Title: Studies of insect temporal trends must account for the complex sampling histories inherent to many long-term monitoring efforts

This comment on Crossley *et al.* 2020. No net insect abundance and diversity declines across US Long Term Ecological Research sites. *Nature Ecology & Evolution* 4: 1368-1376. DOI: 10.1038/s41559-020-1269-4 ***has been published in Nature Ecology and Evolution as a Matters Arising, where it can be cited as:***

Welti, E.A.R., A. Joern, A.M. Ellison, D.C. Lightfoot, S. Record, N. Rodenhouse, E.H. Stanley, & M. Kaspari. 2021. Studies of insect temporal trends must account for the complex sampling histories inherent to many long-term monitoring efforts. Nature Ecology and Evolution. DOI: 10.1038/s41559-021-01424-0

Link to published version:
https://www.nature.com/articles/s41559-021-01424-0

Response by Crossley *et al.*:
https://www.nature.com/articles/s41559-021-01429-9

Authors

Ellen A. R. Welti[1*], Anthony Joern[2], Aaron M. Ellison[3], David C. Lightfoot[4], Sydne Record[5], Nicholas Rodenhouse[6], Emily H. Stanley[7], Michael Kaspari[8]

Author Affliations
[1] Senckenberg Research Institute and Natural History Museum Frankfurt, Gelnhausen, Germany
[2] Division of Biology, Kansas State University, Manhattan, KS, USA
[3] Harvard Forest, Harvard University, Petersham, MA, USA
[4] Museum of Southwestern Biology, Biology Department, University of New Mexico, Albuquerque, NM, USA
[5] Department of Biology, Bryn Mawr College, Bryn Mawr, PA, USA
[6] Biological Sciences, Wellesley College, Wellesley, MA, USA
[7] Center for Limnology, University of Wisconsin-Madison, Madison, WI, USA
[8] Geographical Ecology Group, Department of Biology, University of Oklahoma, Norman, OK, USA
[*]corresponding author; email: ellen.welti@senckenberg.de

Abstract

In a recently published study, Crossley et al. (2020, Nature Ecology & Evolution, "No net insect abundance and diversity declines across US Long Term Ecological Research sites")[1] examine patterns of change in insect abundance and diversity across US Long-Term Ecological Research (LTER) sites, concluding "a lack of overall increase or decline". This is notable if true, given mixed conclusions in the literature regarding the nature and ubiquity of insect declines across regions and insect taxonomic groups[2–6]. The data analyzed, downloaded from and collected by US LTER sites, represent unique time series of arthropod abundances. These long-term datasets often provide critical insights, capturing both steady changes and responses to sudden unpredictable events. However, a number of the included datasets are not suitable for estimating long-term observational trends because they come from experiments or have methodological inconsistencies. Additionally, long-term ecological datasets are rarely uniform in sampling effort across their full duration as a result of the changing goals and abilities of a research site to collect data[7]. We suggest that Crossley et al.'s results rely upon a key, but flawed, assumption, that sampling was collected "in a consistent way over time within each dataset". We document problems with data use prior to statistical analyses from eight LTER sites due to datasets not being suitable for long-term trend estimation and not accounting for sampling variation, using the Konza Prairie (KNZ) grasshopper dataset (CGR022) as an example.

Unsuitable datasets to estimate long-term observational trends

Several of the LTER datasets included in Crossley et al. (2020) either document experiments which have confounding treatment effects or they are too variable in sampling methods to allow for comparison of samples across time. Additionally, in one case, Lepidopteran outbreak dynamics with long intervals (10-13 years) at Hubbard Brook limit power to detect meaningful trends without extremely long-term data[8]. Datasets from Cedar Creek include arthropods collected in plots with nitrogen addition, herbivore exclosures, and manipulated plant diversity. All three of the datasets from Harvard Forest included in Crossley at al.'s analysis have large methodological inconsistencies over time and one dataset documents ants collected in a canopy manipulation experiment, including one treatment where trees were girdled to simulate hemlock woolly adelgid (*Adelges tsugae*) infestation of the hemlock trees years prior to the arrival of the invasive insect to the area. One dataset from North Temperate Lakes documents the responses of two crayfish species in a lake where one species was being experimentally removed. With a few exceptions for partial components of these datasets (e.g. control plots in the arce153 Cedar Creek dataset), these data are inappropriate for estimation of long-term observational species trends.

Not accounting for sampling variation: Konza grasshoppers as a case in point

The KNZ CGR022 dataset documents grasshopper species abundances on 15 KNZ watersheds, and spans 1982-present (up to 2015 included in Crossley et al. 2020). Crossley et al. analyze time series of individual species from each dataset (the number of "Time trends" in their Table 1). However, regardless of variant sampling effort, they regularly sum all individuals within LTER datasets to yield a single value of abundance for a given species and year. This is the case for KNZ grasshoppers, and most other included datasets (number of "Sites" in their Table 1). Importantly, sampling effort at KNZ and other LTER sites was not constant. At KNZ, variation occurred in the number of samples per watershed and the number of watersheds in which grasshoppers were collected per year (Fig. 1). Most notably, 6 bison-grazed watersheds were added to KNZ sampling in 2002. Changes in sample numbers over time are documented in the online metadata (http://lter.konza.ksu.edu/content/cgr02-sweep-sampling-grasshoppers-konza-prairie-lter-watersheds).

