**Reviewer’s comments**

1. The ms is improved, but I still have some clarifications, and some minor edits suggested (e.g. typos). The tone could still be improved by more careful use of language. One issue is that the BII values used are older and have been superceded. It is worth thinking about how this matches the newer versions.

We recognise that some of the BII values have been superseded by the maps produced by de Palma et al (2019) and Hill et al (2018). However, de Palma et al (2019) only calculate BII for tropical forest regions and neither Hill et al (2018) nor de Palma et al (2019) have made the data that underly their maps of BII available. As such, we cannot carry out any analysis to see how closely the BII in these versions matches the estimates of Newbold et al (2016).

1. Around lines 20-22. Make it clear that 100% means intact (‘natural’ abundance levels) and that less means reduced abundance relative to pristine levels.

We agree with the reviewer and have incorporated an explanation of how to interpret the BII (see lines 21-23).

1. Adding the % land cleared in Madagascar would also be useful (as well as lemur drops, given they are one group only and subject to hunting as well as habitat loss).

Unfortunately figures on Madagascan deforestation are unclear and disputed, particularly for long-term habitat conversion, but also for deforestation levels since World War 2 (Scales 2014). Likewise, the Red Listing process (IUCN, 2019) does not attempt to provide precise levels of population decline. No-one, however, suggests Madagascan conversion is anywhere close to being just 10% of original habitat; and habitat loss is accepted as a major driver in the declines of all 34 lemur species thought to have declined by ≥30% (sometimes far more) over the past 3 generations.

1. Line 28-29: Please clarify that MBI looks at plant biomass – in the response document it is clear, but not well described in the actual letter.

We have now modified these lines to fit with our previous description of the BMI in our response to the reviewer (see lines 29-31).

1. Also, particularly in relation to the comparison of BII and BMI below, the timeframes over which these changes are measured matters (last 100 years? How do you deal with Europe where change is over 500+ years?).

Neither the BII (Newbold et al 2016) nor BMI (Erb et al 2018) use a historical baseline for comparison. Newbold et al (2016) modelled the impact of land-use change and other pressures on biodiversity relative to primary vegetation under minimal human pressure. Erb et al (2018) compared estimates of current vegetation biomass to modelled vegetation biomass under current climatic conditions in the absence of humans. The similarity of these approaches means that BII and BMI should be directly comparable.

1. Paragraph in lines 28-38: there is some missing punctuation and some of these sentences aren’t clear. For example, remove “Despite these differences at a global scale” – it doesn’t add much. “However, biomass and abundance metrics measure different attributes of biodiversity”, relating to for example plant biomass versus species abundance (particularly non-plants).

We have added the missing punctuation and cut the “Despite these differences at a global scale” (see line 34). We have also made substantial changes to this section based to the reviewer’s comments that also help to improve the clarity and flow of the document (see lines 34-43). As a point of clarification to the reviewer, BII aims to estimate biodiversity change for all taxa in an area – including plants.

1. Line 39, it is not ideal to start a paragraph with ‘however’. It isn’t necessary there – you can just remove it.

We have changed this ‘however’ to ‘in practice’ (see line 44).

1. In fact, you can remove most of your ‘however’s, as they often don’t add much – carefully review them (the one in line 59 can also go). That would improve the tone somewhat while still emphasising the problems and caution with which the BII (and indeed any indicator) should be used. Similarly, in line 60, change “To be credible” to “to improve credibility”.

We have modified lines 59 and 60 according to the reviewer’s suggestion (see lines 65-66).

1. Line 39: “the two indices exhibit very limited agreement” – it isn’t clear how you define this. The grey band in Fig1 is quite narrow – perhaps widen it, given how hard it is to see the grades of red and blue. This would help define the space where there is a satisfactory level of agreement. What is not?

The grey area represents where there is approximately less than a 10% point difference between BMI and BII. Using the R package we used for this analysis it is not possible to widen the grey area. However, to improve the tone of this section we have removed ‘very’ from line 44, and have removed our statement about ‘very limited agreement’ (lines 48-49).

1. The comparisons with other indicators are useful, but make for a long paragraph – can you break it up?

We have now split this paragraph into two .

1. In the last paragraph, while listing ideal features of the BII is good, just stating that the values should be lower in certain areas is not useful. It should address what is causing that – data bias (as per lines 65-66). I would suggest bringing that up higher, and reducing the more didactic statements about what the values should be where (lines 61-65) – you’ve already covered that in the earlier paragraphs. In the last paragraph, I would recommend focussing only on how the data and process could improve outcomes, not on the patterns themselves, as you’ve discussed them above.

Respectfully we disagree with the reviewer here. Firstly, while we have our suspicions about why patterns of BII calculated by Newbold et al (2016) are not what we would expect, we cannot be certain about which of these are most important. These potential problems include, but are not limited to, biases in the populations sampled in the PREDICTS project database, biases in global land-cover data used, mismatches between the land-covers assigned to individual abundance estimates and used to extrapolate abundance estimates globally, errors in the database built by the PREDICTS project, and use of an arithmetic mean rather than a geometric mean value of current relative to pristine population abundance. Secondly, we think it is vital that we provide a yardstick against which the performance of a revised version of the BII could be assessed. The danger, if we do not include these details, is that in the future a new version of BII will be produced claiming to have solved any problems that we have raised, without subjecting it to any of our suggested tests.

**References**

De Palma, A., Hoskins, A., Gonzalez, R.E., Borger, L., Newbold, T., Sanchez-Ortiz, K., Ferrier, S., Purvis, A. (2019) Annual changes in the Biodiversity Intactness Index in tropical and subtropical forest biomes, 2001-2012 *bioRxiv* 311688; doi: https://doi.org/10.1101/311688

Erb, K. H., Kastner, T., Plutzar, C., Bais, A. L. S., Carvalhais, N., Fetzel, T., ... & Pongratz, J. (2018). Unexpectedly large impact of forest management and grazing on global vegetation biomass. *Nature*, *553*(7686), 73.

Hill, S.L.L., Gonzalez, R., Sanchez-Ortiz, K., Caton, E., Espinoza, F., Newbold, T., Tylianakis, J., Scharlemann, J.P.W., De Palma, A., and Purvis, A. (2018) Worldwide impacts of past and projected future land-use change on local species richness and the Biodiversity Intactness Index. *bioRxiv* (2018): 311787.

IUCN (2019) The IUCN Red List of Threatened Species. Version 2018-2. http://www.iucnredlist.org. Downloaded on 19/02/2019.

Newbold, T., et al. (2016) "Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment." Science 353.6296, 288-291.

Scales (ed) (2014) Conservation and environmental management in Madagascar, Routledge