# Response to reviwer

Kazuya NISHINA

January 30, 2017

Dear anonymous reviewer 2

Thank you for sharing your time to review our manuscript and especially for careful reading. We'd like to respond the individual comment one by one as follows;

> The link to FAOSTAT, which necessarily occurs early and prominently through the manuscript, does not resolve. This represents a major barrier to all readers and users. We need a more reliable and permanent link. Either the authors need to convince FAO to deposit a fixed snap shot of the relevant version under a doi at a reliable site (perhaps too much to ask of any specific users) or the authors should include a version of the FAOSTAT data that they used as a component of the Pangaea resource or deposited under some other data authority. Seriously, it makes no sense to cite this source, and almost the entire paper becomes moot, if a reader has no reliable mechanism to start from the same sources. The FAO link must work reliably now and again two years from now. (As a comparison, the links to FAOSTAT global fertiliser data in ESSD-2016-35, which those authors do not represent as their primary source, do work!)

We agreed on your meaning. In fact, for us, it is difficult to follow the updates in FAOSTAT in detail. But, we cannot find the license to re-distribute FAOSTAT data in public. So, at this time, we don't share the original database in open database. However, we would like to inform that we will share the original data when users contact us. FAOSTAT URL in the citation doesn't work now. But, in the manner of website citation, the date of access and URL should be same in the actual access situation. So, we didn't change the citation in the revised manuscript.

P3L90-92 This statement implies that FAOSTAT includes time series of farmer data (number of farmers?, number, size and types of farms?) at sub-national resolution to allow the authors to successfully manage the changes in national frameworks. We need more information here, to document what the authors used and how the process worked.

Sorry for this. This is typo. The right word is "former" (not farmer). We didn't use the farmer data in this study. So, we revised this in the revised manuscript.

P4 Description and application of the Amelia data imputation package. The authors provide a careful and useful description of what they assumed and how they proceeded. This user also found the documentation for the Amelia R package adequate and helpful. One wonders, however, whether the application of the imputation idea generally and the specific Amelia code to geographic and temporal patterns of fertiliser data represents a unique and creative solution or a misuse. One could argue that fertiliser application data in fact represent social data (a deliberate human intervention) and therefore the use of Amelia seems quite appropriate? Do the authors know of any other applications of Amelia to these types of more geophysical data sets?

P4L112 - What are 'panel' data and why do they fit better with the Amelia assumptions?

The structure of FAO dataset is typical country-years dataset (multi-national statistics with time-series). Amelia is developed for dataset with time units for each of N cross-sectional entities such as countries, where often T < N (Honaker and King, 2010). FAOSTAT seem to be typical case. So, Amelia is appropriate for this type of missing data. In addition, already, Amelia imputation methods have been widely applied not only panel data (e.g., IMF, World Bank, and including FAOSTAT) but also in natural science (as cited in the manuscript) in the previous studies. We thought that the most important thing to the validity to use Amelia in this study is Amelia methods being based on the multi-variate distribution, which enable us to think free from the cause-effect relationship among the covariate. So, we explained the basic concept in detail in the manuscript.

Panel data is a technical term in statistics and econometrics. The term panel data (or longitudinal data) refers to multi-dimensional data frequently involving measurements over time.

P4L115 - 'EM' I suppose this acronym refers to expectation-maximization (as in line 98, same page) but the authors should have defined it there or here?

Thanks. We added the complete expression for EM in the revised manuscript.

P4L119 - Do each of these covariants also come from FAOSTAT?

We added the citation for each item.

P4L122 - "in each was" in each what? Nation?

We added "nation" in this sentence. Thanks.

P4L124 - dividing the fertiliser consumption data for each nation for each year by the maximum fertiliser consumption value for that country for the entire period? (Also, I understand why the authors did this scaling but I believe this manipulation to give only lower than existing values from the imputation process deserves mention as part of the processing uncertainties later.)

Yes, we did so. As you pointed out, this procedure limited the upper by the existing values in the national fertilizer consumption. This procedure is on the grounds that the missing values are likely to be found in the early period (during 1961-2010), whereas the statistic in recent year are usually filled in almost all countries. Actually, N fertilizer consumption increase with time. So, this procedure could not cause inappropriate underestimation in N fertilizer consumption in each country.

