Review of 2008-2023 GMS main season data

Version 1.0: 20/08/2025

Steve Roberts, Centre for Biostatistics, University of Manchester

1 Contents

> ⓘ Feedback
>
> Comments and feedback welcome to steve.roberts@manchester.ac.uk or SteveRoberts@live.co.uk

2 Background

This document aims to provide a summary of what data is available from the GMS, with a particular focus on what is needed for a reliable analysis of trends. The purpose is to inform subsequent analysis of the data.

- Descriptive statistics and plots summarising aspects of the data
- Plots to check for data issues and characteristics that will determine the analysis

3 Dataset Description

This report is based on 2 datasets:

- All records from the main summer scheme (main.Rd)
- A subset selected for the analysis of trends (trends.Rd) - traps with 3 or more years worth of data.

Details of the inclusion/exclusion criteria and cleaning etc in the “About the GMS data” Word document.

The data are analysed by trap not garden - some gardens run multiple traps or change dthe trap type over time. We focus on the data for each trap.

Numbers in each dataset are given in *Table 1* below:

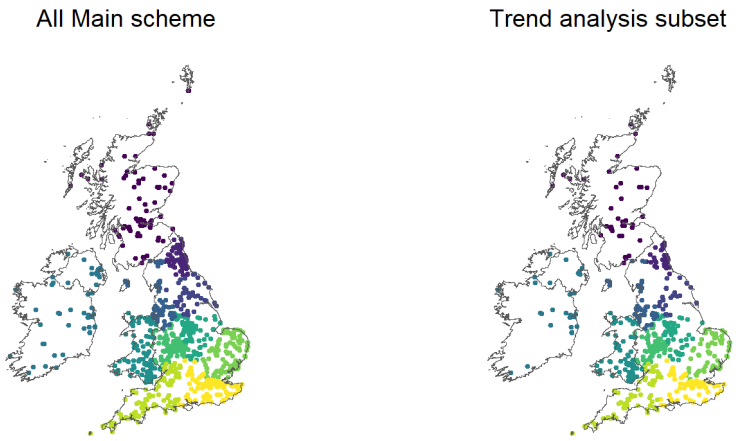
Table 1: Dataset Characteristics

	Full Data	Main Summer scheme	Trend analysis dataset
Records	2,325,096	2,192,675	1,864,894
Gardens	1085	1013	611
Traps	1533	1347	671
Moths	8,189,241	7,791,219	6,742,085
Species	457	455	439
StartYear	2003	2008	2008
EndYear	2024	2023	2023
TrapYears	6341	5354	4439
Nights	212,851	181,323	151,121
EmptyTraps	28,836 (13.5%)	14,846 (8.2%)	11,646 (7.7%)
MissedWeeks	15425 (6.8%)	11421 (5.9%)	8683 (5.4%)

3.1 Trap Locations

For full summer scheme (LH plot) and for those in the trend analysis dataset (RH plot)

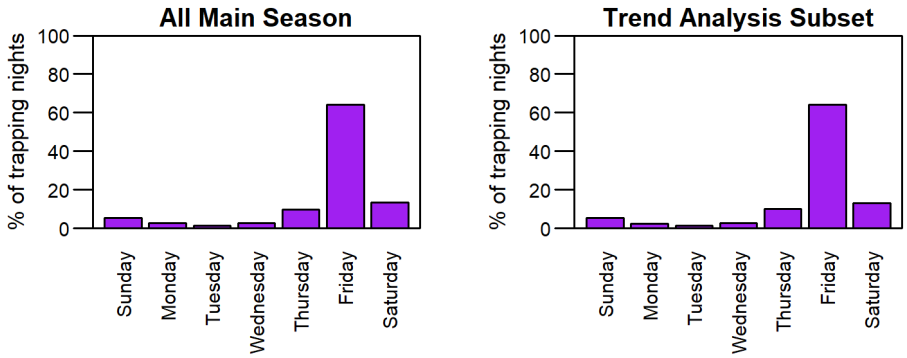
Figure 1: Trap locations



- Long-term recorders a bit more localised with few in S and W Ireland and a hole in Cumbria and few in N Scotland
- But not bad!

