We thank the reviewer for the praise and the many helpful suggestions, which have largely been followed. Below our reply to the specific comments, set in red font color.

Braakhekke et al., use an Earth System Model to constrain the drivers of N leaching from 1901 to 2006. The manuscript is well written, interesting and contributes to a better understanding of the control of N leaching. Please find below a few comments that should be addressed before publication.

1) Atmospheric N deposition: theory and forcing

1.1. I am not really familiar with atmospheric N deposition, but from reading the manuscript it seems like the authors refer to nitrous oxide deposition, since N deposition is the highest in the most populated/polluted areas. I think a line or two properly defining atmospheric N deposition and its control could help better understand the paper.

Section 2.1.2 has been expanded with an explanation of N deposition and biological N fixation:

1.2. It is stated that the forcing for atmospheric N deposition is taken for years 1850-1860, whereas the other forcings (climate and CO2) are taken from year 1901. I guess there is not much difference between 1860 and 1901 for N deposition or at least much less than between 1901-2006, but a) where do estimates of N deposition for 1860 come from? And b) why would you take 1860 instead of 1901?

a) The N deposition data for the complete simulation period, including the spin-up, was obtained from the ACCMIP dataset Lamarque et al. (2013).

b) The aim of the spin-up is to obtain the ecosystem state under "pre-industrial" conditions, i.e. as little influence by humans as possible. For climate and CO_2 concentration anthropogenic effects are relatively small before 1900, compared to 1900-2000, but for N deposition regionally quite strong changes occurred before 1900 due to land-use change (biomass burning). Hence, we chose to use the first time period in the ACCMIP dataset, 1850–1860, both for the spin-up and the simulations where N deposition was held constant.

Section 2.2.1 has been modified to clarify the above two points.

1.3. Following on these 2 comments, I would suggest to restructure a bit section 3.1.1. as follow: The authors could start by describing Figure 1 and explaining the origin/controls of N deposition (natural vs anthropogenic effects, then describe Figure 2 and differences for each biome.

Done.

2) Comparison with previous estimates (Section 3.1.3) Beusen estimates seem to be significantly higher than LPJ-GUESS in Equatorial regions and southern tropics. The sentence L.13-15 is very unclear to me.

The relevant paragraph has been rewritten and is hopefully now more clear.

L. 4: "higher productivity in cold and dry regions at other latitudes." To me it seems more like in the "mid latitudes". East Australia and South Brazil/Argentina are not really cold-dry regions. I am not really convinced by this section 3.1.3. I understand the authors want to try and compare their estimates with previous studies, but here a very rough comparison is made without going really into the reasons for these differences.

The line the reviewer referred to has been modified according to the reviewer's suggestion. An in-depth discussion on the reason of the mismatch for GPP is difficult without a detailed analysis including site observations and other data, which is outside of the scope of this study. Presumably, re-calibration would be needed to improve the fit. The aim of the comparisons with other data sets is to see whether LPJ-GUESS predictions are sufficiently realistic for the purpose of this is study, which—in our view—is the case, despite mismatch for GPP.

3) Climate section Due to the different impacts of temperature and precipitation changes on terrestrial productivity and soil processes, this section is a bit more difficult to follow. I would suggest the authors try to use terms which indicate the direction of the change: e.g. warmer conditions increase N mineralization...

We assume the reviewer is referring to section 4.1.2. The relevant paragraph has been slightly modified to meet the reviewer's request.

Due to the complicated relationship between climate and N leaching, as described in this section, it is not always clear what is the exact mechanism behind the predicted responses. Therefore, we prefer to formulate the discussion as we did: first mention the overall effect of climate change on N leaching and then discuss the likely reasons behind this response.

Minor comments:

- P2, L. 15: "with increasing N input, the capacity of ecosystems to retain N decreases.." that statement surprises me. I can understand that with increasing N input, leaching increases, but the capacity to retain N does not necessarily changes.
  "N input" has been changed to "N availability", which is more accurate.

- P4, L6. Add "be" between can and found.
  Done.

- P6, L. 27, "we"
  Done.

- P6, L.29: remove "to" between by and "the fraction"
  Done.

- Notation of "N leaching : N input ration", I don't; really like that notation as ":" also denotes a ratio. I would suggest to modify to "N leaching/N input" or "N leaching to N input ratio".
  All instances have been changed to "N leaching to N input ratio".

- P8, L. 20: add "is" after variability
  Done.

- Section 3.2.1, L. 18: N deposition and pCO2 increased but not climate, please modify.
  Done.

- P11, L.15, year of citation for Cleveland et al., is missing

Since we already listed the publication year three lines previous, we deliberately omitted it in this sentence. The phrase has been changed from “the Cleveland et al. N fixation-AET relationship” to “the N fixation-AET relationship of Cleveland et al.”

- Figures: not sure that showing the 99% quantile is the best way to go as areas with very large changes do not come out. You could also use a nonlinear color scale.
  The main reason to introduce the cutoff of the color axes in the maps are the extremely high N deposition and N leaching rates in Indonesia, as discussed in sections 3.1.1–3. These high values may not be realistic, and are of less relevance to this study since we are more interested in large-scale patterns. Since using a non-linear scale makes it more difficult to infer the absolute numbers from the graphs, we prefer to keep the color scales as they are.

- Figure 8 and 9: I find it a bit confusing that panel c does not go all the way to 90S. It would be better visually to have the latitudes of the 3 plots match: i.e. if c stops at 60S, then c does not have to be shown over the whole vertical plot and would stop at 60S in panels a and b).
  For all three panels in both figures (and the new figure), the lower limit of latitude is at 60°S.