P5 Section 2.3 - Here the authors provide a useful description of their process for downscaling and for dealing with double cropping. The final statement of the Section, e.g P5L149 seems vague. Do they mean that they applied the double crop weighting for each appropriate crop region for the entire 1961 to 2010 time period or for each time period of double cropping as specified in the SAGE data base?

Thank you for your suggestion. We revised as follow;

> The double cropping regions were determined from global crop use intensity (CUI) map developed by Siebert et al. (2010). We defined the double cropping regions in our map as those where the CUI was greater than 1.3. In this region, we doubled the crop area (i.e., the weighting $W_{ij}$ were set to be 2 in the double cropping region) for the entire 1961 to 2010 period.

P5L152 - millet oats should be millet, oats?

We revised it. Thank you.

P6L162 - second fertiliser application set to 30 days after initial. However, in Figure 1 on page 16, the authors clearly say 45 days after first fertiliser. The difference of 15 days probably does not make an impact on annual fertiliser usage but the authors should clarify?

Thank you for your comments.
This is typo. "45 days" is correct. I'd like to revise it in the revision.

P6L171 - "one digit inflation for just one year" Does these mean what we might otherwise call 'order of magnitude', e.g. plus or minus a factor of 10?

Thanks. We replaced "order of magnitude" instead of "one digit".

P7L194 - I do not know what the authors mean by "peaky" in this context. Can they give a more precise description?

We used "discontinuous change" in the revised manuscript.

P7 Section 3.1 - Here the authors should provide readers and potential users with a more thorough assessment of strengths and weaknesses of imputation approach and of use of Amelia imputation tool. On the one hand, only 16% of countries had missing data and fertiliser use by those countries accounted for an even smaller fraction of total global use. For these reasons the authors conclude that the outcome of the imputation process seems "reasonable". On the other hand, we know that the authors deliberately constrained the imputation process to only produce univariate outcomes - values lower than existing. And we know, as the authors admit and as Figure 3 clearly demonstrates, that missing data did not occur randomly, either in time or geographically. Under these operational constraints, did the large number of iterations (1000) and the use of independent co-variants (GDP, population) in the imputation process reduce or offset the non-random or univariate biases? We need some assurances, or at least a more quantitative assessment, from the authors. Perhaps they know of other applications of Amelia to real world examples that can help us understand the reliability of the outcomes in this case? Reliable outcomes for "all countries"?

Not "only 16% of countries had missing data". 16% of total dataset (totally N = 9811) is missing in national N fertilizer consumption (actually, 73 countries have missing data in N fertilizer during 1961–2010). From your comment, we recognized the sentence, "The countries reported with missing data accounted for a small fraction of N fertilizer consumption compared with global total fertilizer use.", mislead reader to underestimate the importance of missing data. Actually, 5 Tg-N in 2000 of global total N fertilizer consumption was from missing data in our dataset. And this is not the reason to assure the quality of imputation. So we removed this sentence in the revised manuscript.

In the revised manuscript, we explained the effect of transformation in the revised manuscript as follow;

> In the imputation procedure, we used logistic transformation after the scaling N fertilizer consumptions by the observed maximum value of each country for the period 1961–2010. This procedure has the potential to underestimate the total N fertilizer consumption in each country; however, the missing values are likely to occur in the early period (during 1961-2010), whereas statistics for recent years are usually available for almost all countries. In addition, the N fertilizer consumption increases with time. Thus, this procedure can avoid the imputed values being an unreasonable underestimation.

As in the introduction, Amelia had good track records for the imputation of missing data for panel data in various study fields (We added more citation for application of Amelia in the introduction of the revised manuscript). At least, using Amelia, we imputed missing data in an objective, straightforward manner. They are also sufficient excuses to use Amelia to impute missing data.

However, we should agree this method is not perfect solution. So, to share the issues in our application, we added some explanation and discussions for the failure to impute missing data in our study, as follow;

> However, in some countries (e.g., Somalia and Uzbekistan in Fig. 3), the imputed data did not smoothly follow the time-series of national N fertilizer consumption. This was partially because of the abrupt changes in the observed values of those time-series. Changes unrelated to the covariants (i.e., population, crop area) might reduce the accuracy of imputation of missing data. In

fact, we occasionally observed some artifacts even in the reported values (e.g., the sequence of equal values) for the developing countries. In the process of accounting and compiling various datasets, there are inherent uncertainties in the national statistics (Leip, 2010; Winiwarter and Muik, 2010). Such uncertainties could also affect the quality of imputed data; hence, it should be noted that our dataset includes this inevitable uncertainty.

