Improving the quality of Lepidoptera records available via the NBN Gateway

Contract report BC10-a to BioD Services Limited

R. Fox, L. Hill and M.S. Parsons

2011

Butterfly Conservation Report No. S11-06

This report was produced by Butterfly Conservation under the Defra/NBN Trust contract 2008-2011 to improve data provision, management and coordination in the National Biodiversity Network.

Butterfly Conservation
Company limited by guarantee, registered in England (2206468)
Registered Office: Manor Yard, East Lulworth, Wareham, Dorset. BH20 5QP
Charity registered in England and Wales (254937) and in Scotland (SCO39268)

www.butterfly-conservation.org

Fox, R., Hill, L. and Parsons, M.S. 2011. *Improving the quality of Lepidoptera records available via the NBN Gateway*. Butterfly Conservation, Wareham. (Butterfly Conservation Report No. S11-06).

Contents

Fox, R., Hill, L. and Parsons, M.S. 2011. *Improving the quality of Lepidoptera records available via the NBN Gateway*. Butterfly Conservation, Wareham. (Butterfly Conservation Report No. S11-06).

Summary

- The National Biodiversity Network (NBN) seeks to ensure relatively complete coverage of butterfly and macro-moth records on the NBN Gateway and to improve the quality of these data.

- Butterfly Conservation is the organiser of national recording schemes for butterflies (Butterflies for the New Millennium, BNM) and macro-moths (the National Moth Recording Scheme, NMRS) in the UK and is the major contributor of Lepidoptera records to the NBN Gateway.

- A standardised approach to the essential and desirable data fields for Lepidoptera records is recommended, together with information on handling verification queries and issues surrounding the flow and quality control of data.

- Baseline information has been developed for taxon-specific rules to assess the spatial, temporal and identification veracity of moth and butterfly records via the recently-developed NBN data validation software. Taxa have been graded according to inherent difficulty of identification and information derived from existing sources to define currently acceptable geographical distributions, flight periods and appropriate year ranges. These rules will require regular revision.

- Used together, the rules will enable records to be prioritised for scrutiny, which should be undertaken by the appropriate County Moth Recorders and BNM Local Co-ordinators.

- Our recommendation is that all Lepidoptera records should be subject to verification by County Moth Recorders and BNM Local Co-ordinators, as appropriate, prior to uploading to the NBN Gateway or being made available to potential data users via other means.

- Issues of data access are discussed and species that should be considered ‘sensitive’ in the context of data sharing are defined.

- The current framework provided by the BNM and NMRS schemes is fit for purpose. However, the sustainability of the schemes requires substantial ongoing funding and there is great scope to streamline the provision of funding from institutional data users. Other ways in which the framework could be improved are discussed.

1. Introduction

1.1. Background

The National Biodiversity Network (NBN) seeks to facilitate the flow of biological recording data from a large and diverse group of organisations involved in collecting such information to an even wider community of potential users of data. The main tool employed by the NBN to achieve this goal is the NBN Gateway, an interactive website. Although the Gateway already hosts over 60 million species records contributed by over 140 different organisations, there is seen to be scope to improve both the coverage (temporal and geographical) and the quality of the data. The NBN's Data Access Group has identified a set of priority taxa for which work should be undertaken to ensure relatively complete coverage on the Gateway. The target taxa are vascular plants, Lepidoptera (specifically butterflies and macro-moths), lichens and birds.

In taking forward work on these priority groups, the NBN has sought to improve the efficiency of quality checks and data flows associated with records of these taxa. The NBN has developed a data checking software tool to improve quality control. Holders of local datasets could use this tool to highlight records that might need either further scrutiny by a taxon specialist or more information from the recorder. This tool needs to have a suite of rules, relevant to each taxon group, against which records can be checked. Once such rules have been developed and integrated into the software tool it can be distributed to organisations and individuals who collate datasets for these taxa, together with general guidance about recording and data flow within each species group.

As the Lepidoptera is one of the priority groups, the required work to develop rules for the NBN's software tool and to provide general guidance on recording and data flow has been undertaken by Butterfly Conservation. Butterfly Conservation is a UK charity dedicated to the conservation of Lepidoptera (butterflies and moths) and has formidable expertise in this area both nationally and internationally. Butterfly Conservation's staff have the specialist skills and knowledge in species identification, life-history and ecology, biological recording issues and systems, and database management and analysis.

The gathering of records on butterflies has been a core activity for Butterfly Conservation since its inception in 1968. The organisation has harnessed the enthusiasm, skills and knowledge of thousands of volunteers and members of the public to collate huge datasets on butterflies, and more recently on moths. Butterfly Conservation runs the 'national' recording schemes for both butterflies (Butterflies for the New Millennium, BNM) and macro-moths (the National Moth Recording Scheme, NMRS), each of which covers the whole of the UK plus the Isle of Man and the Channel Islands (Asher *et al*. 2001; Fox *et al*. 2011). The butterfly scheme also includes the Republic of Ireland via a partnership with the Dublin Naturalists' Field Club.