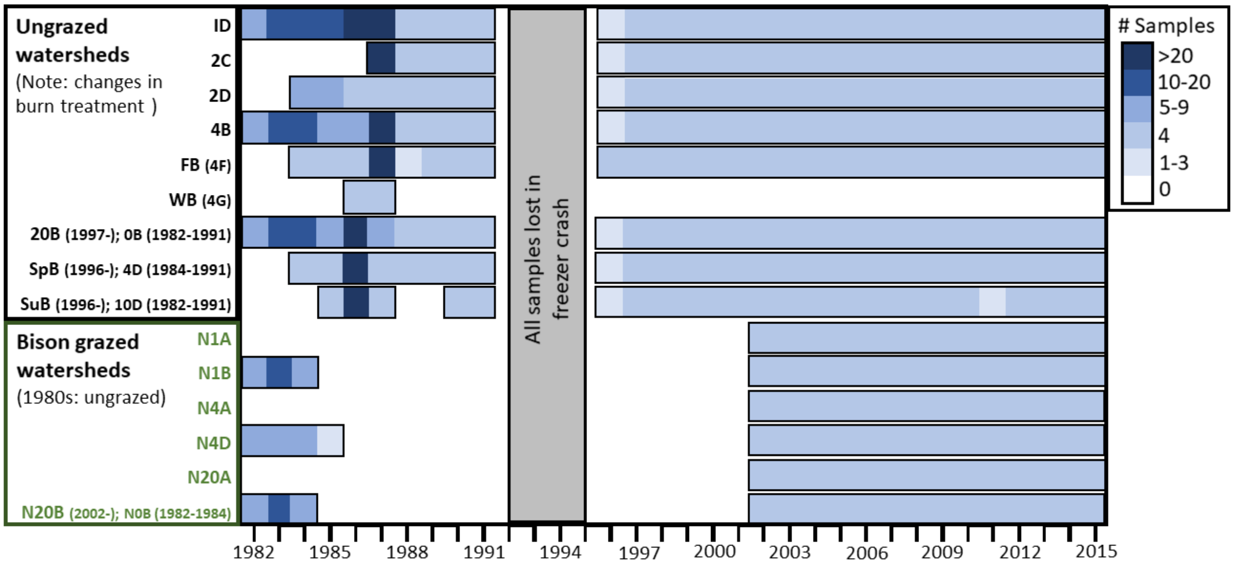


Figure 1. The complex history of sampling of the KNZ grasshopper dataset. The KNZ grasshopper dataset (CGR022) exhibits high variance both in number of watersheds sampled per year (number of bars per year) and number of samples collected within each watershed each year (depicted in color). Other complexities include the tragic loss of four years (1992-1995) of sampling due to a freezer crash, changes in sampling month, changes in watershed burn frequencies, and the reintroduction of bison in the 1990s to six of the later-sampled watersheds.

Accounting for sampling effort and data structure matters (*see also* Supplementary Information: Supplementary Fig. 1). At KNZ, bison-grazed watersheds support higher grasshopper abundances and species richness[9,10]. In a recent analysis using the CGR022 dataset, to account for this change in sampling effort, data were combined only from watersheds collected in the same years (e.g. by splitting samples from grazed watersheds into a separate time series) and abundances within each watershed and year were divided by the number of samples. Analysis of the data structured in this way showed a >2% annual decline in grasshopper abundance, with only one common species increasing[11]. Crossley et al., in contrast report most grasshopper species increased in abundance from 1982-2015. The authors of Crossley et al. (2020) note the discrepancy with both this study[11] and another[3], and suggest it is "driven by falling numbers of just two once-dominant species… whereas many other formerly rare species have become more abundant and both evenness and species richness have increased". However, we believe the discrepancy arises because Crossley et al. did not account for variable sampling effort, including KNZ's incorporation of additional, more diverse grazed habitats midway in the time series. Similar errors, where data structure was not accounted for, are evident in 17 of the 19 datasets which we examined and were included in Crossley et al. (2020)'s results.

Conclusion

We have thus far been able to confirm issues with data from 8 of the 13 LTER sites (comprising 60% of Table 1's "Time trends") included in Crossley et al. (2020). We note that this is not a comprehensive assessment, as we have only included errors from datasets of which either we ourselves are the PIs or we have been able to confirm with the corresponding LTER PIs and information managers. The eight sites are: Baltimore, Cedar Creek, Central Arizona-Phoenix, Harvard Forest, Hubbard Brook, Konza Prairie, North Temperate Lakes, and Sevilleta. We provide details on dataset unsuitability, mistakes in not accounting for sampling effort, and several coding errors in the Supplementary Information.

Given these mistakes, we urge skepticism regarding Crossley et al. (2020)'s general conclusion of no net decline in insect abundances at US LTER sites in recent decades. Although their goal is laudable, both the use of unsuitable datasets and not taking sampling effort into account generate erroneous estimates of population change. Recently, a study reporting widespread collapse of rainforest insect populations at the LTER site Luquillo necessitated a similar correction[5]. We echo those authors, when they suggest that scientists can avoid errors by reading corresponding metadata. Contacting in advance (or even including as authors) the data providers/field biologists are additionally good practices that ensure appropriate use of the data. Like the ecology they document, it is important to take into account that long-term monitoring efforts by LTERs and similar institutions are themselves complex and full of history.

Author Contributions

E.A.R.W., S. R., A.J., and M.K. conceived the idea for the paper. E.A.R.W. wrote the first draft. A.M.E., D.C.L., S.R., N.R., and E.H.S. identified further errors in the Crossley et al. online data. All authors significantly contributed to revisions.

Acknowledgements

Shannon LaDeau, Stevan Earl, Susan Barrott, Elizabeth Borer, and Steve Pennings aided in identifying errors in Crossley et al.'s online data. Karl Roeder provided comments on an early draft of this manuscript. Jeff Taylor aided in identifying changes in Konza watershed names. We thank all LTER information managers and PIs who help keep online metadata updated. We thank Pam Montz, John Haarstad, Amanda Kuhl, Matthew Ayres, Richard Holmes, and the numerous others who did the hard work to generate these long-term datasets. NSF DEB-1556280 to MK and Konza Prairie LTER NSF DEB-1440484 supported this work.

Competing Interests

The authors declare no competing interests.

Data Availability

KNZ grasshopper abundance data are available from the Long-Term Ecological Research Data Portal (https://doi.org/10.6073/pasta/7b2259dcb0e499447e0e11dfb562dc2f). Citations for the additionally described LTER datasets are provided in the Supplementary Information.