P7,8 Section 3.2. The authors provide a careful, detailed and very useful comparison with Potter 2010 based on year 2000. I note that authors for ESSD-2016-35 made a very similar comparison. This section would represent a very good place for the authors of ESSD-2016-24 to compare their outcomes to ESSD-2016-35? On first glance, the total numbers for N use seem very comparable? (As mentioned earlier, I believe ESSD editors should ensure that this request to authors of ESSD-2016-24 should apply equally to authors of ESSD-2016-35.)

Thank you for your recommendation. We add the comparison with ESSD-2016-35 in the revised manuscript.

P8, Section 3.3. How does or should the inclusion of these $NO_3$ and $NH_4$ data improve our understanding of temporal and geographic patterns of N fertiliser use. The authors of ESSD-2016-24 could make a few clear statements of the value of $NO_3$ and $NH_4$ data compared to total N approach in ESSD-2016-35?

This is our next research topic. The purpose and significance of the inclusion of the ratio in the N fertilizer input are described in introduction.

We prepared the map of the ratio of $NH_4^+$ and $NO_3^-$ in national N fertilizer input in PANGAEA. This is utilized also in the other N fertilizer map. So, we thought that there are small significance in the mention only of ESSD-2016-35.

P8L261 to P9L264. The textual description of temporal changes in $NO_3$ and $NH_4$ use seem substantially in contradiction to Figure 6, and Figure 6 seems inconsistent with Figure 10. From the text here and Figure 10 we learn that "the total amount and fraction of $NH_4$ increased consistently". But looking at Figure 6, and particularly at the global average portrayed in

Figure 6, we must conclude that the ratio of NH4 input to total N input has stayed above 0.8 for the entire 1961 to 2010 period and with only a very narrow variation across those decades? Has this reviewer interpreted the text or the figures incorrectly? Do we in fact have a contradiction inherent in the data as presented?

Thank you for this comment. We mistook the evaluation of global value in Fig. 6. We have just took average in the ratio among regions (shown in Fig. 6). So, this is not actual values in global. The values in Fig. 10 is correct. We removed this incorrect global value from Fig. 6 and replaced the correct ones (see below) in the revised manuscript.

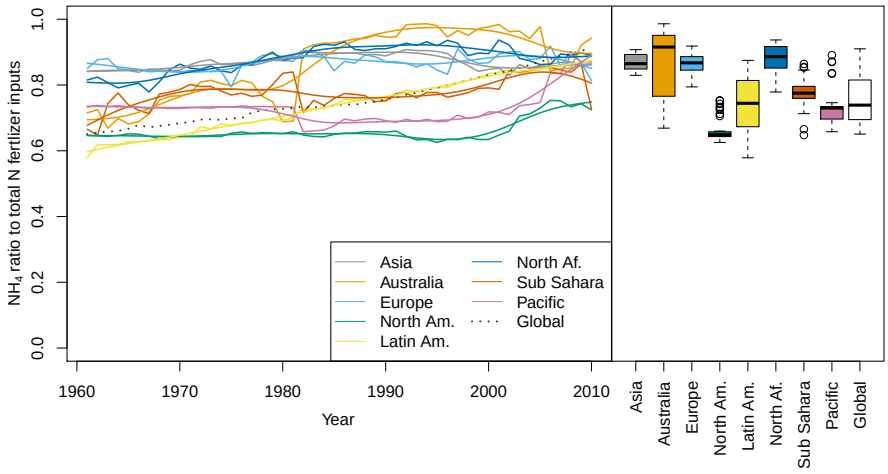


Fig.A Revised figure for the global average

P9L269 - Technically, Figure 11 does not show these total numbers, unless one can integrate the latitudinal data by eye. The authors might consider adding the totals to each panel of Figure 11. In both Figure 10 and Figure 11 the authors should make clear that the values represent the cumulative sum of N inputs while the colours indicate the proportions of NH4 and NO3. I find this overall section quite interesting but the authors might add a sentence or two about the implications of this comparison, to better explain to users why they the users should take an interest in these two sources of reactive nitrogen?