Through these recording schemes Butterfly Conservation has gathered over 7.9 million butterfly distribution records and 12 million records of macro-moths. Almost all of these, 19 million records to date, have been made available by Butterfly Conservation via the NBN Gateway. This is equivalent to 31% of all records of all taxa on the Gateway! Butterfly Conservation is the largest provider of data to the Gateway and has also provided the largest single dataset (the NMRS dataset).

1.2 Objectives

The main objectives of the project were to produce the following:

Objective 1: Guidance on what information records should contain

Produce guidance on which attributes are essential and which are desirable for butterfly and macro-moth records.

Objective 2: Rules for quality assuring records in the NBN data validation software tool
Four separate rules will be developed and specified for each of c.1000 taxa in the Lepidoptera as part of this work: identification difficulty, geographical range, flight period, and year range.

A classification system for identification difficulty of butterfly and macro-moth species in the UK will be developed. Categories will be defined according to the difficultly of identification combined with the level of the recorder's expertise and need for additional evidence to substantiate a record.

Baseline acceptable distributions will be defined for each species to provide geographical rules. This will be done at the 10km x 10km grid square resolution for the UK and provided in a suitable format for inclusion in the NBN validation tool. The recent recorded distributions of each species in the BNM project and NMRS will be used to generate baselines.

Temporal rules for the adult life-cycle stage of each taxon will be developed. These will relate to the main months of the current flight period(s) of each macro-moth and butterfly species in the year. Temporal rules will not cover egg, larval or pupal stages.

Temporal rules governing the acceptable year range for records of each taxon (i.e. species that were first discovered in or first colonised the UK in a certain year and species that became extinct in the UK in a certain year).

Objective 3: Processes for verifying records
Guidelines will be developed for data collators in how to handle records that fall outside the verification rules defined in the work. Comment will also be provided on the NBN's suggested framework for classifying the status of records (i.e. into the categories Correct, Considered correct, Requires confirmation, Considered incorrect, Incorrect, and Unchecked).

Objective 4: Identify criteria for sensitive records
A review of each of the butterfly and macro-moth species will be undertaken to identify those taxa that should be considered sensitive. The criteria used to justify the selection will be stated.

Objective 5: Proposals for addressing gaps in the current framework
Any shortfalls in the current framework to support the collation and verification of Lepidoptera records in the UK will be identified. Where necessary and appropriate, proposals to address these shortfalls will be presented.

As is clear in these objectives, this project addresses issues relating to the verification and flow of Lepidoptera records at a national (i.e. UK) scale. Additional or differing issues may occur at the local level and these are not addressed in detail in this report.

2. Methods
Work to achieve these objectives was undertaken during the period November 2010 – March 2011 by Butterfly Conservation staff Richard Fox (Surveys Manager), Les Hill (Database Manager) and Mark Parsons (Head of Moth Conservation).

All resident butterfly and macro-moth species were included in the assessment of rules for Objective 2, including species that are currently regionally extinct in the UK. Moth species included those c.900 species in 19 Lepidoptera families that are traditionally regarded in the UK as being macro-moths (i.e. Hepialidae, Cossidae, Zygaenidae, Limacodidae, Sesiidae, Lasiocampidae, Saturniidae, Endromidae, Drepanidae, Thyatiridae, Geometridae, Sphingidae, Notodontidae, Thaumetopoeidae, Lymantriidae, Arctiidae, Ctenuchidae, Nolidae and Noctuidae). Species that have only occurred in the UK as very rare accidental imports

were excluded, as were all species whose UK records are considered doubtful. However, species that occur naturally as migrants were included, as were adventive species that have become established in the UK. The total number of species was c.960.

In addition, other taxa recognised by the NBN species dictionary were included in the assessments. These included subspecies and forms, as well as aggregated taxa, where two or more very similar species are recorded without being distinguished to species level due to identification difficulties. Such taxonomic entities were included because they are regularly recorded by moth recorders and therefore would occur in datasets being checked by the NBN validation tool.

Species names follow the Recorder 6/NBN Species Dictionary.

3. Results

3.1 Objective 1: Guidance on what information records should contain

For the purposes of the national recording schemes run by Butterfly Conservation, the following data fields are considered essential and desirable for each record. It is recognised that other collators of Lepidoptera records, including County Recorders and Local Environmental Records Centres that participate in the national scheme, may set their own criteria for information. Our recommendation is that such variation should always maintain, as a minimum common standard, the essential data fields outlined here.

