, Edinburgh.

Attribute information sheet

Principal habitat component: Ponds

Chemical attributes Phosphorus (C1); Nitrogen (C2); Physical attributes Semi-natural land-use (P3); Shading (P4); Grazing intensity(P5); Biological attribute PSYM (B1)

Rationale for inclusion:

- ***C1 Phosphorus and C2 Nitrogen*** - Water quality is a fundamental characteristic that supports natural biological communities and ecological processes. Potential pollutants include nutrients, sediment and chemical contaminants but nutrient pollution is a common cause of impact on ponds.
- ***P3 Semi-natural land-use*** – Ponds can be impacted by land use in the immediate vicinity and further away, particularly with respect to water quality. Adjacent terrestrial habitats are also ecologically critical in allowing species inhabiting ponds to fulfil their full life cycles.
- ***P4 Shading, P5 Grazing intensity*** - Over-shading of ponds has led to a loss of early successional ponds and pond species at the landscape scale. However, shaded ponds in woodland landscapes are also of value. Heavy grazing intensity and shading have been found to have considerable impacts on ponds and have been highlighted in recent declines. However, there is not a simple relationship where a single level of grazing or shading will lead to good quality ponds. Instead, for biodiversity to thrive a variety of levels of grazing and shading are required across the pond resource. Grazing leads to areas of open habitat and even bare mud where there is poaching. Whilst this is not necessarily desirable across the whole of the habitat resource it is a natural and essential element for many species.
- ***B1 PSYM*** - Distinct habitats have characteristic species assemblages that contribute to overall biodiversity. Having a characteristic assemblage is an indicator that the habitat/ecosystem is functioning naturally, as long as the observed assemblage can be reliably compared to the assemblage that would occur under natural environmental conditions.

Source data: Data on ponds will be collected through a new national pond surveillance programme in the future, but in order to illustrate the use of these attributes Countryside Survey data from 2007 has been used to populate the wheel diagrams where possible.

Outline description of dataset including spatial coverage, representativeness, limitations:
Countryside Survey (CS) consists of a field survey of 591 1km x 1km sample squares spread across England, Scotland and Wales, undertaken approximately every eight years. The CS has mapped and counted the ponds that occur in the 1x1 km survey squares since 1984. Presence of new ponds, lost ponds and changes in pond shape or size are all logged. A detailed description of methods used to map ponds can be found in the CS Field Mapping Handbook (CS Technical Report No.1/07).

The first CS survey of pond quality was undertaken in 1996 as part of the Lowland Pond Survey (LPS96). It assessed pond condition using physico-chemical attributes and plant assemblages of ponds. LPS96 was the first survey to introduce a definition of CS ponds, and to specifically distinguish seasonal ponds, which naturally dry out in summer, from ponds which have been drained and are permanently dry, and which can be regarded as 'lost' (these definitions were used in all subsequent CS surveys). The LPS96 was restricted to the lowlands of England, Scotland and Wales. The survey strategy was designed to maximise compatibility between LPS96 and earlier CS data gathered in 1984 and 1990. Pond quality was assessed relative to high quality National Pond Survey reference sites. In total, surveys were undertaken at 150 1 km x 1 km lowland squares; this included 136

squares which contained ponds and 14 “non-pond” squares. In each square, all ponds that were present (n=377) were surveyed in detail to provide ecological data.

CS2000 reported the number and size of ponds in all squares but did not include assessments of pond condition or quality. CS2007 was the first to assess both pond numbers and pond quality across the whole of the British countryside including upland areas. A detailed assessment of pond condition was made for one randomly selected pond in each square containing a pond. Detailed condition assessments were made for a total of 149 ponds in England.

Data ownership and licensing restrictions (if any): Countryside Survey data owned by UK Centre for Ecology & Hydrology. Free licensing with some restrictions (e.g. no re-distribution of data) but distribution of derived product is permitted. Arrangements for the new national pond surveillance programme are yet to be finalised.

Data transfer arrangements: Future data transfers to be arranged when new surveillance data are available.

Frequency of source data update/data transfer: The new pond surveillance programme will be based on a rolling programme which should allow updates to these attributes every year, but an update frequency of 3-5 years would be appropriate.

