Science
AAAS

Supplementary Materials for

Rapid butterfly declines across the United States during the 21st century

Collin B. Edwards *et al*.

Corresponding author: Collin B. Edwards, edwards.evoeco@gmail.com

Science **387**, 1090 (2025)
DOI: 10.1126/science.adp4671

The PDF file includes:

Materials and Methods
Fig. S1 to S5
References

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S10

Materials and Methods

Data sources

We included butterfly count data from three types of butterfly surveys (table S1). First, we used data from the North American Butterfly Association (NABA), which is the longest running volunteer-based systematic count of butterflies in the world (*11*). We included NABA "survey circle" data across the contiguous United States (U.S.); these data represent one or multiple trips that volunteers carry out in one or more locations within a 15-mile (24.1-km) diameter circle. All butterflies seen are identified and counted and counts across trips on the same day are summed. Total party minutes spent surveying are also summed across trips (representing the cumulative time spent across parties, not volunteers). Second, we used data from the Massachusetts Butterfly Club, which carries out organized field trips and records individuals' reports across the state in which participants identify and record butterflies seen. Third, we included repeated transect data from a number of monitoring programs. Many of these programs employ "Pollard walks," a standardized protocol in which an observer walks a systematic path, usually once a week, and counts all butterflies seen for each species (*39*) or for one or a few targeted species (hereafter, "targeted surveys"). A few programs employ distance-based sampling, a repeated transect method similar to Pollard walks that accounts for the reduced probability of observing butterflies at greater distances (*40*). Large-scale and long-term Pollard walk datasets have been collected in an east-west transect across California (*41*), across Ohio and Illinois, and from 14 other regional programs, all represented in PollardBase (https://pollardbase.org). Some programs employing repeated transect methods report time spent by observers on each survey; others do not. We note that the choice of sampling locations is non-random, and available data provide a biased representation of types of land use and climate (e.g., no surveys were established in parking lots).

Data processing

Data cleaning

Each observation ("record") consists of a butterfly species count, site identity (including latitude and longitude), and survey date; in many cases observations also include information about survey effort. Surveys represent unique sampling events, and typically consist of multiple records each representing the count of a different species seen on that survey. Sites generally represent unique locations that were visited for multiple survey events. For most programs employing repeated transect methods, site identity represents a specific route that was walked multiple times per year for many years. For NABA records, "sites" represent the unique 15-mile diameter circle; while generally the same circle was surveyed across multiple years, individual survey events could be based on trips located in different parts of the circle.

We used latitude and longitude to identify US Fish and Wildlife Service regions associated with each observation and removed observations that were outside of the contiguous U.S. For one monitoring program (Texas Butterfly Monitoring Network), latitude and longitude were only provided at the county centroid level.

In carrying out data validation, we identified and resolved issues in the data. Some programs report separate counts for each section of a survey, which leads to multiple records per survey for individual species. We summed these counts to create a single count per species per survey. Some surveys with the same site name were duplicated, but reported with different latitude and longitude values; generally the differences in latitude and longitude appeared to reflect differences in rounding decisions. We reconciled longitude and latitude and then removed duplicate surveys. Some surveys were duplicated but with different reported survey effort; we reconciled survey effort by averaging the "duration" or "party minutes" fields (see below), and then removed duplicate surveys. We also removed surveys reporting negative or 0 minutes spent observing, and surveys that reported more than 12 party-hours of surveying because both are likely errors.

Harmonizing taxonomy

Programs typically reported butterflies using scientific names, sometimes to the subspecies level. To integrate data across programs, we resolved taxonomic conflicts using the North American Butterfly Monitoring Network (NABMN) taxonomic framework (hereafter, "NABMN codes"). NABMN codes were updated when necessary to accommodate changes in both North American Butterfly Association's Checklist of North American Butterflies (*42*) and the Catalogue of the Butterflies of the U.S. and Canada (*43*). Subspecies were aggregated to the species level for our analyses, where species are defined using the NABMN framework. The NABMN framework recognizes four subspecies – *Limenitis arthemis arthemis* (white admiral), *Limenitis arthemis astyanax* (red-spotted purple), *Lycaeides melissa samuelis* (Karner blue), and *Boloria improba acrocnema* (Uncompahgre' dingy fritillary) – which we analyzed at the species level for consistency with programs that did not identify to subspecies level. We also resolved a range of spelling inconsistencies across data sets.

Integrating data structures

Most programs use a presence-only reporting protocol in which, for any given monitoring event (e.g., observation trip or survey day), only species that are observed are recorded. Consequently, if a species was not recorded in a survey, we assume that it was not detected during that survey event, and we treated that as an observation with an "implicit" zero count for that species. For other programs (the Shapiro data set and targeted repeated transect counts), counts of zero are reported for species not seen. To recreate implicit zeroes when combining the data for our analysis, we had to overcome two challenges. First, implicit zeroes are only meaningful within the geographic range of the species of interest. For example, if we recreate implicit zeroes across the U.S. for a species found only in the state of Florida, the vast majority of the "data" for this species will be implicit zeroes in regions in which the butterfly would never be found. For this reason, we only incorporated implicit zeroes for monitoring events within the extent-of-occurrence polygons generated for each species. The estimated extent-of-occurrence polygons were generated by combining expert-drawn range maps (*44*) with environmental niche models (*21*). Second, in some cases observations were reported with identification above the species level (e.g., "unidentified in X genus" or "unidentified in X family"). In total, records with identification at the genus, subfamily, or family level represented only 2.7% of the total available data. However, in surveys that included such records, the absence of a species of interest from a survey might not truly correspond to zero individuals if an individual was observed but only identified to genus or family. We took a conservative approach and only created implicit zeroes for monitoring events in which the species of interest was not reported and there were no individuals identified above the species level that could have belonged to the species of interest.

The extent-of-occurrence polygons did not use the same NABMN taxonomic classifications we used in our study, and at the time of analysis there were two taxa with points of conflict. *Ochlodes yuma* (the Yuma skipper) did not have a unique range map, and instead shared an extent-of-occurrence polygon with *Ochlodes sylvanoides* (the woodland skipper); we used the *Ochlodes sylvanoides* extent-of-occurrence polygon for both taxa. *Phoebis statira* (formerly *Aphrissa statira*; the statira sulphur) had two distinct extent-of-occurrence polygons associated with it, and we excluded this species from further analysis.

Integrating measures of effort

Effort was reported either as duration (in minutes, from single-observer repeated transect count programs) or total party minutes (from NABA). In some cases, effort was not reported but the same route was sampled every survey for a given site (for some Pollard-type programs). We accounted for duration and party minutes separately, as duration represents time spent by a single surveyor while party minutes represents time spent by one or more surveyors who could be surveying together. We also included a random site-level effect to account for variation across survey locations. For computational purposes, we normalized duration and party minutes separately to each have a mean of one. Observations without reported effort were presumed to have average effort (i.e., given an effort value of one after the normalization step). The Massachusetts Butterfly Club surveys did not report effort and did not resample the same sites, and thus we were not able to account for variation in effort across these surveys.

Analysis overview

We focused our study on evaluating changes in abundance between the years 2000 and 2020 because 14 (40%) of the programs started data collection by 2000 and 88% of all available surveys were conducted on or after the year 2000 (figure S4D). In modeling trends in total butterfly abundance as well as trends for individual species, we worked with three distinct measures of trends, which we outline here and summarize in table S2. Fitted models directly estimated "**regional rates of change**" (i.e., growth rates, a common component of continuous time population ecology model) as coefficients. We further integrated regional rates of change and measures of relative abundance ("regional abundance index", below and table S2) in each region to calculate "**yearly-"** and **"overall rates of change"**. Rates of change are the instantaneous rates of an exponential model (i.e., population growth rates), with negative values corresponding to decreasing trends and positive values corresponding to increasing trends. Rates of change are convenient mathematically but often not intuitive to interpret; we therefore focused on presenting "**annual percent change,**" which we calculated by exponentiating the corresponding rate of change (i.e., the annual percent change for a region was the exponentiation of the regional rate of change). The annual percent change is closely linked to the lambda growth parameter common in discrete-time population ecology models (e.g., a lambda of 1.1 = +10% annual increase, and a lambda of 0.85 = -15% annual decrease). For models with trends from more than one region, the yearly rate of change and annual percent change varied across years. We addressed this in two ways. First, our calculation of the overall rate of change accounted for variation across years, and the exponentiation of this term (the overall annual percent change) was equivalent to the geometric mean of each year's annual percent change. Additionally, both for precision and to highlight the changes across the duration of our study, we calculated the "**cumulative percent change.**" This was an estimation of the total change in abundance we observed across the study period and was calculated by comparing model predictions from 2000 to model predictions from 2020.

Modeling trend in total butterfly abundance

We estimated the change in total butterfly abundance by first summing all individuals counted in each survey (including individuals that were not identified to the species level) to create a single abundance count for each survey. We excluded targeted surveys (table S1) for this analysis as these overrepresent very rare butterflies that are targets of conservation efforts and intensive sampling. (However, we did include targeted surveys in species-level models – see below). Trends in butterfly abundance can have distinct regional patterns; therefore, we modeled separate regional trends and then integrated our estimates into a single trend for the contiguous U.S.. To account for spatial variation, we separated our data into the seven U.S. Fish and Wildlife Service Regions, relabeling "Pacific" to "Pacific Northwest" for clarity.