References

1. Crossley, M. S. *et al.* No net insect abundance and diversity declines across US Long Term Ecological Research sites. *Nature Ecology & Evolution* (2020) doi:10.1038/s41559-020-1269-4.
2. Hallmann, C. A. *et al.* More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS One* **12**, e0185809 (2017).

3. van Klink, R. *et al.* Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science* **368**, 417–420 (2020).
4. Lister, B. C. & Garcia, A. Climate-driven declines in arthropod abundance restructure a rainforest food web. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E10397–E10406 (2018).
5. Willig, M. *et al.* Populations are not declining and food webs are not collapsing at the Luquillo Experimental Forest. *Proceedings of the National Academy of Sciences* **116**, 12143–12144 (2019).
6. Pilotto, F. *et al.* Meta-analysis of multidecadal biodiversity trends in Europe. *Nature Communications* **11**, 3486 (2020).
7. Didham, R. K. *et al.* Interpreting insect declines: seven challenges and a way forward. *Insect Conservation and Diversity* **13**, 103–114 (2020).
8. Martinat, P. J. & Allen, D. C. Saddled prominent outbreaks in North America. *Northern Journal of Applied Forestry* **5**, 88–91 (1988).
9. Welti, E. A. R. *et al.* Fire, grazing and climate shape plant-grasshopper interactions in a tallgrass prairie. *Funct. Ecol.* **33**, 735–745 (2019).
10. Bruckerhoff, L. A. *et al.* Harmony on the prairie? Grassland plant and animal community responses to variation in climate across land-use gradients. *Ecology* **101**, e02986 (2020).
11. Welti, E. A. R., Roeder, K. A., de Beurs, K. M., Joern, A. & Kaspari, M. Nutrient dilution and climate cycles underlie declines in a dominant insect herbivore. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 7271–7275 (2020).

Supplementary Information for Matters Arising:

Studies of insect temporal trends must account for the complex sampling histories inherent to many long-term monitoring efforts

Authors: Ellen A. R. Welti, Anthony Joern, Aaron M. Ellison, David C. Lightfoot, Sydne Record, Nicholas Rodenhouse, Emily H. Stanley, Michael Kaspari

LTER data use policies and the importance of metadata

The National Science Foundation funded Long-Term Ecological Research Network (LTER) network data access and user policies are available at: https://lternet.edu/data-access-policy/. While the LTER network strives to make research data publicly available, LTER also urges users of LTER datasets to contact the PIs of datasets with questions about methodology, and encourages data users to collaborate with the data authors. We ask readers to read metadata and communicate (or even collaborate) with the PIs of publically available datasets that you intend to use for publication. Additionally, we acknowledge that data comprehension is a two-way street and urge data providers to include comprehensive, clear, and updated metadata when publishing their data. Following these guidelines improves our ability to conduct good solid science.

Description of data use errors in Crossley et al. (2020)[1]

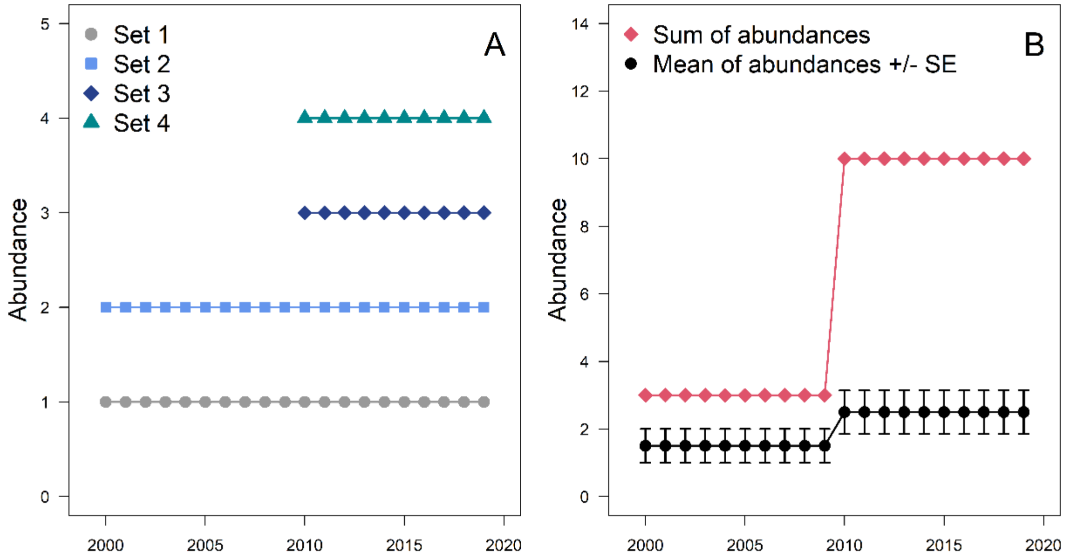
We have two major concerns regarding data use in Crossley et al. (2020). The first error is the use of datasets or parts of datasets not suitable for addressing the question of how arthropod species are changing over time.

The second error we have noted in Crossley et al. (2020) is the use raw annual sums of individuals for entire LTER datasets, which, combined with variation in sampling effort and location, produced unreliable estimates of arthropod temporal trends (Fig S1, summing example). It is evident that Crossley et al. (2020)'s analysis did not account for sampling variation because:
1) they state that they considered all included datasets to have invariant sampling effort,
2) complex datasets were considered one time series in their analyses, and
3) for many datasets, the sum of all species abundances in their online data (https://datadryad.org/stash/dataset/doi:10.5061/dryad.cc2fqz645) equals the total individuals collected within entire LTER datasets suggesting no division by the sample number.

