

# Validation (and verification)

---

David Roy

UKCEH & University of Exeter



UK Centre for  
Ecology & Hydrology



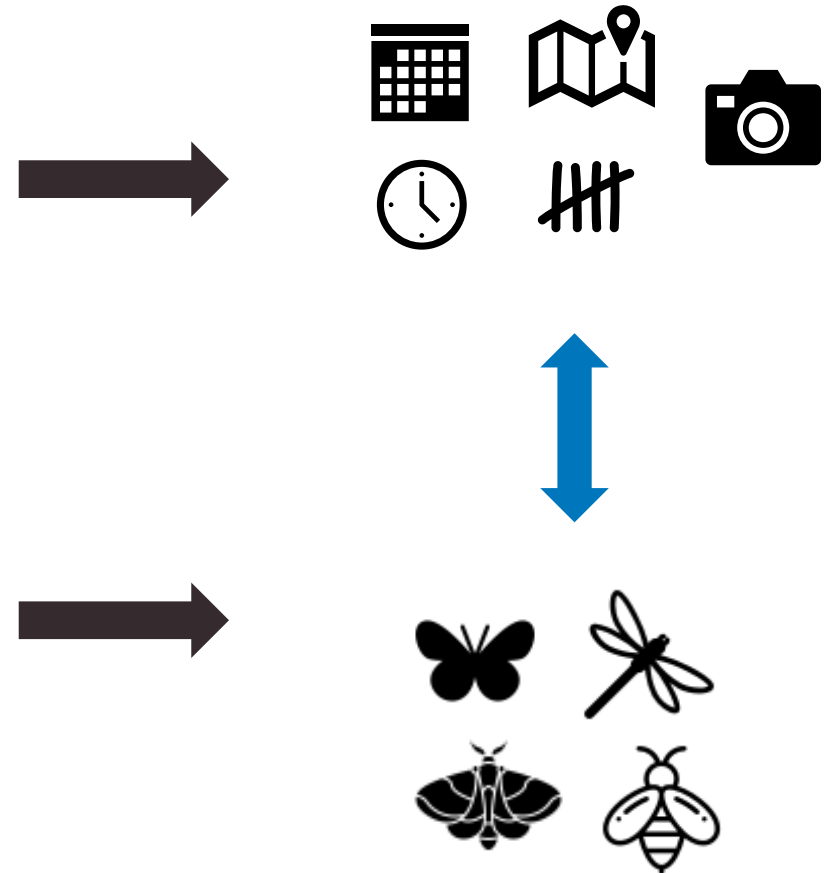
University of Exeter  
Cornwall



# Validation and verification are important for ensuring good data quality

Data **validation**: carrying out standardised, often automated checks on the “completeness”, accuracy of transmission and **validity** of the content of a **record**.

Data **verification**: ensuring the accuracy of the identification of the things being **recorded**.



# What approaches do we use in eBMS online



- Main national portal for biodiversity monitoring in the UK since 2012
- Used by UKBMS
- Indicia open source software
- ~4 million observations per year, with ~2 million photos
- ~24k users per year
- Over 700 expert verifiers

## Crowd sourcing

Wisdom of the crowd

Wisdom in the crowd



# What approaches do we use

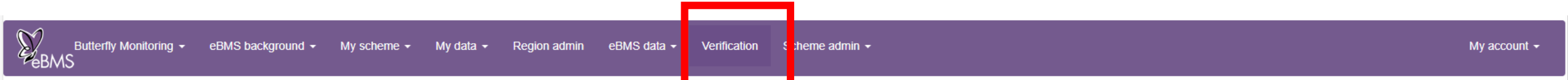


Group	Records (m)	Verified (%)	Verifiers (2024)	Photos (m)
All	29.6	74	700	10.7
Butterflies - all	12.8	94	133	0.7
<i>Butterflies - ukbms</i>	9.9	97		
Moths	3.8	83	138	1.5



# Tools to support data review and verification

<https://butterfly-monitoring.net/verification>



Verification page link

## Verification

View Translate

The most important functions inside the verification system are explained below. For more information read the [iRecord manual](#). This page shows all the records registered based on the filter applied (country, species group, survey...), and that need review.

There is an **automatic check** of all the records based on the flight period and distribution of species, so you see green or red thumbs showing the result of this automatic check. If you hover the mouse over the red thumb symbol it explains why it was rejected.