We fit survey counts with a generalized additive model (GAM) to estimate trends in abundance for each region (20), using the package mgcv (*45*) in the programming language R (*46*). Our model included fixed effects for region, monitoring program, and region by year interaction (box 1). To account for variation in effort, we included separate fixed effects for party minutes and duration, using interactions with indicator variables to apply the associated coefficients only to observations that included reported party minutes or duration, respectively. To account for seasonal variation in activity – which could differ across regions – we included thin-plate regression spline smoothers across day of year for each region with the default maximum flexibility for smoothing splines in mgcv (k = 10), with a shared smoothing penalty (e.g., model *S* of *47*: shared smoothing penalty, no global trend). To account for variation among sites and to capture site-associated effort of Pollard walks, we included site as a random effect. For computational efficiency and to improve model convergence, we pooled all sites with ten or fewer surveys into the same site identity. This arrangement meant the model did not attempt to estimate separate site coefficients for sparsely monitored sites, something that sometimes prevented model convergence. We modeled individual counts with a negative binomial distribution using a log link function and fit the model using fast REstricted Maximum Likelihood (fREM). For computational efficiency, we used the `bam` function of mgcv, an alternative to the `gam` function that is optimized for working with large data sets.

To obtain regional rates of change, we relied on a convenient property of log-link models. Just as the natural logarithm of an exponential growth curve is a line, a linear term in a log-link model represents a relationship of exponential increase or decrease. The interaction of year and region in our model corresponds to separate exponential growth or decline estimates for each region; the estimated coefficients of the interaction and their uncertainty are the regional rates of change and their uncertainty. Note that our model represents regional trends as simple exponential growth or decline; see the section "Why only linear trends?" below for additional discussion.

We calculated the yearly rate of change as the weighted average of regional rates of change, using the proportion of butterflies in each region for the weights. To determine the proportion of butterflies in each region, we calculated a metric of butterfly abundance that was comparable across years and regions. Because our model accounted for seasonal variation in butterfly abundance, there was no way to directly estimate a single value of butterfly abundance within a year. Instead, we calculated the "site-level abundance index" for a given year numerically using model predictions; we then scaled by geographic area to calculate the "regional abundance index" for that year, which we used as weights. We also summed regional abundance indices in each year to calculate the "yearly abundance index" for additional comparisons. Note that the proportion of butterflies in each region (i.e., relative values of regional abundance indices) changed across years; as time went on, the region with the highest growth rate had an increasing proportion of that year's butterflies, leading to an increased yearly rate of change.

To calculate our site-level abundance index for a single region and year, we calculated the area under the curve of predicted counts across a biologically relevant period of the year ("window of activity"). We defined the window of activity for a region by the quantile of surveys with non-zero counts across day of year (DOY): the window began at the DOY of the 0.005 quantile and ended at the DOY of the 0.995 quantile. To obtain model predictions, we defined values for all variables included in the model. If held constant, variables other than year and region – effort, site, monitoring program – would not impact comparisons across years and regions but were still required values for the model to predict butterfly counts. We chose to use the NABA monitoring program, the average of site effect across time and space, and average effort for predictions. From the fitted model, we predicted butterfly counts at 0.1-day intervals across the window of activity for each region and each year and integrated the area under these predictions using the trapezoid method to produce a site-level abundance index. This index is analogous to estimated observable butterfly-days from models with average effort (e.g., *24*). To calculate the regional abundance index, we multiplied the site-level abundance indices calculated for each region by the region's area (in km^2). Under the assumptions that: 1) all sites within a region are equally good indicators of butterfly activity within a year, 2) the entire region is well represented by monitoring sites, and 3) detection is constant across space and time, the abundance index is proportional to the total number of butterflies in each region. These assumptions are consistent with those required for similar analyses (e.g**.,** *16*). We calculated yearly abundance indices by summing regional abundance indices for each year (line in Fig. 1B), and we calculated the cumulative % change in abundance by comparing the yearly abundance indices of 2000 and 2020. We note that the regional and yearly abundance indices do not reflect absolute abundance but can be used as a relative measure of butterfly abundance and can be compared across years when calculated from the same fitted model. Because abundance indices are used for relative comparisons, rescaling the values by multiplying by a constant does not change any of our calculations or results. Thus, when plotting the abundance index in Fig. 1B, we divided by a constant to put the abundance indices on a similar scale to the average survey counts.

Because there was no clear way to estimate uncertainty for changes in our yearly abundance index, we instead calculated and measured uncertainty for the overall rates of change and the annual percent change in abundance. Because the yearly rates of change (weighted average of regional rates of change, using as weights the regional abundance indices in each year) varied across years, we calculated the overall rate of change as the arithmetic mean of each of the estimated yearly rates of change from 2000 to 2020. Exponentiating this value gave an annual percent change that was equivalent to the geometric mean across annual percent changes. This is because the geometric mean $\bar{X}_{geometric}$ of some variable $\mathbf{X} = x_1, \ldots, x_n$ is equal to $e^{\mathrm{E}[\log(X)]}$ and so the log of the geometric mean of $\mathbf{X}$ is the arithmetic mean of the log values of $\mathbf{X}$. To determine the uncertainty for the overall rates of change, we extracted the standard error associated with regional rates of change from the fitted GAM model and propagated this uncertainty to yearly and overall rates of change using the delta method as implemented in the R package msm (*48*). In doing so, we assumed the covariance in propagated error between years was zero. We calculated annual percent changes in abundance by exponentiating the overall rate of change.

Aggregating observations of butterflies of different species into a single count provided a useful estimate of the overall trend in butterfly abundance. However, this approach could mask the dynamics of most butterfly species if the most abundant species had unusual trends. We thus assessed the robustness of our overall trend estimate by repeating the above analyses after removing all data from the most common species in each region (greatest total count summed across sites and years). We obtained qualitatively similar results when fitting our model to these truncated data, with the exception of the regional estimate for the Pacific Northwest (table S4). In the Pacific Northwest, the most prevalent species was *Nymphalis californica* (California

Tortoiseshell), a highly irruptive species which spiked in abundance within the region (and our data) in 2018 and 2019. See main text for full results.

Modeling trends in individual species

Species-level full model

We estimated changes in abundance individually for 356 species with sufficient data (i.e., ten distinct years with non-zero counts and ≥30 total non-zero counts). For the 306 species that also had at least 10 unique sites with non-zero observations and at least 60 total surveys with non-zero counts in one or more regions, we estimated species trends regionally with the "full" model (table S3). For each species, we filtered our data to the extent-of-occurrence polygons calculated for that species (*21*) and imputed implicit zeros (see above) for all appropriate surveys. We then fit the same model used to estimate the trend in overall butterfly abundance with the following terms: fixed effects of region, region by year interaction, monitoring program, thin plate regression smoother across day of year for each region, fixed effects of duration and party minutes interacting with indicator variables, and random effect of site (box 1). Because these models involved less data than the total abundance model, we were able to treat more sites independently, and only pooled sites with two or fewer surveys into the same site identity. For some species we had data for only a single region, a single monitoring program, or a single (or no) effort type (duration or party minutes) reported. In those cases, we simplified the model accordingly.

For computational reasons, we fit models with 6,000 or more observations using the \`bam\` function with the fREML method; for species with fewer data points we instead fit using \`gam\` with REstricted Maximum Likelihood (REML). In a few cases, individual species did not converge under the above treatment, and either required the use of \`gam\` instead of \`bam\` or required additional data filtering to ensure model convergence. This amended approach was generally necessary when one or more monitoring programs were present but very poorly represented in the data, having no or very few non-zero counts. Additional filtering was necessary for model convergence for 10 species (table S5). For 10 species, the model fitting process failed to reach convergence for the smoothing penalty, resulting in a warning. This may reflect that the smoothing term required additional maximum flexibility (more knots) for optimal fitting. However, the smoothing term did not directly contribute to our estimation of rates of change or changes in abundance index, and we chose to use a consistent modeling framework (including smoother flexibility) across all species. Thus, we retained these 10 species for all analyses.

We calculated site-level, regional, and yearly abundance indices using the same methods as in our model of trends in total butterfly abundance, with a few exceptions. A few species were not observed in NABA circle counts, and we used a different monitoring program as the baseline when calculating abundance indices. Because the baseline monitoring program was the same when producing predictions across years and regions for a given species, the choice of monitoring program did not change the relative values of our abundance indices. To calculate the regional abundance index from the site-level abundance indices for individual species, instead of multiplying by the total area in the region, we multiplied by the total area (in km^2) of the extent-of-occurrence polygons for that species in the corresponding regions. Yearly abundance indices were the sum of regional abundance indices of that year; as with the trends in the total butterfly abundance model, we calculated the cumulative percent change in abundance by comparing the yearly abundance indices of 2020 and 2000. Regional and annual abundance indices are comparable among regions and years within species, but do not represent absolute abundances of butterflies and are not comparable across species. However, proportional changes in abundance (e.g., cumulative percent changes) are comparable across species (table S5).

We calculated species-level overall rates of change and annual percent change in abundance using the same methods as in the total butterfly abundance model. For species present in only one region, the regional rate of change extracted from the GAM model was simply the overall rate of change.