Essential data fields:

- *Species name* – all records must clearly have a taxon name associated with them (although this may be an aggregate of several difficult to identify species). All UK butterflies and macro-moths have vernacular names as well as scientific names and either is sufficient, although scientific names are preferred (especially for moths). Although provision of either the scientific or vernacular name is essential, it is desirable to have both as part of a record (see under Desirable data fields). Names should follow the NBN's Species Dictionary managed by the Natural History Museum (http://nbn.nhm.ac.uk/nhm/). However, it is recognised that not all recorders or recording software packages will conform to the NBN Species Dictionary and data compliers will have to resolve cases where species names do not match up.

- *Grid reference* – one of the principle purposes of biological recording is to locate species or habitats spatially. This can then provide information on species' distribution and be used to inform land-use and conservation decision making. Ordnance Survey grid references provide a reliable way to locate sightings in geographical space in the UK. A maximum spatial resolution of 10km x 10km grid square (two-figure grid reference) is acceptable, but records should ideally have a spatial reference of a 100m grid square (six-figure grid reference) or 1km grid square (four-figure grid reference). For species of conservation interest, more precise 10m (eight-figure) grid references are useful for the exact locations of the sighting or any breeding colony.

- *Site/location name* – the provision of a site name, ideally one that is recognisable from an Ordnance Survey map, is important as a cross reference to the grid reference. Grid reference errors are frequent and a site/location name can enable these to be rectified easily without the need to refer back to the recorder. Full postal addresses should not be used as site names as these, when matched with recorder name, constitute personal data under the Data Protection Act.

- *Recorder name* – the name of the recorder (i.e. normally the observer) of the sighting is essential, in part for verification purposes and in part to define the ownership of that record. It is acceptable for records to be anonymous if no recorder name can be located,

but this should be exceptional. The preferred formats for names are: R.B.Powell or Jackie Lee. The following formats are also satisfactory: Robert B.Powell, Mr J.C.Jones, Dr Jim Hopkirk, Mrs Jackie E.Lee. Multiple recorders' names can be combined, separated with a comma, semi-colon, forward-slash, backward-slash, plus sign and either 'and' or '&' before the last name, using any of the above formats e.g. "R.B.Powell;Jackie Lee;Dr Jim Hopkirk". The titles Col / Dr / Miss / Mr / Mrs / Ms / Prof / Rev can be accommodated by the NMRS.

- *Determiner name* – if the species recorded has been identified by a different person to the Recorder then this person's name should also be provided as the 'Determiner', in a separate data field. There should only be a single determiner in the Determiner field, formatted as per Recorder names; however, there can be any number of other determiners listed in the Comments field as required.

- *Date* – is an essential part of a record and should be the date that the species was seen. If species are being recorded at a light trap run overnight, then the date given for the records should be the date when the trap was switched on, not the date when the trap was opened and sorted through. If light traps are run over several consecutive nights before the contents are recorded, then a date range should be applied to the records. The Date field should be in, or exportable in, the format *dd/mm/yyyy* e.g. 12/07/2007. It is important to use four digits for the year to avoid confusion between centuries. Recording the exact date is preferable but records at month and year, year only or year range resolution are acceptable. If the date is a month only, use the format *00/mm/yyyy*; if the date is a year only, use *00/00/yyyy*. Year ranges should be recorded as *yyyy->yyyy* (i.e. with a dash followed by a greater-than sign between) and exact date ranges as *dd/mm/yyyy-dd/mm/yyyy* (i.e. two dates with a dash between).

- *Stage* – it is important to distinguish between life-cycle stages i.e. egg, larva, pupa or adult. Verification procedures and analyses rely on this information. In addition, knowledge of the locations of immature stages is important for conservation as it identifies probable breeding locations or colonies.

Desirable data fields:

- *Species code* – UK Lepidoptera are allocated a Bradley (Bradley and Fletcher) number. This can be useful in cases of doubt as cross-referencing information in conjunction with the scientific and/or vernacular name.

- *Species name* – while it is essential that all Lepidoptera records include the name of the species, each UK species of butterfly and macro-moth has both a scientific and a vernacular name. Providing both with records is desirable to help eliminate nomenclature confusion e.g. **White-pinion Spotted** *Lomographa bimaculata* and **White-spotted Pinion** *Cosmia diffinis* and where species with multiple sub-species are allocated different vernacular names i.e. **Campanula Pug** *Eupithecia denotata denotata* and **Jasione Pug** *Eupithecia denotata jasioneata*.

- *Confirmation* – some species identifications are 'confirmed' by a third party, other than the recorder and determiner, i.e. someone else gave a second opinion and agreed with the identification. In this case the name of the confirmer can be put in a Comment field to provide additional evidence of the species identification.

- *Quantity* – the quantity or abundance of a recorded species is useful and potentially important information. For example it can help to identify breeding colonies or important sites for species of conservation concern. If no abundance is recorded we recommend the use of "Present". It should be noted that the use of "0" (zero) in the abundance field is a problematic issue in biological recording. The MapMate recording software, which is widely-used by butterfly and moth recorders and record collators, utilises "0" to indicate a

positive record that has no abundance information, whereas the Recorder 6 software uses “0” to indicate a negative record (i.e. recording was undertaken but no individuals of the taxon were seen).