Form of attribute:

- ***C1 Phosphorus and C2 Nitrogen*** – The current form of these attributes uses nitrate and phosphate levels because that is what has been historically recorded by Countryside Survey. The new pond surveillance programme may do likewise due to practicality and cost considerations. Assessment of Total Phosphorous and Total Nitrogen would be preferable in standing waters because substantial amounts of nutrients are locked up in biomass (particularly in the growth season). This said, previous studies have shown that such data can give a reasonable impression of the water quality in ponds, particularly when used in conjunction with biological attributes.
- ***P3 Semi-natural land-use*** – The current form of the attribute uses riparian land use at 5 metres from the pond edge, because that is the most suitable observation recorded by Countryside Survey. Whilst the 5m data will provide an insight to the immediate surroundings, ponds can also be impacted by land use further away. Countryside Survey also recorded land use 100m away, but this will have less ecological influence and is harder for surveyors to accurately assess. For the final version of this attribute land use at 15m is proposed, which needs to be considered in the design of the new pond surveillance programme.
- ***P4 Shading and P5 Grazing intensity*** – These attributes compare recorded levels of shading and grazing with an ideal distribution of shading and grazing across the pond resource as a whole, The further the current distribution of grazing and shading levels are from this ideal the lower the class score. This is a different approach to other attributes in the assessment framework, which assess the naturalness of individual ponds. As this provides a single class for the whole pond resource for these two attributes they cannot be visualised in the same way as other attributes in the wheel diagram. The inner grey-scale wheels are not relevant and are blocked out in solid grey, but a value can be generate for the mean attribute score – this then feeds into the physical component score and the overall naturalness score as normal.

Data processing method for generating attribute output:

C1 Phosphorus and C2 Nitrogen – Values are taken for each pond and classified according to the class boundaries in the next section.

P3 Semi-natural land-use – This attribute is based on the CS 2007 broad habitat dataset, which provides land-use percentage breakdowns within 5 m and within 100 m of pond edge. Land-use percentages within 5 m were used. Land-use categories considered to be semi-natural were retained: trees/woodland, heathland/moorland, unimproved grassland, rank vegetation, streams/ditches, ponds/lakes, rock/stone/gravel, fen/marsh/flush, bracken, montane. For each pond, the percentage values for these categories were summed, then these summed percentages were re-classified as:

Attribute	Class 1	Class 2	Class 3	Class 4	Class 5
P3	100%	90-100%	65-90%	33-65%	<33%

P4 Shading and P5 Grazing intensity - An ideal distribution of shading and grazing levels has been produced based on the literature and expert opinion on ponds in England. The production of an ideal distribution of shading and grazing relies on knowledge of the current trends in pond condition in relation to grazing and shading and knowledge of the current distribution of levels of shading and grazing. For example we know that a loss of grazing can lead to a decline in pond quality, but that excessive levels of grazing could also lead to decline. So, using the distribution of grazing intensity from Countryside Survey data from 2007, the ideal would be to increase the amount of light to moderately grazed ponds and reduce the number of non-grazed and heavily grazed ponds.

Process for P4 Shading:

Step 1: The percentage shading due to tree cover (100% means whole pond shaded) recorded at each surveyed pond is re-classified into five classes as follows:

Attribute	Class 1	Class 2	Class 3	Class 4	Class 5
P4	0%	<=10%	10-40%	40-80	>80%

Step 2: The percentage of surveyed ponds in each Class is calculated for the whole resource.

Step 3: These percentages are compared to the ideal percentages in each Class, i.e. 30% of the resource in Class 1, 30% in Class 2, 25% in Class 3, 10% in Class 4, 5% in Class 5. The sum of absolute differences between actual and ideal percentage allocations was calculated.

Step 4: The whole resource is assigned a single class based on the summed differences as follows: <=10%, Class 1, 10-20%, Class 2, 20-30%, Class 3, 30-40%, Class 4, >40%, Class 5.

Process for P5 Grazing:

Step 1: Grazing intensity data is recorded in CS2007 as 0-5 discrete values (0 no grazing). Values 0 and 1 were merged as Class 1, and other values were assigned to the matching class (e.g. intensity 2 is Class 2).