We note that Crossley et al. did consider quantifying sampling effort as they include a column in their online data to tabulate the number of observations (called "n.obs"). However, the corresponding author of Crossley et al. (2020), Dr. Michael Crossley, informed us that n.obs was never used in their analyses. We further note that even if abundances had been divided by n.obs, this may not appropriately account for changes in sampling effort/location because:
1) based on Crossley et al. (2020)'s R code, n.obs does not always capture sample observations correctly (e.g. watershed is not included in the calculation of n.obs for the Konza grasshopper dataset and n.obs is incorrectly listed as "1" for all rows of Crossley et al.'s online data for both Central Arizona-Phoenix pitfall datasets, the Cedar Creek grasshopper dataset, the Hubbard

Brook White Mountains Region caterpillar dataset, the Sevilleta grasshopper dataset, and the Sevilleta pitfall dataset), and

2) if changes in sampling location correspond to the gain/loss of subsites/sampling times (e.g. seasons) that support different species and abundances, and the subsites/ sampling times vary in years sampled, it is not appropriate to average samples (Fig S1, averaging example).



Supplementary Figure 1. Example of how errors in trend estimation can arise when not accounting for sampling effort and when combining datasets covering different temporal periods. This example uses four sets of time series which can represent either different subsites, different sampling times within the year, or different sampling methods. The slope of abundances over time for all sets = 0; however, abundances vary across sets with the two sets in which sampling began only in 2010 having higher abundances than the two sets sampled across the full sampling interval of 2000-2019 (A). Summing does not account for variation in sampling effort over time (adding more sets) while both summing and averaging do not account for combining datasets from different temporal periods. ***Both summing and averaging abundances across the four sets creates bias prior to use of any statistical approach***, in this example resulting in artificially positive trends (B). Correct approaches include 1) estimating trends for each of the four sets separately, 2) combining only sets sampled for the same temporal periods, or 3) excluding sets so that remaining, analyzed sets cover the same temporal periods (e.g. excluding sets 3 & 4).

We document specific data use errors in Crossley et al. (2020) below. Where the assumption of invariant sampling error occurred, we provide the raw numbers of individuals from each dataset to allow others to check our work. We include information only where either we ourselves are the PIs of these datasets or we have been able to confirm errors with PIs and information managers from corresponding LTERs. LTER sites are listed in alphabetical order and include Baltimore (pg. 3), Cedar Creek (pg. 4), Central Arizona-Phoenix (pg. 5), Harvard Forest (pg. 6), Hubbard Brook (pg. 9), Konza Prairie (pg. 10), North Temperate Lakes (pg. 11), and Sevilleta (pg. 13).

Baltimore

1) **Mosquito dataset (knb-lter-bes.3500.100)**[2]
   https://doi.org/10.6073/pasta/14f78bf8f3c87f0a56d5e0bbdfd25c6a

   A) No correction was made for variation in sampling effort (number of sampling weeks) and changes in sampling locations across years. Crossley et al.'s calculation of "n.obs" reflects the number of sampling weeks, but was not used in their analyses. To allow others to check our work, we note that the sum of all individuals in the Crossley et al. (2020) online time series data (32,831 individuals) is higher than the total individuals collected from the entire knb-lter-bes.3500.100 dataset (32,329). This discrepancy is at least in part due to data corrections by the LTER site occurring after the data was downloaded by Crossley et al.; however, the high number of individuals in Crossley's online data demonstrate no correction for sampling effort.

Cedar Creek

2) **Grasshopper dataset (ghe014)[3]**
https://doi.org/10.6073/pasta/239b3023d75d83e795a15b36fac702e2

   B) No correction was made for variation in sampling effort and changes in sampling locations which are documented in the metadata (https://www.cedarcreek.umn.edu/research/data/methods?e014; see Table: Supplemental Old Fields Grasshopper Sampling for description of missing months and fields sampled within years). The sum of all individuals in the Crossley et al. (2020) online data time series (52,116 individuals) is the same as the total individuals collected from the entire ghe014 dataset, indicating no correction for sampling effort.

   C) This dataset is not correctly linked in Supplementary Table 1 and incorrectly described as a nitrogen addition and fire experiment.

3) **Arthropod "Sweep1" dataset (arce153)[4]**
https://doi.org/10.6073/pasta/a79b1120729dffc992897de58a2c5408

   A) This dataset is not appropriate to answer questions about general insect trends, since it is an experiment including nitrogen addition treatments and herbivore exclosures. While analyzing control plots alone would be appropriate, data from all experimental plots was included in Crossley at al. (2020).

   B) This dataset is not correctly linked in Crossley et al. (2020)'s Supplementary Table 1.

4) **Arthropod "Sweep2" dataset (aage120)[5]**
https://doi:10.6073/pasta/4c1795e6769bf78e3c947e92db75eef6

   A) No correction was made for variation in sampling effort and changes in sampling locations which are documented in the metadata (https://www.cedarcreek.umn.edu/research/data/methods?e120). Samples collected per year vary with sampling month and range from 1-3 samples. The sum of all individuals in the Crossley et al. (2020) online data time series (151,227 individuals) is the same as the total individuals collected from the entire dataset (after subtracting the 44,027 unidentified "undet undet" individuals). While not used in analyses, calculation of "n.obs" in Crossley et al. (2020)'s R code does not include plot number, only month and year of observation.

   B) This dataset is not appropriate to answer questions about general insect trends, since it is an experiment with treatments having different levels of plant diversity (ranging from 1-16 seeded plant species).

   C) This dataset is not correctly linked and incorrectly described in Crossley et al. (2020)'s Supplementary Table 1.

Central Arizona-Phoenix

1) **Arthropod sweep dataset (knb-lter-cap.652.2)**[6]
https://doi.org/10.6073/pasta/0669ee6a71b24abb1ae3827f4ee77f6d

No correction was made for variation in sampling locations which are documented in the metadata (https://data.sustainability.asu.edu/cap-portal/metadataviewer?packageid=knb-lter-cap.652.2). The sum of all individuals in the Crossley et al. (2020) online data time series (34,316) is a similar number to the total individuals (34,323) in the entire dataset. There is 1 individual listed as unidentified but we cannot account for the discrepancy of the 6 remaining individuals. While not used in the analysis, the calculation of "n.obs" in Crossley et al. (2020)'s R code does not include subsite, only sample date.

