

*Supplement to:*

**Comment on Krüger, L., Decreasing trends of chinstrap penguin breeding colonies in a region of major and ongoing rapid environmental changes suggest population level vulnerability. *Diversity* 2023, 15, 327**

**W. Chris Oosthuizen\*, Murray Christian and Mzabalazo Ngwenya**

Centre for Statistics in Ecology, the Environment and Conservation, Department of Statistical Sciences, University of Cape Town, Cape Town, 7701, South Africa

\*chris.oosthuizen@uct.ac.za / w.chris.oosthuizen@gmail.com

## Supplement text S1

Historical breeding population count data are sparse and sporadic for most chinstrap penguin colonies in the Antarctic Peninsula. Much of the available data is summarized in the Mapping Application for Penguin Populations and Projected Dynamics (MAPPPD) (Humphries et al. 2017). However, data processing choices (“researcher degrees of freedom”) can affect the amount of MAPPPD data used in analysis. Researchers make many decisions during data processing (e.g., whether to include possibly unreliable data, what constitutes unreliable data, etc.) that may influence their results (e.g., Schweinsberg et al. 2021). These decisions are often judgement calls, presumably with the goal of reaching robust conclusions. Here we highlight two judgement calls relevant to this current study. Krüger (2023) removed nest counts where data fields (e.g., day or month of count) contained missing data (indiscriminate use of `na.omit` in R code). At Penguin Island, for example, nest counts from six breeding seasons (all with good accuracy) were not included in Krüger (2023) because ‘day’ of the count was unknown (Table S1.1). Given the sparse data (Figure S1.1), the current study opted to include these counts (rationale: even when ‘day’ and ‘month’ were available, we did not select counts given specific days or months; so, no information is lost when we include counts with missing ‘day’ or ‘month’). In contrast, we opted to exclude other counts that were used in Krüger (2023). For example, we excluded all counts with very high uncertainty (“to an order of magnitude”; MAPPPD level 5). Our general aim was to use as much of the available data as possible but being cautious to avoid using highly uncertain counts and to limit, to some degree, extrapolations far from observed data. Figures S1.2 and Figure S1.3 show the structure of the data we used for the analysis presented in the article’s main text. Supplementary text 6 briefly explores the impact of different data processing choices on estimates of population change.

Table S1.1. MAPPPD data for Penguin Island. Krüger (2023) excluded counts with unknown (NA) fields, and thus did not consider count data prior to 2010 for Penguin Island (grey shading). These counts were included in the current study, but any counts with an ‘accuracy’ of 5 would have been removed.

site_name	day	month	year	season_starting	penguin_count	accuracy	count_type
Penguin Is.	NA	12	1979	1979	7058	1	nests
Penguin Is.	NA	12	1980	1980	7581	2	nests
Penguin Is.	NA	12	1980	1980	8794	1	nests
Penguin Is.	NA	1	2000	1999	3774	1	nests
Penguin Is.	NA	12	2000	2000	3296	1	nests
Penguin Is.	NA	12	2003	2003	2672	1	nests
Penguin Is.	NA	12	2008	2008	4161	1	nests
Penguin Is.	6	12	2010	2010	3017	1	nests
Penguin Is.	12	1	2013	2012	1545	3	nests
Penguin Is.	16	1	2018	2017	2252	2	nests
Penguin Is.	19	12	2017	2017	2537	1	nests

## References

- Humphries, G. R. W., Naveen, R., Schwaller, M., Che-Castaldo, C., McDowall, P., Schrimpf, M., & Lynch, H. J. (2017). Mapping application for penguin populations and projected dynamics (MAPPPD): data and tools for dynamic management and decision support. *Polar Record* 53: 160-166.
- Krüger, L. (2023). Decreasing Trends of Chinstrap Penguin Breeding Colonies in a Region of Major and Ongoing Rapid Environmental Changes Suggest Population Level Vulnerability. *Diversity* 15: 327.
- Schweinsberg, M., Feldman, M., Staub, N., van den Akker, O. R., van Aert, R. C., Van Assen, M. A., et al. (2021). Same data, different conclusions: Radical dispersion in empirical results when independent analysts operationalize and test the same hypothesis. *Organizational Behavior and Human Decision Processes* 165: 228-249.