Species-level simplified model

Several species (50 species) had insufficient data to meet our criterion for the full model with regional-level analysis, but still had ten or more unique years with non-zero counts and 30 or more surveys with non-zero counts (table S3). Often these were rare or at-risk species that were heavily surveyed in only a few locations. Generally, these species had very small ranges, which precluded the need to account for regional variation. We fit data for these species with a simplified model identical to the species-level full model described above, except (a) we included a single smoother across day of year instead of one per region, and (b) we included a main effect of year instead of a year by region interaction (box 1). As with the full model, we fit species with 6,000 or more observations using `bam` with fREML, using `gam` with REML for all others. Several species with more than 6,000 observations required the use of `gam` to ensure convergence and three species required additional data filtering to ensure model convergence (table S5). For two species, model fitting failed to converge for the smoothing penalty, suggesting the smooth term may have been insufficiently flexible. As with the species-level full model, we chose to use a consistent modeling framework rather than customize model flexibility for individual species, and we report results for these two species and include them in further analyses. We calculated the abundance indices, rates of change, annual percent change, and cumulative percent change as above; because the species-level simplified model fits a single effect of year, we extracted overall rate of change and associated uncertainty directly from the model (i.e., the same method as the species-level full model for species present in only one region).

Expert review of population trends

After calculation of population trends, each species was reviewed by at least two experts from among the authors (T. Wepprich, Oregon Department of Forestry, Salem, OR also served as an expert for some species). Species were flagged for incongruence with trends available from external sources or if visual inspection found that 1-3 individual survey events appeared to drive an estimated trend which was not representative of most observations. Fourteen species were removed at this step: *Amblyscirtes belli* (Bell's Roadside-Skipper), *Amblyscirtes elissa* (Elissa Roadside-Skipper), *Amblyscirtes texanae* (Texas Roadside-Skipper), *Boloria chariclea* (Arctic Fritillary), *Callophrys polios* (Hoary Elfin), *Chlosyne acastus* (Sagebrush Checkerspot), *Cymaenes tripunctus* (Three-spotted Skipper), *Heliopetes laviana* (Laviana White-Skipper), *Hesperia metea* (Cobweb Skipper), *Nastra neamathla* (Neamathla Skipper), *Neominois ridingsii* (Ridings' Satyr), *Poanes aaroni* (Aaron's Skipper), *Problema bulenta* (Rare Skipper), and *Siproeta stelenes* (Malachite) After removing these species, we were left with 301 species fit with the species-level full model, and 41 species fit the species-level simplified model; we used these 342 species in all further analyses.

Richness calculation

We created a richness index to represent the number of species that could reasonably be found in each region. From each of the 301species for which we had sufficient data to fit separate regional trends and for which NABA counts could be predicted, we used the species-level full model (see previous section) to predict counts for a NABA survey with average effort at an average site on the day of year with highest activity for that species. If the predicted count was ≥1, we considered the species 'likely observable'. We then calculated species richness as the total number of likely observable species in the region-year of interest. Uncertainty in these richness estimates was not calculated.

Why only linear trends?

The models we employed captured only exponential increases or decreases in each region (i.e., linear trends on a log scale). Where possible, ecologists – including many of the authors here – seek to fit population models with greater flexibility. Models that incorporate nonlinearities can provide key insights into population behavior, especially in the context of changing environmental drivers like climate, pesticide use, land use, and habitat management. We initially explored a range of model options for this study, and eventually recognized it was not feasible to capture non-linear trends using a general framework that would produce comparable estimates across species. Species varied greatly in their number of detections/observations within the data, with many not having sufficient data to support more complex models. Modeling nonlinear trends would have required either removing hundreds of species or fitting a series of models with differing levels of complexity, limiting our ability to make comparisons across species. In using different model structures, any estimated differences in trends across species could have been due to true biological processes, or simply to the differing constraints of the various model structures. This issue is compounded by data availability differences among regions for individual species, such that some species would require different model structures for each region. Although incorporating a flexible structure in our models would have been ideal, the inclusion of nonlinear trends increases the complexity (e.g., calculating uncertainty in estimates) and likelihood of errors (e.g., achieving model convergence) while simultaneously complicating the process of expert validation. Future work should consider the possibility of incorporating nonlinear trends into butterfly analyses as additional data becomes available.

Post-hoc analyses

False discovery rate correction

By fitting separate models to hundreds of species and reporting individual P-values, we increased the likelihood of encountering false positives (species for which an apparently significant non-zero trend was in fact not significant). Because of our low statistical power, we focus on reporting the exact P-values in main text. However, we also calculated corrected P-values based on the false discovery rate correction method of (*49*). We found that 87 species had significant overall trends at the $P < 0.05$ level using the more stringent P-value calculations compared with 123 species reported in the main text (table S5). We further found that the ratio of declining to increasing species was similar, 79 declining and 8 increasing species (9.9:1 ratio) using the more stringent P-value calculations compared to 114 declining and 9 increasing species (12.7:1 ratio) reported in the main text.

Regional comparisons

Several studies have found Northern hemisphere species have higher population growth rates at the northern edge of their range compared to the southern edge, consistent with a warming climate and latitudinal variation in thermal constraints (*14*, *22*). To test for this pattern within our data, we identified pairs of regions adjacent to one another along a north-south gradient: Southeast and Northeast regions, Southeast and Midwest regions, Southwest and Mountain Prairie regions, and Pacific Southwest and Pacific Northwest regions (Fig. 1). For each region-pair, we used only species present in both regions. We fit linear regression models, using species' estimated regional rates of change as the response variable and including as predictors a fixed effect for region geography (north vs south) and a fixed effect of species identity. We weighted each observation (i.e., each species-region) by the precision matrix (inverse of the variance-covariance matrix) of the regional rates of change to account for uncertainty. We included a fixed effect of species in this model to create a paired t-test; this approach is possible because we only included species with growth rates in both regions. We fit each region pair this way and tested for significant differences using marginal hypothesis testing as implemented in the R package car (*50*).

Trait association

We hypothesized seven key traits might be associated with butterfly population trends: 1) body size (wingspan in cm), 2) degree of host-plant specialization, 3) voltinism, 4) life stage of overwintering individuals, 5) affinity for moist habitats, 6) association with human-dominated habitats, and 7) affiliation with different types of canopies. Larger butterflies have sometimes been found to have more negative or downward-trending population trends relative to smaller butterflies (*1*); conversely, larger butterflies are often more mobile, which may allow them to move in response to climate change or other disturbances (*51-52*). Butterflies that are specialists on one or a small number of food plants are generally expected to be less resilient to variable environments (*53*), and butterflies that are host plant specialists have experienced more negative population trends (*54*). At-risk butterfly species are disproportionately univoltine (one generation per year) (*12*), and multivoltinism (more than one generation per year) has sometimes been associated with greater population trends (*14*); we expected multivoltine species to have higher trends than univoltine species. Species that overwinter as eggs have been more vulnerable to declines in regional studies (*12, 55*). Dry regions and dry years have previously been associated with more negative butterfly population trends (*16-17, 56*), and we expected species with lower moisture affinity to have more negative population trends. Given the ongoing restructuring of natural habitats in the Anthropocene (*6*), we expected species with higher affinity for human-dominated habitats to have higher population growth rates. Species of different canopy affiliations may be more or less vulnerable to habitat loss according to regional land cover changes, and canopy generalists are expected to be less likely to show declines than canopy specialists.

We obtained wingspan, voltinism, diapause stage, host plant specialization, and habitat affinities from LepTraits 1.0, a database of global butterfly traits (*57*). When the LepTraits 1.0 species consensus trait indicated variation, we referenced source records from North American field guides to simplify trait codes. Similarly, we extracted additional trait records from North American field guides to fill gaps from LepTraits 1.0. We defined host-plant specialists as species that feed on no more than one plant family. Ordinal designations for habitat affinity were translated to a +2 to -2 integer scale based on the strength and direction of a species' association with mesic habitats, disturbance, canopy cover, and canopy specialization ("Very strong" = 2 or -2; "weak" = 1 or -1; "both", "varies", and "no evidence" = 0) .

We conducted post-hoc analyses relating species overall rates of change to species-level traits. We used weighted linear models with species rates of change as the response variable and individual species traits as predictors in separate models, weighted by the precision matrix of species rates of change. We then evaluated the support for traits as predictors of rates of change using marginal hypothesis testing (table S6). In the case of canopy affinity, which was represented by two variables, we included both in the same model. In the cases of voltinism and overwintering stage, our model identified the traits as a significant predictor, but each trait was made up of categorical variables. We report the significance of the marginal hypothesis testing, and then present the qualitative findings of post hoc comparisons implemented using the R package emmeans (*58*). We also fit a model with all trait terms to identify the maximum possible variation in rates of change explained by traits; this model included traits that were not significant predictors of growth rate, and thus may overestimate the variation explained by traits.