3.2 Recording Day

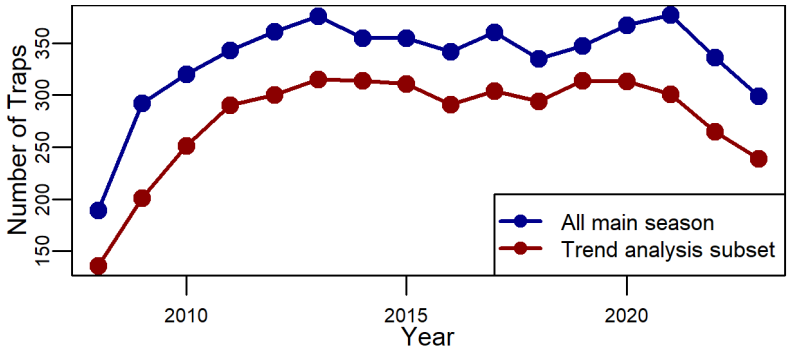
Figure 2: Day of week traps were run



- 64.2% recorded on Friday night (main scheme dataset)
- 64.1% recorded on Friday night (trend analysis dataset)
- (Note some recorders choose a different day than Friday - may be worth looking at deviations from some definition (mode?) of usual day for each trap)

3.3 Traps per year

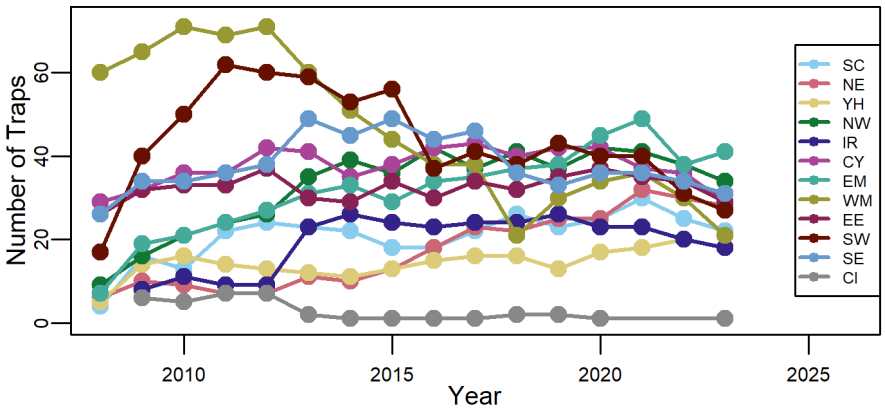
Figure 3: Number of traps recording each year



- Note numbers falling off in recent years!

And by region…

Figure 4: Number of traps recording each year (All main scheme)

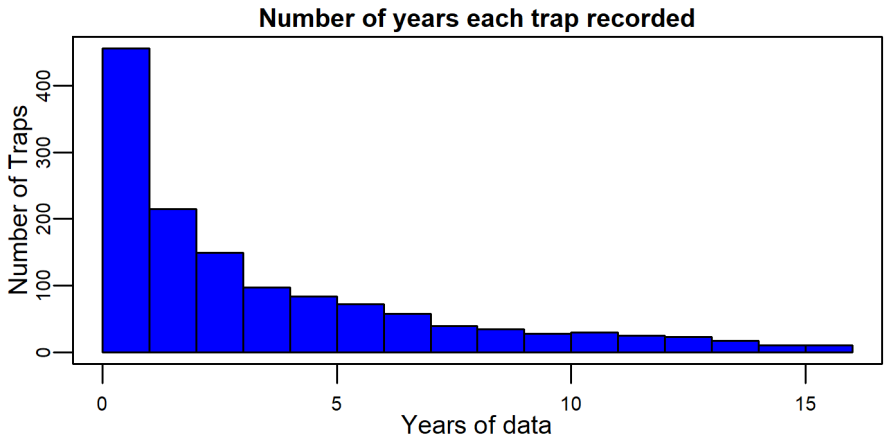


- Early years dominated by West Midlands

- CI actually has very few traps in any year and nothing before 2009
- Slow start in some regions in first couple of years - three only start in 2008
- 2009/10 before get good coverage

3.4 Length of recording

Figure 5: Number of years each trap recorded (main summer dataset)



