Developing an interactive platform for insect biodiversity meta-analyses using R Shiny

Grace Skinner

August 2022

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science at Imperial College London. Submitted for the MSc in Computational Methods in Ecology and Evolution.

Contents

Declaration

* Was the data provided to you or did you collect or assemble it?
* Were you responsible for data processing or cleaning, if required?
* Were any mathematical models developed by you or by your supervisor?
* What role, if any, did your supervisor play in developing the analyses presented?

Data collected, meta-meta-analysis process and initial models by Christina Raw

Abstract

Introduction

Insects constitute the world’s most species-rich taxonomic group (Homburg *et al.*, 2019; Cardoso *et al.*, 2020), representing extensive unique and overlapping ecological functions. Pollination by Hymenoptera, Lepidoptera, and Diptera strongly influences crop yields and profits (Habel *et al.*, 2019) with approximately three-quarters of crop species dependent on pollination to some extent (Klein *et al.*, 2007). Insects contribute to biological pest control and organic matter recycling, and constitute a major link in the food chain between primary producers and consumers (Sánchez-Bayo and Wyckhuys, 2019). The loss of these services threatens global ecosystem health and food security (Potts *et al.*, 2016), particularly apparent in the global south where reliance on insect pollinated crops is high (Dicks *et al.*, 2021).

Despite their invaluable contributions, insects are underrepresented in long-term biodiversity studies compared to vertebrates (Outhwaite *et al.*, 2020; Wagner *et al.*, 2021b). A review by Titley *et al.* (2017) found an even split between biodiversity papers reporting on vertebrates and invertebrates, a substantial difference to the proportions of species in existence where 95% are invertebrates. Possible reasons for this include the viewing of insects as pests and disease vectors (Lawton *et al.*, 1998; Miličić *et al.*, 2021), difficulty in identifying individuals to species level, and issues monitoring populations given their inconspicuous nature and high annual variation (Fox *et al.*, 2019). The imbalance has improved recently, with a ten-fold increase in the number of insect decline papers between 2000 and 2010 (Eggleton, 2020), though the disparity between the one million insects described and the 5.5 million thought to exist (Stork, 2018) means studies continue to suffer from a lack of knowledge. Therefore, while we may be experiencing the sixth mass extinction in evolutionary history (Dirzo *et al.*, 2014), it proves difficult to assess the extent to which insects are affected.

Although more long-term data is needed, current consensus is that insect decline is sufficient to warrant further study and action (Montgomery *et al.*, 2020). Sánchez-Bayo and Wyckhuys (2019)'s speculations that 40% of insect species could go extinct in the next few decades have been prominent in bringing the issue to the forefront, though fundamental issues—particularly their biased literature search strategy, vote-counting methodology, and unjustified global extrapolation of findings—have been discussed (Simmons *et al.*, 2019; Saunders *et al.*, 2020). A similarly alarming seasonal decline of 76% of total flying insect biomass was observed in Germany (Hallmann *et al.*, 2017). Nevertheless, some insect populations are stable or increasing (Boyes *et al.*, 2019; Wagner *et al.*, 2021b) such as the 5.5% increase in UK terrestrial insect occupancy (1970-2015) reported by Outhwaite *et al.* (2020). Crossley *et al.* (2020) also found no overall trend in US insect abundance or diversity, though this study has been criticised for using datasets with inconsistent sampling methodologies (Welti *et al.*, 2021).

A variety of reasons contribute to the contrasting results, one being geographical variation on multiple scales. Trends in carabids markedly differed across the UK from 50% declines in northern moorlands to 50% increases in southern downlands (Brooks *et al.*, 2012). On a global scale, most research to date has been restricted to human-dominated landscapes in Europe and North America. Recognising this, researchers have begun searching for biodiversity change elsewhere including the tropics (Lister and Garcia, 2018; Wagner *et al.*, 2021b) and the Arctic (Loboda *et al.*, 2018; Gillespie *et al.*, 2020).

Conflicting results may also be due to temporal or taxonomic variation. Ollerton *et al.* (2014) calculated that the highest extinction rates for British bee and flower-visiting wasps occurred in the 1920s to 1950s, likely coinciding with intensification of agricultural practice. Among taxa, Biesmeijer *et al.* (2006) found that bee richness decreased in 52% of British cells, but no significant changes for hoverflies. Additionally, the choice of biodiversity metric may impact conclusions. It is possible to observe decreases in species richness but not biomass, potentially explained by smaller species showing stronger declines than larger ones (Homburg *et al.*, 2019). Likewise, using only biomass or abundance, as done by Van Klink *et al.* (2020), can result in researchers overlooking species richness changes or the replacement of sensitive species with tolerant ones (Jähnig *et al.*, 2021). Overall, this variation indicates that caution must be taken when generalising results especially as temporal, taxonomic, or geographical distance increases.

It is an additional challenge to understand the drivers of insect biodiversity change. The threats are primarily human-caused (Eggleton, 2020) and include land-use change, climate change, habitat loss and fragmentation, pollution, and invasive species (Cardoso *et al.*, 2020). Of these, land-use—particularly agricultural expansion and intensification—has been widely discussed (Newbold *et al.*, 2014; Newbold *et al.*, 2016; Newbold *et al.*, 2018; Seibold *et al.*, 2019; Gillespie *et al.*, 2022). Natural habitats in the vicinity of agricultural land may also experience inflated rates of insect decline due to reduced dispersal ability in a fragmented habitat or increased pesticide exposure (Seibold et al., 2019). Moreover, effects of land-use vary geographically and taxonomically. Millard *et al.* (2021) found pollinator species richness in non-tropical areas to be significantly higher in minimal-intensity cropland than primary vegetation, in contrast to the decrease observed in tropical regions. Among insect orders, Engelhardt *et al.* (2022) observed decreases in butterfly, but not grasshopper nor dragonfly habitat specialists. This is potentially because butterflies possess a higher quantity of specialised taxa, the existence of which are associated with high quality habitats (Poniatowski *et al.*, 2018).