Phylogenetically corrected trait association

Phylogenetic regressions are sometimes used instead of conventional regression models because species are not independent, as closely related species tend to have common traits (*59*). We repeated our trait analysis using phylogenetic regression models. Species trends and traits were mapped onto a phylogeny of North American butterflies (*60*). Of the species for which we estimated species-specific rates of change, 97% were present in this phylogeny and thus could be included in this analysis. For each trait, we used the `pgls.SEy` function (*61*) in the ‘phytools’ package (*62*) to generate both a null model with no predictors for growth rate and a model with the given trait as a predictor for growth rate. The `pgls.SEy` function allows the inclusion of estimated error in the response variable; we included the estimated error associated with species rates of change. For each trait, the null model and trait model were fit only to the data for which that trait was known. For each trait, we compared the two models using AIC; as the null model was always the better fit, we did not pursue further statistical tests (table S8).

Software

All analyses were performed in R 4.2.2 (*46*). We used the following key packages: the tidyverse suite (*63*) to process data; mgcv (*45*) to fit GAMs; msm (*48*) to implement the delta method; sf (*64*), sp (*65*), rgdal (*66*) and terra (*67*) to carry out spatial data processing; car (*50*) for marginal hypothesis testing; emmeans (*58*) for post hoc comparisons; phylolm (*68*), phytools (*62*), and ape (*69*) for phylogenetically corrected trait association analyses; tidyterra (*70*), ggplot2 (*71*) and patchwork (*72*) to generate figures. Spatial data processing implemented in the rgdal package used the open-source software GDAL (*73*).

BOX 1

To calculate trends in total butterfly abundance and the abundance for most individual species (the “species-level full model”), we estimated rates of change in each region, as well as terms to account for seasonality of activity (e.g., phenology), differences across regions, monitoring programs, sites, and differences in survey effort.

We used a negative binomial model:

$$count_{j,k,t} \sim NB(\mu_{j,k,t}, \theta)$$

in which the count (for either a species or all butterflies) at location *j* on day of year *k* and year *t* comes from a negative binomial distribution with a mean $\mu_{j,k,t}$ and an overdispersion parameter θ. We estimated the mean using a log-link function:

$$log\left(\mu_{j,k,t}\right) = f_{region}(DOY) + \beta_r \times region_j + \beta_{r,t} \times region_j \times year_t + \beta_p \times program_j + \beta_{ed} \times duration_{j,k,t} \times I_d + \beta_{pm} \times partyminutes_{j,k,t} \times I_{pm} + \epsilon_s$$

where f_{region} is the smoother for seasonality for the given region, *DOY* is the day of year (from 1 to 366), *duration* and *partyminutes* are the two possible reported measures of survey effort, I_d and I_{pm} are indicator variables to identify observations in which effort was reported as duration or party minutes respectively, and ϵ_s is a random effect of survey site. In the programming language R, we write this model as:

```
    bam(count ~ -1 + s(doy, by = region, bs = "tp", k = 10)
+ region + year:region +  program + duration:Id
+ partyminutes:Ipm + s(site, bs = "re"), family = "nb")
```

For those species with sparse data (“species-level simplified model”), we simplified the model by removing the region terms to produce a single range-wide model:

$$log\left(\mu_{j,k,t}\right) = f(DOY) + \beta_t \times year_t + \beta_p \times program_j + \beta_{ed} \times duration_{j,k,t} \times I_d + \beta_{pm} \times partyminutes_{j,k,t} \times I_{pm} + \epsilon_s.$$

In the programming language R, we write this model as:

```
    bam(count ~ s(doy, bs = "tp", k = 10) + year + program +
duration:Id + partyminutes:Ipm +
                  s(site, bs = "re"), family = "nb")
```

Figure S1. Outline of our modeling process. (**A**) For each of our models, we fit survey counts with generalized additive models (GAMs), from which we calculated regional abundance indices and regional rates of change. We then used these values to calculate annual percent change at the regional and national level and cumulative percent change. This process was simplified for species-level models when the species was present in only one region and in cases when species data were limited. (**B**) Regional abundance indices were calculated using the site-level abundance indices scaled by the area of the range in each region (species models) or the area of each region (total butterfly abundance). (**C**) We report trends from three types of models: total butterfly abundance, species-level full model (301 species), and species-level simplified model (41 species). We identified regional trends in richness using only species-level full models. We used the results from all the species-level models for post-hoc analyses to identify regional differences in trends and the associations between trends and species traits.

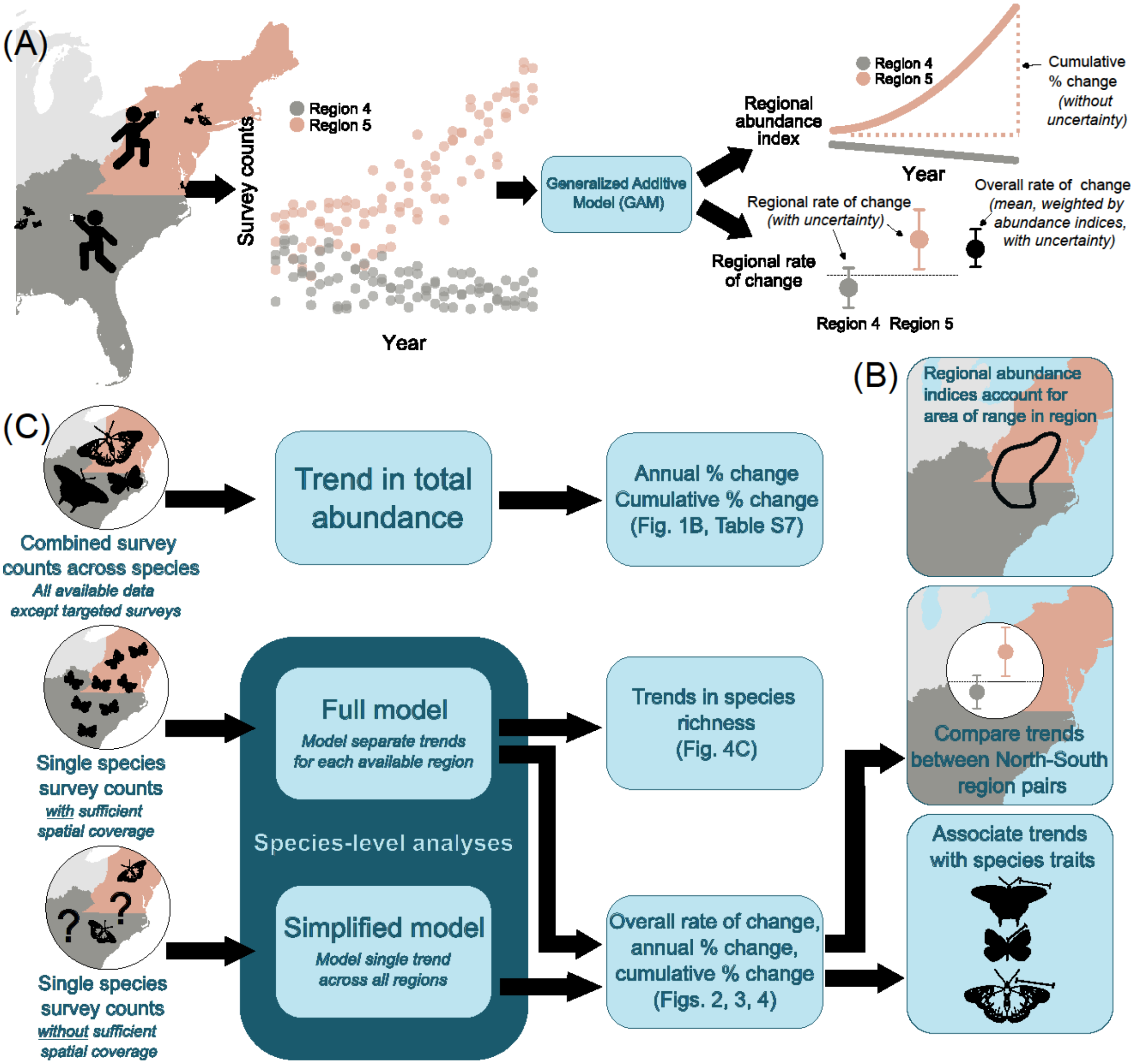

Figure S2. Phenology curves for six species. Examples of the fitted smooths included in species-level models to control for the seasonality of butterfly activity. Black curve shows the fitted model; rugplot along the x axis shows density of surveys. (**A-C**) The smooths for the three species with the highest growth rates among those species identified as increasing; (**D-F**) the smooths for the three species with the most negative growth rates among those species identified as declining. Smooth shape represents a balance between biology and statistical limitations: *Poanes melane* shows clear bimodal activity curve reflective of its two generations per year, while *Appias drusilla* data was so sparse that there was insufficient support for a less linear curve. In the limiting case, sparse data would produce a completely linear phenology term, as was used to control for phenology in (16). The smooths used to account for phenology did not vary across years.