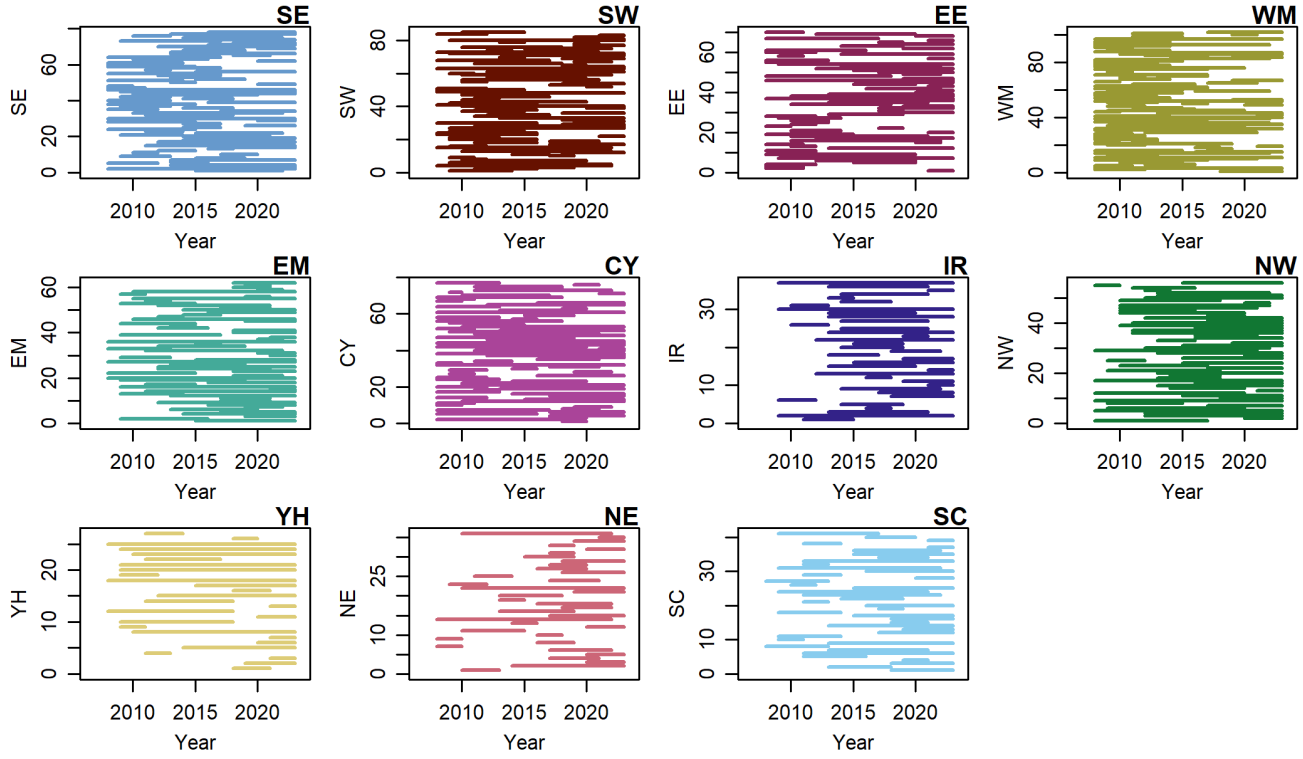
456 (34%) traps with only 1 years recording
891 Traps with 2 or more years
671 (50%) traps with only 1 or 2 years recording
676 Traps with 3 or more years
430 Traps with 5 or more years

- For analysis only those with 2 years or more recording provide any useful information
  - Currently using a minimum of 3 years for the trend analysis subset

3.5 Patterns of recording

Here I plot a line showing the time period each trap was active. Separate panels for each region. Plotted in roughly South to North order. Each horizontal line represents one trap with the line extending over the range of years that trap was in the scheme. This illustrates that there are very few traps with data for the whole period and that trends over time will have to be derived by piecing together many shorter-term trends. It is also apparent that some regions were rather later getting started so care will be needed in interpreting trends as the regional coverage varies over time.

Figure 6: Duration of recording for each trap by region (trend dataset)



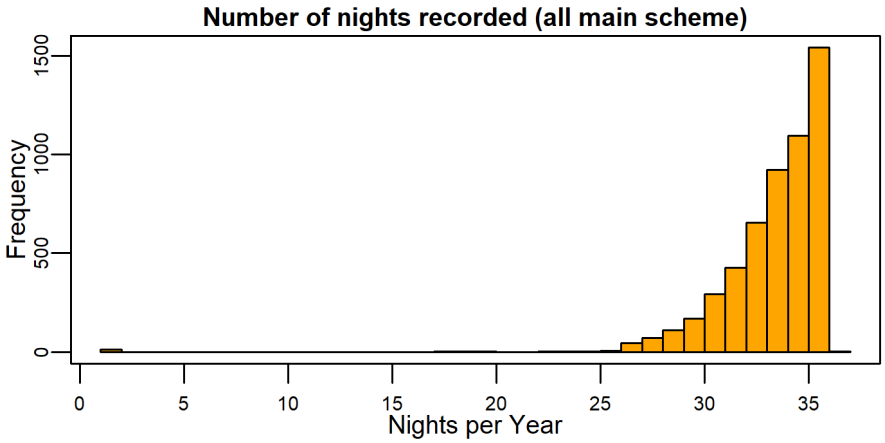
- CI really doesn't have the coverage for meaningful analysis - dropped.
- Ireland a bit slow starting and nothing for 2008.
- NE a bit patchy.