To untangle the trends and drivers, we must synthesise information across the literature to provide scientists, media, and policy-makers with the best available evidence obtained through the aggregation of multiple studies. A necessary step in achieving this is collecting studies, as done by EntoGEM (Haddaway *et al.*, 2020; Grames *et al.*, 2022). This mapping of relevant literature allows easy identification of the distribution of evidence and access to information on study methodology.

Once sufficient data exists to answer a proposed hypothesis, there are several approaches available for drawing conclusions across studies. Synthetic analyses, in which models are built based on collated primary data, are common in papers utilising PREDICTS (Hudson *et al.*, 2017), a database designed to explore how biodiversity responds to land-use (Newbold *et al.*, 2014; Gray *et al.*, 2016). Alternatively, meta-analyses can be used to quantitatively summarise results across multiple studies in a replicable process, answering pre-defined questions (Arnqvist and Wooster, 1995; Gurevitch *et al.*, 2018). Effect sizes are calculated from primary studies, weighted according to study size, and then fed into a statistical model to determine an overall effect size and an associated confidence measure.

Building upon drawing conclusions across studies, the natural advancement is ensuring reviews do not remain static. This is the concept of a living review (Elliott *et al.*, 2017), in which results are updated as new evidence becomes available, allowing decisions to be based upon the current body of evidence. This is done by the Metadataset website and its dynamic meta-analysis R Shiny app, which enables users to browse a living database for relevant data—currently in the fields of invasive species and cover crops—and then perform a meta-analysis on that subset (Shackelford *et al.*, 2021).

As increasing numbers of meta-analyses are completed, we can perform meta-meta-analyses, enabling further increases in the accuracy of estimations (through larger sample size) and increased chance of detecting variables that significantly influence effect size. As I define it here, meta-meta-analyses is the process of analysing multiple meta-analytic studies together by combining effect sizes for each of these.

Here, I present a new interactive platform for running meta-meta-analytic models of insect biodiversity change, built using R Shiny. The platform is designed around insect biodiversity meta-analyses, though repurposing it for other means is highly feasible. The app reads in data from a living review of meta-analyses stored within Google Sheets. Using this data, the user can run a custom model based on their hypotheses of insect biodiversity change. Estimations of change are specific to the threats and biodiversity metrics—and in future other variables such as location and taxonomic group—the user has chosen to investigate. As new meta-analyses are conducted, users can upload these results, with the figures then reactively updating to these changes.

The creation of the insect biodiversity meta-analytic platform is motivated by the need to make best use of existing and future data, helping to improve our understanding of the challenging field of insect biodiversity change. To achieve this, I aimed to build a platform that is interactive, reactive, transparent, and efficient. The purpose aligns well to, and is planned to contribute towards, the GLobal Insect Threat-Response Synthesis (GLiTRS) project, which aims to synthesise many lines of evidence of insect biodiversity trends. In the short-term, the GLITRS team is the target audience for the app though in the longer term, my hope is that this platform can assist insect biodiversity practitioners in conveying key messages to decision makers, as well as the public on how insect declines can be mitigated.

Methods

The first section of the methods introduces each of the qualities that I aimed the platform to possess, and lays out how I designed and built the platform around these characteristics. The second describes the meta-analytic models of insect biodiversity change utilised in the development of the platform.

*Important characteristics for developing a insect biodiversity meta-analytic platform*

The four qualities I wanted to build into the platform were interactivity, reactivity, transparency, and efficiency. Any platform developed to allow users to investigate and analyse data must be interactive. By allowing the user to interact with the data, models, and their corresponding outputs, the tool has increased usefulness compared to an individual primary paper which reports the specific analyses run by the researchers.

The R Shiny package (Chang *et al.*, 2017) is ideal for building interactivity into the platform due to its purpose of enabling the development of interactive web-based applications in R. Shiny was chosen to take advantage of the analytical and graphical capabilities of R familiar amongst conservation biologists. It also removed the need to use HTML, CSS, or JavaScript often used for developing websites, though these can be used in Shiny apps for customisability of the user interface. The apps link well to, and work well with, databases as well as being easy to share with others—the user simply needs the web address of the deployed Shiny app. I chose Shiny over the Python equivalent Dash (Hossain *et al.*, 2019) due to Shiny requiring significantly less code than Dash to produce an comparable output. I also did not require the greater computational efficiency provided by Python.

Interactivity is a core feature of the app: the main function of the platform is to allow users to run their own custom model based on their hypotheses, in this case to investigate the effect of agricultural systems on biodiversity. At present, the user can filter the data to be included in the model based on the biodiversity metric used by the studies with the model then running in real-time.

Additional interactivity stems from the ability of the user to contribute to the platform by providing data from their own meta-analysis. To do so, the user must input their name, which must only contain letters, spaces, and hyphens to prevent code injection where malicious executable input is entered and then mistakenly run by the app. The file to be uploaded must be a comma-separated values file to maintain compatibility with other data spreadsheets. It must contain data appropriate for insect biodiversity meta-analyses such as the log response ratio and agricultural systems. Finally, it must not be a duplicate of a pre-existing dataset. Upon passing these checks, the user is presented with a preview of their dataset and can then upload it to Google Sheets.

If a platform is to be interactive, then it must also be reactive to those decisions made by the user. Achieving this involves the platform being reactive to new data. The data is remotely stored within Google Sheets, from which it is read into the insect biodiversity change Shiny app using the googlesheets4 package (Bryan, 2021). Each time the app is launched, all existing data stored, including any recently added data, is utilised by the models run within it. Remote storage is necessary due to local storage being insufficient in allowing users access to data uploaded by previous users. If a user were to upload a file to the Shiny app, but not to remote storage, the file would only be only available locally. If a new user loads the app, they will likely be directed to a different server where this file does not exist. Google Sheets was chosen due to the long-lived access tokens available for authentication, compared to short-lived tokens which exist for Dropbox and its r2drop package (Ram and Yochum, 2020). Google Sheets also provides a free basic storage plan, compared to the limited free usage options offered by Amazon S3.