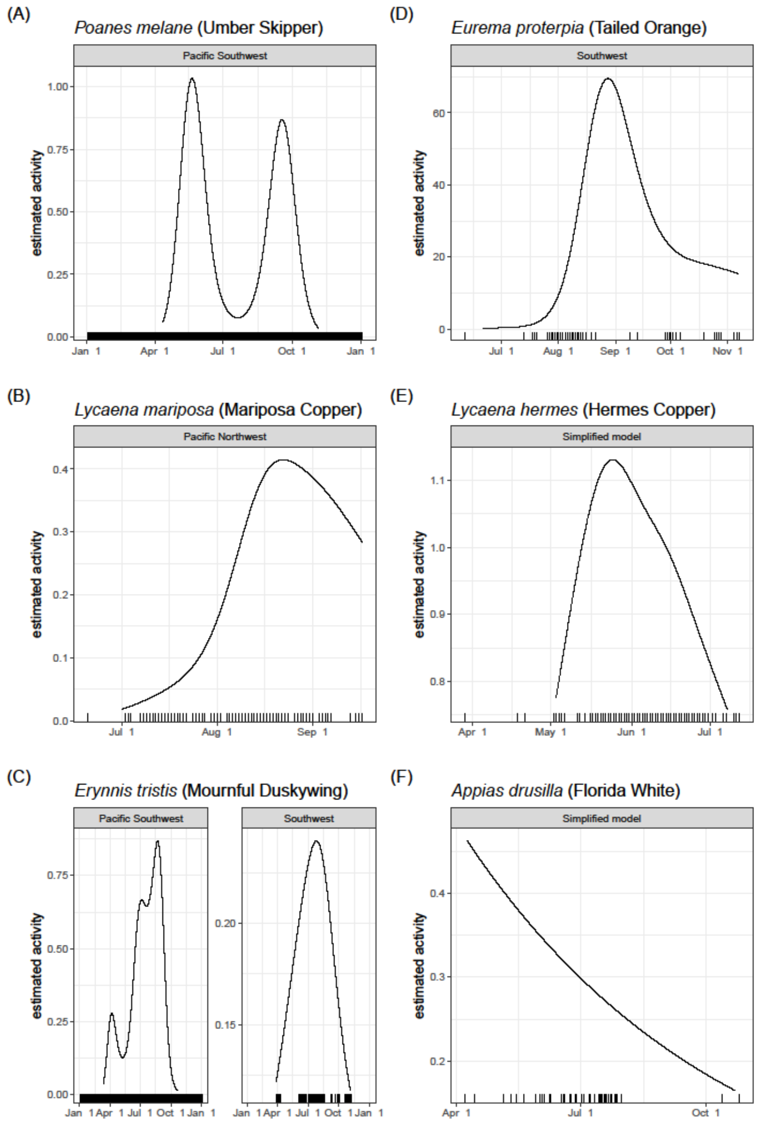

Figure S3. Butterfly monitoring varied by region and time. Bars show the number of surveys in each region (columns) and each time period (rows; 5-6 years), binned by week. Note the difference in y scales: the Midwest region is home to several of the largest monitoring programs and had substantially more surveys than any other region.

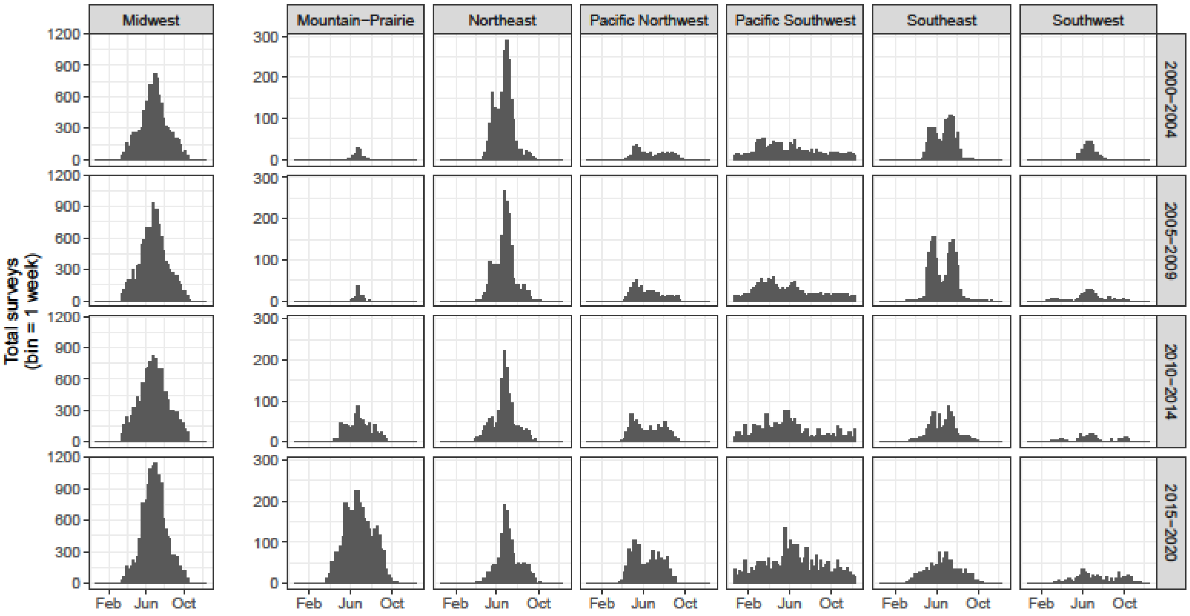

Figure S4: Common threats to butterflies have increased for the past century while the preponderance of butterfly monitoring data is available only within the past two decades. (A) Pounds of agricultural pesticide active ingredient reported in the US for 21 selected crops (*74*). (B) Temperature anomalies in the U.S., measuring the difference in degrees Fahrenheit between yearly mean temperature and the average from 1901-2000 (*75*). (C) Change in land use measured by percent of grassland lost since 1850 (*76*). (D) Number of butterfly surveys available per year from the 35 monitoring programs used in this study. Dotted line shows the year 2000, the first year of our analysis; threats in A-C had already reached elevated severity at that time.

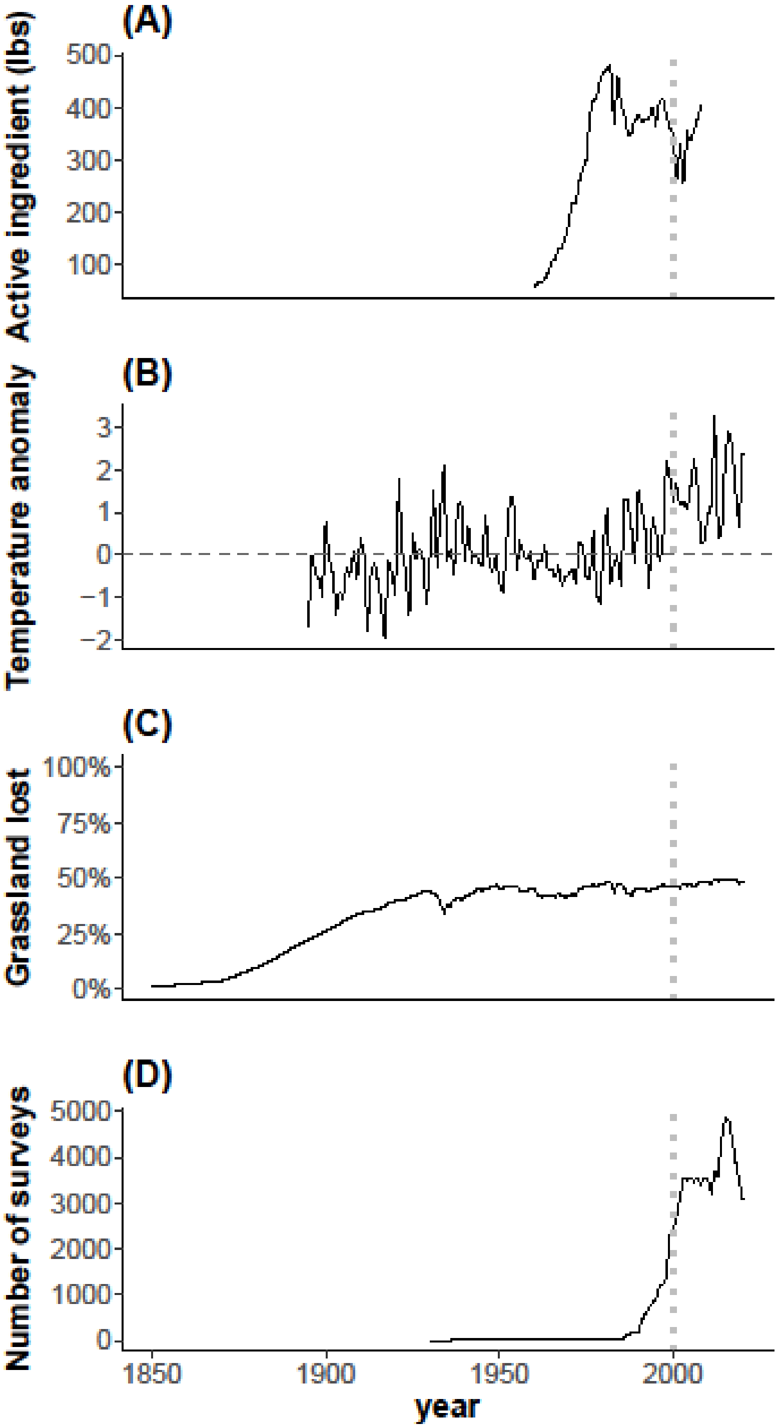

Figure S5. Histogram of species rates of changes in abundance over the study period (2000-2020), shown on a linear scale. Species with rates of change that significantly differed from zero (P < 0.05) are labeled “declining” (114 species) and “increasing” (9 species), respectively; remaining species that changed by at least +/-10% from 2000 to 2020 were labeled as “possibly declining” and “possibly increasing”. The median species declined in abundance by 41.5% across the study period. This figure is an alternative to Fig. 2, which presents the histogram on the log scale. Because declines are bounded at -100% while increases have no bounds, the log scale allows for easier interpretation of the bulk of the data.