2) **Ground arthropod pitfall central Arizona-Phoenix dataset (knb-lter-cap.41.16)**[7]
https://doi.org/10.6073/pasta/f8aef1bde862f13b48aaf4c3b104dabd

It is likely that no correction was made for variation in sampling effort and changes in sampling locations which are documented in the metadata (https://data.sustainability.asu.edu/cap-portal/metadataviewer?packageid=knb-lter-cap.41.16). While we cannot account for the discrepancy between the number of individuals in the full pitfall dataset (2,563,183 individuals) and the number in the Crossley et al. (2020) online data time series (2,529,604 individuals, 98% of those in the full dataset), considering the high variability in subsite number and location per year in this dataset we remain concerned that sampling effort and location were not accounted for.

3) **Ground arthropod pitfall McDowell dataset (knb-lter-cap.643.2)**[8]
https://doi.org/10.6073/pasta/6ce5de2c3251607d5c939c66d9dccee0

No correction was made for variation in sampling effort and changes in sampling locations which are documented in the metadata (https://data.sustainability.asu.edu/cap-portal/metadataviewer?packageid=knb-lter-cap.643.2). The sum of all individuals in the Crossley et al. (2020) online data time series (22,360 individuals) is the same as the total individuals collected from the entire dataset from (after subtracting the 1 unidentified "Unknown" individual).

Harvard Forest

1) **Harvard Forest Hemlock Removal Experiment Ant dataset (knb-lter-hfr.118.30)[9] https://doi.org/10.6073/pasta/7a6b956fb0960d7fe8bb048b1fe26956**

A) Sampling effort differed among years for the Harvard Forest Hemlock Removal Experiment (HF-HeRE) dataset analyzed by Crossley et al. (2020). These ants were collected within a long-term experiment in which there were four plot types: two controls (intact hemlock and intact mixed hardwood) and two canopy manipulations (hemlocks girdled and logged). While logged canopy manipulation is characteristic of the landscape, the girdled plots simulate the effects of hemlock woolly adelgid (HWA) on trees four years before any HWA was documented in the area.

B) In the Crossley et al. (2020) analysis this dataset is coded with Locales: "ants.pitfall", "ants.bait", "ants.hand", and "ants.litter" that represent pitfall trapping, bait sampling with cookies and tuna fish, hand collections, and sieved litter samples of ants from 2003-2015. Sampling effort differed among years for these different sampling methods coded as "Locales" in the Crossley et al. (2020) analysis. All four sampling methods (i.e., "Locales" equal to "ants.pitfall", "ants.bait", "ants.litter", and "ants.hand" were sampled in June, July, and August from 2003-2005, in July and August in 2006, and in July only from 2007-2008. From 2009-2015, only pitfall traps were set within the HF-HeRE. The number of pitfall traps (sample number) set from 2003-2012 was 25 traps total (situated in a 10 m × 10 m array). In 2012, a deer and moose exclosure was set up within the experimental plots of the HF-HeRE and an additional 10 m × 10 m array of 25 pitfall traps was set up within the exclosure (i.e., pitfall trap sampling effort doubled from 2012-2015 relative to the number of pitfall traps from 2003-2011. We note that Crossley et al. (2020) do account for the different sampling methods and that the baits, litter, and hand samples were only collected from 2003-2008, but other differences in sampling effort were not accounted for. The "n.obs" (number of observations) in Crossley et al. (2020)'s online data is set to one for all rows in the dataset, so it does not account for differences in numbers of samples per year.

C) For the Harvard Forest ant data, Crossley et al. (2020) treat the number of ants collected by pitfall, bait, and litter samples as raw abundances, which may overestimate abundance of ants if they happen to occur nearby colonies with actively foraging workers[10].

2) **Nantucket ant dataset (knb-lter-hfr.147.21)[11] https://doi.org/10.6073/pasta/3493424abf9fc36eac7b62b732e4ea55 (hf147-09-nantucket-sites-2004-09.csv)**

This dataset contains ants sampled with pitfall traps in two bogs and surrounding forests in 2000 combined with ants sampled from upland habitats from 2004–2009 by a variety of methods and at different intensities and sites. It also includes "velvet ants", a group of ant-mimicking wasps, which were identified only to family (Mutillidae). These data were collected to assess relationships of ant diversity with habitat and management regime[12] and cannot be used to analyze temporal trends within a site. No correction was made for this variation in sampling effort and changes in sampling locations, all of which are documented in the metadata (https://portal.lternet.edu/nis/metadataviewer?packageid=knb-lter-hfr.147.21). The sum of all individuals in the Crossley et al. (2020) online data time series (32,146 individuals) is the same as the total individuals collected from the entire dataset (after subtracting the 9 individuals with year listed as "NA" and 2 individuals with species code listed as "NA"). While not used in analysis, the calculation of n.obs in Crossley et al. (2020)'s R code does not include subsite ("site") or collection method, only community type (habitat description), month, and year of observation.

3) Tick dataset (knb-lter-hfr.299.3)[13] https://doi.10.6073/pasta/ b29a97941c11ddf45540ea30066fde35

A) These data are collected with student time sheets for payroll to raise awareness of tick bites for students in the Harvard Forest Summer Research Program in Ecology. The tick survey is voluntary, has variable response rates each year depending on the group of students, and generally shows a decline in collection intensity during the summer as students increasingly fail to report weekly data. The summer of 2019 also had a much lower response rate because the program switched to using digital, rather than paper, time sheets. While not used in analysis, the calculation of n.obs (for this dataset coded as "n.y1") in Crossley et al. (2020)'s R code is the sum of hours reported by the tick survey, which is the number of hours worked during the day when the student found a tick on their body. This number does not represent the response rate of the survey, which would need to be accounted for to address differences in samples per year.