- *Method* – information about the method of recording e.g. light trap, daytime observation, dusking etc. can be provided. If a light trap, the bulb type can be specified (e.g. MV or actinic).

- *Sex* – the gender of observed insects can be recorded. Any one record should only contain one gender. Separate records should be created in order to record separate quantities of males and females.

- *Comment* – the presence of a comment field is very useful for any other relevant comments. These might include whether there is supporting evidence for a record in the form of a photo or voucher specimen, whether genitalia determination has been carried out, or if anyone else saw and confirmed the identification. This is extremely useful for rare or difficult species. The comment field can also be used to flag up unverified data of any kind, by adding the phrase “Dubious record”.

3.2 Objective 2: Rules for quality assuring records in the NBN data validation software tool

Four separate sets of rules were developed for use with the NBN validation software. For each of c.1000 taxa (species, sub-species and aggregates) of butterflies and macro-moths, information was provided to define rules of identification difficulty, geographical range, flight period, and year range.

Classification of macro-moths and butterflies according to identification difficulty
UK macro-moths and butterflies were classified for identification difficulty into four categories (grades). Categories were defined according to the difficulty of species identification combined with the necessary level of the recorder’s expertise and need for additional evidence to substantiate a record. The classification was produced as a three stage process. First, a group of taxonomic and recording scheme experts on the staff of Butterfly Conservation reviewed the category definitions and criteria used in existing verification classifications produced by County Moth Recorders and local moth groups in various parts of the UK. From this a first draft of the number of required grades and their definitions was produced. Two of the authors (RF and MP) then independently scored all the macro-moth and butterfly taxa against these draft definitions. These draft classifications were then compared and reviewed by the third author (LH) and a meeting was held to find consensus for those taxa that had been classified into different grades. The definitions of each grade were refined as part of this second stage. Finally, the classification was reviewed by an external expert and the comments received were taken into account in the production of the final version of the classification.

The classification produced applies at the UK level (i.e. England, Northern Ireland, Scotland and Wales) and can also include the Isle of Man, but does not include the Channel Islands or Republic of Ireland. Immigrant as well as resident species were included.

The classification of each species was based purely on identification difficultly, not on likelihood of occurrence in a particular place or time of year. It should be recognised however, that location, habitat, time of year and other factors, also play important parts in the correct identification of species. Butterfly Conservation strongly recommends that the identification difficulty classification is used in conjunction with rules addressing spatial and temporal occurrence for the purposes of verification of moth and butterfly records (as in the case of the data validation software being developed by NBN).

UK macro-moths and butterflies were classified into the following four grades according to identification difficulty. **In all cases, it is assumed that the insect being identified is in**

good condition (excessive wear of wing scales or other damage will make identification more difficult) **and that the recorder has a basic level of competence** (e.g. ability to use a field guide). The latter is important as almost no lepidopteran taxa can be identified with complete reliability by members of the public with no training, experience or access to materials to aid identification (e.g. field guides, leaflets etc.).

Grade 1: Species that can be relatively easily identified even by inexperienced recorders with the use of up-to-date field guides. Such records would be accepted without the need for additional evidence as long as they conform to the normal flight period and geographical range of the species.

Grade 2: Species where care is needed for identification. Should not present any difficulty for experienced recorders, but inexperienced recorders may be expected to provide additional evidence in the form of photographs of reasonable quality that show as many of the identification features as possible or, alternatively, a specimen. Even experienced recorders may be expected to provide additional evidence if the record is outside of the normal flight period or geographical range of the species.

Grade 3: Species that are difficult to identify. Even experienced recorders may be expected to provide additional evidence, for example a specimen, particularly if the record is outside of the normal flight period or geographical range of the species.

Grade 4: Species that can only be identified following critical assessment, often involving dissection and examination of genitalia. A specimen should be retained for confirmation.

In a parallel project, Butterfly Conservation has produced a book covering the identification of difficult macro-moth species (Townsend *et al*. 2010). This provides up-to-date information and detailed identification features and techniques for each of the macro-moths assigned to Grade 4 of the identification difficulty classification, as well as many of the species in Grade 3.

It is recommended that the classification is reviewed at regular intervals and amendments made in response to user feedback and changes to the UK Lepidoptera fauna (e.g. due to taxonomic revisions, species colonisations etc.).

It should also be noted that such classifications can be, and have been, produced at more restricted geographical scales (e.g. county or country level). There are benefits to such an approach as some 'similar species' may not be present at the local scale, thus simplifying identification and verification. While the UK classification produced here can be of use for record verification it should not replace or supersede local classifications that are more closely adapted to the Lepidoptera fauna of a particular area. Through the NMRS, Butterfly Conservation has encouraged the production of such local classifications as best practice and will continue to do so.