The application is also designed to be reactive to the choices made by the user. Upon running a model, the user is presented with their results conveying the effects of agricultural systems on biodiversity. The user can then alter the figure to plot percentage change, which can be more intuitively understood than the log response ratio, converted using the formula

( 1 )

In the case where the default model is run on all available data, there are additional options for user input including a choice of agricultural systems to plot, and whether to re-scale the x axis as different agricultural systems are selected. As the user selects and de-selects different options, the results are reactively filtered using dplyr (Wickham *et al.*, 2019) and plotted using ggplot2 (Wickham *et al.*, 2019).

I wanted to incorporate transparency into the platform both in terms of inputs and outputs. Transparency of inputs is the concept that the user is aware of the origins of the data. The app provides details of each study used including the number of agricultural systems studied and total number of data points. Additionally, the user is able to choose a paper for which they wish to see further breakdown of the agricultural systems studied. These features allow the user to gauge the size of any analyses they may conduct, and can highlight any studies that may have a potentially influential effect. Utilising the maps package (Becker *et al.*, 2021), the user is also presented with a world map layered with data points for the set of studies for which we have geographic coordinates to indicate the geographical representativeness of the data.

The outputs of the model are also designed to be transparent. A table displays coefficients extracted from the model and the frequency of instances for each agricultural system that the model is based on. Additionally, the user can choose to display descriptions of the agricultural systems and download the R model summary, or table of coefficients. If the model does not successfully run due to insufficient data, the user is prompted to select additional or alternative biodiversity metrics and re-run the model.

Finally, for the platform to be useful, it needs to be efficient. This means models need to be run, and files uploaded within a reasonable time.

*Assessing the characteristics of the insect biodiversity meta-analytic platform*

The reactivity of the platform is tested by assessing the extent to which results differ between models. I compare the output of the default model (run on all data available) and a custom model where only data collected using diversity as a biodiversity metric is included. The reasoning behind this is due to the impact of biodiversity metric choice on conclusions (Hillebrand *et al.*, 2018). For example, a decrease in biomass does not necessarily equate to a decrease in species richness (Jähnig *et al.*, 2021).

To test the efficiency, I ran a sequence of 10 models with increasing numbers of rows of data (between 30 and 470) using the app and recorded the time taken to produce the outputs using the proc.time() function (R Core Team, 2013). Each model was run 10 times and an average time calculated. The same approach was taken to test uploading data using a set of files of increasing file size (from 59 KB to 563 KB). Times to upload data to the Shiny app, and to Google Sheets were recorded.

*Leveraging meta-analytic models of insect biodiversity change*

The app is prototyped using data collected for the purpose of conducting a meta-meta-analysis on the effect of agricultural systems on biodiversity, thus is composed of effects from separate meta-analyses. The data is useful in achieving the aim of assembling a platform for insect biodiversity meta-analyses, though in itself is not the main focus of the project. Not every meta-analytic study included was exclusively focused on insects, thus also represents plants, fungi, and bacteria. At this stage in development, I do not view this as an issue because once sufficient insect data is available, the non-insect data points can be filtered out.

The user can either run a default or a custom model. The default model uses all available data whereas a custom model uses data filtered based on user choice. The models are robust linear mixed-effects models, fitted using the rlmer function from the robustlmm package (Koller, 2016). This model was chosen due to its ability to handle variation and outliers in data, thus enabling superior fitting compared with traditional linear mixed-effects models fitted with the lme4 package (Bates *et al.*, 2015). Additionally, robust models are employed over metafor models (Viechtbauer, 2010) (routinely utilised in meta-analysis) as the meta-analytic studies in the data underpinning the app insufficiently reported statistics such as variance necessary for metafor.

The model compares biodiversity of agricultural systems to the reference conventional agricultural system in terms of the log response ratio. The model is fitted using the formula

( 2 )

where agricultural system is a fixed effect and paper identification and control agricultural system are random effects, accounting for the non-independence of the data. Control agricultural system is included as a random effect due to the log response ratio in part depending on the baseline agricultural system the treatment agricultural system is compared to.

The log response ratio is a popular effect size in ecology due to its ability to quantify proportionate change between treatments (Hedges *et al.*, 1999) and it is more robust to non-independence than the standardised mean difference (Noble *et al.*, 2017). Non-independence results in the standard deviation of the data being lower than that of a dataset with independent data points. As the standardised mean difference, but not the log response ratio, is calculated using standard deviation it is therefore more affected by non-independence. The logarithm of the response ratio is taken to ensure symmetry between positive and negative change (Hedges *et al.*, 1999). For example, if one study found insect biodiversity doubled in response to switching to an organic agricultural system (two times the biodiversity), but another found it halved (0.5 times the biodiversity), in a meta-analysis these combine to give a conclusion of no effect, which succeeds if the natural log is taken. Without doing so, the average calculates to 1.25.

The Shiny app presented here was built using the shiny R package version 1.7.1 (Chang et al., 2017) within R version 4.2.0 (R Core Team, 2013), which was also used to test the efficiency of the insect biodiversity meta-analytic platform. The app is deployed to be hosted on the shinyapps.io service via the rsconnect package (Atkins *et al.*, 2022). The shinyapps.io service is free and it is fast and simple to deploy an application, making it ideal for developing a app such as this. The app is available at https://r26dnk-grace-skinner.shinyapps.io/meta\_meta\_analysis/, and the code is available at https://github.com/gls21/CMEE\_NHM\_Project/tree/master/christina\_mma/meta\_meta\_analysis.

Results

*Typical use of the insect biodiversity meta-analytic platform*

The process of a typical use of the insect biodiversity meta-analytic platform is summarised in figure 1. Upon opening the app, the user is presented with an ‘introduction’ tab which provides information on the data sources used within the app, such as details on the number of agricultural systems studied and the total number of data points supplied by each paper (figure 2). The user can select a paper for which they wish to see a breakdown of the agricultural systems studied, and view a world map to gauge the geographical representativeness of the data. Currently, 27 out of 334 data points have geographic coordinates available enabling them to be plotted.

**Figure 1.** Flow diagram summarising a typical use of the insect biodiversity meta-analytic platform to run meta-meta-analytic models and upload a new data set.

Table

Description automatically generated

**Figure 2.** Screenshot of a table included in the ‘introduction’ tab of the insect biodiversity meta-analytic platform which gives information on the number of agricultural systems studied, and the total number of data points provided by each study used within the app.