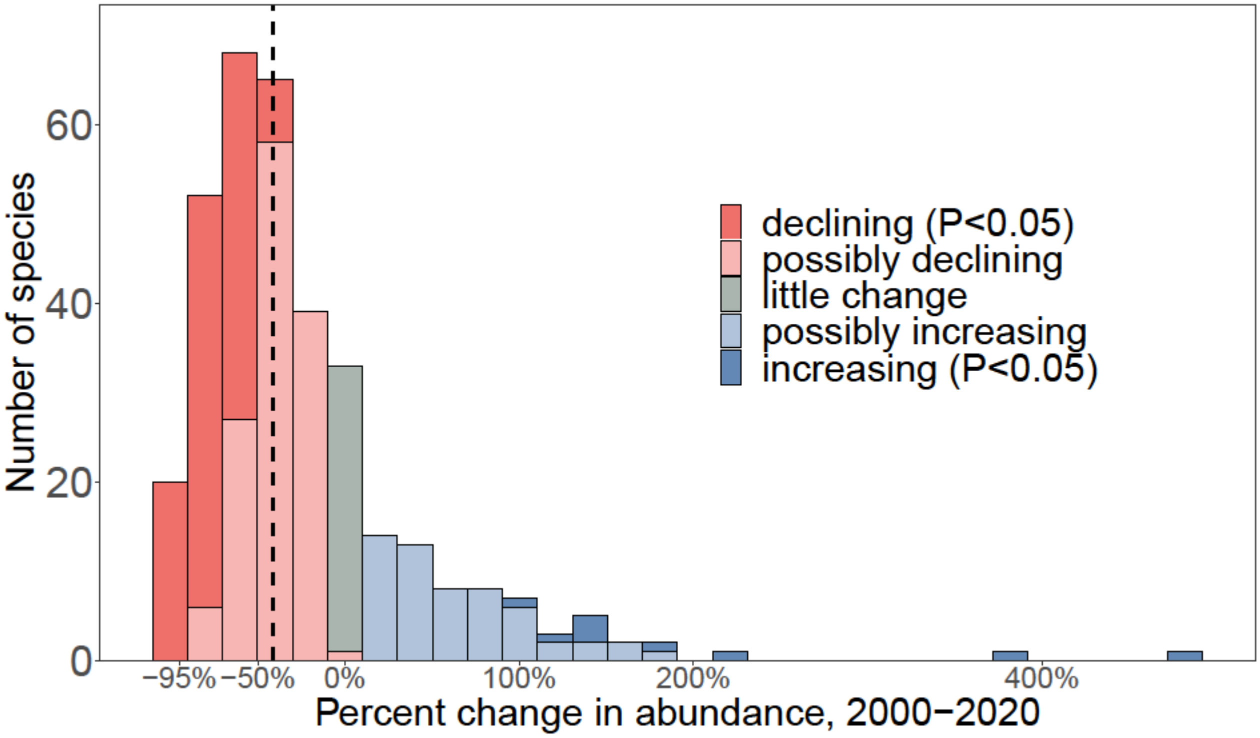

Table S1.

Summary of data sources. Detailed information on each `Monitoring program` that contributed data to our study. `Data label` denotes the label used in the data files (the `source` column) if it differs from the listed monitoring program. `Data collectors` distinguishes data collected by career scientists (“Researcher(s)”) or by volunteers (“Volunteer scientists”). `Targeted` identifies programs that recorded only individuals from one or a few target species (“yes”) versus programs that recorded all butterflies seen (“no”). In `Monitoring approach`, “repeated transect” refers both to Pollard-walks and more elaborate transect methods, such as distance sampling; for details on circle counts and field trips, see *19*. `Number surveys` lists all surveys available from the data sources in the contiguous U.S. for the study period (2000 to 2020).

Table S2.

Summary of key analysis terms. The estimated abundance indices are metrics of relative abundance for a species (or total butterfly abundance). These indices are comparable within a model (e.g., to evaluate a species’ trend), but not across models (e.g., to compare abundance between two species in a given year).

Table S3.

Data used to fit species-level models. A summary of the volume of data available for each species included in our analyses. “# surveys (direct observation)” provides number of non-implicit records used in model fitting; “# surveys (implicit zeroes)” provides the number of implicit records used, and “% implicit zeroes” gives the percent of data used that were implicit zeroes. “# sites” gives the number of unique sites in the data used, # monitoring programs gives the number of monitoring programs represented in the data used (see table S1), and “Total butterflies seen” shows the sum of counts across all observations used.

Table S4.

Trends in total butterfly abundance for each region. Comparison of trends in overall butterfly abundance as estimated in the main manuscript (“original”) to an identical model fit to data in which the most abundant species was removed from each region (“without most common”). “Total” row represents the total changes for the contiguous U.S.

Table S5.

Abundance trends for each of the 342 species that were estimated using species-level models. `Trend` is defined as in Figs. 2 and 4: species with rates of change that significantly differed from zero ($P < 0.05$) are labeled “declining” (82 species) and “increasing” (8 species), respectively; remaining species that changed by at least +/- 10% from 2000 to 2020 are labeled as “possibly declining” and “possibly increasing”. `Estimate` reports the rate of change, which is the natural log of the annual per capita growth rate (i.e., the population growth rate). `SE` is the standard error associated with the estimated rate of change. `95% CI` is the 95% confidence interval of the rate of change as calculated from the estimate and standard error. `Pval` provides the P-value associated with the rate of change, either calculated directly from the model (species-level simplified model, species-level full model for species present in a single region) or based on the estimate and standard error. `Adjusted Pval` gives P-value after correcting for the false discovery rate. `Type of model` distinguished species for which we fit separate trends for each region (“full”) and species for which we fit a single trend (“simplified”). `mgcv method` identifies the function used to fit each species. The final columns identify species that required additional data filtering for model convergence and cases when the fitted model provided a warning about the estimation of the smoothing penalty. The smoothing penalty only impacts estimated seasonality of activity (phenology) and not trend estimates.

Table S6.

List of species trends by butterfly family. Percent of species in each family that are declining, stable, and increasing. Species with rates of change that significantly differed from zero ($P < 0.05$) are labeled “declining” (81 species) and “increasing” (8 species), respectively; remaining species that changed by at least +/-10% from 2000 to 2020 were labeled as “possibly declining” and “possibly increasing”. “Little change” reports the species with estimated cumulative change between -10% and +10%. Parentheses show the number of species.

Table S7.

Summary of trait analyses. Results of marginal hypothesis testing (`Anova()` from the car package; *50*) of linear mixed models that included family as a random effect and weighted species rates of change by the inverse of the associated uncertainty. For categorical traits, we present coefficient estimates for each category calculated using the emmeans package (*58*), and list all pairwise contrasts with $P < 0.1$ calculated using the Tukey method.

Table S8.

Summary of phylogenetically-corrected trait analyses. Statistics comparing a null model that includes only phylogeny with a model including phylogeny and the specified trait, weighting species by the precision matrix of estimated rates of change. The number of species included in each pair of models differs due to available trait data. In all cases the null model had the best support, with ‘dAIC’ reporting the difference in AIC between the null and trait models.

Table S9.

Comprehensive list of species regional trends for the 319 species fit with the species-level full model. Columns are defined as in Table S5, with trend, rate of change and cumulative change now referring to estimates for individual regions rather than the overall U.S estimate.

Table S10.

Number of surveys for each study year (2000-2020). The number of surveys used, across all available monitoring programs, in the trend model to estimate changes in total butterfly abundance and to generate Figure 1B.

References and Notes

1. R. Dirzo, H. S. Young, M. Galetti, G. Ceballos, N. J. B. Isaac, B. Collen, Defaunation in the Anthropocene. *Science* **345**, 401–406 (2014). doi:10.1126/science.1251817 Medline

2. G. Ceballos, P. R. Ehrlich, A. D. Barnosky, A. García, R. M. Pringle, T. M. Palmer, Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Sci. Adv.* **1**, e1400253 (2015). doi:10.1126/sciadv.1400253 Medline

3. G. Ceballos, P. R. Ehrlich, P. H. Raven, Vertebrates on the brink as indicators of biological annihilation and the sixth mass extinction. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 13596–13602 (2020). doi:10.1073/pnas.1922686117 Medline

4. D. Goulson, The insect apocalypse, and why it matters. *Curr. Biol.* **29**, R967–R971 (2019). Medline

5. R. van Klink, D. E. Bowler, K. B. Gongalsky, A. B. Swengel, A. Gentile, J. M. Chase, Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science* **368**, 417–420 (2020). doi:10.1126/science.aax9931 Medline

6. D. L. Wagner, E. M. Grames, M. L. Forister, M. R. Berenbaum, D. Stopak, Insect decline in the Anthropocene: Death by a thousand cuts. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2023989118 (2021). doi:10.1073/pnas.2023989118 Medline