Geographical ranges of butterflies and macro-moths
Baseline acceptable geographical distributions for each taxon were defined and provided to Graham French of NBN in spreadsheet format for conversion into text files that can be used in the NBN's validation software. Records falling outside of the defined distribution would be flagged for further verification.

Baseline distributions were defined from existing recent (1990-2009) records in the BNM and NMRS databases held by Butterfly Conservation. These sources represent reasonably accurate modern assessments of the true distribution of each taxon, although some species and geographical areas are under-recorded. Thus it will be important to revise the baseline distribution rules over time as more records are gathered from currently un-recorded or under-recorded areas. The year 1990 was chosen as the starting point for the modern distributions because recording has been particularly intense over this period (thus coverage

is good) and because of the need to represent rapidly declining species appropriately in the NBN validation software. If all records from historical sources were used, potentially important new records with a high need for verification might be overlooked if they came from areas where the species had occurred previously but was long extinct.

The baseline distributions were defined at the 10km x 10km grid square resolution for the UK, Isle of Man and Channel Islands simply as a list of those grid squares for which there were records of a species in the 1990-2009 period. The recorded distributions of many Lepidoptera species are changing rapidly; with severe declines in distribution for some while others expand their ranges (Asher *et al*. 2011; Fox *et al*. 2006; Fox *et al*. 2011; Hill *et al*. 2010; Warren *et al*. 2001). Therefore, it will be necessary to update the baseline distributions on a regular basis in order to maintain the efficiency and effectiveness of the verification rules.

For resident species, the spatial rule will work simply in the NBN validation software – records from 'new' 10km grid squares will be highlighted, those from already occupied grid squares will not. This will apply to all species regardless of their status as common or rare and is a sensible approach because there is no pressing need to undertake special verification on records even of rare species if they are from well-known, established colonies.

The situation is different, however, for immigrant species. The spatial occurrence of immigrant species is highly variable and, therefore, the use of past sightings to inform verification decisions is usually inappropriate except in the very broadest geographical sense (e.g. an immigrant normally seen of the south coast of England being recorded in Scotland). For the baseline distributions developed in this contract, common and rare immigrant species were dealt with differently. Common immigrants (e.g. **Red Admiral** *Vanessa atalanta*, **Painted Lady** *V. cardui*, **Humming-bird Hawk-moth** *Macroglossum stellatarum* and **Silver Y** *Autographa gamma*) were not included at all. Thus records of common migrants, which have the potential to occur anywhere in the UK (in good years, at least) will not be checked by the NBN validation software. Scarce and rare migrants, where it would be important to check each and every record, are included in the rules but no 10km square list of existing records has been provided. This means that the NBN verification tool should highlight all records of these species.

There is an issue about species that have not yet been recorded in the UK. Clearly these have not been included in the development of the geographical baselines for validation rules and, thus, the software will not flag up records of new species. This is unfortunate as species 'new to the UK' are a regular occurrence for macro-moths (far less so for butterflies) and such records should be picked out for detailed verification. At least six new macro-moth species have been recorded in Britain for the first time since 2000 and have now colonised the country (Parsons 2010). Some of these are now being recorded regularly across large areas of Britain. Several others have been recorded for the first time as immigrants, but have not become established, and this trend appears to be increasing, perhaps due to climate change (Sparks *et al*. 2005). In this situation, it is important to revise the geographical baseline distributions regularly so as to include new species recorded in the UK or to refine the NBN validation software so that it flags up records of species that are not in the NBN Species Dictionary.

Flight periods of macro-moths and butterflies

Temporal rules for the adult life-cycle stage of each taxon were developed. These were in the form of a start month and end month for each species, between which records of adult butterflies or moths should be regarded as normal occurrences for verification purposes. The period defined therefore covers the main months of the flight period(s) of each macro-moth and butterfly species in the year, including periods when adult insects may be in hibernation. The temporal rules developed do not include any other life-cycle stages.

Flight periods were extracted from field guides and other texts (Asher *et al*. 2001; Skinner 2009; Thomas and Lewington 2010; Waring *et al*. 2009). Only recent references were used so as to ensure that the flight periods defined were appropriate for modern records. The flight periods of many Lepidoptera species have changed in recent decades, both in the UK and elsewhere, almost certainly in response to climate change (Altermatt 2010; Forister and Shapiro 2003; Roy and Sparks 2000). In the vast majority of cases, the change has been for earlier and longer flight periods (as well as additional flight periods for some species), thus the temporal rules we defined from modern flight period information will tend to encompass species' historical flight periods as well.

It is normal for individuals to vary considerably in emergence time within and between populations, as well as in different years. The flight periods given in field guides aim to be indicative of the timing for the main bulk of the population rather than inclusive of all individuals of that species. Therefore, having extracted the start and end months for each taxon from the reference books, we then buffered the flight period by one month at each end to include unusually early or late flying individuals. Undoubtedly a few individuals will still occur outside the temporal rules that we have defined, but these would be worthy of further verification checks.