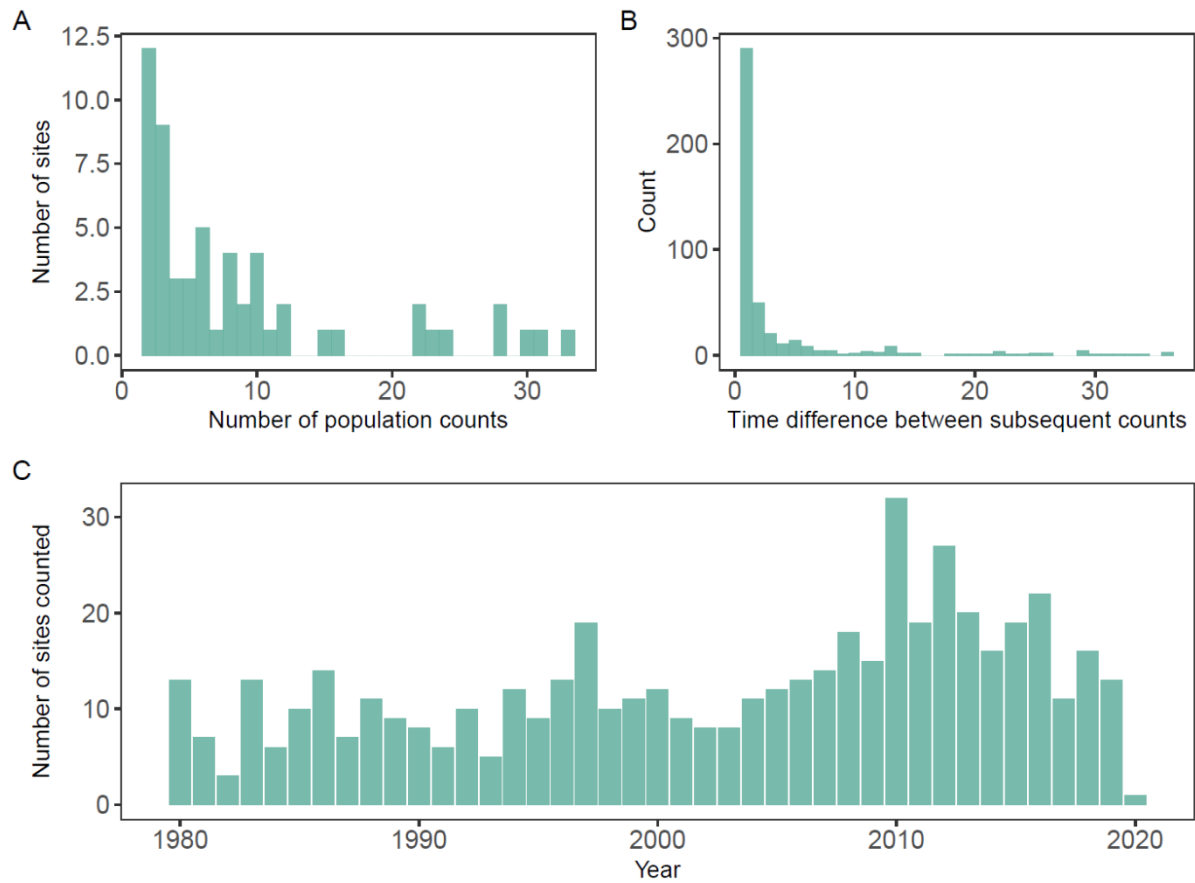


Figure S1.2. Distribution of chinstrap penguin nest count data (1980-2019) analysed in this paper (“dataset 3”). (A) In total, 57 sites met data processing criteria (see below). Most sites were only counted twice. (B) While year-on-year counts at a site were common, decades lapsed between successive counts at some sites. (C) Counts were made in all years, but sampling was unbalanced at the site-level (Figure S1.3).