7. M. S. Warren, D. Maes, C. A. M. van Swaay, P. Goffart, H. Van Dyck, N. A. D. Bourn, I. Wynhoff, D. Hoare, S. Ellis, The decline of butterflies in Europe: Problems, significance, and possible solutions. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2002551117 (2021). doi:10.1073/pnas.2002551117 Medline

8. J. F. Tooker, K. A. Pearsons, Newer characters, same story: Neonicotinoid insecticides disrupt food webs through direct and indirect effects. *Curr. Opin. Insect Sci.* **46**, 50–56 (2021). doi:10.1016/j.cois.2021.02.013 Medline

9. E. M. Grames, G. A. Montgomery, C. Youngflesh, M. W. Tingley, C. S. Elphick, The effect of insect food availability on songbird reproductive success and chick body condition: Evidence from a

systematic review and meta-analysis. *Ecol. Lett.* **26**, 658–673 (2023). doi:10.1111/ele.14178 Medline

10. P. Cardoso, T. L. Erwin, P. A. V. Borges, T. R. New, The seven impediments in invertebrate conservation and how to overcome them. *Biol. Conserv.* **144**, 2647–2655 (2011). doi:10.1016/j.biocon.2011.07.024
11. D. Taron, L. Ries, "Butterfly monitoring for conservation" in *Butterfly Conservation in North America: Efforts to Help Save our Charismatic Microfauna*, J. C. Daniels, Ed. (Springer, 2015), pp. 35–57.
12. G. A. Breed, S. Stichter, E. E. Crone, Climate-driven changes in northeastern US butterfly communities. *Nat. Clim. Chang.* **3**, 142–145 (2013). doi:10.1038/nclimate1663
13. T. Wepprich, J. R. Adrion, L. Ries, J. Wiedmann, N. M. Haddad, Butterfly abundance declines over 20 years of systematic monitoring in Ohio, USA. *PLOS ONE* **14**, e0216270 (2019). doi:10.1371/journal.pone.0216270 Medline
14. J. P. Michielini, E. B. Dopman, E. E. Crone, Changes in flight period predict trends in abundance of Massachusetts butterflies. *Ecol. Lett.* **24**, 249–257 (2021). doi:10.1111/ele.13637 Medline
15. N. B. Kucherov, E. S. Minor, P. P. Johnson, D. Taron, K. C. Matteson, Butterfly declines in protected areas of Illinois: Assessing the influence of two decades of climate and landscape change. *PLOS ONE* **16**, e0257889 (2021). doi:10.1371/journal.pone.0257889 Medline
16. M. L. Forister, C. A. Halsch, C. C. Nice, J. A. Fordyce, T. E. Dilts, J. C. Oliver, K. L. Prudic, A. M. Shapiro, J. K. Wilson, J. Glassberg, Fewer butterflies seen by community scientists across the warming and drying landscapes of the American West. *Science* **371**, 1042–1045 (2021). doi:10.1126/science.abe5585 Medline
17. M. S. Crossley, O. M. Smith, L. L. Berry, R. Phillips-Cosio, J. Glassberg, K. M. Holman, J. G. Holmquest, A. R. Meier, S. A. Varriano, M. R. McClung, M. D. Moran, W. E. Snyder, Recent climate change is creating hotspots of butterfly increase and decline across North America. *Glob. Change Biol.* **27**, 2702–2714 (2021). doi:10.1111/gcb.15582 Medline
18. P. Cardoso, S. R. Leather, Predicting a global insect apocalypse. *Insect Conserv. Divers.* **12**, 263–267 (2019). doi:10.1111/icad.12367
19. Materials and methods are available as supplementary materials.

20. P. Rothery, D. B. Roy, Application of generalized additive models to butterfly transect count data. *J. Appl. Stat.* **28**, 897–909 (2001). doi:10.1080/02664760120074979

21. E. M. Grames, M. Flynn, C. B. Edwards, W. E. Thogmartin, J. Glassberg, M. L. Forister, Data from: Integrated range maps for North American butterflies derived from expert opinion and relative predicted habitat suitability, Figshare (2024). doi:10.6084/m9.figshare.22747928

22. V. Shirey, N. Neupane, R. Guralnick, L. Ries, Rising minimum temperatures contribute to 50 years of occupancy decline among cold-adapted Arctic and boreal butterflies in North America. *Glob. Change Biol.* **30**, e17205 (2024). doi:10.1111/gcb.17205 Medline

23. N. S. Cobb, L. F. Gall, J. M. Zaspel, N. J. Dowdy, L. M. McCabe, A. Y. Kawahara, Assessment of North American arthropod collections: Prospects and challenges for addressing biodiversity research. *PeerJ* **7**, e8086 (2019). doi:10.7717/peerj.8086 Medline

24. B. Van Deynze, S. M. Swinton, D. A. Hennessy, N. M. Haddad, L. Ries, Insecticides, more than herbicides, land use, and climate, are associated with declines in butterfly species richness and abundance in the American Midwest. *PLOS ONE* **19**, e0304319 (2024). doi:10.1371/journal.pone.0304319 Medline

25. M. L. Forister, B. Cousens, J. G. Harrison, K. Anderson, J. H. Thorne, D. Waetjen, C. C. Nice, M. De Parsia, M. L. Hladik, R. Meese, H. van Vliet, A. M. Shapiro, Increasing neonicotinoid use and the declining butterfly fauna of lowland California. *Biol. Lett.* **12**, 20160475 (2016). doi:10.1098/rsbl.2016.0475 Medline

26. N. L. Haan, B. G. Iuliano, C. Gratton, D. A. Landis, "Chapter Five - Designing agricultural landscapes for arthropod-based ecosystem services in North America" in *The Future of Agricultural Landscapes, Part II*, vol. 64 of Advances in Ecological Research Series, D. A. Bohan, A. J. Vanbergen, Eds. (Academic Press, 2021), pp. 191–250.

27. J. R. Pecenka, L. L. Ingwell, R. E. Foster, C. H. Krupke, I. Kaplan, IPM reduces insecticide applications by 95% while maintaining or enhancing crop yields through wild pollinator conservation. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2108429118 (2021). doi:10.1073/pnas.2108429118 Medline

28. R. E. Bonoan, E. E. Crone, C. B. Edwards, C. B. Schultz, Changes in phenology and abundance of an at-risk butterfly. *J. Insect Conserv.* **25**, 499–510 (2021). doi:10.1007/s10841-021-00318-7

29. C. B. Edwards, C. B. Schultz, S. P. Campbell, C. E. Fallon, E. H. Henry, K. C. King, M. Linders, T. Longcore, D. A. Marschalek, D. Sinclair, A. B. Swengel, S. R. Swengel, D. J. Taron, T. Wepprich, E. E. Crone, Phenological constancy and management interventions predict population trends in at-risk butterflies. *J. Appl. Ecol.* **61**, 2455–2469 (2024). doi:10.1111/1365-2664.14735

30. C. Haaland, R. E. Naisbit, L. F. Bersier, Sown wildflower strips for insect conservation: A review. *Insect Conserv. Divers.* **4**, 60–80 (2011). doi:10.1111/j.1752-4598.2010.00098.x

31. A. M. Franco, J. K. Hill, C. Kitschke, Y. C. Collingham, D. B. Roy, R. I. Fox, B. R. Huntley, C. D. Thomas, Impacts of climate warming and habitat loss on extinctions at species' low-latitude

range boundaries. *Glob. Change Biol.* **12**, 1545–1553 (2006). doi:10.1111/j.1365-2486.2006.01180.x

32. E. R. Zylstra, L. Ries, N. Neupane, S. P. Saunders, M. I. Ramírez, E. Rendón-Salinas, K. S. Oberhauser, M. T. Farr, E. F. Zipkin, Changes in climate drive recent monarch butterfly dynamics. *Nat. Ecol. Evol.* **5**, 1441–1452 (2021). doi:10.1038/s41559-021-01504-1 Medline
33. R. S. Vose, S. Applequist, M. Squires, I. Durre, M. J. Menne, C. N. Williams Jr., C. Fenimore, K. Gleason, D. Arndt, Improved historical temperature and precipitation time series for US climate divisions. *J. Appl. Meteorol. Climatol.* **53**, 1232–1251 (2014). doi:10.1175/JAMC-D-13-0248.1
34. J. J. Hellmann, R. Grundel, C. Hoving, G. W. Schuurman, A call to insect scientists: Challenges and opportunities of managing insect communities under climate change. *Curr. Opin. Insect Sci.* **17**, 92–97 (2016). doi:10.1016/j.cois.2016.08.005 Medline
35. A. R. Whiteley, S. W. Fitzpatrick, W. C. Funk, D. A. Tallmon, Genetic rescue to the rescue. *Trends Ecol. Evol.* **30**, 42–49 (2015). doi:10.1016/j.tree.2014.10.009 Medline
36. J. J. Lawler, D. D. Ackerly, C. M. Albano, M. G. Anderson, S. Z. Dobrowski, J. L. Gill, N. E. Heller, R. L. Pressey, E. W. Sanderson, S. B. Weiss, The theory behind, and the challenges of, conserving nature's stage in a time of rapid change. *Conserv. Biol.* **29**, 618–629 (2015). doi:10.1111/cobi.12505 Medline
37. C. B. Edwards, E. F. Zipkin, E. H. Henry, N. M. Haddad, M. L. Forister, K. J. Burls, S. P. Campbell, E. E. Crone, J. Diffendorfer, M. R. Douglas, R. G. Drum, C. E. Fallon, J. Glassberg, E. M. Grames, R. Hatfield, S. Hershcovich, S. Hoffman-Black, E. A. Larsen, W. Leuenberger, M. J. Linders, T. Longcore, D. A. Marschale, J. Michielini, N. Neupane, L. Ries, A. Shapiro, A. B. Swengel, S. R. Swengel, B. Van Deynze, J. Wiedmann, W. E. Thogmartin, C. B. Schultz, Data and code for "Butterflies are declining rapidly in the United States during the 21st century", Version 1, Figshare (2024). doi:10.6084/m9.figshare.27934629
38. E. H. Henry, C. B. Edwards, V. Shirey, J. S. Pippen, D. Waetjen, M. L. Forister, E. A. Larsen, C. B. Schultz, J. Michielini, N. Brockman, K. J. Burls, R. G. Drum, M. Gatch, J. Glassberg, N. Hamlett, S. Hershcovich, C. Le, S. McGaffin, J. Meilinger, L. Richter, R. Rochefort, C. Schelz, A. M. Shapiro, K. Sullivan, D. J. Taron, W. E. Thogmartin, A. Walker, A. Westphal, J. Wiedmann, I. U. Wilcockson, J. Zaspel, L. Ries, Data for "Twenty years (2000 - 2020) of