As with the other verification rules, it will be important to update the flight period information over time due to ongoing phenological shifts driven by climate change.

Year ranges for macro-moths and butterflies
Temporal rules were also defined to cover the acceptable year range for records of each taxon. The default start year was 1700 and default end year was not specified (to signify the present year). Taxa that were first discovered in the UK after 1700, or have colonised since that date, were given a start year appropriate to the discovery/colonisation event. Similarly, taxa that have become extinct in the UK were given an end year to their acceptable year range, equating to the year of known or assumed extinction. Thus potentially important records of species recorded either before the first published occurrence or after the last known occurrence will be highlighted by the NBN validation software.

As with each of the sets of rules, the year range temporal rule will become outdated over time as new species are recorded in the UK for the first time or species become extinct. Three macro-moth species are considered to have become extinct in the UK since the year 2000 (Parsons 2010).

3.3 Objective 3: Processes for verifying records

Dealing with records that fall outside the verification rules
Both the NMRS (macro-moth) and BNM (butterfly) recording schemes operated by Butterfly Conservation place key volunteers, termed County Moth Recorder or BNM Local Co-ordinator as appropriate, at the heart of the verification process. The final decision about the acceptance or otherwise of a record rests with them. Thus, the process for dealing with records that fall outside of the verification rules defined by this contract and used in the NBN validation software is to refer to the County Recorder for moth records or the BNM Local Co-ordinator for butterfly records. Contact details for these expert volunteers are available online or by contacting Butterfly Conservation.

The County Moth Recorders are listed on the Moths Count project website at http://www.mothscount.org/text/57/county_moth_recorders.html and the BNM Local Co-ordinators are listed on the main Butterfly Conservation website at http://www.butterfly-conservation.org/text/251/recording_contacts.html.

At the local 'county' level, the Recorder/Co-ordinator may have additional specific procedures in place for dealing with verification issues. For example, some local butterfly and moth recording groups have verification panels to whom the County Recorder/Local Co-ordinator

can refer difficult cases. Many also have local sets of verification rules that they will apply to records (as mentioned in section 3.2). In addition, many County Recorders/Local Co-ordinators will refer difficult records on to acknowledged national experts in an informal way. In Scotland, this process has been made slightly more formal for macro-moth recording, with an informal panel of experts (currently Mark Young, Roy Leverton and Keith Bland) willing to provide advice and opinions to County Moth Recorders.

It is recommended that brief information about the outcome of a verification decision is recorded, where the original details of the record are amended, and this can be done using the comment field of the record.

Classification for defining the status of records
The classification that has been proposed by the NBN previously is considered to be a sensible and reasonable approach for use with Lepidoptera records, namely:

- Correct
- Considered correct
- Requires confirmation
- Considered incorrect
- Incorrect
- Unchecked

However, it must be recognised that while this is a recommended approach, in reality, many local compilations of data do not have the capacity to flag records except by using the comments field. Recording software systems commonly used by the volunteer recording community do not have verification status fields and, even if they did, there would be substantial problems in maintaining such flags when data are moved from one format to another, as frequently occurs with hierarchical recording schemes.

In both the BNM and NMRS recording schemes, which operate by the accumulation of many independent local datasets, the current normal practice is for all records to be considered correct. Records requiring confirmation or regarded as incorrect are often removed from the dataset. Where such records do occur in local datasets a diverse range of phrases have been used to signify their status. This diversity makes it difficult for such records to be located and flagged on importation into the central NMRS database. In response, the NMRS recording scheme has recently encouraged County Moth Recorders to adopt a standardised approach to flag up unverified data of any kind, by adding the phrase “Dubious record” to the comments field. The use of the categories proposed by the NBN will now be considered as an alternative approach.

3.4 Objective 4: Identification of sensitive records

Sharing of data is one of the most challenging aspects of the collation and management of biological records. It is a complex legal situation involving the ‘copyrights’ of recorders and compliers of various local and national datasets, as well as the legislation and regulations dealing with Freedom of Information, Environmental Information Regulations and Data Protection.

Butterfly Conservation supports the seven NBN Data Exchange Principles and these form the basis for data policies and procedures relating to the butterfly and macro-moth recording schemes. Butterfly Conservation aims to maximise the use of these biological data, in a controlled and managed way, for the conservation of moths, butterflies and their habitats in line with Butterfly Conservation’s charitable objectives. Nevertheless, with recording schemes reliant on the participation and good will of thousands of volunteers, data access is inevitably an issue involving compromise. It must be recognised that recorders hold widely disparate, even polarised, views on this topic and scheme organisers have to seek a compromise that will include the majority, for the sake of generating comprehensive and useful datasets. Following extensive consultation with volunteer recorders, County Recorders, local environmental records centres, and other organisations who compile the

local data sets that make up the BNM and NMRS national datasets, public access to records has been limited to the 2km x 2km square resolution (and coarser). Access to records at capture resolution is available only in specific circumstances.