Step 2: The percentage of surveyed ponds in each Class is calculated for the whole resource.

Step 3: These percentages are compared to the ideal percentages in each Class, i.e. 70% of the resource in Class 1, 10% in Class 2, 10% in Class 3, 10% in Class 4, 0% in Class 5. The sum of absolute differences between actual and ideal percentage allocations was calculated.

Step 4: The whole resource class is assigned a single class based on the summed differences as follows: <=10%, Class 1, 10-20%, Class 2, 20-30%, Class 3, 30-40%, Class 4, >40%, Class 5.

B1 PSYM - The PSYM scores created from Countryside Survey data were used for this component as described in Report JP016.

Naturalness class boundaries: (see Form of attribute section above)

Attribute robustness: The robustness of these attributes will be determined by the scale and representativeness of the new pond surveillance programme. Countryside survey sampled a small number of ponds and was limited about what it could say about the pond resource as a whole. The robustness of attributes P4 and P5 are also influenced by the definition of ideal shading/grazing regime for the pond resource as a whole. As data will be provided from the new pond surveillance programme rather than Countryside Survey there is a risk that that the illustrative figures used will no longer represent the ideal distribution, due to differences in the sampling design between the NCEA and Countryside Survey. As the new surveillance programme design becomes clear and the first data are collected the ideal distribution may need to be refined to take account of this.

Storage location for raw dataset and processed data: Data from CS2007 are stored on the Environmental Information Data Centre (EIDC). The storage location of data from the new pond surveillance programme is yet to be determined. Summarised processed data are stored in a structured spreadsheet of naturalness scores attached to the B6 information pack. A geodatabase is being prepared that will indicate spatial variations in naturalness scores across England, as far as this can be shown with available datasets.

Attribute information sheet

Principal habitat component: Ponds

Hydrological attribute Artificial influences (H1); Physical attributes Natural pond base (P1); Natural shoreline (P2) and Cross-cutting attribute Landscape connectivity (CC1)

Rationale for inclusion:

- ***H1 Artificial influences*** - Naturally fluctuating water levels support distinctive biological communities with traits adapted to these conditions. Water levels influence hydrological connectivity, both vertical and lateral.
- ***P1 Natural pond base –*** Natural pond substrates are critical to the development of submerged and marginal plant assemblages and their dependent fauna, as well as fauna that exploit the substrate itself. Ponds can be lined with various materials that interfere with this inter-relationship between natural substrates and flora/fauna.
- ***P2 Natural shoreline*** - A natural transition of aquatic and wetland vegetation in the marginal zone of ponds is a critical habitat component for both characteristic plant species and associated fauna. These transitions (or hydroseres) also play an important role in wave dissipation and the harbouring of animals that maintain clearwater conditions that is vital for submerged plant growth and visual predators, all generating ecological resilience .
- ***CC1 Landscape connectivity*** - Pond numbers have declined dramatically in recent history. Sufficient pond density is needed to provide adequate habitat extent but also landscape-scale refugia and stepping stones for a range of aquatic and terrestrial biota that are associated with ponds and other freshwater habitats.

Source data: Data on these attributes has yet to be sourced.

Outline description of dataset including spatial coverage, representativeness, limitations: Not appliable at present. Data on attributes H1, P1 and P2 should be provided by the new national pond surveillance programme. For attribute CC1, an NCEA pilot citizen science project is currently underway to assess methods of collecting suitable data. The results will be considered in the next stage of B6 development.

Data ownership and licensing restrictions (if any): Arrangements for the new national pond surveillance programme are yet to be finalised.

Data transfer arrangements: To be arranged when data sources are in place.

Frequency of source data update/data transfer: The new pond surveillance programme will be based on a rolling programme which should allow updates to these attributes every year, but an update frequency of 3-5 years would be appropriate.

Form of attribute: To be arranged.

Data processing method for generating attribute output: To be arranged.

Naturalness class boundaries: To be arranged.

Attribute robustness: This will be assessed when the new surveillance programmes are established.

Storage location for raw dataset and processed data: To be arranged.