butterfly monitoring data across the contiguous United States", Version 1, Figshare (2024). doi:10.6084/m9.figshare.27934602

39. E. Pollard, A method of assessing the abundance of butterflies in Monks Wood National Nature Reserve in 1973. *Entomol. Gaz.* **26**, 79–88 (1975).

40. S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, L. Thomas, *Introduction to Distance Sampling* (Oxford Univ. Press, 2001).

41. A. M. Shapiro, Art Shapiro's Butterfly Site (2020); https://butterfly.ucdavis.edu/.

42. North American Butterfly Association, Checklist of North American Butterflies (2016); https://naba.org/butterfly-names-checklist/.

43. J. P. Pelham, A Catalogue of the Butterflies of the United States and Canada (2023); https://butterfliesofamerica.com/US-Can-Cat.htm.

44. J. Glassberg, *A Swift Guide to Butterflies of North America* (Princeton Univ. Press, ed. 2, 2017).

45. S. N. Wood, N. Pya, B. Säfken, Smoothing parameter and model selection for general smooth models (with discussion). *J. Am. Stat. Assoc.* **111**, 1548–1575 (2016). doi:10.1080/01621459.2016.1180986

46. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2023).

47. E. J. Pedersen, D. L. Miller, G. L. Simpson, N. Ross, Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ* **7**, e6876 (2019). doi:10.7717/peerj.6876 Medline

48. H. C. Jackson, Multi-state models for panel data: The msm Package for R. *J. Stat. Softw.* **38**, 1–29 (2011). doi:10.18637/jss.v038.i08

49. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* **57**, 289–300 (1995). doi:10.1111/j.2517-6161.1995.tb02031.x

50. J. Fox, S. Weisberg, *An R Companion to Applied Regression* (Sage, ed. 3, 2019).

51. M. Kuussaari, M. Saarinen, E.-L. Korpela, J. Pöyry, T. Hyvönen, Higher mobility of butterflies than moths connected to habitat suitability and body size in a release experiment. *Ecol. Evol.* **4**, 3800–3811 (2014). doi:10.1002/ece3.1187 Medline

52. C. B. Schultz, B. G. Pe'er, C. Damiani, L. Brown, E. E. Crone, Does movement behaviour predict population densities? A test with 25 butterfly species. *J. Anim. Ecol.* **86**, 384–393 (2017). doi:10.1111/1365-2656.12609 Medline

53. C. E. Dumoulin, P. R. Armsworth, Environmental stochasticity increases extinction risk to a greater degree in pollination specialists than in generalists. *Oikos* **2022**, e09214 (2022). doi:10.1111/oik.09214

54. M. S. Warren, J. K. Hill, J. A. Thomas, J. Asher, R. Fox, B. Huntley, D. B. Roy, M. G. Telfer, S. Jeffcoate, P. Harding, G. Jeffcoate, S. G. Willis, J. N. Greatorex-Davies, D. Moss, C. D. Thomas,

Rapid responses of British butterflies to opposing forces of climate and habitat change. *Nature* **414**, 65–69 (2001). doi:10.1038/35102054 Medline

55. E. A. Larsen, M. W. Belitz, G. J. Di Cecco, J. Glassberg, A. H. Hurlbert, L. Ries, R. P. Guralnick, Overwintering strategy regulates phenological sensitivity and consequences for ecological services in a clade of temperate North American insects. *Funct. Ecol.* **38**, 1075–1088 (2024). doi:10.1111/1365-2435.14543

56. M. L. Forister, E. M. Grames, C. A. Halsch, K. J. Burls, C. F. Carroll, K. L. Bell, J. P. Jahner, T. A. Bradford, J. Zhang, Q. Cong, N. V. Grishin, J. Glassberg, A. M. Shapiro, T. V. Riecke, Assessing risk for butterflies in the context of climate change, demographic uncertainty, and heterogeneous data sources. *Ecol. Monogr.* **93**, e1584 (2023). doi:10.1002/ecm.1584

57. V. Shirey, E. Larsen, A. Doherty, C. A. Kim, F. T. Al-Sulaiman, J. D. Hinolan, M. G. A. Itliong, M. A. K. Naïve, M. Ku, M. Belitz, G. Jeschke, V. Barve, G. Lamas, A. Y. Kawahara, R. Guralnick, N. E. Pierce, D. J. Lohman, L. Ries, LepTraits 1.0 A globally comprehensive dataset of butterfly traits. *Sci. Data* **9**, 382 (2022). doi:10.1038/s41597-022-01473-5 Medline

58. R. Lenth, emmeans: Estimated Marginal Means, aka Least-Squares Means, R package version 1.8.9 (2023); https://CRAN.R-project.org/package=emmeans.

59. A. Grafen, The phylogenetic regression. *Philos. Trans. R. Soc. London Ser.* **326**, 119–157 (1989). doi:10.1098/rstb.1989.0106 Medline

60. C. Earl, M. W. Belitz, S. W. Laffan, V. Barve, N. Barve, D. E. Soltis, J. M. Allen, P. S. Soltis, B. D. Mishler, A. Y. Kawahara, R. Guralnick, Spatial phylogenetics of butterflies in relation to environmental drivers and angiosperm diversity across North America [Dataset], Dryad (2021). doi:10.5061/dryad.00000002j

61. A. R. Ives, P. E. Midford, T. Garland Jr., Within-species variation and measurement error in phylogenetic comparative methods. *Syst. Biol.* **56**, 252–270 (2007). doi:10.1080/10635150701313830 Medline

62. L. J. Revell, phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012). doi:10.1111/j.2041-210X.2011.00169.x

63. H. Wickham, M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H.

Yutani, Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686 (2019). doi:10.21105/joss.01686

64. E. Pebesma, R. Bivand, *Spatial Data Science: With Applications in R* (Chapman and Hall, 2023).
65. R. Bivand, E. Pebesma, V. Gomez-Rubio, *Applied Spatial Data Analysis with R* (Springer, ed. 2, 2013).
66. R. Bivand, T. Keitt, B. Rowlingson, rgdal: Bindings for the 'Geospatial' Data Abstraction Library, R package version 1.6-7 (2023); https://cran.r-project.org/src/contrib/Archive/rgdal/.
67. R. Hijmans, terra: Spatial Data Analysis, R package version 1.7-55 (2023); https://CRAN.R-project.org/package=terra.
68. L. s. T. Ho, C. Ané, A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst. Biol.* **63**, 397–408 (2014). doi:10.1093/sysbio/syu005 Medline
69. E. Paradis, K. Schliep, ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019). doi:10.1093/bioinformatics/bty633 Medline
70. D. Hernangómez, tidyterra: tidyverse Methods and ggplot2 Helpers for terra Objects, Zenodo (2023). doi:10.5281/zenodo.6572471
71. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
72. T. Pedersen, patchwork: The Composer of Plots. R package version 1.1.3 (2023); https://CRAN.R-project.org/package=patchwork.
73. GDAL/OGR contributors, GDAL/OGR Geospatial Data Abstraction software Library. Open Source Geospatial Foundation, Zenodo (2023). doi:10.5281/zenodo.5884351
74. J. Fernandez-Cornejo, R. F. Nehring, C. Osteen, S. Wechsler, A. Martin, A. Vialou, Pesticide Use in U.S. Agriculture: 21 Selected Crops, 1960-2008, USDA-ERS Economic Information Bulletin Number 124. SSRN 2502986 (2015). doi:10.2139/ssrn.2502986
75. NOAA National Centers for Environmental information, Climate at a Glance: National Time Series, published September 2023; https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/national/time-series [accessed 9 November 2023].
76. X. Li, H. Tian, C. Lu, S. Pan, Four-century history of land transformation by humans in the United States (1630-2020): Annual and 1km grid data for the HIStory of LAND changes (HISLAND-US). *Earth Syst. Sci. Data* **15**, 1005–1035 (2023). doi:10.5194/essd-15-1005-2023