To investigate how different agricultural systems impact biodiversity, the user can go to the ‘Agricultural systems models’ tab. The user can click ‘Run default model’ to run a model on all available data and view the results in graphical and tabular form. They can also choose to plot specific agricultural systems and plot percentage change instead of the log response ratio with the figures and tables updating in response to these selections. Finally, the user can view definitions of the agricultural systems and download the R model summary and table of results if they wish to locally store them. After studying the default model output, the user may decide to run their own custom model based on their specific hypothesis. The app currently allows the user to filter data based on the biodiversity metric it was collected with. The user can make their selections and then run a model using their selected data.

Following exploration with the platform, the user may decide to conduct their own meta-analysis. Once completed, they can upload their data to the platform so future models can be run including this data. In the ‘Upload data’ tab (figure 3), the user can upload a file using the ‘Browse’ button , and view a preview upon completion of checks. They can then click the ‘Upload to database’ button to upload the data to the remote storage database. If the user wishes to run models including their data, they can refresh the app and re-run the models.

Graphical user interface, text, application

Description automatically generated

**Figure 3.** Screenshot of part of the ‘Upload data’ tab in the insect biodiversity meta-analytic platform which shows the process of uploading a new meta-analytic data set. The user enters their name, uploads the data to the Shiny app and views a preview, and then uploads the data set to the remote storage database and receives a success message.

*Assessing the characteristics of the insect biodiversity meta-analytic platform*

Firstly, I tested the reactivity of the platform by comparing the results of a default model (run on all available data) and a custom model which was run only on data collected using diversity as a biodiversity metric (figure 4). The sustainable agricultural system has the largest positive effect on biodiversity compared to the conventional agricultural system in both the default (log response ratio = 0.37, t = 6.48, data points = 10) and custom model (log response ratio = 0.41, t = 8.33, data points = 8).

The transgenic agricultural system is not significantly different to the conventional agricultural system in the default model results (log response ratio = -0.21, t = -1.05, data points = 106). However, it can be concluded to have significantly less biodiversity than the conventional agricultural system in the custom model results (log response ratio = -0.23, t = -2.41, data points = 38).

Chart

Description automatically generated

(a)

Chart, box and whisker chart

Description automatically generated

(b)

**Figure 4.**  Screenshots taken from the agricultural systems models tab in the insect biodiversity meta-analytic platform displaying the outputs of (a) the default model, which is run on all available data, and (b) a custom model, which is run only on data collected using diversity as the biodiversity metric.

I tested the efficiency of the app in terms of the time taken to run models, and the time taken to upload data to the platform. The following results are generated with an internet upload speed of 94 Mb. In terms of running models, I ran a sequence of models on data with 30, 110, 302, 412, and 470 rows. In general, the larger the data set, the longer it took to fit the model (figure 5), with the time ranging from 2.4 (standard error = 0.05) to 17.2 (standard error = 0.54) seconds on average for 30 and 470 rows of data, respectively. The time taken to run the model on 412 rows of data was shorter (12.8 seconds [standard error = 0.12]) than for 470 rows of data (17.2 seconds [standard error = 0.54]).

Chart, scatter chart

Description automatically generated

**Figure 5.** The time taken (seconds) to fit a robust linear mixed-effects model to the data generally increased as the number of rows of data is increased. The error bars represent standard error of the mean.

To test the efficiency of uploading data, I recorded the time taken to upload a sequence of files of sizes 59, 113, 172, 227, 285, 342, 399, 456, 509, and 563 KB to the Shiny app and to Google Sheets remote storage. There was no substantial change in the time taken to upload the files to the Shiny app as the file size increased, with the average time ranging from 3.5 (standard error = 0.08) to 4.0 (standard error = 0.15) seconds. Contrastingly, the time taken to upload the same files to Google Sheets remote storage increased with file size from 2.1 seconds (standard error = 0.03) for a file of size 59 KB to 9.8 seconds (standard error = 0.11) for a file of size 563 KB.

Chart, scatter chart

Description automatically generated

**Figure 6.** Time taken (seconds) to upload a meta-analytic data set to the Shiny app does not noticeably increase as the file size (KB) increases. Time taken (seconds) to upload the same data set to Google Sheets remote storage increases as file size (KB) increases. The error bars represent standard error of the mean.

Discussion

The Shiny app introduced here represents a platform on which we can build a living review of insect biodiversity meta-analytic data. Using this data, the platform enables interactive models to be run based on user hypotheses, thus improving our knowledge of insect biodiversity change drivers, which research to date has strived to achieve, but failed to fully understand. Although the app was designed to investigate insect biodiversity change, it has potential for applicability in other fields, especially where the data shows high heterogeneity. The design of the app lends itself to be easily adaptable to analyse another database as it is coded so that the outputs are dependent upon the inputs.

The insect biodiversity meta-analytic platform has a number of uses. Firstly, it can aid decision-making processes in alleviating insect biodiversity loss. Secondly, the platform can also be used to highlight areas where research is lacking, potentially indicated by a model not successfully running due to insufficient data. Finally, the platform may indicate interesting findings worth investigating further, perhaps in a new meta-analysis—the results of which could eventually be incorporated into the app.

As shown through the testing of the reactivity of the insect biodiversity meta-analytic platform, the variables chosen by the user influence the results, which demonstrates the requirement for a platform in which you can model and visualise the data interactively and reactively. The results change due to the geographic, taxonomic, and temporal variation of insect biodiversity, along with differences due to the biodiversity metric used. This makes it unlikely that literature will exist for the specific combination of variables the user is interested in. The insect biodiversity meta-analytic platform meets this need by enabling the user to run models specific to their hypotheses, and providing evidence on how insect biodiversity is influenced by particular circumstances. This is achieved by giving the user access to a growing collection of data and the tools needed to analyse it without the need to code in R.