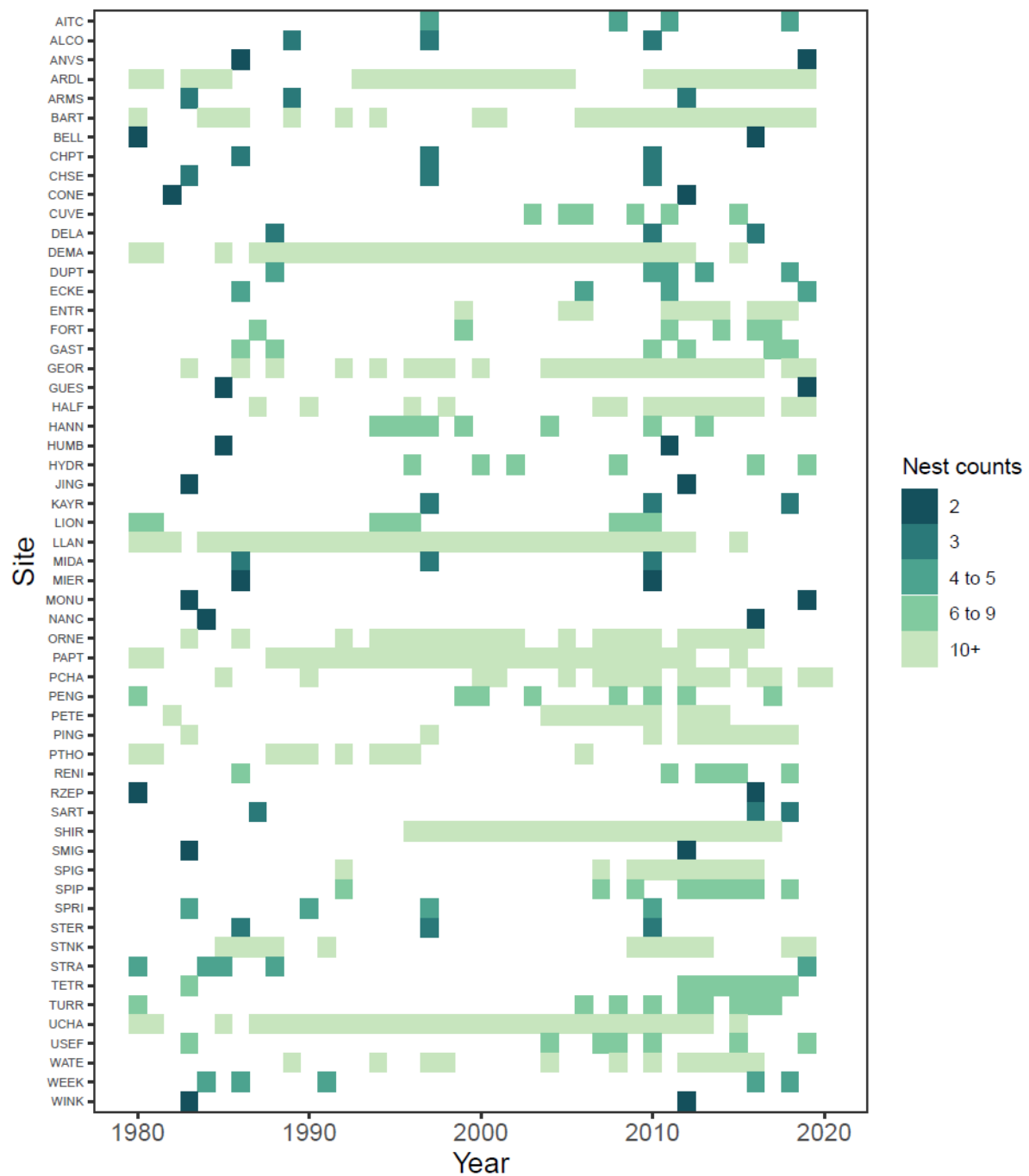


Figure S1.3. Heatmap of the time series structure of the MAPPPD data analysed in this paper. We made the following data processing choices for “dataset 3”: (1) exclude counts with very high uncertainty (“to an order of magnitude”; MAPPPD level 5); (2) retained nest counts where ‘day’ and ‘month’ of count was unknown; (3) retained ‘true zeros’ (counts with zero nests); (4) limit data to sites with at least one count since 2004 (within 15 years of 2019), and at least one count prior to 2005 (within 15 years of 1990).

## Supplementary text 2

Krüger (2023) restricted initial (exploratory) analysis to colonies that declined between their first and last counts, and reported that 46 % of these colonies have decreased by more than 75 % (Figure 2.1 A). The correct estimate based on Krüger (2023)'s input data and analysis is 20 %, as the 46 % erroneously refers to a decrease of over 55 % (not 75 %) between the first and last count. Figure S2.1 B gives the corrected figure stemming from Krüger (2023)'s input data.