You can use the different functions of the verification menu to find better records. For example, you can click on the white arrows of each column to order it, like the Check column to see the rejected records. You can search directly by species, survey method, location, recorder, etc., on the searching boxes and you can use the filters on the top for filter by picture or record status. As other eBMS tables on the website, you can change what you see on this table using the tool symbol on the top.

**How to correct records.** First, select one of the records you want to edit, it will highlight the whole row in blue. Then buttons on the right appear to use for correcting a record. The most common one is the label button, to redetermine a record, changing the species name. You can add a comment or not.

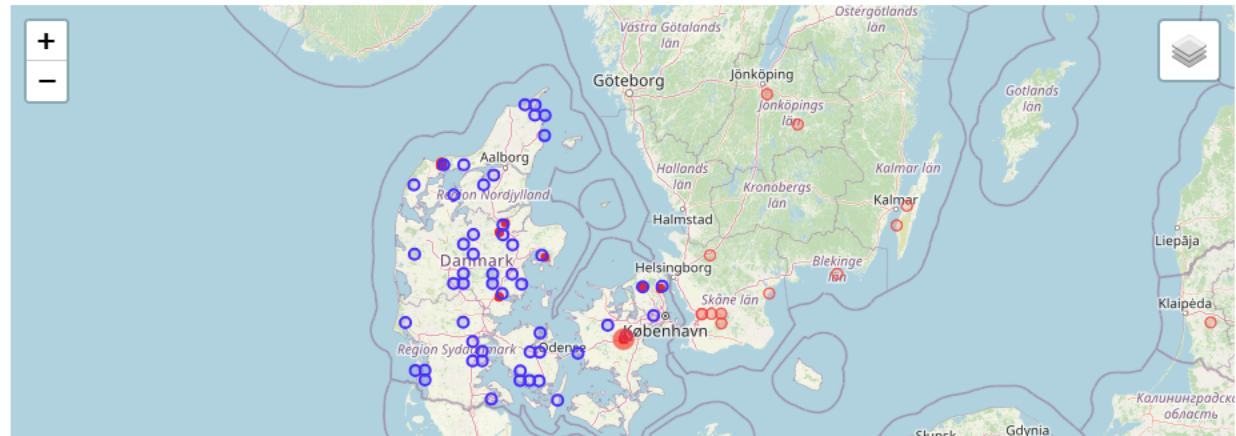
**Validate the records.** On the first column appears a clock symbol for all records waiting for a review. To change that you can select several records at once and apply the review action. First, click on the first symbol on the right, then a column of tick squares will be available. You can select the whole page in one click, selecting the first square and then, you can select the action you want in the upper part. You can do it one by one as well, but this way is faster. Meaning of the symbols: First double green tick marks for records with pictures; one green tick for records considered correct (good location, time of the year); orange tick, plausible but not able to corroborate (unconfirmed); red cross unable to verify; white cross record incorrect.

Search:  Context: Denmark Andreas verifier Status: Pending

Photos: -no filter-

Stored filters Filter: Select stored filter... Apply filter Reset Create a filter

ID	Src	Species	Map ref.	Date	Recorder	Determiner	Media	Checks	Last updated	Sample ID	Verifier name
4056096 3	118 562	Aeshna grandis	57.06085 N 8.8623E	2011-08-18	Nicholas Bell				31/03/2025 08:20	2878877 1	
4056096 2	118 562	Aeshna grandis	57.06116 N 8.86408E	2011-08-18	Nicholas Bell				31/03/2025 08:20	2878876 9	