Currently, the user can filter the data and run a custom model based on their choice of biodiversity metric. This is an important variable to give the user choice over due to literature showing that the biodiversity metric used can impact conclusions with observed declines in one metric not necessarily corresponding to equivalent declines in another (Hillebrand *et al.*, 2018). For example, Homburg *et al.* (2019) only found declines in species richness—but not biomass or abundance—of carabids implying the loss of smaller, rarer species led to a significant effect on species richness, but not biomass nor abundance. This could explain the difference in results between the default and custom models, where transgenic agricultural systems only have significantly less biodiversity in the model run on data collected with diversity as the biodiversity metric. Though it must be considered that this could be influenced by the reduced number of data points for transgenic agricultural systems in the custom (38) compared to the default model (106).

The insect biodiversity meta-analytic platform is efficient for running models and uploading data, a feature necessary for the platform to be useful. A model with 470 rows of data took on average 17.2 seconds to run, and a data set of size 563 KB took an approximate total time of 14 seconds to upload to the Shiny app and to Google Sheets remote storage, both reasonable for the user to be satisfied with the speed of completing these tasks.

*Comparison to existing tools*

Comparing the app presented here to those previously created, the biggest similarities occur between the insect biodiversity meta-analytic platform and a tool called dynamic meta-analysis, created alongside the Metadataset website (Shackelford *et al.*, 2021). After choosing an intervention in the dynamic meta-analysis Shiny app, the user can perform sub-group analyses in a similar manner to the Shiny app introduced here where custom results are based on a filtered subset of data. However, dynamic meta-analysis uses the rma.mv function from the metafor package (Viechtbauer, 2010), taking one or more effect sizes and corresponding variances calculated from each primary study as an input. Meanwhile, the insect biodiversity meta-analytic platform runs robust models from the robustlmm package (Koller, 2016) using multiple effect sizes collected from each primary study included in each meta-analysis study. The insect biodiversity meta-analytic platform is unable to use metafor due to the meta-analysis studies not reporting the necessary statistics (including variance) needed for metafor.

At this stage, the dynamic meta-analysis app (Shackelford *et al.*, 2021) provides the users with more filters and options on ways to analyse the data than the platform introduced here. For example, the dynamic meta-analysis app provides the option of meta-regression (also called mixed-effects models) where different subsets are analysed whilst accounting for effects of other variables, which is considered more powerful than sub-group analysis due to the model including all of the data and controlling for variation that may be driven independent of the fixed effects. For the insect biodiversity meta-analysis platform, additional variables such as sampling location could be included as covariates to test for differences in model coefficients.

The dynamic meta-analysis app remains a work in progress with the usage example discussed in Shackelford *et al.* (2021) not replicable using the app in its current state because the version described in the paper differs from the prototype app available for public use. On the other hand, anyone is able use the insect biodiversity meta-analytic platform by using the web address of the deployed app and following the process outlined in the results section. Additionally, the stand-out feature of the insect biodiversity meta-analytic platform is the ability of the user to upload their own data directly from within the app. This will allow the app to continue to grow independently through the addition of data by its users, rendering it increasingly useful in the future with minimal input from the initial platform developer. Comparatively, the dynamic meta-analysis app depends on the identification of new data sources and their addition to the database by the person or team responsible for its upkeep.

*Expanding the functionality*

The app is limited by the quantity and quality of available data, though the issue of quantity will reduce as more studies are added. Deficiencies in the quality—particularly the reporting of statistics such as variance necessary to complete meta-analyses with metafor (Viechtbauer, 2010)—have been highlighted (Hedges *et al.*, 1999; Stewart, 2010; Gurevitch *et al.*, 2018). However, the robust model approach taken by the Shiny app is less susceptible to this due to it not depending on the reporting of these statistics that are necessary for metafor.

The custom model that can be run by the user is currently based upon a choice of biodiversity metric. I plan to expand the range of variables a user can choose as more data is collected. At present, filtering by multiple variables would likely result in limited data quantity, reducing the chance of a model successfully running. Ultimately, including a choice of sampling location, taxonomic group studied, and date of sampling would be ideal for investigating geographic, taxonomic, and temporal variation in results. This proves especially difficult with the current dataset due to the majority of data points not reporting these variables, as well as inconsistent location descriptions; some studies report multiple continents, others individual countries. It also means the map included in the introductory tab undersells the geographic representativeness of the insect biodiversity meta-analytic platform.

I also intend to expand the insect biodiversity meta-analytic platform in terms of the different threats that can be investigated to increase the usefulness of the app. In addition to investigating the effect of agricultural systems on biodiversity, the user will have the choice to analyse the influence of other variables such as land-use and intensity, climate variables, or species traits. Of these, land-use has been commonly linked with insect biodiversity change (Newbold *et al.*, 2016; Seibold *et al.*, 2019; Gillespie *et al.*, 2022). Meanwhile, there is less convincing evidence for the effect of climate change. Conclusions drawn by Lister and Garcia (2018) that biomass declines in a Puerto Rican rainforest were due to climate change have been challenged by (Willig *et al.*, 2019) because Lister and Garcia (2018) failed to account for drought or hurricanes in addition to temperature related aspects of climate change, neither did they adjust abundance data according to sampling effort. For species traits, declines are more frequently reported in species which are rare (Powney *et al.*, 2019; Outhwaite *et al.*, 2020), habitat or dietary specialists (Biesmeijer *et al.*, 2006; Boyes *et al.*, 2019; Wagner *et al.*, 2021a), small (Homburg *et al.*, 2019), or poor dispersers (Cardoso *et al.*, 2020). Therefore, including these variables in the insect biodiversity meta-analytic platform will provide an approach to untangling the effects of these using a living body of evidence.

In future, the insect biodiversity meta-analytic platform could provide multiple approaches to analysing data. Firstly, it can continue to analyse the effect sizes from each primary study within each meta-analysis using robust models. Secondly, it could be an option to combine the overall effect sizes calculated for each meta-analysis (one effect size per meta-analysis) using metafor. This would be dependent on sufficient reporting of variance statistics alongside effect sizes, and it would only be feasible to combine and analyse meta-analyses that investigated similar interventions.

*Scaling up*

The insect biodiversity meta-analytic platform relies upon future meta-analyses following a standardised protocol and data entry sheet. With these in place, the data collected from different studies becomes more comparable and requires less data re-structuring before incorporation into the app. A protocol for the use of the insect biodiversity meta-analytic platform will also need to be assembled to avoid the issues of cherry picking, where users run multiple models but only report findings which support their hypotheses and ignore those that do not.