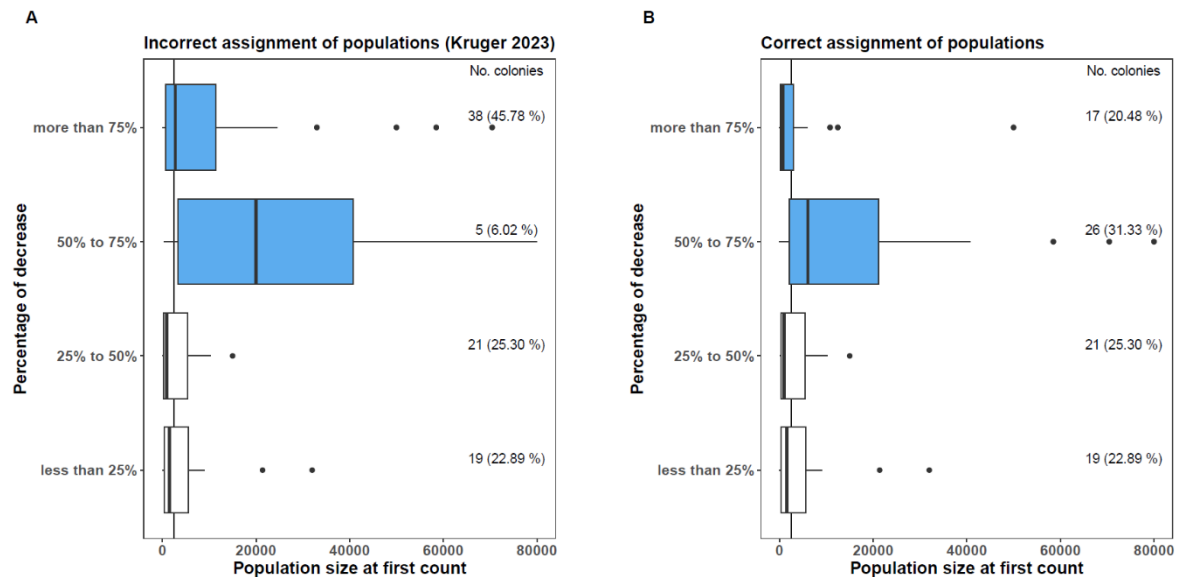


Figure S2.1. Correction of Figure 2 in Krüger (2023) (blue box plots indicate changes). The figures show the percentage change in decreasing chinstrap penguin breeding populations (first count vs. last count) in the Antarctic Peninsula.

It can be risky, in general, to diagnose multi-year population trends based on a comparison of counts in two years. Unfortunately, many chinstrap populations have only been counted twice (or, only two counts remain in processed input data). Thus, two counts often represent the only available information from which to make inference. Having such little data can be limiting. In Figure S2.2, the long-term population trend tends to positive (blue line), but the comparison of counts made in the first and last year is negative (red line). Though there are few such cases in the data analysed by Krüger (2023), it suggests that sparse data and binary data summaries (population increase or decrease) may not always capture longer-term trends (e.g., population fluctuations around a stable point).

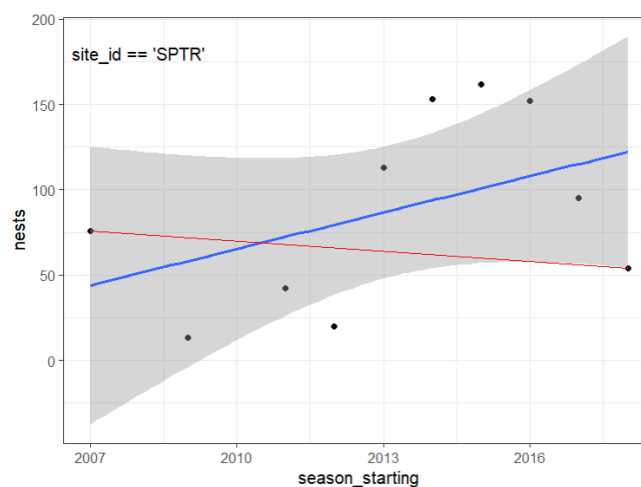


Figure S2.2. Nest counts from the Spigot Peak Tripod (SPTR) chinstrap penguin colony. The red line connects the first and the last count (negative slope) while the blue regression line (and grey shading confidence region) represents the positive long-term trend at this colony.

### Supplementary text 3

Table S3.1. Descriptions of survey accuracy, and number of chinstrap penguin nest count estimates per quality flag in MAPPPD data analysed by Krüger (2023).

Quality flag	Description <sup>1,2</sup>	Assumed accuracy	Number of nest counts	Proportion of Krüger (2023) input data
1	Individuals counted	$\pm < 5 \%$	291	61 %
2	Individuals counted per area and extrapolated over colony area	$\pm 10 \%$	16	3 %
3	Accurate estimate	$\pm 10 - 15 \%$	29	6 %
4	Rough estimate	$\pm 25 - 50 \%$	68	14 %
5	Guesstimate	Order of magnitude	75	16 %

<sup>1,2</sup> Croxall and Kirkwood (1979); Strycker et al. (2020).