# Tools to support data review and verification

**Verification grid**

**Top-level filters**

**Extra filter options**

**Map pane**

**Record details pane**

**Verification controls**

Media	ID	Species	Common name	Location	Date	Recorder	Determiner	Check	Day of year	Month	Life stage
	1312851	Bombylius major	Bee Fly	London N11 3HT	31/03/2020	Penny Marlon	Penny	0	91	3	adult
	1312844	Bombylius major	Bee Fly	Penparcau Hwb Bee Garden	31/03/2020	[recorder names hidden for demo purposes]		0	91	3	
	1312843	Bombylius minor	Heath Bee-fly	Hove	01/04/2020			0	92	4	not recorded
	1312795	Bombylius discolor	Dotted Bee-fly	Newhaven: Fakholme Road	31/03/2020			0	91	3	adult
	1312790	Bombylius major	Bee Fly	Titchmarsh LNR	31/03/2020			0	91	3	not recorded
	1312784	Bombylius major	Bee Fly	Thrapston Lakes	31/03/2020			0	91	3	adult
	1312781	Bombylius major	Bee Fly	Titchmarsh LNR	31/03/2020			0	91	3	adult
	1312777	Bombylius major	Bee Fly	Adjacent to river, along path, Okehampton	31/03/2020			0	91	3	adult
	1312769	Bombylius major	Bee Fly	SO18 9DF	31/03/2020			0	91	3	adult

Showing 1 to 30 of 65

Details | Comments | Recorder experience

**ID**[status][checks] **1312851** | ⌵ | ⌵

**Accepted name** Bombylius major Linnaeus, 1758

**Common name** Bee Fly

**Name as entered** Dark-edged Bee-fly

**Location** London N11 3HT

**Vice County** Middlesex

**Grid ref** TQ28099228

**Date seen** 31/03/2020

**Recorder**

**Determiner**

**Dataset** iRecord :: iRecord general data

**Sample comment** We found it in our kitchen nearly drowned in a vase. Gave it sugar water, warmed it up and after 30 minutes it flew off. Very satisfying.

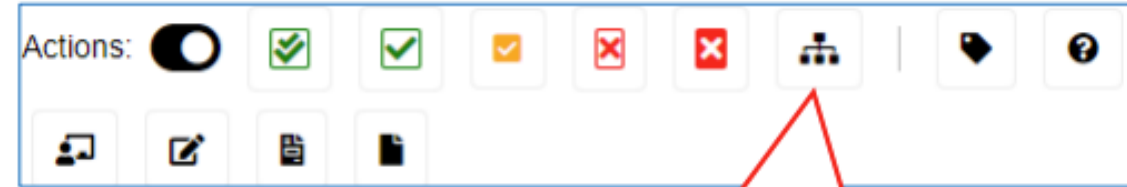
**Occurrence comment**

**Additional occurrence attributes**

**Submitted on** 01/04/2020 12:28

**Last updated on** 01/04/2020 12:28

# Tools to support data review and verification



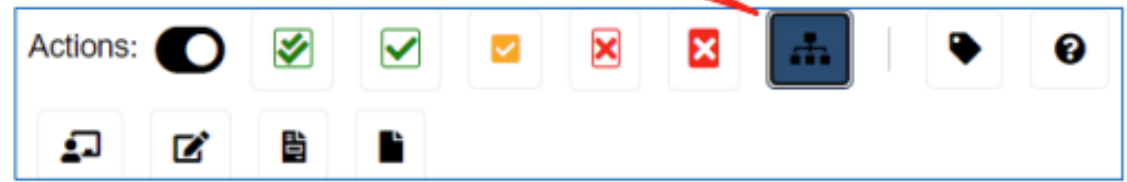
The new "sample bulk verification" button appears to the right of the main verification buttons

New function for verification page:  
sample bulk verification button



It remains greyed out for 'normal' records, and can only be applied to records from a set of 'child samples' such as a transect

If you click on any one record from within a set of child samples the button becomes active



# Tools to support data review and verification

Here is a set of records from a (non-genuine!) timed count of Swallowtail butterflies – each record from the count appears to be identical



First, click on the sample bulk verification button to activate it, and avoid having to verify each individual Swallowtail record



Next, click on the appropriate verification button to accept any one record

Your decision will be applied to all records of that one species in the sample; you can add a comment if needed (which will be applied to each individual record)