Methods will need to be implemented to allow the app to sustain efficiency as increasing numbers of users run models simultaneously. To achieve this, I will need to deploy the insect biodiversity meta-analytic app using a service such as ShinyProxy instead of shinyapps.io. shinyapps.io has limitations on the number of usage hours, and multiple users often share the same R process. This means users have to wait for other user’s models to run before their own is run, lengthening the time taken to run their model, and ultimately limiting the number of users that can reasonably use the app simultaneously. ShinyProxy overcomes these issues by providing a separate session for each user, thus enabling multiple users at once without a reduction in modelling speed. ShinyProxy is open-source, though it requires more effort to deploy a Shiny app using this service than shinyapps.io, and additionally require some knowledge of docker containers.

Furthermore, the insect biodiversity meta-analytic platform also needs to remain efficient as more data is uploaded. Part of this will require a move away from using Google Sheets as the database—more accurately defined as a spreadsheet program—and instead use a relational database, which has greater capacity, better security, and faster speed especially when multiple users are accessing the database. Larger data sets also increase the time taken to run robust models. Nevertheless, it becomes increasingly less necessary to run a robust model as the size of the data set increases due to variation and outliers becoming less influential. Therefore, the user would benefit from sustained efficiency if the platform was coded to run a standard linear mixed-effects model when the amount of data meets a threshold that confidently exceeds the requirement for running a robust model.

The insect biodiversity meta-analytic platform introduced here is a tool allowing practitioners the opportunity to take full advantage of the data we currently have, and will have in the future as further meta-analyses are conducted. With far-reaching applicability to other fields, the platform aids scientists in disentangling complex patterns and drivers of trends by enabling custom models to be run on an living review database. The interactive, reactive, transparent, and efficient platform provides a simple, effective, and convincing approach to deliver evidence to decision makers and the public on how to curtail insect biodiversity loss.

Data and Code Availability

Acknowledgements

NHM, Andy, Joe, Christina, rest of team at NHM.

References

Arnqvist, G. and Wooster, D. (1995) Meta-analysis: synthesizing research findings in ecology and evolution. *Trends in Ecology & Evolution.* 10 (6), 236-240.

Atkins, A., McPherson, J. and Allaire, J. (2022) *rsconnect: Deployment Interface for R Markdown Documents and Shiny Applications* (Version R  package version 0.8.26 ) <https://CRAN.R-project.org/package=rsconnect>.

Bates, D., Maechler, M., Bolker, B. and Walker, S. (2015) *Fitting Linear Mixed-Effects Models Using lme4* Journal of Statistical Software.

Becker, R., Wilks, A., Brownrigg, R., Minka, T. and Deckmyn, A. (2021) *\_maps: Draw Geographical Maps\_* (Version R package version 3.4.0) <https://CRAN.R-project.org/package=maps>.

Biesmeijer, J. C., Roberts, S. P., Reemer, M., Ohlemuller, R., Edwards, M., Peeters, T., Schaffers, A., Potts, S. G., Kleukers, R. and Thomas, C. (2006) Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. *Science.* 313 (5785), 351-354.

Boyes, D. H., Fox, R., Shortall, C. R. and Whittaker, R. J. (2019) Bucking the trend: the diversity of Anthropocene ‘winners’ among British moths. *Frontiers of Biogeography.* 11 (3).

Brooks, D. R., Bater, J. E., Clark, S. J., Monteith, D. T., Andrews, C., Corbett, S. J., Beaumont, D. A. and Chapman, J. W. (2012) Large carabid beetle declines in a United Kingdom monitoring network increases evidence for a widespread loss in insect biodiversity. *Journal of Applied Ecology.* 49 (5), 1009-1019.

Bryan, J. (2021) *googlesheets4: Access Google Sheets using the Sheets API V4* (Version R package version 1.0.0) <https://CRAN.R-project.org/package=googlesheets4>.

Cardoso, P., Barton, P. S., Birkhofer, K., Chichorro, F., Deacon, C., Fartmann, T., Fukushima, C. S., Gaigher, R., Habel, J. C. and Hallmann, C. A. (2020) Scientists' warning to humanity on insect extinctions. *Biological conservation.* 242, 108426.

Chang, W., Cheng, J., Allaire, J., Xie, Y. and McPherson, J. (2017) Shiny: web application framework for R. *R package version.* 1 (5), 2017.

Crossley, M. S., Meier, A. R., Baldwin, E. M., Berry, L. L., Crenshaw, L. C., Hartman, G. L., Lagos-Kutz, D., Nichols, D. H., Patel, K. and Varriano, S. (2020) No net insect abundance and diversity declines across US Long Term Ecological Research sites. *Nature ecology & evolution.* 4 (10), 1368-1376.

Dicks, L. V., Breeze, T. D., Ngo, H. T., Senapathi, D., An, J., Aizen, M. A., Basu, P., Buchori, D., Galetto, L. and Garibaldi, L. A. (2021) A global-scale expert assessment of drivers and risks associated with pollinator decline. *Nature ecology & evolution.* 5 (10), 1453-1461.

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. and Collen, B. (2014) Defaunation in the Anthropocene. *Science.* 345 (6195), 401-406.

Eggleton, P. (2020) The state of the world's insects. *Annual Review of Environment and Resources.* 45, 61-82.

Elliott, J. H., Synnot, A., Turner, T., Simmonds, M., Akl, E. A., McDonald, S., Salanti, G., Meerpohl, J., MacLehose, H. and Hilton, J. (2017) Living systematic review: 1. Introduction—the why, what, when, and how. *Journal of clinical epidemiology.* 91, 23-30.

Engelhardt, E. K., Biber, M. F., Dolek, M., Fartmann, T., Hochkirch, A., Leidinger, J., Löffler, F., Pinkert, S., Poniatowski, D. and Voith, J. (2022) Consistent signals of a warming climate in occupancy changes of three insect taxa over 40 years in central Europe. *Global Change Biology.*

Fox, R., Harrower, C. A., Bell, J. R., Shortall, C. R., Middlebrook, I. and Wilson, R. J. (2019) Insect population trends and the IUCN Red List process. *Journal of Insect Conservation.* 23 (2), 269-278.