### References

Croxall, J. P., & Kirkwood, E. D. (1979). The distribution of penguins on the Antarctic Peninsula and islands of the Scotia Sea. British Antarctic Survey.

Krüger, L. (2023). Decreasing Trends of Chinstrap Penguin Breeding Colonies in a Region of Major and Ongoing Rapid Environmental Changes Suggest Population Level Vulnerability. *Diversity* 15: 327.

Strycker, N., Wethington, M., Borowicz, A., Forrest, S., Witharana, C., Hart, T., & Lynch, H. J. (2020). A global population assessment of the Chinstrap penguin (*Pygoscelis antarctica*). *Scientific Reports* 10: 19474.

#### Supplementary text 4

Trace plots show the sequence of parameter values generated by the Markov Chain Monte Carlo (MCMC) algorithm as it iteratively explores the posterior distribution of parameters. Trace plots of Krüger (2023)'s variance-covariance matrix of the random effects revealed high autocorrelation, which indicate slow mixing and inefficiency in the MCMC chain (Figure S4.1A). In contrast, trace plots indicated good mixing in the variance-covariance matrix of the random effects obtained from our revised model (Figure S4.1B).

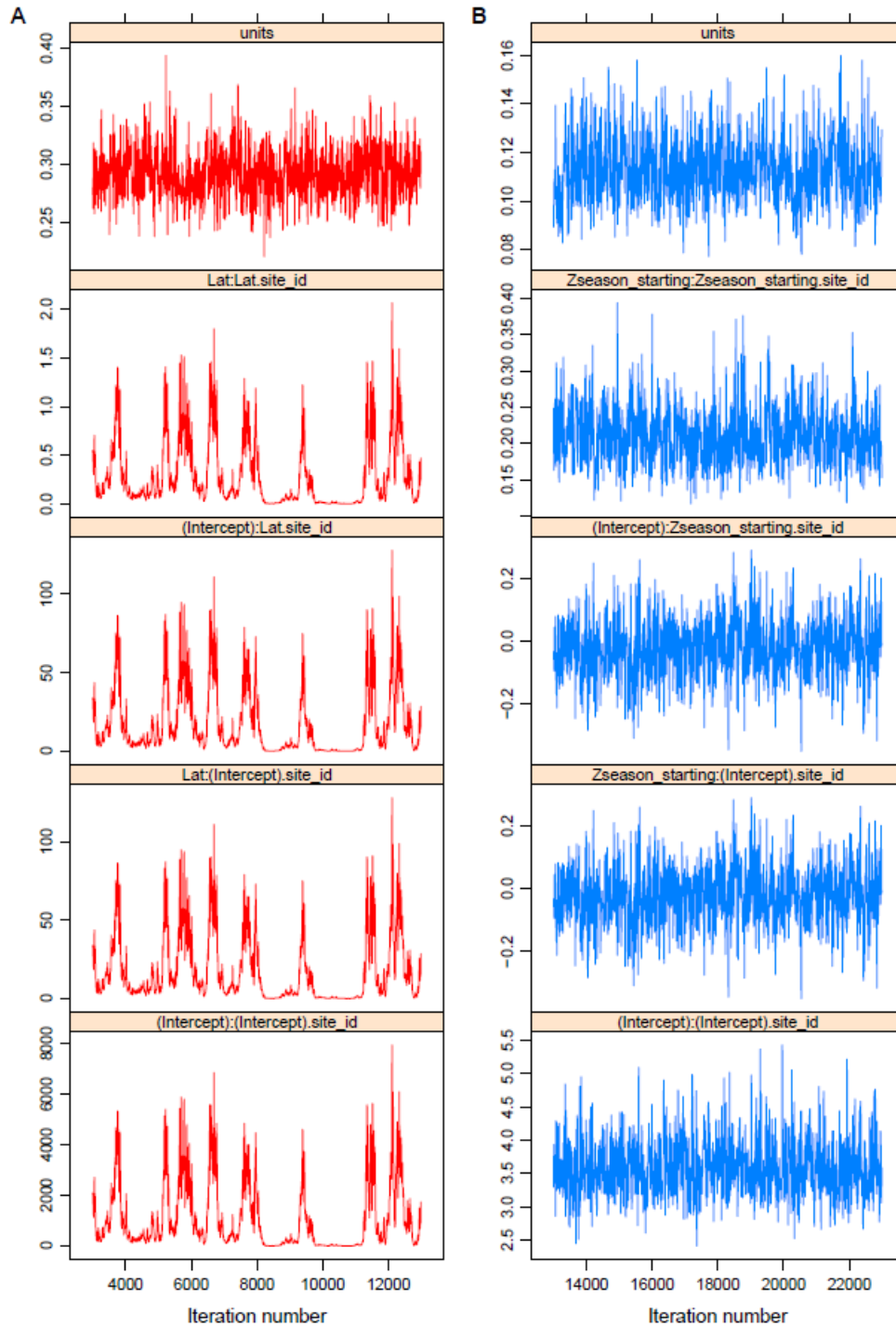


Figure S4.1. (A). Trace plots of Krüger (2023)'s variance-covariance matrix of the random effects. (B). Trace plots of the variance-covariance matrix of the random effects from our revised model.



## Supplementary text 5

In Figure S5.1, we compare model predictions obtained from simulated data using the Krüger (2023) GLMM model specification (Figure S5.1A), and a (simple) random intercept model (Figure S5.1B). To show equivalence, these predictions are not derived from posterior predictive simulation (`interval="prediction"` in `predict.MCMCglmm`, as used by Krüger (2023)). Instead, we specified `interval="confidence"` for credible intervals, but only plot the posterior mean here.

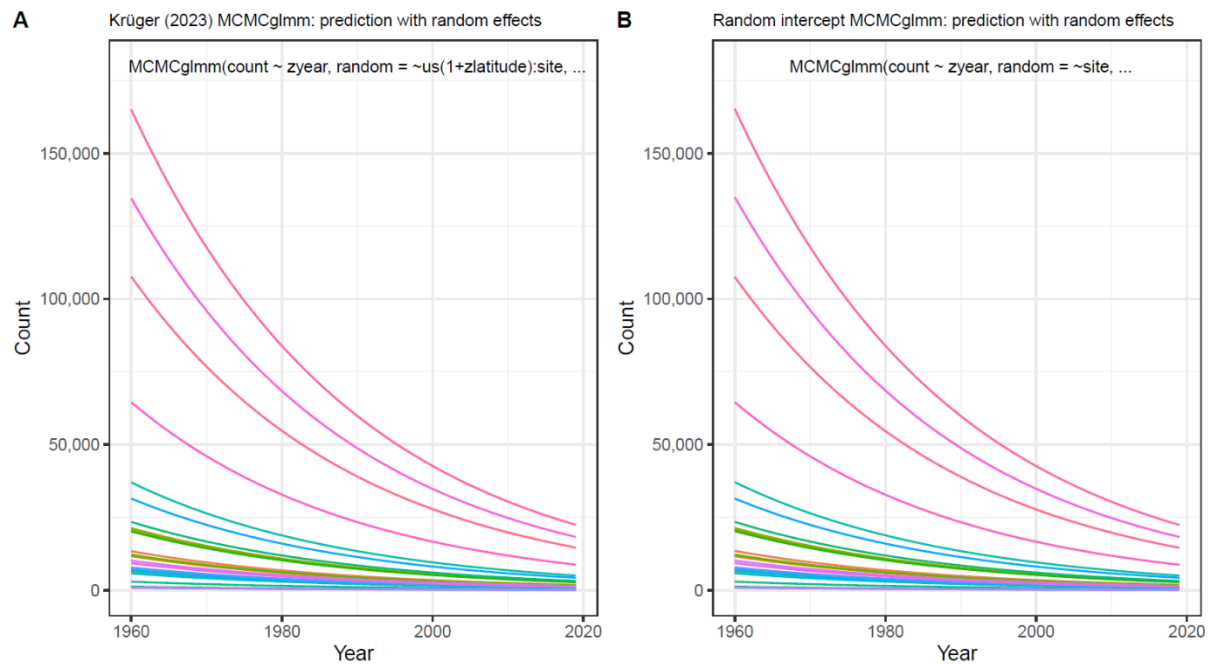


Figure S5.1. Comparison of model predictions obtained from simulated data using (A) the Krüger (2023) GLMM model specification, and (B) a (simple) random intercept model, indicating equivalence. These predictions include the respective random effects in the prediction.