B) For the tick data, Crossley et al. (2020) analyze 30 separate time series based on the locations of collection, but it is not clear how those locations were delineated. Many of the "location.names" from this dataset have overlap as they are filled in with text by students in the survey form. For instance, the "location.names" of "Harvard Forest" in the survey overlaps with many possible locations listed by students (e.g., "greenhouse", "Prospect Hill", "Shaler Hall"). Thus, it is not appropriate to analyze these data as separate time series as they refer, in some instances, to the same general location.

4) Carnivorous plant prey dataset (knb-lter-hfr.111.16)[14] https://doi.org/10.6073/pasta/cb95637eda0f96c3fdbd1a97e632c7b7

These data were from a global review of arthropod prey spectra of carnivorous plants[15]. None of the data were collected at Harvard Forest (and most were collected on other continents), and for each carnivorous plant species, "year" indicates the year the data were published and no time-series (repeat collection) was observed or implied by the data or discussed in the review. Although these data were not included in the final analysis of Crossley et al. (2020), rows for these data are listed in Crossley et al. (2020)'s online data (https://datadryad.org/stash/dataset/doi:10.5061/dryad.cc2fqz645; External_Database_S1_PerSpecies_Abundance_LTER_annotated.csv) and all abundance values are listed as zero. The rows corresponding to this dataset (lines 28497 – 36898: 8401 records) were inaccurately included in Crossley et al. (2020)'s count of 82,777 observations (the number of rows in their online data: External_Database_S1_PerSpecies_Abundance_LTER_annotated.csv), comprising >10% of the stated number of observations.

5) We also note more generally that for the Harvard Forest datasets that the environmental data are all for the Harvard Forest site in Petersham in central Massachusetts, but the Nantucket dataset should report different environmental data as it was collected from an island off eastern Massachusetts that has very different climate from central Massachusetts. Furthermore, the locations in the tick dataset, which are each given a different time series should also have location specific environmental data as locations of data collections were variable (e.g., most in western MA at Harvard Forest, but some in Connecticut; Cambridge, MA; etc.).

Hubbard Brook

1) **Lepidoptera datasets (knb-lter-hbr.82.8)**[16]
https://doi.org/10.6073/pasta/5d2a8c67c5a3278032b2b14d66c09a7f

A) Sampling effort differed among plots and years for one of the two Hubbard Brook datasets and was not accounted for in the analysis by Crossley et al. (2020). The first dataset was coded as Locale: "Lepidoptera1" and represents visual counts of caterpillars on one plot at Hubbard Brook from 1986-2018, while the second dataset was coded as Locale: "Lepidoptera2", spans 1986-1995, and represents three different plots in the White Mountains Region that are located outside of the Hubbard Brook valley. Consistent sampling effort occurred throughout the "Lepidoptera1" time series; however, sampling effort for the "Lepidoptera2" dataset differed among years and months within years (ranging from 1 – 10 counts per month). The sum of all individuals in Crossley et al. (2020)'s online data (4,030 individuals) is the same as the total individuals collected from the entire White Mountains Region caterpillar ("Lepidoptera2") dataset, demonstrating no correction for sampling effort.

B) The caterpillar populations documented in these data exhibit outbreaks at long intervals (e.g., 10-13 years apart[17]), limiting the ability of trend analysis to detect meaningful trends with time series of shorter lengths (10-33 years for Hubbard Brook data used in Crossley et al. [2020]). A sample of 10 years duration from a population that experiences 10-13 yr pseudo-cycles is likely to provide a misleading indication of long-term trend in abundance[18] and such trend tests will have very low power due to the small sample size and inflated variance[19,20].

C) The Hubbard Brook datasets are described in Crossley et al. (2020)'s Supplementary Table S1 in a confusing manner that does not make it clear there are two datasets. The "time operational" of 1986-1997 in Table S1 for these datasets does not correspond to either dataset, as the "Lepidoptera 1" dataset spans 1986-2018, while the "Lepidoptera 2" dataset spans 1986-1995.

Konza Prairie

1) **Grasshopper dataset (CGR022)**[21]
https://doi.org/10.6073/pasta/7b2259dcb0e499447e0e11dfb562dc2f

A) No correction was made for variation in sampling effort and changes in sampling locations which are documented in the metadata (http://lter.konza.ksu.edu/content/cgr02-sweep-sampling-grasshoppers-konza-prairie-lter-watersheds). The sum of all individuals in the Crossley et al. (2020) online data time series (121,229 individuals) is the same as the total individuals collected from the entire CGR022 dataset from 1982-2015 (after subtracting the 459 unidentified "unknown" individual grasshoppers), indicating no correction for invariant sampling. While n.obs was not used in analysis, the calculation of n.obs in Crossley et al. (2020)'s R code does not include watershed, only month, day, and the replicate code within the watershed ("a" or "b"). Standardization to account for variation in sampling effort should done by dividing by the number of samples (e.g. not by day, as the number of samples varies with day of collection).

B) Three taxa (Tettigoniidae, Oecanthinae, and Gryllidae, the non-Acrididae Orthoptera) included in Crossley et al. (2020)'s analysis for the full duration (1982-2015) were only recorded in the KNZ dataset starting in 2013. Prior to 2013, these taxa occurred in samples, but no record was kept of their counts. Including these taxa which were recorded only at the end years of the time series creates a bias toward a positive community trend.

2) **Gall insects (CGP01)**[22]
http://dx.doi.org/10.6073/pasta/b2ac9e918a66dbbb18c7a6b39dc1efab

No correction was made for variation in sampling locations and plant species sampled which are documented in the metadata (https://portal.edirepository.org/nis/metadataviewer?packageid=knb-lter-knz.27.11). The sum of all individuals in the Crossley et al. (2020) online data time series (27,819 galled stems is the same as the total galled stems in the entire CGP01 dataset. While not included in the analysis, the calculation of n.obs (for this dataset coded as "n.y1") in Crossley et al. (2020)'s R code does not include watershed or account for the different plant species sampled and only accounts for the number of sampled stems.