Gillespie, M. A., Alfredsson, M., Barrio, I. C., Bowden, J. J., Convey, P., Culler, L. E., Coulson, S. J., Krogh, P. H., Koltz, A. M. and Koponen, S. (2020) Status and trends of terrestrial arthropod abundance and diversity in the North Atlantic region of the Arctic. *Ambio.* 49 (3), 718-731.

Gillespie, M. A., Baude, M., Biesmeijer, J., Boatman, N., Budge, G. E., Crowe, A., Davies, N., Evans, R., Memmott, J. and Morton, R. D. (2022) Landscape-scale drivers of pollinator communities may depend on land-use configuration. *Philosophical Transactions of the Royal Society B.* 377 (1853), 20210172.

Grames, E. M., Montgomery, G. A., Boyes, D. H., Dicks, L. V., Forister, M. L., Matson, T. A., Nakagawa, S., Prendergast, K. S., Taylor, N. G. and Tingley, M. W. (2022) A framework and case study to systematically identify long‐term insect abundance and diversity datasets. *Conservation Science and Practice.* e12687.

Gray, C. L., Hill, S. L., Newbold, T., Hudson, L. N., Börger, L., Contu, S., Hoskins, A. J., Ferrier, S., Purvis, A. and Scharlemann, J. P. (2016) Local biodiversity is higher inside than outside terrestrial protected areas worldwide. *Nature communications.* 7 (1), 1-7.

Gurevitch, J., Koricheva, J., Nakagawa, S. and Stewart, G. (2018) Meta-analysis and the science of research synthesis. *Nature.* 555 (7695), 175-182.

Habel, J. C., Samways, M. J. and Schmitt, T. (2019) Mitigating the precipitous decline of terrestrial European insects: Requirements for a new strategy. *Biodiversity and Conservation.* 28 (6), 1343-1360.

Haddaway, N. R., Grames, E. M., Boyes, D. H., Saunders, M. E. and Taylor, N. G. (2020) What evidence exists on conservation actions to conserve insects? A protocol for a systematic map of literature reviews. *Environmental Evidence.* 9 (1), 1-8.

Hallmann, C. A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H. and Hörren, T. (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PloS one.* 12 (10), e0185809.

Hedges, L. V., Gurevitch, J. and Curtis, P. S. (1999) The meta‐analysis of response ratios in experimental ecology. *Ecology.* 80 (4), 1150-1156.

Hillebrand, H., Blasius, B., Borer, E. T., Chase, J. M., Downing, J. A., Eriksson, B. K., Filstrup, C. T., Harpole, W. S., Hodapp, D. and Larsen, S. (2018) Biodiversity change is uncoupled from species richness trends: Consequences for conservation and monitoring. *Journal of Applied Ecology.* 55 (1), 169-184.

Homburg, K., Drees, C., Boutaud, E., Nolte, D., Schuett, W., Zumstein, P., von Ruschkowski, E. and Assmann, T. (2019) Where have all the beetles gone? Long‐term study reveals carabid species decline in a nature reserve in Northern Germany. *Insect Conservation and Diversity.* 12 (4), 268-277.

Hossain, S., Calloway, C., Lippa, D., Niederhut, D. and Shupe, D. (2019) *Visualization of bioinformatics data with dash bio*, *Proceedings of the 18th Python in Science Conference*

Hudson, L. N., Newbold, T., Contu, S., Hill, S. L., Lysenko, I., De Palma, A., Phillips, H. R., Alhusseini, T. I., Bedford, F. E. and Bennett, D. J. (2017) The database of the PREDICTS (projecting responses of ecological diversity in changing terrestrial systems) project. *Ecology and evolution.* 7 (1), 145-188.

Jähnig, S. C., Baranov, V., Altermatt, F., Cranston, P., Friedrichs‐Manthey, M., Geist, J., He, F., Heino, J., Hering, D. and Hölker, F. (2021) Revisiting global trends in freshwater insect biodiversity. *Wiley Interdisciplinary Reviews: Water.* 8 (2), e1506.

Klein, A.-M., Vaissiere, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C. and Tscharntke, T. (2007) Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B: Biological Sciences.* 274 (1608), 303-313.

Koller, M. (2016) robustlmm: an R package for robust estimation of linear mixed-effects models. *Journal of statistical software.* 75, 1-24.

Lawton, J. H., Bignell, D. E., Bolton, B., Bloemers, G., Eggleton, P., Hammond, P. M., Hodda, M., Holt, R., Larsen, T. and Mawdsley, N. (1998) Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest. *Nature.* 391 (6662), 72-76.

Lister, B. C. and Garcia, A. (2018) Climate-driven declines in arthropod abundance restructure a rainforest food web. *Proceedings of the National Academy of Sciences.* 115 (44), E10397-E10406.

Loboda, S., Savage, J., Buddle, C. M., Schmidt, N. M. and Høye, T. T. (2018) Declining diversity and abundance of High Arctic fly assemblages over two decades of rapid climate warming. *Ecography.* 41 (2), 265-277.

Miličić, M., Popov, S., Branco, V. V. and Cardoso, P. (2021) Insect threats and conservation through the lens of global experts. *Conservation Letters.* 14 (4), e12814.

Millard, J., Outhwaite, C. L., Kinnersley, R., Freeman, R., Gregory, R. D., Adedoja, O., Gavini, S., Kioko, E., Kuhlmann, M. and Ollerton, J. (2021) Global effects of land-use intensity on local pollinator biodiversity. *Nature communications.* 12 (1), 1-11.

Montgomery, G. A., Dunn, R. R., Fox, R., Jongejans, E., Leather, S. R., Saunders, M. E., Shortall, C. R., Tingley, M. W. and Wagner, D. L. (2020) Is the insect apocalypse upon us? How to find out. *Biological conservation.* 241, 108327.