Once we start to look regionally, the coverage is not great and different time periods have different contributions from different regions. This will need to be considered in the analysis and interpretation.

3.6 Recording Nights

Plot of the number of nights recorded for each trap/year:

Figure 7: Number of nights recorded (Main season dataset)

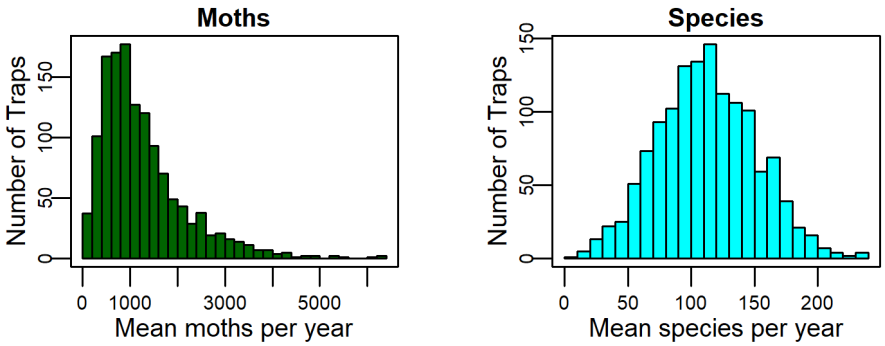


- Note: 12 Trap/Years with very few (<10) nights recording
- Note: 17 Trap/Years with few (<25) nights recording
- Need to exclude these! (currently gone for a minimum of 25 nights for inclusion in the trend dataset)
- But the vast majority have complete or near complete weeks.

3.7 Moth numbers per trap

Histograms showing the distributions of the mean number of moths per year and mean number of species accross traps. Main summer dataset, but excluding the few traps with <25 nights recording.

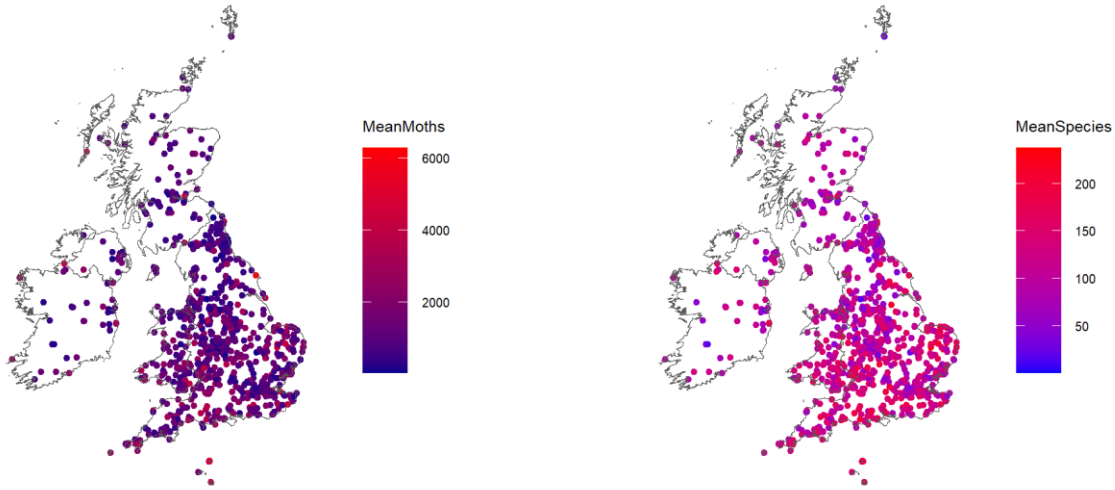
Figure 8: Mean number of moths and species per year for each trap



- Range 34 to 6284 moths per year.
- Range 1 to 237 species per year.
- Note we have selected traps with reasonably complete coverage (>25w) - low numbers are real!
- A vast range. Need to be careful that any analysis isn't dominated by the few very productive gardens.

