General comments

I really appreciated the comprehensive and complete answer of the authors, I would like to thank them for the time they dedicated to answer my comments and the serious of their work. I think the modification done by the authors have improved a lot the quality of the manuscript. I don't have major comments left that I did not raise. However, I would like to bring last thoughts about the observer effect that author used.

Thanks to the authors for their explanations in their answer that have clarified my mis-understanding of their observer effects construction. I still think that this index might capture other effects than the observer effects. Some years can be exceptionally bad for bird reproduction in general, often due to climatic factors. The observer effects proposed here could remove these "bad years" from the population trend and absorb them, while their increasing/decreasing frequency would be very informative and should be included in the population trend. Could authors justify with some supplementary figures the threshold of 25%? How variable is the summed abundance across years for a given site? Is there any sites followed during several years by the same participant? If yes, these sites could be used to analyse the expected variability across year.

The figure 5 shows that there is an important temporal shift in observer effects from negative to positive, how can authors explain this? Did participants change behaviour? Did the kind of participants change over time? This is likely to affect the resulting population trends.

The editor and authors will decide if it is relevant to discuss again this, but since authors aim to propose a general method to analyse bird counts data, I think the solution they propose should be adapted to most of the possible datasets, and I am not sure this observer effect is.

In addition to this comment, the manuscript would need some minor modifications to be ready for publications: figure caption could be more detailed, authors used a lot of acronyms, homogenisation of the terms used to describe variables ("regions", "landscape", etc.). Also, I did not miss anything I think there is an error in the formula of the population trend (see my comment below about line 695).

Specific comments

Lines 131-132: "*Moreover, TRIM is restricted to categorical covariates, requiring climate or landscape composition covariates to be transformed into categories (Bogaart et al. 2020).*" If I got it well it is now what authors do also (landscape is a categorical variable), so it is a common limit between TRIM and their model. For this reason i would raise that point in discussion rather than in introduction.

Line 181: results are given for both regions together (25 to 36) but the maximum number of sites was given per region (43 and 47), which is a bit confusing.

Lines 190-194: I think I got what authors did here but it took me a bit of time. The verb "merge" is not really explicit. I think the lack of clarity comes from the superposition of old and new methods version. The biogeographic regions occupy a big place in Figure 1 and beginning of the methods but they are not used in the analyses. This is confusing. I would rewrite the methods and presentation of the dataset focusing on variables that will be used in the analyses. For example, here I would say something like:

"Site were distributed into six natural regions (Fig. 1) and located or not in a Metropolitan area. We combined these two layers of information to produce a landscape categorical variable with 7 levels, all sites within the metropolitan area being classified as "urban" while other were classified according to the natural region they belong."

Lines 218: Here landscape is re-used again for other variable than the one defined line 192. I would avoid that and use another word for one of them. Either stick to natural region in the first place, or use "habitat" in the second place.

Figure 2: Caption says biogeographical region while actual what is plotted is the "landscape categorical variable" defined by authors in lines 192-194. Also, acronyms should be defined in the figure caption. The abbreviation for Continental is still in German (kon).

Lines 280-301: Thanks to the authors for their explanations in their answer and the details about that observer index. I still think that this index might capture other effects than the observer effects. Some years can be exceptionally bad for bird reproduction in general, often due to climatic factors. A deviation from 25% of the mean could thus be expected? If yes, the observer effects proposed here would remove these "bad years" from the population trend and absorb them, while their increasing/decreasing frequency would be very informative and should be included in the population trend. Could authors justify with some supplementary figures the threshold of 25%? How variable is the summed abundance across years for a given site? Is there any sites followed during several years by the same participant? If yes, these sites could be used to analyse the expected variability across year.

It might be that the datasets authors analyse here do not present these variations, but since they aim to propose a general method to analyse bird counts data, I think the solution they propose should be adapted to most of the possible data.

The figure 5 shows that there is an important temporal shift in observer effects from negative to positive, how can authors explain this? Did participants change behaviour? Did the kind of participants change over time? This is likely to affect the resulting population trends.

Line 693: "*Based on pairwise differences of simulated abundances…*" Do authors mean predicted abundances?

Line 695: I f I did not miss anything this formula means that the trend is a difference of density? So, it is not a decline/increase per unit of time?

Most often to calculate a trend we divide by the duration of the temporal window $\frac{N_{ij}-N_{i(j-\Delta t)}}{\Delta t}$. Authors could also calculate growth rates, that is a bit more meaningful for comparison among species that do not have the same abundance: $100 \times \left(\left(\frac{N_{ij}}{N_{i(j-\Delta t)}} \right)^{\frac{1}{\Delta t}} - 1 \right)$.

Line 754: is that acronym (DDA) used after that? The authors defined a lot of acronyms along the methods and could try to limit this number. If an acronym is not used afterwards, there is no need to define it. Otherwise, the reader might try to keep all of them in memory thinking that it will be useful for latter and waste focusing abilities.

Figure 4: it would be nice to have the unit of the population trends on the y-axis. Especially because the formula used in the methods is not clear about how authors calculate the trends.

Figure 6: I think figure caption could be a bit more explicit for the reader, and could be used as reminder of the meaning of the acronyms, for example, instead of "*correlation of annual indices*…" authors could say "Correlation between our annual index of abundance (EAS) and the one form the German Common Bird Monitoring scheme (Mhb)…"

Figure 2.1 of the supplementary materials: Authors present the coefficient associated with the effect of PC1, PC2 and PC3. However, interpreting a polynomial function from its coefficient is extremely hard.

Even if the polynomial is only of order 2, I would instead or in addition of presenting the coefficient, present the predicted polynomial in itself (abundance as a function of PC1, for average level of other variables), it would be more informative. It would also be informative to remind the reader which variable compose essentially each of the PCA axes.