ID	Src	Species	Common name	Location	Map ref.	VC no.	Date	Recorder	Determiner	Media	Checks	Last updated
29886620	231685	Papilio machaon	Swallowtail	TEST ONLY	SP820062	24	11/04/2023	Harvey, Martin C.				11/04/2023 11:39
29886619	231685	Papilio machaon	Swallowtail	TEST ONLY	SP820062	24	11/04/2023	Harvey, Martin C.				11/04/2023 11:39
29886618	231685	Papilio machaon	Swallowtail	TEST ONLY	SP820062	24	11/04/2023	Harvey, Martin C.				11/04/2023 11:39
29886617	231685	Papilio machaon	Swallowtail	TEST ONLY	SP820062	24	11/04/2023	Harvey, Martin C.				11/04/2023 11:39
29886616	231685	Papilio machaon	Swallowtail	TEST ONLY	SP820062	24	11/04/2023	Harvey, Martin C.				11/04/2023 11:39
29886615	231685	Papilio machaon	Swallowtail	TEST ONLY	SP820062	24	11/04/2023	Harvey, Martin C.				11/04/2023 11:39
29886614	231685	Papilio machaon	Swallowtail	TEST ONLY	SP820062	24	11/04/2023	Harvey, Martin C.				11/04/2023 11:39
29886612	231685	Papilio machaon	Swallowtail	TEST ONLY	SP821062	24	11/04/2023	Harvey, Martin C.				11/04/2023 11:39
29886611	231685	Papilio machaon	Swallowtail	TEST ONLY	SP820062	24	11/04/2023	Harvey, Martin C.				11/04/2023 11:39
29886610	231685	Papilio machaon	Swallowtail	TEST ONLY	SP820062	24	11/04/2023	Harvey, Martin C.				11/04/2023 11:39







Set status to Accepted as considered correct

This verification decision will also be applied to other records of the same taxon within the parent sample (e.g. within the transect or timed count)!

Add the following comment:



# Tools to support data review and verification

Symbol	Meaning
	Passes all rule checks; <u>OR</u> No rule checks available for this species
	Fails at least one rule, in which one or more of the three following reasons will be displayed:
	Fails on identification difficulty (hover over icon to see details)
	Fails due to being outside the known range
	Fails due to being outside the flight/activity period
	Records from this source do not have rule checks applied

# Tools to support data review and verification

Search:  Status:  Photos:

Stored filters\* Filter:

**What** Select a list of species or species groups to include

**Where** Records in Denmark

**When** Define a date range for records to include

**Who** Define whose records to include

**Record ID** Select records by record ID

**Quality** Exclude Not accepted (all) Distribution check failed

**Source** Select records based on source website, survey or input form

To save these filter settings for future use, give the filter a name then click Save Filter.

Filter name:

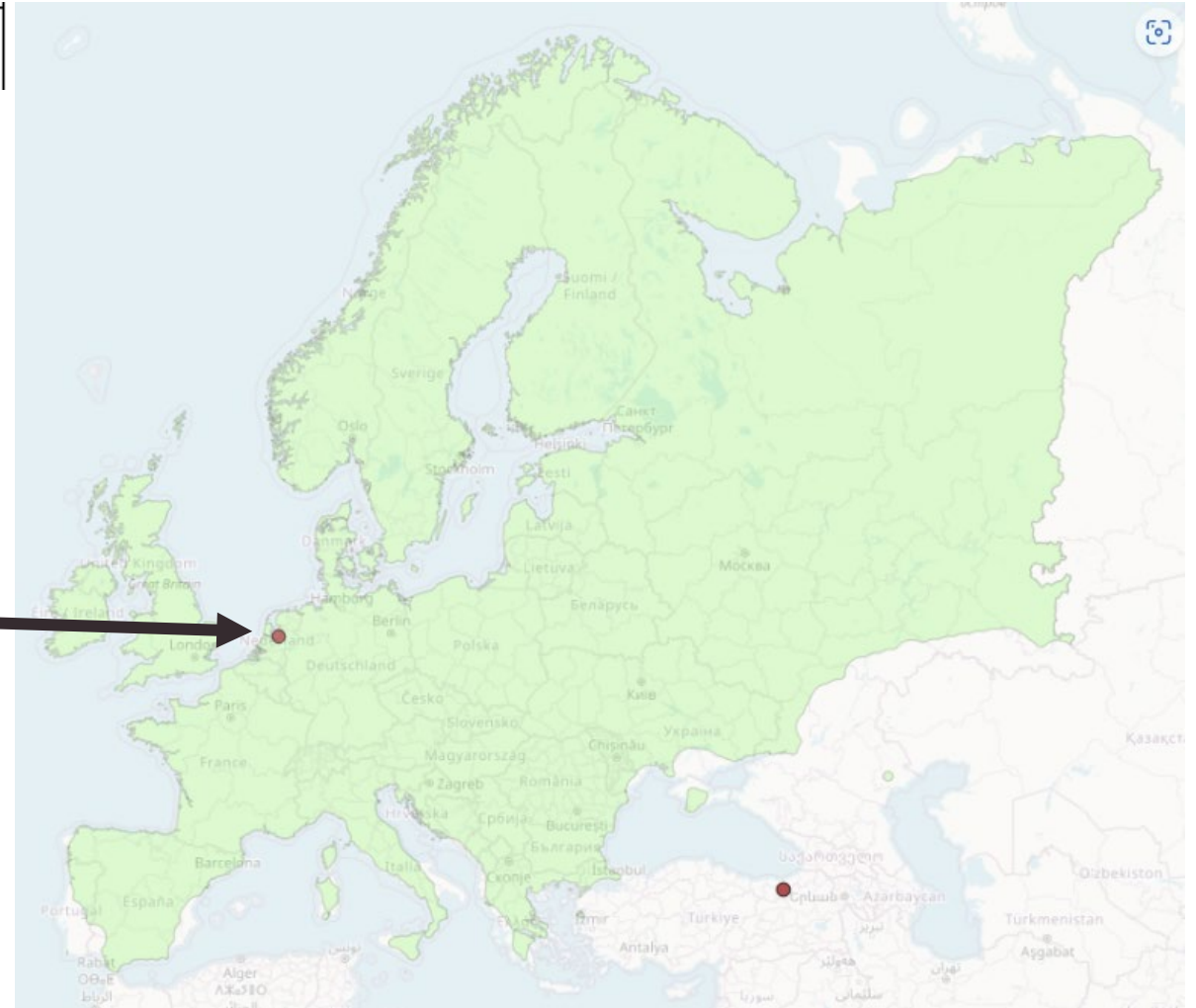
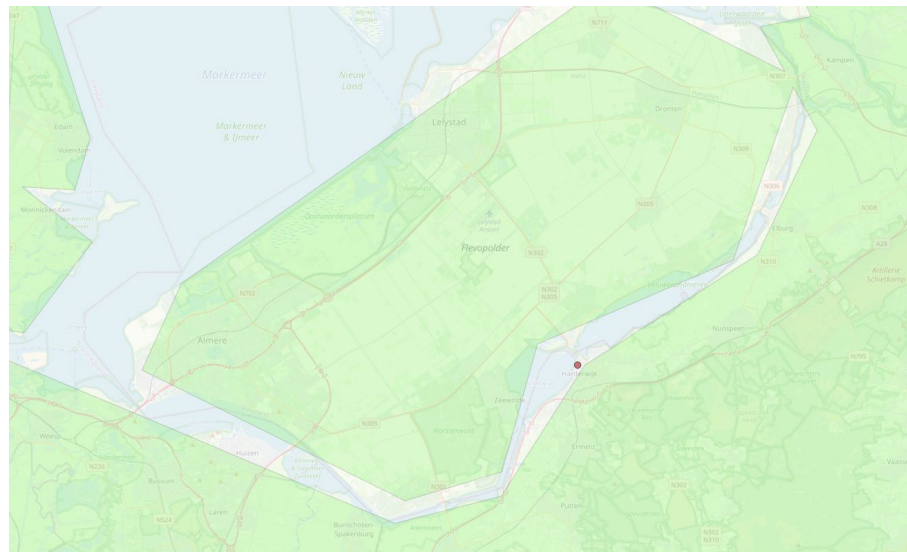
ID	Src	Species	Map ref.	Date	Determiner	Media	Checks	Last updated	Sample ID	Verifier name
<input checked="" type="checkbox"/>	<a href="#">118 565</a>	<i>Polyommatus icarus</i>	56.91856N 8.43681E	2023-07-28			<input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	19/09/2024 08:56	23104771	Bennetsen Boe, Andreas
<input checked="" type="checkbox"/>	<a href="#">118 565</a>	<i>Polyommatus icarus</i>	56.91882N 8.4371E	2023-07-28			<input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	19/09/2024 08:56	23104769	Bennetsen Boe, Andreas
<input checked="" type="checkbox"/>	<a href="#">118 565</a>	<i>Polyommatus icarus</i>	56.91882N 8.4371E	2023-07-28			<input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	19/09/2024 08:56	23104768	Bennetsen Boe, Andreas
<input checked="" type="checkbox"/>	<a href="#">118 565</a>	<i>Polyommatus icarus</i>	56.91924N 8.4371E	2023-07-28			<input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	19/09/2024 08:56	23104766	Bennetsen Boe, Andreas
<input checked="" type="checkbox"/>	<a href="#">118 565</a>	<i>Polyommatus icarus</i>	56.91878N 8.43699E	2023-07-28			<input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	19/09/2024 08:56	23104764	Bennetsen Boe, Andreas
<input checked="" type="checkbox"/>	<a href="#">118 565</a>	<i>Polyommatus icarus</i>	56.91878N 8.43699E	2023-07-28			<input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	19/09/2024 08:56	23104763	Bennetsen Boe, Andreas
<input checked="" type="checkbox"/>	<a href="#">118 565</a>	<i>Polyommatus icarus</i>	56.91879N 8.43701E	2023-07-28			<input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	19/09/2024 08:56	23104762	Bennetsen Boe, Andreas
<input checked="" type="checkbox"/>	<a href="#">118 565</a>	<i>Polyommatus icarus</i>	56.91874N 8.43697E	2023-07-28			<input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	19/09/2024 08:56	23104760	Bennetsen Boe, Andreas
<input checked="" type="checkbox"/>	<a href="#">118 565</a>	<i>Polyommatus icarus</i>	56.91874N 8.43697E	2023-07-28			<input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	19/09/2024 08:56	23104759	Bennetsen