Plotting the number of moths and species (average per year) by location:

Figure 9: Moth Numbers



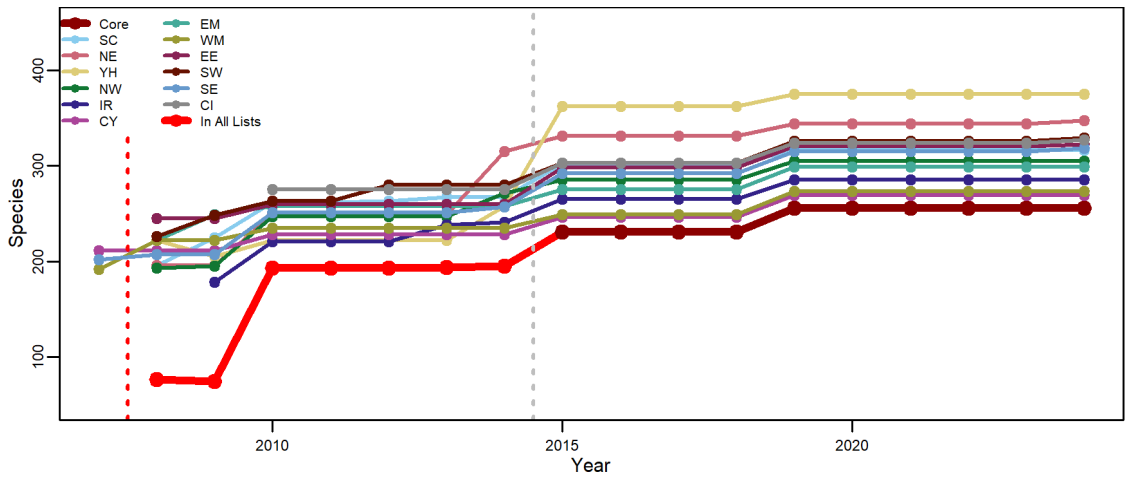
- Mainly an urban/rural split?

4 Recording forms

Excel file SpeciesLists.xlsx documents the species reciorded in each region and each year.

Figure 10 below shows the number of species in each regional form along with the "core" set defined in 2015 (dark red) and a set of species that are on the list for all regions in that year (light red):

Figure 10: Numbers of species on recording forms



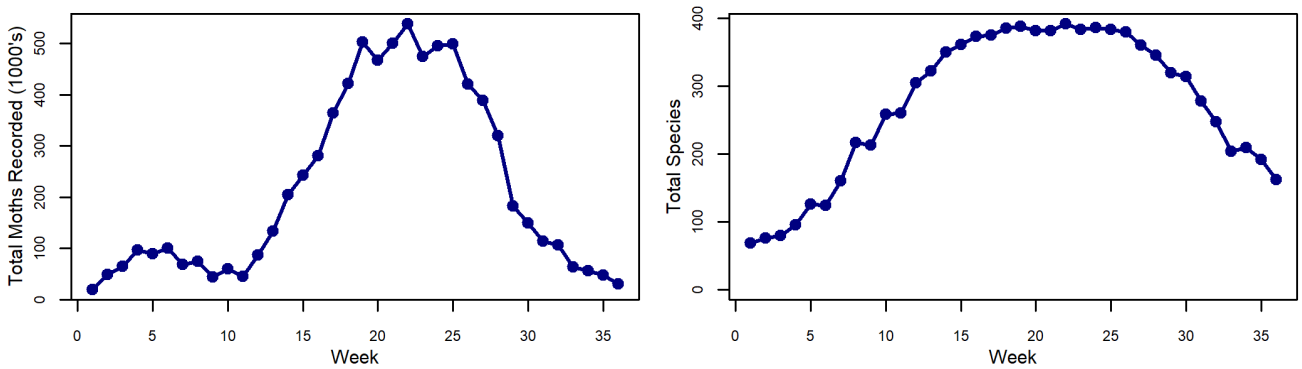
The number of species recorded has increased dramatically over the years.

- There was a big increase in the number on all lists in 2010. We do however note that most moths on the 2010 common list were recorded in nearly all regions in 2008-9
- A review of the lists in 2015 increased the number recorded in all regions slightly, and introduced a core list which was used in all regions. From 2015 onwards recorders could opt to use the core list or the appropriate regional list.