In Figure S5.2 we show why we did not use posterior predictive simulation (`interval="prediction"`) to compare the models above. When prediction intervals and marginalized random effects are specified (this was the syntax implemented by Krüger (2023)) the predictions are highly variable ('wiggly') (see Figure S5.2). Also, note that the predictions differ each time we predict using the same GLMM as object (Figure S5.2). Most (but not all) of this variation is removed when we specify prediction intervals but include random effects in the prediction (Figure S5.3).

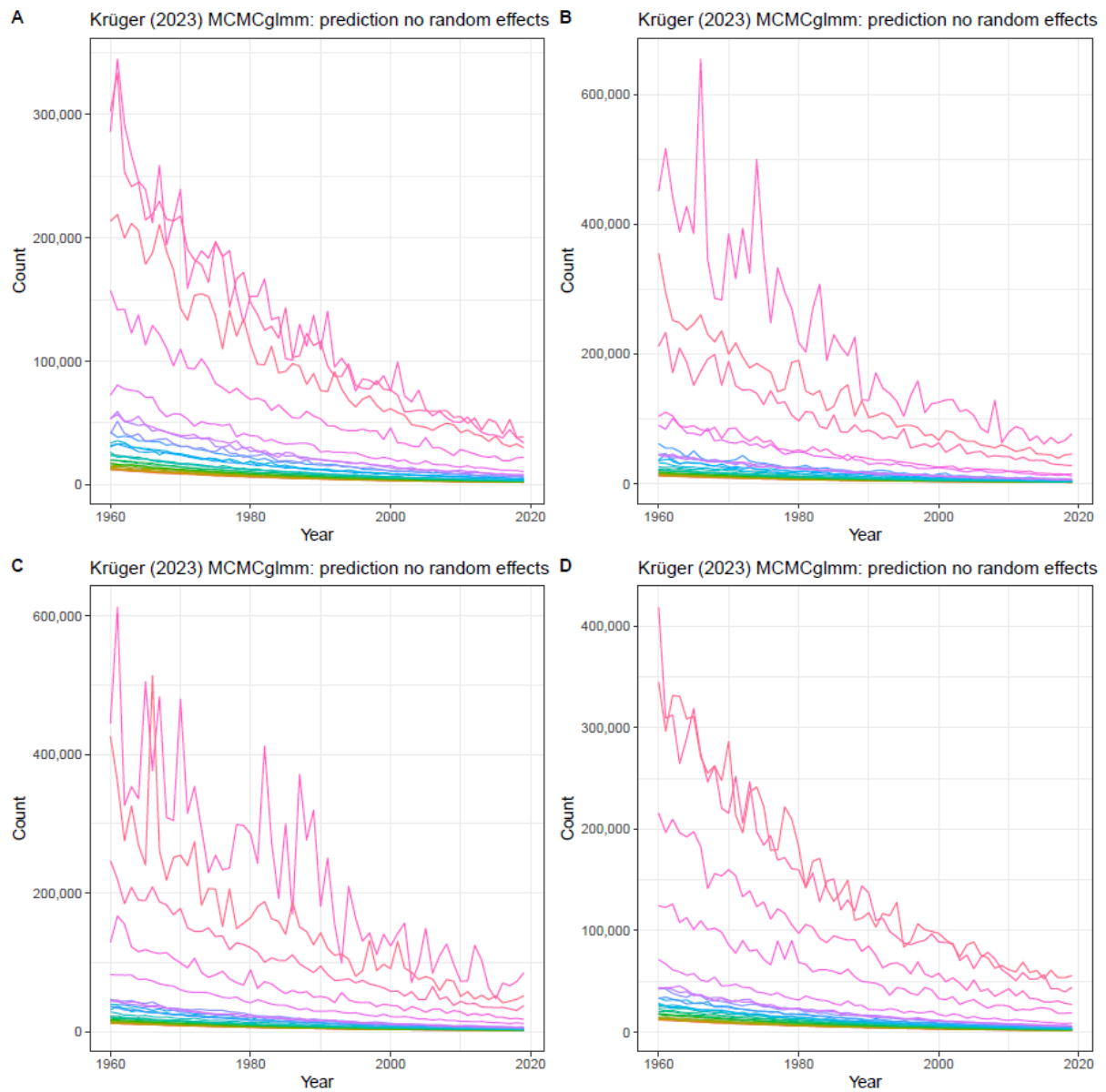


Figure S5.2. Comparison of model predictions obtained from simulated data using the Krüger (2023) GLMM. All predictions were obtained using the same data and fitted GLMM (we did not re-run the GLMM). The `predict.MCMCglmm` syntax for each of the four figures are the same as in Krüger (2023): `predict(mc_Kr, newdata=df, type="response", marginal=mc_Kr$Random$formula, interval="prediction", posterior="mean")`

