

BIS Validation and Verification Policy & Procedures

Background:

- BIS aims to ensure the data that it disseminates is of the highest possible quality and therefore aims to validate and verify every biological record it receives.
- Validation is a procedure that ensures the minimum standards needed to make a biological record are adhered to.
- Verification is a procedure to ensure that the record has been attributed to the correct taxon and that species identification by the observer has been accurate. This process may be carried out by a determiner who could be a local or national species expert.

Policies and implementation notes:

BIS ensures that all data is validated during the data input or import process

- BIS sets the minimum standard for a valid biological record to be that it clearly includes:
 - **WHAT was seen?** A valid species name is required (scientific or common).
 - **WHERE was it seen?** A valid grid reference or detailed location description (from which a valid grid reference can be determined) is required.
 - **WHEN was it seen?** A date is preferred, but date ranges and even vague dates (including months or years can also be accepted).
 - **WHO saw it?** A full recorders' name is preferred, although partial names, initials or organisations are accepted.
- A range of different methods of data validation are employed to check the accuracy of incoming records. These are described in more detail in the BIS verification procedure in Appendix 1.

BIS aims to ensure as much data as possible is verified by local or national experts

- The quality of taxonomic identifications is critical to the operations of BIS. BIS users expect accurate and reliable data, therefore all data provided must be of the highest quality possible and the reliability of all data must be known.
- BIS assigns and maintains a verification level for every record in its database. These are standardised to agree with LERC outputs from Aderyn and eMapper. The verification levels are:
 - 0 = Incorrect
 - 1 = Unassessed
 - 2 = Unconfirmed
 - 3 = Correct
- BIS's data verification system utilises, wherever possible, local expertise and existing data verification procedures eg iRecord, WiReD and LERC App records are open to registered experts to verify on-line through iRecord.
- BIS makes records available to VC recorders by taxon group and geographical area through Aderyn. Registered users can filter for unassessed records, download and mark up any incorrect records to be updated in the BIS database.
- BIS recognises that the verification process can be very time-consuming for volunteer recorders and that therefore records may remain unverified for a considerable period. As part of its standard practice BIS chooses to make unverified data available in reporting unless data providers or verifiers raise a specific objection. Unverified records are shown as Unassessed (1) in Aderyn reporting.

- As part of its programme of continuous improvement, BIS is currently prioritising work to ensure that the number of unverified records in its database is reduced. A target could be to clear the backlog of unassessed historical records so that there will be no unassessed data in our database (except records which have been held for less than 12 months).
- A range of different methods of data verification are employed to check the accuracy of taxon identifications associated with incoming records. These are described in more detail in the BIS verification procedure in Appendix 1

Appendix 1 BIS Validation and Verification procedures

Validation Procedures

All records received by BIS are entered into the Recorder 6 data base which has its own validation checks to ensure the necessary components of the record (who, what, where and when) are present and in the correct format. Casual records may be entered via BIS WiReD which includes validation checks and exposes them to verification rules from Record Cleaner and verification by registered experts registered on iRecord. Other records may be entered directly from paper format or imported either from an electronic spreadsheet or from an exported database eg Recorder or MapMate using the 'import wizard'. This is used mainly for large datasets received from vice-county recorders.

Paper records

- Grid reference - before entry records are checked that the given grid reference matches the location name on an OS 1:10,000 map, in GIS.
- Duplication – records are checked that they are not already in the database
- After entry into R6 all the records are checked by the operator for entry errors and the checkbox in R6 ticked for each record.
- Where resources allow, data entry is re-checked by another operator for errors, especially if the original operator is a volunteer or newly trained. Initially this may be 100% for new operators and then lessened to around 10% when an operator has gained more experience.
- After entry, the records can be mapped within R6 to check again for outliers and possible incorrect grid references.
- Copies of the original paper data are kept on file and signed off by the operator when entered in R6 and by the person that checked the entry.
- Alternatively, pages can be scanned, and the file linked to the Survey in Recorder 6.