5 Moths Numbers and Species by week

Total numbers of moths and species recorded in each GMS week in the full summer GMS dataset:

Figure 11: Total numbers of Moths and species recorded by GMS week

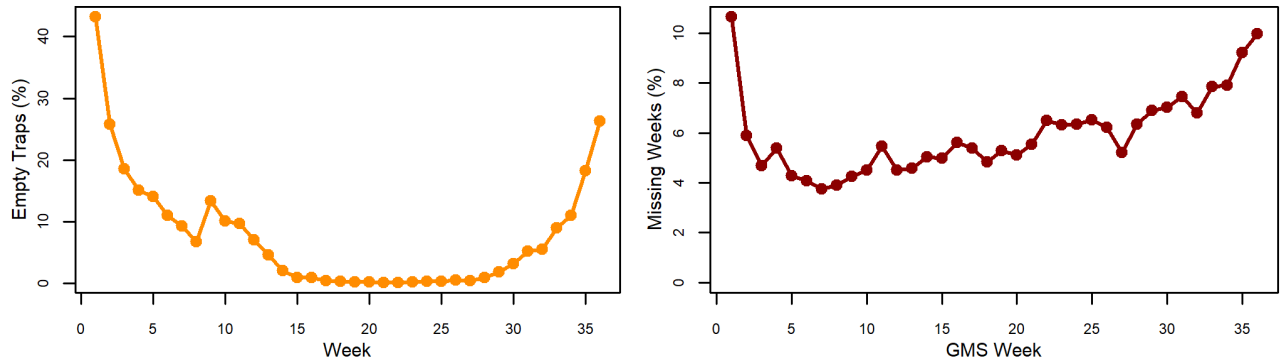


- Note the May gap!

5.1 Empty Traps and missing weeks

(Based on all main season data)

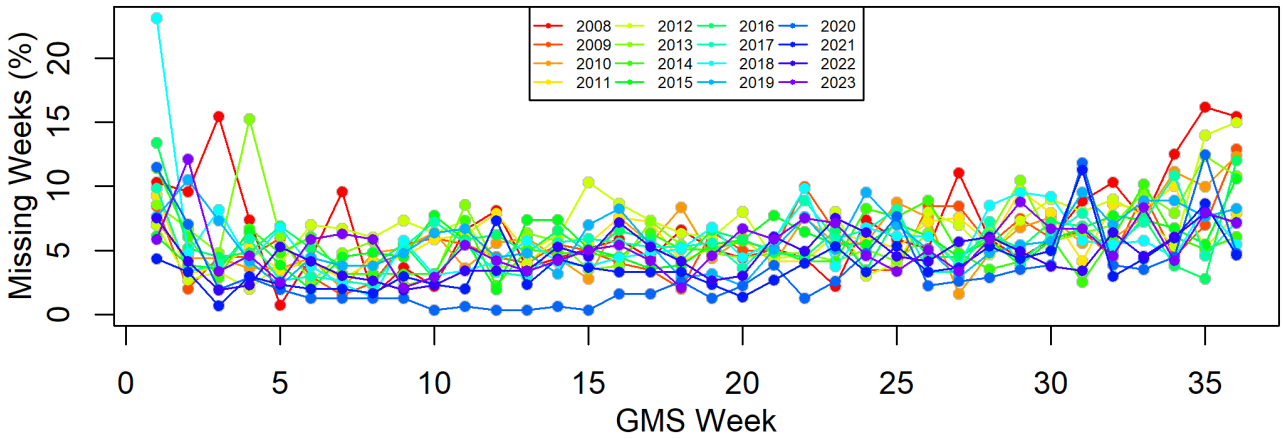
Figure 12: Empty traps and Missing (traps not run) weeks



- Empty traps follows the moth numbers as would expect - very high proportion in first week!
- Most missing weeks at ends - traps less likely to be run at ends of season

Check missing weeks in individual years...

Figure 13: Missing (traps not run) weeks for each year



- 2018 has a lot missing the first week - no obvious reason why!

6 Numbers of each species

Plots of the number of moths of each species recorded in the trend analysis subset.