North Temperate Lakes

1) **Benthic macroinvertebrate dataset (knb-lter-ntl.11.34)**[23]
https://doi.org/10.6073/pasta/1bad728523ce4c39ade38fa666a59aee

A) Likely due to program R being case sensitive, the time series for Sparkling Lake which was coded both "SP" and "sp" was accidentally split into two time series with "sp" having non-zero values only in 2016-2017. However, Crossley et al. (2020) considered Locale "sp" a separate time series spanning 1981-2017. Based on the taxa listed from "sp", this locale can only pertain to this dataset.

B) No correction was made for variation in sampling effort and changes in sampling locations. While we have not been able to identify why there is a discrepancy between the number of individuals in the Crossley et al. (2020) online data time series (126,041 individuals) and those in the full dataset (140,100 individuals), it is evident that sampling effort changes were not accounted for because the full time series (1981-2017) was included in Crossley et al. (2020)'s analysis, even though some lakes did not have sampling in all years.

2) **Pelagic macroinvertebrate dataset (knb-lter-ntl.13.32)**[24]
https:// doi:10.6073/pasta/50e2f7b297046aaf01b77b46a011b6da

A) While listed in Crossley et al. (2020)'s Table 1 and Supplemental Table 1, these data are not included in Crossley et al.'s online data. This dataset documents 5 taxa/ life stages, coded as "BYTHOTREPHES", "CHAOBORUS LARVAE", "CHAOBORUS PUPAE", "LEPTODORA", "MYSIS", none of which occur in Crossley et al.'s online data with the exception of documentation of the genus *Chaoborus*, but coded as "CHAOBORU" and originating from the North Temperate Lakes benthic macroinvertebrate dataset (knb-lter-ntl.11.34).

B) The link provided in Crossley et al. (2020)'s Table S1 links to a summary version of these data (summary version: https://doi.org/10.6073/pasta/2ebb7f5e89391d3caada53acd8c9a5d7) rather than the raw data.

3) **Crayfish dataset (knb-lter-ntl.3.28)**[25]
https://doi.org/10.6073/pasta/61619e749daf99c71a289dcadafb795c

While included in Crossley et al.'s online time series data ("External_Database_S1_PerSpecies_Abundance_LTER_annotated"), all abundance values are listed as zero. No entries from this dataset are listed in Crossley et al.'s online trend data ("External_Database_S2_time_trends_arthropods_relaxed"), thus these data do not appear to be included in Crossley et al. (2020)'s final analysis.

4) **Crayfish dataset (knb-lter-ntl.217.9)**[26]
https://doi.org/10.6073/pasta/4a22c4b3707f68ba5c03cc3ed70e98b6

A) This dataset has an incorrect link listed in Crossley et al. (2020)’s Supplementary Table 1. We were able to reconstruct which dataset was used by matching total sums between North Temperate Lakes crayfish datasets and finding identical yearly sums between the Crossley et al. (2020) online data and knb-lter-ntl.217.9 for 2001-2010 (both totaling 95,066 individuals for this duration).

B) Crossley et al. (2020) online data for this dataset contains data from 2011, when none exists in the dataset.

C) This dataset is not appropriate to answer questions about general arthropod trends, since it contains data on an experiment of crayfish removal. The dataset documents two species of crayfish, one is an invasive species that was removed from the lake in a whole-lake experiment designed overexploit this species. The second species was a native species that likely experienced competitive release from the removal of the invasive species.

Sevilleta

1) **Grasshopper dataset (sev-106)**[27]
   https://doi.org/10.6073/pasta/c1d40e9d0ec610bb74d02741e9d22576

   No correction was made for changes in sampling locations (termination of the pinyon-juniper [Goat Draw] vegetation type sampling site, and the initiation of a new [Blue Grama] sampling site at SEV in 2002) which are documented in the metadata (https://portal.edirepository.org/nis/metadataviewer?packageid=knb-lter-sev.106.152976). The sum of all individuals in the Crossley et al. (2020) online data time series (36,634 individuals) is the same number as total individuals in the entire sev-106 dataset. The added Blue Grama site had considerably higher numbers of grasshoppers than the old pinyon-juniper site, and this change in sampling location likely inflated the numbers of grasshoppers in Crossley et al. (2020)'s calculation starting in 2002.

2) **Ground arthropod dataset (sev-29)**[28]
   https://doi.org/10.6073/pasta/9e7e6dc9c9d8f72e9e0bca07a1e76ccd

   No correction was made for changes in sampling locations (termination of the pinyon-juniper [Goat Draw] vegetation type sampling site, and the initiation of a new [Blue Grama] sampling site at SEV in 2002) which are documented in the metadata (https://portal.edirepository.org/nis/metadataviewer?packageid=knb-lter-sev.29.175390). Collection of ground arthropods for this dataset also varied in number of traps per collection period/subsite. Some traps in each sample set of 3 subsample traps were often omitted from data tabulation due to individual traps being disturbed by precipitation runoff, or vertebrate animals. Summing omitted subsample traps (missing values, not zeros) would have reduced the sum counts for a line of 3 traps. Crossley et al. (2020)'s online data contains 39,926 individuals while the full sev-29 dataset contains 52,188 individuals identified to genus level. Crossley et al. note in their re-analysis that this discrepancy is due to removing the first three sampling years (1992-1994) to account for variation in trap number, and that this is noted in the R code. It would be advisable to note that data was omitted in additional locations besides R code, such as in Table S1. The added Blue Grama site had considerably higher numbers of ground arthropods than the old pinyon-juniper site, and this change in sampling location likely inflated the numbers of ground arthropods in Crossley et al. (2020)'s calculation starting in 2002.