Newbold, T., Hudson, L. N., Contu, S., Hill, S. L., Beck, J., Liu, Y., Meyer, C., Phillips, H. R., Scharlemann, J. P. and Purvis, A. (2018) Widespread winners and narrow-ranged losers: Land use homogenizes biodiversity in local assemblages worldwide. *PLoS biology.* 16 (12), e2006841.

Newbold, T., Hudson, L. N., Hill, S. L., Contu, S., Gray, C. L., Scharlemann, J. P., Börger, L., Phillips, H. R., Sheil, D. and Lysenko, I. (2016) Global patterns of terrestrial assemblage turnover within and among land uses. *Ecography.* 39 (12), 1151-1163.

Newbold, T., Hudson, L. N., Phillips, H. R., Hill, S. L., Contu, S., Lysenko, I., Blandon, A., Butchart, S. H., Booth, H. L. and Day, J. (2014) A global model of the response of tropical and sub-tropical forest biodiversity to anthropogenic pressures. *Proceedings of the Royal Society B: Biological Sciences.* 281 (1792), 20141371.

Noble, D. W., Lagisz, M., O'dea, R. E. and Nakagawa, S. (2017). *Nonindependence and sensitivity analyses in ecological and evolutionary meta‐analyses.* Wiley Online Library.

Ollerton, J., Erenler, H., Edwards, M. and Crockett, R. (2014) Extinctions of aculeate pollinators in Britain and the role of large-scale agricultural changes. *Science.* 346 (6215), 1360-1362.

Outhwaite, C. L., Gregory, R. D., Chandler, R. E., Collen, B. and Isaac, N. J. (2020) Complex long-term biodiversity change among invertebrates, bryophytes and lichens. *Nature ecology & evolution.* 4 (3), 384-392.

Poniatowski, D., Stuhldreher, G., Löffler, F. and Fartmann, T. (2018) Patch occupancy of grassland specialists: Habitat quality matters more than habitat connectivity. *Biological conservation.* 225, 237-244.

Potts, S. G., Imperatriz-Fonseca, V., Ngo, H. T., Aizen, M. A., Biesmeijer, J. C., Breeze, T. D., Dicks, L. V., Garibaldi, L. A., Hill, R. and Settele, J. (2016) Safeguarding pollinators and their values to human well-being. *Nature.* 540 (7632), 220-229.

Powney, G. D., Carvell, C., Edwards, M., Morris, R. K., Roy, H. E., Woodcock, B. A. and Isaac, N. J. (2019) Widespread losses of pollinating insects in Britain. *Nature communications.* 10 (1), 1-6.

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [*http://www*](http://www)*. R-project. org/.*

Ram, K. and Yochum, C. (2020) *rdrop2: Programmatic Interface to the 'Dropbox' API* (Version R package version 0.8.2.1) <https://CRAN.R-project.org/package=rdrop2>.

Sánchez-Bayo, F. and Wyckhuys, K. A. (2019) Worldwide decline of the entomofauna: A review of its drivers. *Biological conservation.* 232, 8-27.

Saunders, M. E., Janes, J. K. and O’Hanlon, J. C. (2020) Moving on from the insect apocalypse narrative: engaging with evidence-based insect conservation. *BioScience.* 70 (1), 80-89.

Seibold, S., Gossner, M. M., Simons, N. K., Blüthgen, N., Müller, J., Ambarlı, D., Ammer, C., Bauhus, J., Fischer, M. and Habel, J. C. (2019) Arthropod decline in grasslands and forests is associated with landscape-level drivers. *Nature.* 574 (7780), 671-674.

Shackelford, G. E., Martin, P. A., Hood, A. S., Christie, A. P., Kulinskaya, E. and Sutherland, W. J. (2021) Dynamic meta-analysis: a method of using global evidence for local decision making. *BMC biology.* 19 (1), 1-13.

Simmons, B. I., Balmford, A., Bladon, A. J., Christie, A. P., De Palma, A., Dicks, L. V., Gallego‐Zamorano, J., Johnston, A., Martin, P. A. and Purvis, A. (2019) Worldwide insect declines: an important message, but interpret with caution. *Ecology and evolution.* 9 (7), 3678-3680.

Stewart, G. (2010) Meta-analysis in applied ecology. *Biology letters.* 6 (1), 78-81.

Stork, N. E. (2018) How many species of insects and other terrestrial arthropods are there on Earth? *Annual review of entomology.* 63, 31-45.

Titley, M. A., Snaddon, J. L. and Turner, E. C. (2017) Scientific research on animal biodiversity is systematically biased towards vertebrates and temperate regions. *PloS one.* 12 (12), e0189577.

Van Klink, R., Bowler, D. E., Gongalsky, K. B., Swengel, A. B., Gentile, A. and Chase, J. M. (2020) Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science.* 368 (6489), 417-420.

Viechtbauer, W. (2010) Conducting meta-analyses in R with the metafor package. *Journal of statistical software.* 36 (3), 1-48.

Wagner, D. L., Fox, R., Salcido, D. M. and Dyer, L. A. (2021a) A window to the world of global insect declines: Moth biodiversity trends are complex and heterogeneous. *Proceedings of the National Academy of Sciences.* 118 (2).

Wagner, D. L., Grames, E. M., Forister, M. L., Berenbaum, M. R. and Stopak, D. (2021b) Insect decline in the Anthropocene: Death by a thousand cuts. *Proceedings of the National Academy of Sciences.* 118 (2).

Welti, E. A., Joern, A., Ellison, A. M., Lightfoot, D. C., Record, S., Rodenhouse, N., Stanley, E. H. and Kaspari, M. (2021) Studies of insect temporal trends must account for the complex sampling histories inherent to many long-term monitoring efforts. *Nature ecology & evolution.* 5 (5), 589-591.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L. and Hester, J. (2019). *Welcome to the tidyverse. J. Open Source Softw. 4 (43), 1686 (2019).*

Willig, M., Woolbright, L., Presley, S., Schowalter, T., Waide, R., Scalley, T. H., Zimmerman, J., González, G. and Lugo, A. (2019) Populations are not declining and food webs are not collapsing at the Luquillo Experimental Forest. *Proceedings of the National Academy of Sciences.* 116 (25), 12143-12144.