Select a row to view details

# Tools to support data review and verification



Fails due to being outside the known range



# Tools to support data review and verification



Fails due to being outside the flight/activity period



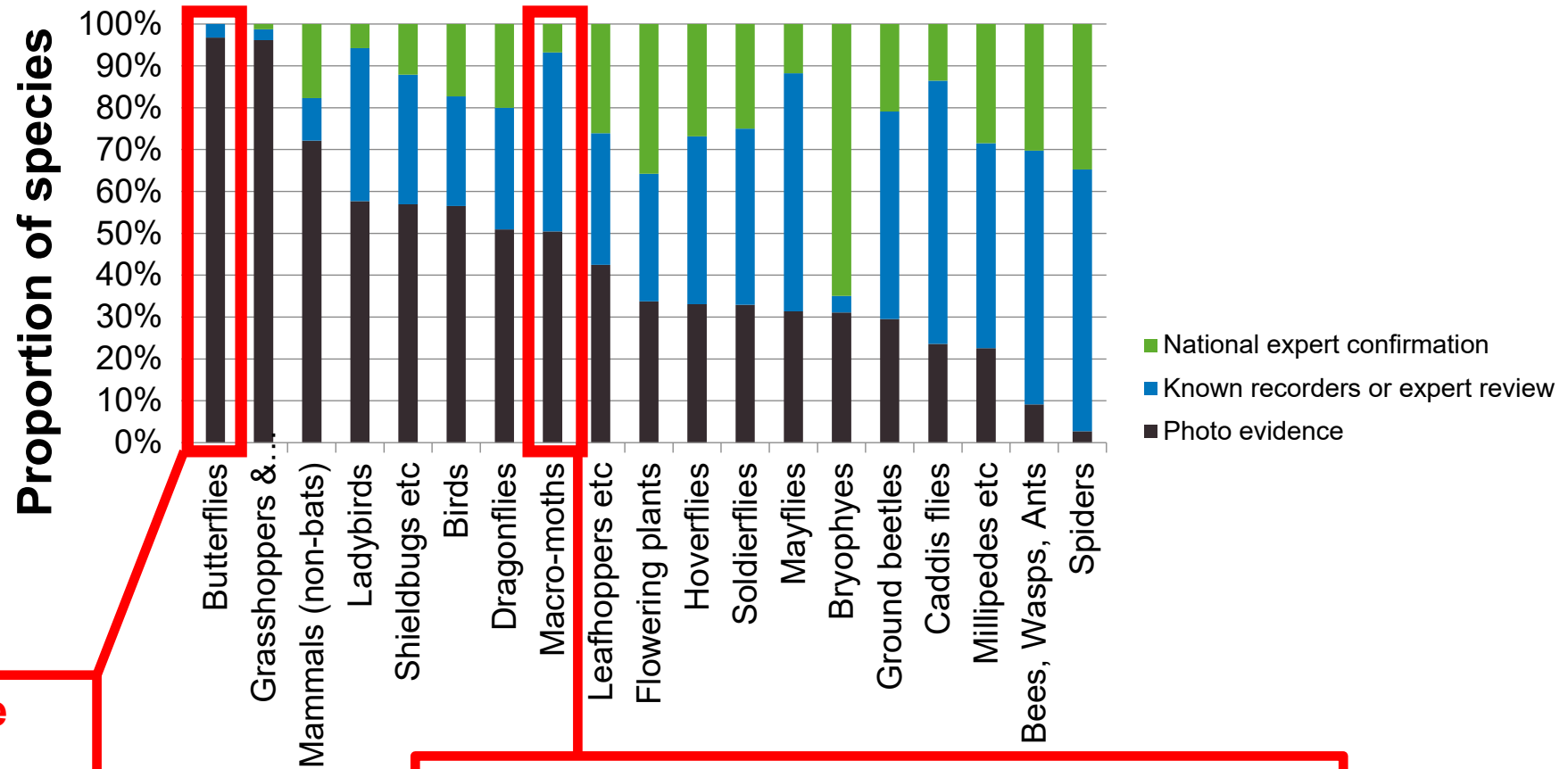
species	region	phen
Aglais ichnusa	Alpine	16/07-31/07
Aglais ichnusa	Mediterranean	16/03-15/04; 16/05-31/07; 16/08-31/08; 16/12-31/12
Aglais io	Alpine	16/03-15/10
Aglais io	Arctic	01/02-15/02; 16/07-31/07; 01/09-31/10; 16/11-30/11
Aglais io	Atlantic	16/03-31/05; 16/06-15/10
Aglais io	BlackSea	16/03-31/03; 16/04-15/06; 01/07-31/07
Aglais io	Boreal	01/01-15/01; 01/04-15/06; 01/07-15/09
Aglais io	Continental	01/01-15/01; 16/03-30/09
Aglais io	Mediterranean	16/02-31/08; 16/09-30/09; 16/10-31/10
Aglais io	Outside	01/06-30/06
Aglais io	Pannonian	16/03-15/05; 01/06-31/08; 16/09-15/10
Aglais io	Steppic	01/04-15/05; 16/06-15/10

