Supporting Online Material for

**Worldwide vulnerability of local pollinator abundance and crop-pollination to land use and climate**

Joseph Millard\*, Charlotte Outhwaite, Lynn Dicks, Jeff Ollerton, Silvia Ceausu, Tim Newbold\*

\*To whom correspondence should be addressed. E-mail: [joseph.millard@nhm.ac.uk](mailto:joseph.millard@nhm.ac.uk); [tim.newbold@ucl.ac.uk](mailto:tim.newbold@ucl.ac.uk)

**This PDF file includes:**

* Materials and Methods
* Figs. S1-S13
* Tables
* References

**Materials and Methods**

**Pollinator dataset construction**

We used the PREDICTs database, and a set of likely pollinating species identified in PREDICTS (see Chapter 3 for a detailed background on the development of this PREDICTS database subset), to model the response of local pollinator abundance to the interactive effect of climate change and agricultural land use. PREDICTS is a global database of local biodiversity records, with a hierarchical structure such that each record is nested according to a series of levels (i.e. ‘source’, ‘study’, ‘block’, and ‘site’; see Chapter 3 Appendix 2 Figure S2.13, and Hudson et al., 2017). Each record in PREDICTS is associated with a land use type (primary vegetation, mature secondary vegetation, intermediate secondary vegetation, young secondary vegetation, plantation, pasture, cropland, and urban) and land use intensity (high, low and minimal), meaning change in biodiversity can be predicted as a function of anthropogenic land-use disturbance (e.g. Newbold et al., 2015; also see Chapter 3 ‘Methods’). Land-use type categories in PREDICTS are defined according to the Representative Concentration Pathway land-use estimates (Hurtt et al., 2011), plus plantation and a gradation of stages according to age within secondary vegetation (Hudson et al., 2014). Land-use intensity for each land-use type is defined according to a series of variables, such as fertiliser and pesticide application, mechanisation, and hunting (see Newbold et al., 2015 for more details).

We identified pollinating species in PREDICTS through a semi-automatic approach combining text-mining, manual inspection, and expert consultation. We first used the stemmed term ‘pollinat\*’ on Scopus to return all abstracts for English language primary research papers. From this set of abstracts, we then used a set of name-entity recognition algorithms to extract all animal species Latin binomial names (see Chapter 2 ‘Methods’; Millard et al., 2020 for details). For each animal genus returned by the name-entity recognition algorithm, we then read the corresponding abstracts searching for evidence confirming that genus as pollinating. Given my set of pollinators identified from Scopus abstracts could only ever be a sample, we then searched for higher-level taxonomic groups, aiming to identify broader groups of animals that we could be confident would pollinate (see Chapter 3 ‘Methods’ for more details). After compiling my list of pollinators from automated text-analysis and manual searching, we then consulted a group of 7 expert pollination ecologists, and removed or added any groups at their suggestion (see Chapter 3 ‘Methods’ for details).

We then built four separate datasets with different subsets of the original set of species in PREDICTS: pollinating vertebrates, non-pollinating vertebrates, pollinating insects, and non-pollinating insects. My set of non-pollinating species are not strictly a set of confirmed non-pollinators, rather a set of species not confirmed as pollinators (see Discussion for an exploration of this limitation). For each of my four data subsets, we then calculated site-level total abundance (the sum of all species sampled abundances at a site). Sampling effort was accounted for by dividing the abundance values for each measurement by the relative sampling effort among sites within a study, rescaled to a maximum value of 1 in each study, as in (Newbold et al., 2014).

