
Formatting and submitting datasets for the National Moth Recording Scheme

Introduction

Exchange of datasets between recorders using the same software systems i.e. MapMate-to-MapMate or *Recorder-to-Recorder* is a relatively straight-forward process which requires little intervention from the users and is not described here.

However, the export of datasets between different systems i.e. MapMate-to-Recorder, requires the exported dataset to be compatible with the data import specification of *Recorder* as used by the National Moth Recording Scheme (NMRS). This requires manual intervention by the NMRS data manager which, dependant on the size and state of the exported dataset can take some time.

The intended audience for this document is the Vice-county Moth Recorder (CMR) network irrespective of how their electronic datasets are stored i.e. MapMate, Microsoft Excel etc.

This document is not intended to tell the reader how to manage their datasets. It merely advises and provides the CMR with informed choices and considered “best practice” solutions on how to manage local datasets, in order to maximise compatibility with the NMRS database and save time.

This document is based in part on material produced by Martin C. Harvey with additional input from Tim Norriss and Ian Thirlwell and thanks are extended to Martin, Tim and Ian for their assistance in preparing this document.

The guidelines within this document may be enhanced or changed from time-to-time and will be posted to the Moths Count website (www.mothscount.org):

- To reflect feedback and constructive ideas from CMRs.
- When better practises are identified importing data into the NMRS database due to changes in software and external procedures.

When to send in your datasets

All CMRs should be aware, for example via eMoth, that the initial call for moth datasets is well under way; however, not all datasets from all vice-counties will be called for at the same time as this is an unmanageable task.

It is expected that the request for subsequent datasets will occur annually commencing around spring by which time the previous year’s data entry may have taken place over winter.

To ensure all additions, amendments and deletions are incorporated into the NMRS database, the full and complete vice-county dataset is required year-on-year.

Data fields required to import to the NMRS

These are the fields or columns in your exported data, some are compulsory and some are recommendations only.

The headings described here are taken from the MapMate Import Specification, the most popular software choice of CMRs; however, the same headings are relevant whatever the choice of electronic storage by the CMR.

If submitting data via Microsoft Excel or Microsoft Access, please ensure the top row is reserved for your headings only. Each subsequent row should consist of a single moth record only.

Taxon	Site	Gridref	VC	Recorder	Determiner	Date	Quantity	Method	Sex	Stage	Status	Comment

Please do not forward datasets in a 'Crosstab' format (similar to the example below). Crosstabs cannot be easily reformatted to a "one record per row" format required to import to a database.

	A local NR	A Wood	A local NR	A Wood
	125W MV	125W MV	125W MV	125W MV
	13/6/2007	15/7/2007	17/7/2007	21/10/2007
2089 Heart and Dart	12	45	16	-
2107 Large Yellow Underwing	9	-	17	13
2125 Setaceous Hebrew Character	-	9	23	14

Important – Non-printable ASCII characters in fields: Some submitted datasets have embedded characters in specifically the 'Site name' and 'Comments' fields i.e. Tabs, Carriage-return/Line Feeds (CRLF). It is highly likely these characters have not been inserted deliberately by recorders but introduced by some unknown means. At present, these characters are being removed during pre-processing stages before being imported into the NMRS database.

Code / Taxon / Vernacular

- **Code** - The Bradley number is voluntary, although can be useful as a cross-referencing tool when used in conjunction with *Taxon* and/or *Vernacular*.
- **Taxon** - The scientific name is compulsory and a requirement.
- **Vernacular** - The English name is voluntary and can be very useful to further assist with data confirmation along with *Code* and *Taxon* to help eliminate nomenclature confusion e.g. *Lomographa bimaculata* (Fabr.) **White-pinion Spotted** and *Cosmia diffinis* (L.) **White-spotted Pinion** and where species with multiple sub-species are allocated different English names i.e. *Eupithecia denotata denotata* (Hb.) **Campanula Pug** and *Eupithecia denotata jasioneata* Crewe **Jasione Pug**.

Site name / OS Grid Reference

Both are compulsory; records submitted with no recognisable OS grid reference are not included in the NMRS database.

Best practice is to ensure all OS grid references irrespective of spatial resolution (10m, 100m, 1km, 2km or 10km) have unique site names to avoid confusion, although it is recognised in practice this isn't always easy to achieve.

The following site name / OS grid reference examples follow the above best practice; however, it may not always be realistic to use varying levels of grid reference precision in the way shown here.

Oak Woods NNR (SY1835)

Oak Woods NNR: West Meadow (SY182353)

Oak Woods NNR: West Meadow: Duck Pond (SY18273538)

Where applicable, use hierarchical site names similar to the above example, starting with the largest area - this makes it easy to combine, extract and analyse records for all Oak Woods NNR sites as per this example. *However, please restrict the site name to no more than 64 characters if possible.*

Please don't set up sites called "Unknown" unless you have to, it is much better to have at least a vague site name for each grid reference; the main element of a site name should at least be recognisable from an Ordnance Survey map, preferably the 1:25,000 scale series.