This current position is continually under review and it is Butterfly Conservation's aim to move towards increased public access to the records (i.e. at finer spatial resolution) over time, in consultation with the recording community and where such access does not threaten harm to the species, their populations and habitats or contravene the consent of individuals and organisations who have volunteered the information.

In accordance with the principles set out above, the data policies covering Butterfly Conservation's recording schemes, and the Environmental Information Regulations, certain records are deemed to be sensitive. Details of these records should not be released if they could lead to the identification of colony locations at finer than 10km x 10km grid square resolution.

All butterfly and macro-moth species have been reviewed and records of the following are deemed sensitive either under the environmental harm exemption or the volunteered information exemption. In some cases only records from certain countries or geographical areas (e.g. specific 10km x 10km grid squares of the Ordnance Survey National Grid) are considered sensitive. This is listed as an 'extent of sensitivity' field in the table below. General reasons for the sensitivities are also given.

Species	**Scientific name**	**Extent of sensitivity**	**Reasons for sensitivity***
Barberry Carpet	*Pareulype berberata*	All records	3
Black-veined Moth	*Siona lineata*	All records	3
Duke of Burgundy	*Hamearis lucina*	10km grid squares SD47, SE58, SE59 and SE88 only	2,3
Fiery Clearwing	*Pyropteron chrysidiformis*	All records	1,2
Fisher's Estuarine Moth	*Gortyna borelii*	All records	1,2,3
Grayling	*Hipparchia semele*	10km grid square SE86 only	3
Large Blue	*Phengaris arion* (formerly *Maculinea arion*)	All records	1,2,3
New Forest Burnet	*Zygaena viciae*	Scotland only	1
Pearl-bordered Fritillary	*Boloria euphrosyne*	10km grid square SE78 only	3
Reddish Buff	*Acosmetia caliginosa*	All records	3
Sussex Emerald	*Thalera fimbrialis*	All records	1,2
Dark Bordered Beauty	*Epione vespertaria*	Scotland only	1,3
Speckled Footman	*Coscinia cribaria*	All records	1,3

* Generic reasons for the justification of the sensitive status of these records are:
1. Threat of harm to individuals and/or populations of the species under the environmental harm exemption.
2. Threat of damage to habitat and/or site under the environmental harm exemption.
3. Protection of landowner/site privacy under the volunteered data exemption.

It must be recognised that the sensitivity of records, as set out above, may change with time, and must be kept under review. It should also be recognised that we have defined sensitivity only from a taxonomic perspective, whereas individual recorders and the County

Recorders/Local Co-ordinators or organisations that compile local datasets may also define particular sites as sensitive irrespective of the taxa recorded there. In particular, this will be where permission to access private land has been given in return for restrictions on the release of records of species found there. Therefore, in addition to the baseline national-level sensitivities detailed above, we recommend that compilers of local datasets take appropriate steps to identify local data sensitivities, whilst striving for openness in accordance with the NBN Data Exchange Principles.

3.5 Objective 5: Gaps in the current framework

The current framework provided by the national recording schemes, Butterflies for the New Millennium and the National Moth Recording Scheme, is deemed to be broadly fit for purpose. This is evidenced by the success of the schemes in generating huge datasets of distribution records, which have proved extremely valuable for research, policy development and the implementation of biodiversity conservation. Nevertheless there are considerable ongoing challenges to maintain and develop the schemes. Butterfly Conservation has identified that three full-time staff posts are required as a minimum in order to run the butterfly and macro-moth recording schemes, including handling the large number of both recorders and records, and providing a comprehensive, accurate and accessible database to support conservation. As a result, substantial annual funding, in the order of £160,000 p.a., is needed to continue the schemes and maintain the flow of records. Although a large amount, this real cost provides an efficient and cost-effective recording framework which generates some £7 million p.a. in volunteered effort. In spite of this, the willingness of institutional users of the data to help fund the recording schemes that generate the data has, until very recently, proved lacking.

There are strong signals that this situation is improving, but far greater co-ordination and alignment between public sector bodies with regard to accessing Lepidoptera records and supporting the ongoing collation and management of data is urgently required.

In addition to the underlying issues of sustainability, there are a number of areas of potential improvement in the flow of records within the framework. These include a commitment on the part of all local and national organisations involved in Lepidoptera recording to standardisation and improvement of data verification through vesting responsibility for all records in the taxonomic authority of County Recorders/Local Co-ordinators. We recommend that Lepidoptera records should not be made available on the NBN Gateway unless they have been verified by the appropriate County Recorder/Local Co-ordinator associated with the BNM or NMRS recording schemes, irrespective of details provided in the metadata associated with such datasets.