**Climate change data**

We used Climatic Research Unit Time Series (CRU TS) v4.03 (Harris et al., 2020) mean daily temperature estimates per month at a spatial resolution of 0.5° to calculate a global standardised temperature anomaly (STA) for the year of each PREDICTS sample, using an approach developed previously (Outhwaite et al., in review). Although extreme temperatures have been shown to predict contemporary changes in biodiversity (Mantyka-Pringle et al., 2011), we used mean temperatures since they provide a measure of the overall change in temperature per month. Mean temperatures have been used in a similar manner for a number of other studies on insect thermal tolerance (e.g. Deutsch et al., 2008; Johansson et al., 2020). To calculate our standardised temperature anomaly, we first calculated a 30-year baseline temperature for the years 1901-1930 as the grand mean temperature over all 360 monthly mean daily temperatures for each cell. For each PREDICTS site, we then extracted the sample end date (i.e. the last date at which that measurement was taken) for the measurements at each site, and then for the month of that end date and the 11 months previously we calculated the grand mean temperature over all monthly mean daily temperatures for that sample. We then calculated a climate anomaly for each site as the difference between the 30-year baseline mean and the mean annual temperature at the time of the sample. We then standardised this climate anomaly by dividing the anomaly at each site by the standard deviation of the baseline, calculated as the standard deviation across monthly mean daily temperatures for the same 30-year baseline period. A standardised temperature anomaly of less than 0 indicates a region that has cooled since the baseline. A value between 0 and 1 indicates a region that has warmed, but current average temperature remains within 1 standard deviation of the variability in baseline temperatures. A value greater than 1 indicates a region in which average warming is 1 standard deviation greater than the variability in the baseline (i.e. it is now experiencing high novel temperatures). Again, using CRU TS v4.03, we also calculated a global standardised temperature anomaly independent of PREDICTS sites (for an average of the 36 monthly mean temperatures in the period 2004-2006), which we mapped globally for context (Figure 4.2). We chose the period 2004-2006 to coincide with the temporal distribution of most records in the PREDICTS database (Hudson et al., 2017).

**Interactive effects of land-use and climate change on pollinator biodiversity**

To model the effects of land-use and climate change on pollinator abundance, we built linear mixed-effects models predicting total abundance as a function of land-use type (primary vegetation and cropland), standardised temperature anomaly, and their interaction. We did not use a generalised linear model with Poisson errors because most recorded measurements are non-integer values (also see Chapter 3, showing that log-linear and negative binomial models for land use intensity yielded similar results). We focussed on primary vegetation and cropland given my interest in the interactive effects of land-use and climate change between natural and anthropogenic land. We chose primary vegetation since it best represents natural vegetation, and cropland since it is both an anthropogenic land-use type and most relevant to crop production that depends on animal pollination. Since the interactive effect of land-use and climate change is likely greater between natural and anthropogenic land than among levels of intensity within a land-use type (Outhwaite et al., in review), for both cropland and primary vegetation we grouped together all levels of intensity. We loge-transformed all total abundance values (adding one because of zero values) to normalise the model residuals. Due to the nested nature of the database (Hudson et al., 2017), we included a random intercept of study identity to account for variation in sampling methods, sampling effort, and broad geographical differences among studies, and a random intercept of spatial block within study to account for the spatial structuring of sites. Random-effects structures were selected to minimise AIC values.

**Potential future risk to crop pollination from land-use and climate change**

Given the importance of insect pollinators to global crop production, we used our model of insect pollinator abundance to predict geographic and temporal patterns of pollination shortfall risk, based on a combination of estimated crop production dependent on animal pollination, and projections of future climate change. We focus on two forms of risk: production risk, and proportional production risk. Production risk is a measure of the total crop production that could be at risk of shortfall. Proportional production risk is a measure of the crop production at risk as a proportion of the total production for a given cell, crop, or country. For both measures of risk, we specifically focussed on insect pollinator abundance given that insects make a bigger contribution than vertebrates to global crop pollination (Klein et al., 2007).

My projections assume that a projected loss in pollinator abundance combined with crop dependence on pollination is a good proxy for risk to crop production. There are three core uncertainties associated with these projections (also see Discussion). First, we do not know if there is a mechanistic link between the interactive effects of land-use and climate change on pollinator abundance. We reason that a significant interactive effect is at least likely however, given prior localised studies demonstrating a synergistic effect of climate change and anthropogenic land use in insects (Outhwaite et al., in review; Zaragoza-Trello et al., 2021). Second, we do not account for changes in the distribution of crops from the effects of climate change alone. Third, it remains unclear how local abundance change will impact crop pollination, and in turn how crop pollination will relate to yield change. Given these uncertainties our projections should not be interpreted as absolute projections of yield loss, but rather as a relative measure of crop production risk. Such a measure of global risk does not yet exist, despite its marked importance to policy-makers (IPBES, 2017).