# Questions for discussion



Fails on identification difficulty (hover over icon to see details)

## 1. Could we classify European butterflies by their ID difficulty?



Almost all uk butterflies can be identified from a photo

Half uk macro-moths can be identified from a photo; most by experienced recorders



# Questions for discussion

## 2. Do we want a photo-led version of verification page

### Verification - photo based

This is the **photo version** of the main verification page. A full guide to its use is available on the [verifiers' resources](#) page. See the similar page for iNaturalist records under the Verify menu.

Search:  Context: Moths; Oxfordshire; Roy, David Status: Accepted Photos: -no filter-

Stored filters Filter: Select stored filter... Apply filter Reset Create a filter



*Eupithecia centaureata*  
Lime-speck Pug  
Ashcroft, Ruth  
01/09/2024  
not recorded  
adult  
👍



*Acleris cristana*  
Tufted Button  
Barber, Nicholas  
23/03/2025  
not recorded  
adult  
👍



*Panolis flammea*  
Pine Beauty  
Barber, Nicholas  
23/03/2025  
not recorded  
adult  
👍



*Selenia dentaria*  
Early Thorn  
Barber, Nicholas  
23/03/2025  
not recorded  
adult  
👍



# Questions for discussion

3. Do we want to include results of species classifier

The screenshot displays a specimen record for *Xanthorhoe designata* (Flame Carpet) by David Short on 30/05/2018. It features three photographs of the moth from different angles. Below the images, a black box with a red warning icon states: "Image classifier conflicts with identification provided". To the right, an "Image classifier results" section shows three options: "Dark-bar Twin-spot Carpet" (53%), "Red Twin-spot Carpet" (25%), and "Flame Carpet" (6%). A red callout box points to the 6% result with the text: "Click on alternative suggestion to redetermine record". At the bottom, an "Actions" bar contains various icons for record management.

Species	Percentage
Dark-bar Twin-spot Carpet	53%
Red Twin-spot Carpet	25%
Flame Carpet	6%