The reliance on volunteer County Recorders/Local Co-ordinators brings many benefits (not least expertise, dedication and cost-effectiveness) but also constraints. There is regular, albeit slow, turnover of these key volunteers. Often efficient handovers take place from an outgoing Recorder/Co-ordinator to the incoming one and there is little or no disruption of data flow. On occasions however, counties are left temporarily without a Recorder/Co-ordinator and the flow of data is disrupted.

Further development of the ability of County Recorders/Local Co-ordinators to handle the flow and volume of records (e.g. through greater use of information technology) and to instigate and manage relationships with other organisations involved with data compilation or data use at the local scale would also be beneficial.

Finally, continued support from local and national organisations to aid the engagement, recruitment, development, mentoring and informal training of recorders by County Recorders/Local Co-ordinators and recording groups would be very valuable, particularly in under-recorded areas of the UK.

4. Conclusions

Butterflies and macro-moth recording is popular and well-organised via national recording schemes that operate from local (county) to UK-levels. As a result, huge numbers of Lepidoptera records are available on the NBN Gateway, both from the national recording schemes and, in some cases, also direct from local data compilers. Generally, data quality is good and, in part through the activities of the national recording schemes, expert volunteer County Recorders/Local Co-ordinators are empowered to verify records at the local level. However, a range of mechanisms to standardise recording and improve the efficiency of data verification are recommended in this report.

In particular, baseline information has been provided to form a series of taxon-specific rules to assess the spatial, temporal and identification veracity of macro-moth and butterfly records in datasets via the recently-developed NBN validation software. This should enable data compilers with limited taxonomic expertise for the group to rapidly generate subsets of records that merit scrutiny from the County Recorder/Local Co-ordinator for verification purposes.

It is hoped that the accuracy of Lepidoptera records at local and national (including NBN Gateway) levels will be improved further as a result.

5. Acknowledgements

The views of the authors provided in this report have been informed and shaped by discussions with numerous members of the biological recording community over many years and we are very grateful for their input. We would like to specifically thank Tony Davis and Roy Leverton for their help in developing and refining the identification difficulty classification. We are also grateful to Graham French for technical support and to Paula Lightfoot for guidance during this contract.

6. References

Altermatt, F. (2010) Climatic warming increases voltinism in European butterflies and moths. *Proceeding of the Royal Society B* **277**: 1281-1287.

Asher, J., Fox, R. and Warren, M.S. (2011) British butterfly distributions and the 2010 target. *Journal of Insect Conservation* **15**: 291-299.

Asher, J., Warren, M., Fox, R., Harding, P., Jeffcoate, G. and Jeffcoate, S. (2001) *The millennium atlas of butterflies in Britain and Ireland*. Oxford University Press, Oxford.

Forister, M.L. and Shapiro, A.M. (2003) Climatic trends and advancing spring flight of butterflies in lowland California. *Global Change Biology* **9**: 1130-1135.

Fox, R., Asher, J., Brereton, T., Roy, D. and Warren, M. (2006) *The state of butterflies in Britain and Ireland*. Pisces Publications, Newbury.

Fox, R., Randle, Z., Hill, L., Anders, S., Wiffen, L. and Parsons, M.S. (2011) Moths Count: recording moths for conservation in the UK. *Journal of Insect Conservation* **15**: 55-68.

Hill, L., Randle, Z., Fox, R. and Parsons, M. (2010) *Provisional atlas of the UK's larger moths*. Butterfly Conservation, Wareham.

Parsons, M. (2010) The changing moth and butterfly fauna of Britain – the first decade of the twenty-first century (2000-2009). *Entomologist's Record and Journal of Variation* **122**: 13-22.

Roy, D.B. and Sparks, T.H. (2000) Phenology of British butterflies and climate change. *Global Change Biology* **6**: 407-416.

Skinner, B. (2009) *Colour identification guide to the moths of the British Isles*. Apollo Books, Stenstrup.

Sparks, T.H., Roy, D.B. and Dennis, R.L.H. (2005) The influence of temperature on migration of Lepidoptera into Britain. *Global Change Biology* **11**: 507-514.

Thomas, J. and Lewington, R. (2010) *The butterflies of Britain and Ireland*. British Wildlife Publishing, Gillingham.

Townsend, M.C., Clifton, J. and Goodey, B. (2010) *British and Irish moths: an illustrated guide to selected difficult species (covering the use of genitalia characters and other features)*. Butterfly Conservation, Wareham.

Waring, P., Townsend, M. and Lewington, R. (2009) *Field guide to the moths of Great Britain and Ireland (second edition)*. British Wildlife Publishing, Gillingham.

Warren, M.S., Hill, J.K., Thomas, J.A., Asher, J., Fox, R., Huntley, B., Roy, D.B., Telfer, M.G., Jeffcoate, S., Harding, P., Jeffcoate, G., Willis, S.G., Greatorex-Davies, J.N., Moss, D. and Thomas, C.D. (2001) Rapid responses of British butterflies to opposing forces of climate and habitat change. *Nature* **414**: 65-69.