We first used a global dataset of crop production (Monfreda et al., 2008), in combination with the animal pollination dependencies reported in Klein et al (2007), to build a map of global crop production dependent on animal pollination for the year 2000. For each crop in Monfreda et al (2008), we adjusted total production for the pollination dependence ratios reported in Klein et al (2007) (essential, 0.95; great, 0.65; modest/great, 0.45; modest, 0.25; little, 0.05; no increase, 0). Given some Monfreda crops are represented by multiple crops in Klein et al (2007), for each Monfreda et al (2008) crop we first calculated the mean pollination dependence ratio among Klein crops. We then adjusted each crop for its pollination dependence and summed dependence-adjusted production for all crops at each cell as

where *PollinationProd* = animal pollination dependent crop production, *Prod* = crop production, *c* = Monfreda crop category, *d* = average animal pollination dependence ratio, and *i* = a given cell. We also calculated the total production in any given cell (independent of animal pollination dependence) for the same set of crops, as

where *TotalProd* = total crop production. In any further work, we assumed that the distribution of pollination-dependent production in 2000 will be representative of that in 2050 (see ‘Discussion’ for an exploration of this limitation). We also did not exclude pollination-dependent production which relies on either vertebrates or non-insect invertebrates. Although there are some crops recorded in Klein et al (2007) as reliant on either vertebrate or non-insect pollinators, after reconciling these crops with Monfreda et al (2008) there are no crop groups that are exclusively dependent on vertebrates or non-insect invertebrates.

We then used ISIMIP temperature anomalies from Frieler et al (2017) to project potential future change in standardised temperature anomaly under three RCP (Representative Concentration Pathway) scenarios (8.5, 6.0, 2.6), using an ensemble mean of the climate models GFDL, HadGEM2, IPSL, and MIROC5. RCP 8.5 represents a worst-case high-emissions scenario, 6.0 a pathway with some degree of mitigation, and 2.6 a pathway with significant reductions in emissions (van Vuuren et al., 2011). ISIMIP temperature anomalies were added onto a historical baseline for the period 1979-2013 (inclusive), calculated as the grand mean over all monthly mean daily CRU temperatures for each cell. For each year, the standardised temperature anomaly was projected across the terrestrial surface of the globe (for the years 2016-2048), using a 3-year rolling average to smooth change in risk over time. For each 3- year projection window, insect pollinator abundance on cropland was predicted according to the model in Figure 4.1 (top left panel) for all cells containing crop production dependent on animal pollination. These abundance values were then expressed as the proportional loss of abundance compared to the abundance expected on croplands that have experienced no warming (i.e. standardised temperature anomaly of 0) as

where *l* = abundance loss, *t* = each 3-year time window, *aw*= predicted abundance on cropland under projected warming, and *a0*=predicted abundance on cropland under no warming. In each cell, animal-pollination-dependent crop production was then adjusted for the percentage reduction in abundance at that time step, before summing production at risk for all cells as

where *ProdRisk* = the total crop production risk (i.e. production risk). To check the influence of single climate models on our projections, we calculated one projection as the average of all models for that RCP scenario, and a set of additional jack-knifed projections, dropping each climate model in turn. For the average of all climate models at RCP 8.5, we also checked the extent to which extrapolating abundance loss beyond the lowest fitted abundance value (i.e. greatest abundance loss) of our PREDICTS models affected my projections. Specifically, we projected change in total pollination-dependent crop production at risk when the greatest loss in predicted abundance is capped at the largest fitted value of standardised temperature anomaly (Appendix 3, Figure S3.2). For my projection in which abundance loss is not capped, we also calculated the percentage of cells where abundance loss was extrapolated beyond the maximum fitted value (Appendix 3, Figure S3.2).

To identify geographic regions in which a high proportion of crop production could be at risk under a worst-case climate scenario (i.e. RCP 8.5), we projected pollination risk for each cell as above (using the average of all four climate models), and then expressed pollination risk as a proportion of all crop production within each cell

where *PropRisk* = the proportion of total crop production at risk (i.e. proportional production risk). A value of 1 therefore indicates a hypothetical region in which all crop production in that cell is dependent on pollination, and predicted insect pollinator abundance loss is 100% (i.e. one would expect a very high risk of pollination shortfall). We projected this dependence risk for 2006 and 2050, and mapped these projections globally. We also plotted the change in risk for the 10 countries with the highest overall risk (calculated as the average over the whole time series), and the change in risk for the 10 countries with the highest rate of change (the difference between the lowest and highest over the whole time series). For each of these high-risk countries, we took the overall risk to be the median of all cells in that country at each time step, onto which we overlaid the 2.5th and 97.5th percentiles as a measure of the variation in risk across the country.