Supplementary Information References

1. Crossley, M. S. *et al.* No net insect abundance and diversity declines across US Long Term Ecological Research sites. *Nature Ecology & Evolution* (2020) doi:10.1038/s41559-020-1269-4.
2. LaDeau, S. Mosquito ovitrap data from Baltimore city and county (2011-2015). (2018) doi:10.6073/PASTA/14F78BF8F3C87F0A56D5E0BBDFD25C6A.
3. Knops, J. Core old field grasshopper sampling: Successional dynamics on a resampled chronosequence. (2018) doi:10.6073/PASTA/239B3023D75D83E795A15B36FAC702E2.
4. Knops, J. Arthropod sweepnet sampling: Interactive effects of deer, fire and nitrogen. (2018) doi:10.6073/PASTA/A79B1120729DFFC992897DE58A2C5408.
5. Tilman, D. Main plots all arthropod insect sweepnet sampling 1996-2000: Biodiversity II: Effects of plant biodiversity on population and ecosystem processes. (2018) doi:10.6073/PASTA/4C1795E6769BF78E3C947E92DB75EEF6.
6. Childers, D. *et al.* Ecological survey of central Arizona: a survey of key ecological indicators in the greater Phoenix metropolitan area and surrounding Sonoran desert, ongoing since 1999. (2018) doi:10.6073/PASTA/0669EE6A71B24ABB1AE3827F4EE77F6D.
7. Grimm, N. & Childers, D. Long-term monitoring of ground-dwelling arthropods in central Arizona-Phoenix, ongoing since 1998. (2020) doi:10.6073/PASTA/F8AEF1BDE862F13B48AAF4C3B104DABD.
8. Earl, S., Grimm, N. & Childers, D. Long-term monitoring of ground-dwelling arthropods in the McDowell Sonoran Preserve, Scottsdale, Arizona, ongoing since 2012. (2019) doi:10.6073/PASTA/6CE5DE2C3251607D5C939C66D9DCCEE0.
9. Ellison, A. Ant assemblages in hemlock removal experiment at Harvard Forest since 2003. (2018) doi:10.6073/PASTA/7A6B956FB0960D7FE8BB048B1FE26956.
10. Gotelli, N. J., Ellison, A. M., Dunn, R. R. & Sanders, N. J. Counting ants (Hymenoptera: Formicidae): biodiversity sampling and statistical analysis for myrmecologists. *Myrmecological News* **15**, 13–19 (2011).
11. Ellison, A. & Gotelli, N. Ant distribution and abundance in New England since 1990. (2018) doi:10.6073/PASTA/3493424ABF9FC36EAC7B62B732E4EA55.
12. Ellison, A. M. The ants of Nantucket: Unexpectedly high biodiversity in an anthropogenic landscape. *Northeastern Naturalist* **19**, 43–66 (2012).
13. Ellison, A. Incidence of ticks and tick bites at Harvard Forest since 2006. (2019) doi:10.6073/PASTA/B29A97941C11DDF45540EA30066FDE35.
14. Ellison, A. & Gotelli, N. Prey capture by carnivorous plants worldwide 1923-2007. (2018) doi:10.6073/PASTA/CB95637EDA0F96C3FDBD1A97E632C7B7.
15. Ellison, A. M. & Gotelli, N. J. Energetics and the evolution of carnivorous plants—Darwin's 'most wonderful plants in the world'. *J Exp Bot* **60**, 19–42 (2009).

16. Ayres, M. & Holmes, R. T. Long-term trends in abundance of Lepidoptera larvae at Hubbard Brook Experimental Forest and three additional northern hardwood forest sites, 1986-2018. (2020) doi:10.6073/PASTA/5D2A8C67C5A3278032B2B14D66C09A7F.

17. Martinat, P. J. & Allen, D. C. Saddled prominent outbreaks in North America. *Northern Journal of Applied Forestry* **5**, 88–91 (1988).

18. Fournier, A. M. V., White, E. R. & Heard, S. B. Site-selection bias and apparent population declines in long-term studies. *Conservation Biology* (2019) doi:10.1111/cobi.13371.

19. White, E. R. Minimum time required to detect population trends: The need for long-term monitoring programs. *BioScience* **69**, 40–46 (2019).

20. Wauchope, H. S., Amano, T., Sutherland, W. J. & Johnston, A. When can we trust population trends? A method for quantifying the effects of sampling interval and duration. *Methods in Ecology and Evolution* **10**, 2067–2078 (2019).

21. Joern, A. CGR02 Sweep sampling of grasshoppers on Konza Prairie LTER watersheds. (2020) doi:10.6073/PASTA/7B2259DCB0E499447E0E11DFB562DC2F.

22. Hartnett, D. CGP01 Gall-insect densities on selected plant species in watersheds with different fire frequencies. (2019) doi:10.6073/PASTA/B2AC9E918A66DBBB18C7A6B39DC1EFAB.

23. Magnuson, J., Carpenter, S. & Stanley, E. North Temperate Lakes LTER: Benthic macroinvertebrates 1981 - current. (2020) doi:10.6073/PASTA/1BAD728523CE4C39ADE38FA666A59AEE.

24. Magnuson, J., Carpenter, S. & Stanley, E. North Temperate Lakes LTER: Pelagic macroinvertebrate abundance 1983 - current. (2020) doi:10.6073/PASTA/50E2F7B297046AAF01B77B46A011B6DA.

25. Magnuson, J., Carpenter, S. & Stanley, E. North Temperate Lakes LTER: Crayfish abundance 1981 - current. (2019) doi:10.6073/PASTA/61619E749DAF99C71A289DCADAFB795C.

26. Carpenter, S. *et al.* Biocomplexity at North Temperate Lakes LTER; Whole lake manipulations: Exotic crayfish removal 2001 - 2010. (2013) doi:10.6073/PASTA/4A22C4B3707F68BA5C03CC3ED70E98B6.

27. Lightfoot, D. Long-term core site grasshopper dynamics for the Sevilleta National Wildlife Refuge, New Mexico (1992-2013). (2015) doi:10.6073/PASTA/C1D40E9D0EC610BB74D02741E9D22576.

28. Lightfoot, D. Ground arthropod community survey in grassland, shrubland, and woodland at the Sevilleta National Wildlife Refuge, New Mexico (1992-2004). (2013) doi:10.6073/PASTA/9E7E6DC9C9D8F72E9E0BCA07A1E76CCD.