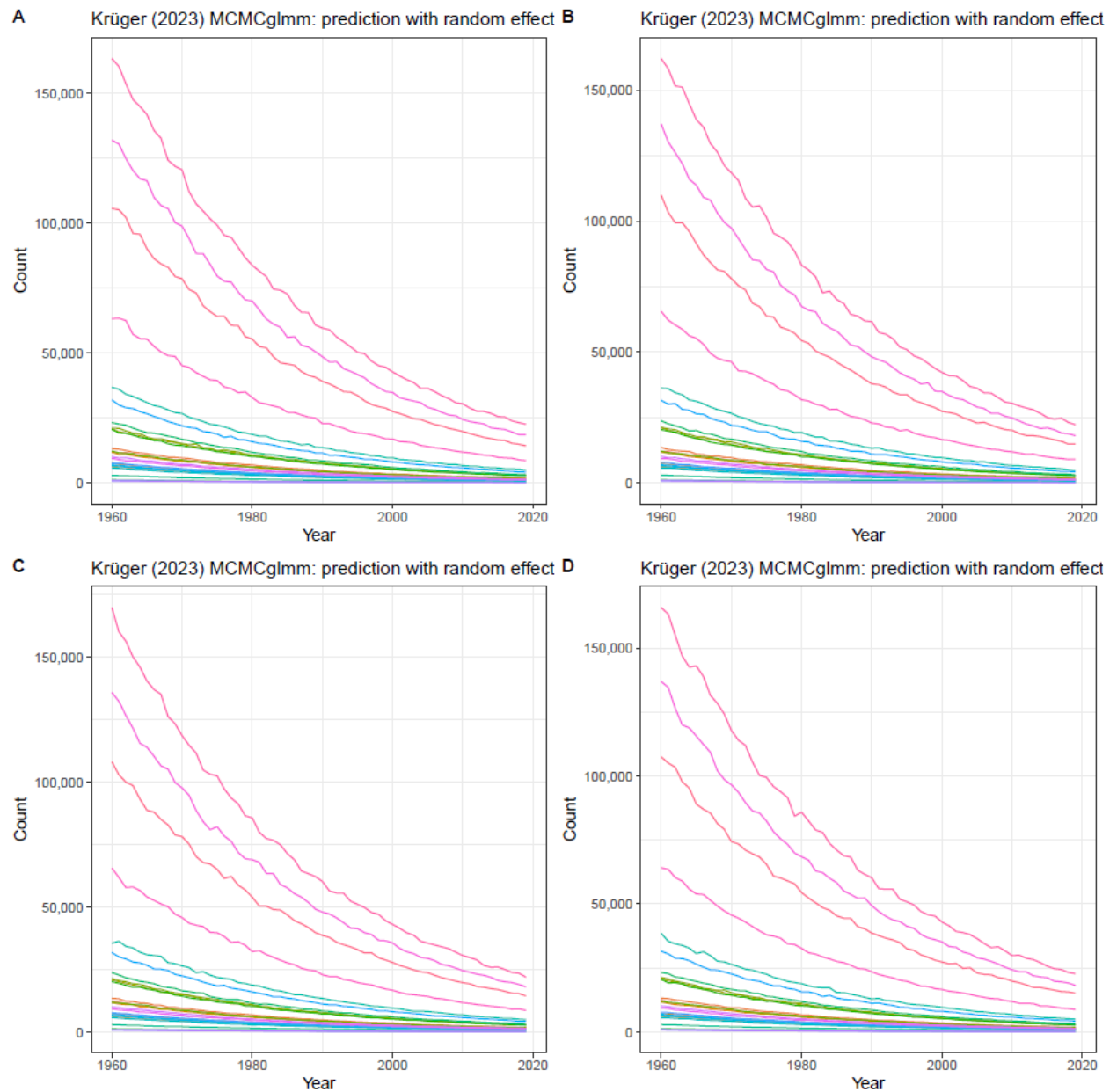


Figure S5.3. Comparison of model predictions obtained from simulated data using the Krüger (2023) GLMM. All predictions were obtained using the same data and fitted GLMM (we did not re-run the GLMM). The `predict.MCMCglmm` syntax for the four figures are: `predict(mc_Kr, newdata=df, type="response", marginal=NULL, interval="prediction", posterior="mean")`.