Using the same approach as above, we also projected proportional production risk for each of the top 20 crops by total pollination-dependent production (i.e. *PropRiskcti*). Again, we took the overall risk to be the median of all cells in which that crop appears at each time step, onto which I overlaid the 2.5th and 97.5th percentiles as a measure of the variation in risk across locations where the crop is grown. We also calculated a change in total production risk for the same top 20 crops (i.e. *ProdRiskct*), which we then adjusted to a percentage of total production for each crop (i.e *PropRiskct*). This measure therefore estimates the total proportion of global production at risk for a given crop. All analyses were carried out in R v4.0.5. All core analyses can be visualised interactively in a Shiny app currently hosted online ([joemillard.shinyapps.io/pollinator\_dependence\_visualisation/](https://joemillard.shinyapps.io/pollinator_dependence_visualisation/)).

**Chart, bar chart, histogram

Description automatically generated**

Fig S1. Total pollination dependent production for the 20 crops with the highest pollination dependent production. Total production values are for the year 2000, taken from Monfreda et al (2008). Pollination dependent production is calculated by multiplying total crop production for each crop by the pollination dependence ratios for that crop, as reported in Klein et al (2007). For any Monfreda crop represented by multiple dependence ratios, I took the pollination dependence to be the mean of the ratios for that crop.

**Chart, histogram

Description automatically generated**

**Fig S2**. Projected change in total production risk under three RCP scenarios (8.5, 6.0, and 2.6), for the average of four climate models (GFDL, HadHEM2, IPSL, and MIROC5). Here circles represent a projection in which abundance loss is extrapolated beyond the maximum standardised temperature anomaly (STA) (as in Figure 4.3), whereas triangles represent a projection in which abundance loss is capped at an STA of 1.58 (2dp). Colour refers to the percentage of cells at each time step that have been extrapolated beyond the maximum STA in Figure 4.3, which in the projection in which abundance loss is capped (triangles) will always be 0.

Diagram

Description automatically generated

**Fig S3.** Change in proportional production risk for each crop, projected under RCP scenario 8.5 from the average of four climate models (GFDL, HadGEM2, IPSL, and MIROC5), for the top 20 crops by total pollination dependent production (ordered by rate of change). For each crop global standardised temperature anomaly was projected for all areas of pollination-dependent cropland to 2050, using the 3-year average approach as described in Figure 4.3. For each value of standardised temperature anomaly, insect pollinator abundance was predicted according to the model in Figure 1. Insect pollinator abundance at each cell at each time step was then expressed as proportional loss compared to cropland that has experienced no warming (i.e. standardised temperature anomaly of 0). Animal pollination-dependent production at each cell for each crop was then adjusted for the predicted proportional loss of insect pollinator abundance, and then summed for all cells at each time step. The sum at each time step was then divided by the total production for that crop and multiplied by 100, giving a percentage of total production at risk.

**Chart, line chart

Description automatically generated**

**Fig S4**. Change in production risk for the average of four climate models (GFDL, HadHEM2, IPSL, and MIROC5) under RCP 8.5, here adjusted to an index for each of 4 data quality subsets. For each coloured line crop production was subsetted according to four levels of data quality reported in Monfreda et al (2008), before the projection was then rerun: 1, county level census data; 0.75, state level census data; 0.5, regional interpolation from census data; 0.25, country level census data; 0, no census data. For each simulation here I converted total production risk to an index by calculating the percentage change between each time point, and then calculating the cumulative product of these percentage changes (starting at an initial index of 1).

Chart, diagram

Description automatically generated

**Fig S5**.

Chart, histogram

Description automatically generated

**Fig S6**.

A picture containing timeline

Description automatically generated**Fig S7**.

A picture containing diagram

Description automatically generated**Fig S8**.

Map

Description automatically generated**Fig S9**.

Chart, line chart

Description automatically generated

**Fig S10**.

Chart, surface chart

Description automatically generated

**Fig S11**.

Chart, scatter chart

Description automatically generated

**Fig S12**.

Chart

Description automatically generated

**Fig S13**.