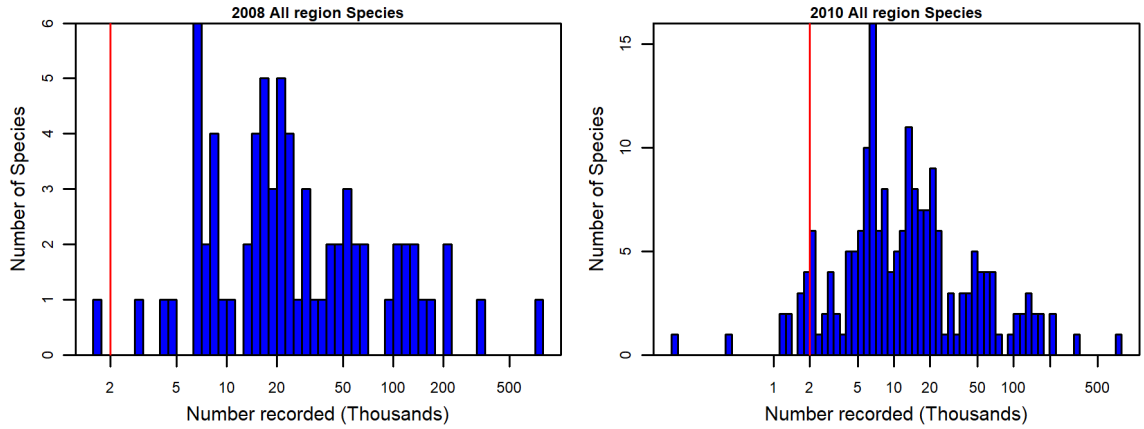
An optimistic sample size calculation would suggest we need a minimum of 2000 or so moths to detect trends. In practice we will need more given the size and complexity of the models.

A first set of species we would have enough data for trend analyses would be those recorded in all regions in 2008 and 2009 (these all remained till 2024, but a few on all 2008 list were not on the Ireland list added in 2009).

The lists increased substantially in 2010 so we have a rather larger set we can look at from 2010-23.

Here are plots of the total number of moths in recorded for both the 2008 and 2010 sets of species (trend dataset), with a line showing the arbitrary 2000 moths as an indicator of the kind of numbers we might need to show trends.

Figure 14: 2008 all region list - Numbers of each species



- 71 species in 2008 list
- 193 species in 2010 list
- 122 additional species
- 13 rarely recorded (<2000) in 2010 set (just 1 in 2008 set)
- So from 2010 we have about twice as many moths we can potentially look at trends in.

7 Some crude time trends

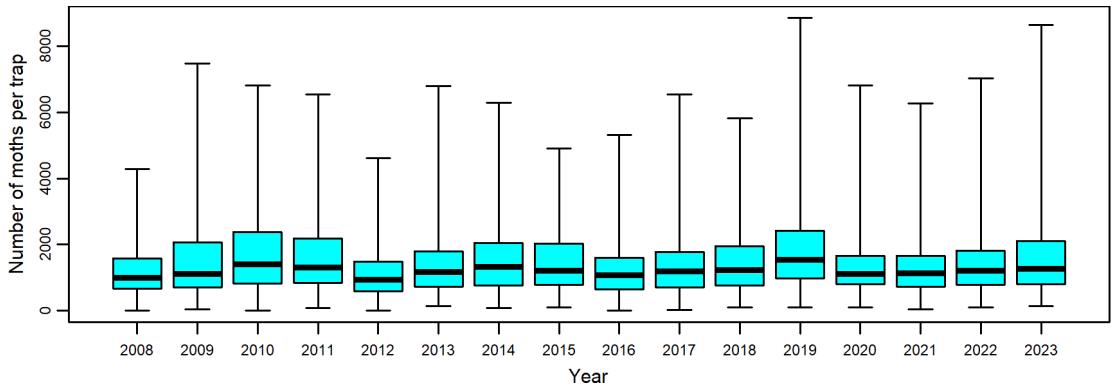
As a sanity check we plot a few things over time. These plots are simply a visualisation of what is there, with no adjustments, and cannot be used to infer any real trends!

- Based on main season dataset

7.1 Numbers of moths per trap per year

Boxplots showing the median, inter-quartile range and full range of the mean number of moths seen by each recorder in each year:

Figure 15: Number of empty traps by year

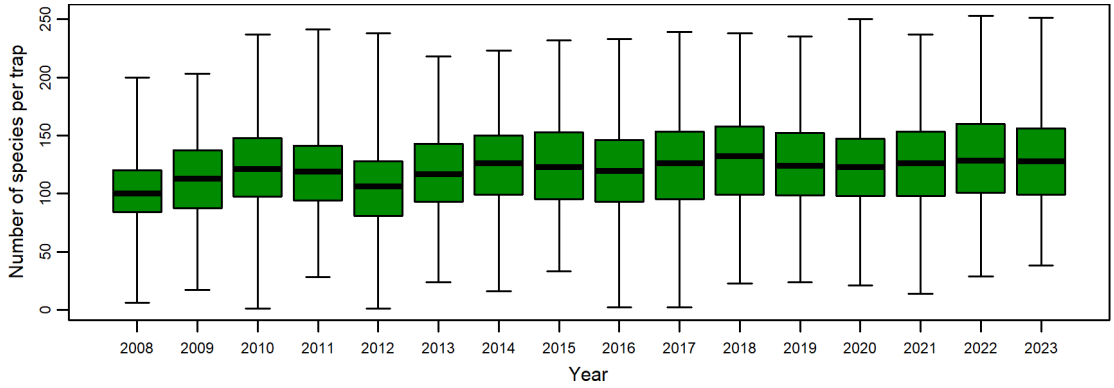


- Rather variable - tempting to try and interpret but we should avoid trying to as recorders vary from year to year!
- low numbers a bit artificial as dataset includes some recorders with little participation

7.2 Numbers of species per trap per year

Similarly the mean number of species per year:

Figure 16: Number of species by year

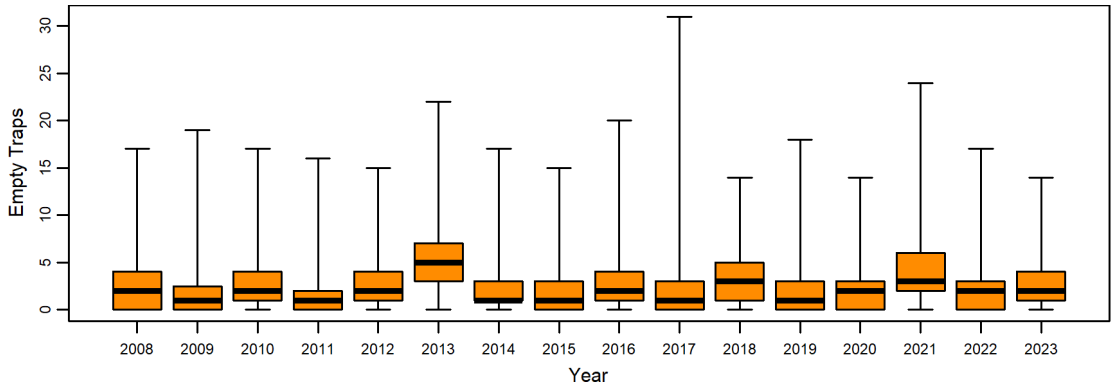


- increase in early years due to increase in size of recording form

7.3 Numbers of Empty traps per year

Boxplots showing the median, inter-quartile range and full range of the mean number of empty traps seen by each recorder in each year:

Figure 17: Number of Empty Traps by year

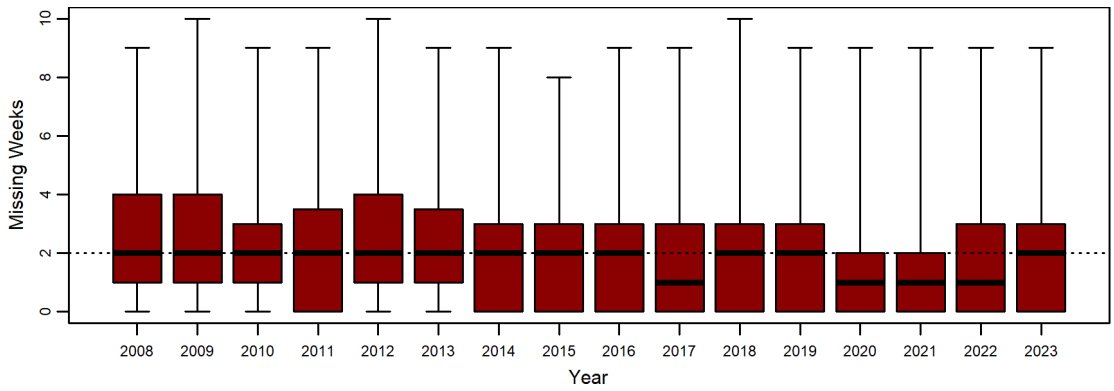


- 2013 was a bad year!
- there are odd traps with very few moths recorded and lots of empty traps

7.4 Missed nights

Boxplots showing the median, inter-quartile range and full range of the number of weeks missed each recorder in each year:

Figure 18: Number of missed nights by year



- Pretty consistent with 2 nights missed per trap in each season - a very few with more