Electronic records

- Grid reference – records within a spreadsheet are mapped in ArcGIS or QGIS to identify any obvious outliers outside the Vice-county specified or the BIS area.
- Spreadsheets usually have to undergo editing to conform to the format needed by R6. An original copy of the original data is always kept and the edits for import are kept on separate worksheet. After import the spreadsheet file is linked to the 'Survey' in Recorder.
- The import process into R6 ensures the taxon name, date, recorder and grid reference are all in the correct format.
- Any taxonomic names changes or discrepancies eg synonyms where new scientific name is not in taxon dictionary, are agreed by the data provider where possible. Where records are recorded under a synonym or as *sensu stricto* or *sensu lato*, which does not match the taxonomic dictionary, the recorded name is entered in the sample comments within R6. Changes are recorded in the edited spreadsheet and highlighted for any future reference.

Verification procedures

All records entered into R6 need to be verified unless they have come from a 'trusted source' such as a VC recorder, a national expert or a recording Scheme or Society. The aim is to get the records verified by the appropriate VC recorder where possible.

- casual records received by BIS on recording forms, email etc are now entered through BIS WiReD as described in the Data Manual, with the appropriate recorders name. Here they can be verified by registered experts through iRecord.
- Other large datasets are imported as spreadsheets into Recorder 6 immediately as unassessed records to ensure they are available for the enquiry service. They are only marked as verified when the VC recorder (VCR) agrees this.

- Since 2018 VCRs have access to LERC Wales data for their area and taxon group via Aderyn. Records can be filtered and downloaded to spreadsheet and verification marked up and returned to BIS. Otherwise they can be sent updates via spreadsheets at agreed intervals.
- All records for appropriate time period should be sent to VCRs for their own use and verification, with unassessed ones highlighted.
- Historical records – although BIS receives data sets from VC recorders it has also received large datasets collated by other organisations such as the Wildlife Trusts and Brecon Beacons National Park Authority. BIS has also mobilised approx 600,000 records from the then CCW up to 2009. A lot of these records were from trusted recorders but have not been formally verified. From 2009 BIS has sent large datasets to VCRs and Schemes & Societies. As they are such large datasets it may take several years before this historical verification is completed.
A log of exported records is held at X:\BIS\LOGS\exports.xls
Original export files are held within relevant folders at X:\Data\exports
- NBN Record Cleaner – it was suggested at the 2017 Recorders Forum that BIS use the Record Cleaner (RC) to help verify historical datasets and cut down number of records that need to be checked by VC recorders. Failed records could be highlighted for further verification. It has proved impractical to assess new spreadsheets supplied by recorders through RC, so they need to be imported into R6 and then checked against Record Cleaner. RC still rejects many taxonomic names from R6 and the process is work intensive. BIS will look at which datasets still need to be verified and whether this is the best procedure to apply. BIS has suggested that the rules from NBN Record Cleaner be integrated into the New Recorder 6.

BIS verification terms to use in Recorder 6

Determination type	Description	Determiner	Determiner role
Correct	Provided by Scheme or Society (S or S) and has undergone verification checks	SorS or named rep.	Specialist
Considered correct	Provided verified by VCR	VCR (unless named by data provider but not original recorder)	County recorder
Probably correct	Collated by VCR but not gone through strict verification checks OR provided by 'trusted provider' eg BBNP ecologist	BIS operator	BIS employee
Unconfirmed	VCR is not sure or needs more evidence	VCR	
Unassessed	Verification has not been attempted or received	BIS operator	BIS employee
Considered incorrect	Determiner has doubt about ID	VCR or BIS operator	
Incorrect	Marked as definitely incorrect eg. Grid ref	VCR or BIS operator	

Standardisation of output verification levels – LERC Wales partners and users have requested that a record verification level should be shown in outputs including reports and GIS layers. The 4 LERCs have agreed to standardise these verification levels as shown in the table below. There are 3 levels of verification, since records failing verification will not be included in any output. This was implemented in 2010 although many BIS records will remain 'unassessed' until they have gone through the VC verification process as described above.

LERC Wales output verification	Output verification integer	BIS record verification R6(Determination type)	Description	R6 integer (Verified flag)
Unassessed	1	Not verified	Verification has not been attempted	0
Unconfirmed	2	Unconfirmed	Determiner cannot confirm or awaiting evidence	0
Considered correct	3	Correct	Provided by SorS and has undergone verification checks	2
Considered correct	3	Considered correct	BIS records verified by VC recorder	2
Considered correct	3	Probably correct	Collated by VCR but not gone through strict verification checks OR provided by 'trusted provider' eg BBNP ecologist	2
		Incorrect	Determiner marks as definitely incorrect or incorrect details eg grid reference	1
		Considered incorrect	Determiner has doubt about ID	1

▪