Important – Data Protection Act: *Please don't use a full postal address as a site name. When matched with a recorder's name this construes as personal data which is covered by the Data Protection Act 1984 (Revised 1998) and requires special considerations under that Act. The NMRS data manager will make best endeavours to ensure personal data doesn't get into the public domain.*

Please use at least a six-figure (100m) grid reference precision for records as this detail is vital for many recording and conservation activities; eight-figure and 10-figure grid references are most welcome where possible; four-figure (1km) or tetrad (2km) grid references for larger sites are acceptable.

Two-figure (10km) grid references are also acceptable, especially for vague sites where site confidentiality is required by landowners, for example.

For species of conservation interest, more precise eight-figure grid references (10m) are useful for the exact locations of the sighting or any breeding colony.

Irrespective of grid reference resolution recorded, by default the highest resolution the NMRS will display records via online distribution maps is 2km.

The NMRS database supports four spatial reference systems:

- OSGB (Ordnance Survey Great Britain)
- OSNI (Ordnance Survey Northern Ireland)
- UTM (Universal Transverse Mercator)
- Latitude and longitude

Recorder (Observer)

Compulsory; Please note: Recorders' names will not be used for any other purpose than correlating a species record to a recorder i.e. defining the ownership of that record for verification purposes, for example, or acknowledgement.

The best formats for names are:

R. B. Powell Preferred format

Jackie Lee Preferred format

Jones, J. Please note: If the *{surname, initial}* format is used, please only use one initial with no titles or forenames.

The following formats are also satisfactory:

Robert B. Powell

Mr J. C. Jones

Dr Jim Hopkirk

Mrs Jackie E. Lee

You can combine multiple recorders' names, 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, Jones, J. and Dr Jim Hopkirk"

Col / Dr / Miss / Mr / Mrs / Ms / Prof / Rev are titles that can be accommodated at the time of writing.

Determiner

The 'determiner' is the person who actually identified the species. Usually this will be the same as the recorder.

If there is a separate determiner e.g. if a specimen has been given to someone else to identify, then their name must be added just as for the recorder. Some people also like to record that a record has been 'confirmed', i.e. R.B. Powell recorded and identified a species, but he got someone else to give a second opinion and confirm it. In this case R.B. Powell would go down as recorder and as determiner, but the name of the confirmer could be put in the *Comment* field to provide additional evidence of the species identification.

In all instances, 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

Compulsory.

This must be in *dd/mm/yyyy* format e.g. 12/07/2007. Please don't use two-digit years i.e. 07 for 2007 as this could also mean 1907, 1807 etc.

If the date is a month only, use the format *00/mm/yyyy*; if the date is a year only, use *00/00/yyyy*.

If your date is a year range use *yyyy->yyyy* (i.e. with a dash followed by a greater-than sign between).

If your date is a date range use *dd/mm/yyyy-dd/mm/yyyy* (i.e. two dates with a dash between).

Important: *If data is submitted with the date formatted in any other way i.e. dd-mm-yyyy, dd-mmm-yyyy, mmm-yyyy or space-delimited etc., it will be very difficult to resolve the dates into the correct format. Please also ensure no future dates are present as these will get rejected during the import process.*

Quantity

Quantity or abundance is voluntary although it is required by the import process; however, abundance will be utilised by the NMRS for data analysis; therefore, please where possible provide this data.

If no abundance is recorded please use "0" (zero) or "Present". Please don't just use "1" to show that the species is present unless it really was only "1" of that particular species recorded.

Method

Voluntary. Use this field to note the method of recording e.g. light trap, daytime observation, dusking etc. If a light trap, please specify the bulb type and where known e.g. 125W MV.

Sex (Gender)

Voluntary; however, if gender is recorded, please annotate accordingly and create separate records for each (if recorded together).

Stage (Life cycle)

Compulsory; i.e. Adult, Larval, Pupal, Exuviae, Not recorded etc.

The NMRS will produce charts and atlases based on mature and immature stages. Knowledge of locations of immature stages is important as it equates to breeding locations, whereas adult moth sightings enable phenology or flight charts.

Comment

Any relevant comment can go here, record whether you have a photo, retained voucher specimen or seen the genitalia, or if anyone else saw and confirmed your identification - this is extremely useful supporting evidence, especially for rare or difficult species, and the more you can put in here the better!

What formats to send your data to the NMRS

- ***MapMate 'sync' (*.sqz)***
- ***CSV file (*.csv)***
- ***Excel file (*.xls, *.xlsx)***
- ***Text file (*.txt)***
- NBN Data (*.xml)
- NBN Access Database (*.zip)
- Dbase file (*.dbf)
- Lotus 1-2-3 file (*.wk1)
- Paradox file (*.db)
- QuattroPro file (*.wq1)

The preferred file formats are in ***Bold/Italic*** type.

...and finally!

Notwithstanding any part of this document, please do not hesitate to contact me if you have any issues or wish to discuss anything to do with sending your records to the NMRS.

This document will be updated periodically and these updates will be made available to download on the Moths Count website <http://www.mothscount.org/>.

Les Hill
Data Manager, National Moth Recording Scheme
Butterfly Conservation
Manor Yard
East Lulworth
Wareham
Dorset
BH20 5QP

T: 01929 406008

E: lhill@butterfly-conservation.org

MapMate is a Registered Trade Mark of Teknica Ltd.

Recorder is a Registered Trade Mark of the Joint Nature Conservancy Council and Dorset Software.