Thank you for your positive suggestion. We added the global total $NH_4^+$ and $NO_3^-$ inputs in both N fertilizer and N deposition for each plot in the revised figure 11. For the $NH_4^+$ and $NO_3^-$, we have already shown colored legends in both figures, however, we added the text explanation in the captions in the revised manuscript.

In Section 3.4, we added a following sentence in the first paragraph to explain the purpose of this comparison.

> N fertilizer and N deposition are the most important sources of disturbance of terrestrial N cycling (Gruber & Galloway, 2008). Although both inputs (finally) comprised of $NH_4^+$ and $NO_3^-$, there has as yet been no quantitative comparison of these two inputs at global.

P9L290 - I think the authors mean that few other sources exist for these kinds of global N data sets. Here the authors might mention ESSD-2016-35 as a comparison?

Thanks. We referred ESSD-2016-35 in this part in the revised manuscript.

P9L292 - Yes, these N data come from national consumption data mapped to crop area, but the process as the authors have described seems quite far from simple.

Indeed, the procedure in developing our map seem to be complex, but N fertilizer rates in each country were determined by simple equation (e.q. 3). For example, our map doesn't consider crop species in N fertilizer rates. So, we wrote "simple" in this sentence.

P10L298 - I think the authors have understated the utility of the NH4 and NO3 data. Those data provide much additional information about national sources and about the time history of use of various forms of N fertiliser.

Thanks.

P10L300 - Yes, the nationally-provided data have uncertainties, but can the authors provide a quantitative estimate to that uncertainty? How do those uncertainties affect the total cumulative use data, e.g. 110 Tg N as in Figure 10. Plus or minus 10%? 20%? In all the figures that follow, only

> Figure 3 displays error bars and then only because the imputation run over such a large number of iterations provides statistical uncertainty information. Either here or in the prior discussion the authors should provide their best estimates of cumulative uncertainty from all sources: original data, the imputation process, the crop area estimate, etc. These uncertainty estimates would prove very helpful for users, especially modellers. The statement here about inverse modelling would absolutely require such uncertainty estimates? I find this overall conclusion weak compared to the large effort the authors have put in to assembling and describing this data set. The authors should emphasise the utility of these data for inverse modelling studies, and perhaps compare the strengths (times series, 0.5 degree resolution) and uncertainties in these numbers to the uncertainties around atmospheric N concentrations. There remains a long and somewhat hidden gap between these N input data and the N2O emissions data used in Thompson et al. Also, Winiwarter, cited earlier in the section, addresses uncertainties in nationally reported greenhouse gas inventories. How do those uncertainties apply or compare here? We need a better summary of uncertainties from all sources.

We absolutely agreed on the importance of your comments, here. How the uncertainty matters in global N cycling is really important issue.

However, we cannot provide the consistent uncertainty framework (and values) for each country and global total consumption from this study, because we just evaluated the uncertainty ranges of imputed values in the missing data. As shown in our dataset, the location of missing data in time-series are highly depending on countries. So, our estimation might understate the uncertainties in N fertilizer use at both global and regional scales. In addition, I did not save the imputed dataset (N = 1000) for the summation (The 95% intervals in the imputed data are not entirely symmetric. We need Monte-Carlo integration to sum up them.).

Also, we'd like to emphasize that the uncertainty issue is beyond the scope of this paper (though this manuscript is not research paper). We need further research to address this issue.

> P16 Figure 1 - The authors should provide explicit reference to the sources of their crop area data, doubling cropping data, etc. Or link more explicitly to the text where they provide those descriptions? This figure needs better documentation.

Thanks. To clarify these resources, we added the references for these data in this caption.

> P17 Figure 2 - For the US, fertiliser application rates rise but total crop area stabilises or declines so total N fertiliser consumption falls slightly - that make sense. But for China, during a time of expanded crop area and increased fertiliser application rates the total fertiliser consumption appears to fall behind, at least for some years? Have I interpreted this plot correctly? Do the authors have an explanation? Mention the similarity to or emphasise the contrast with Figure 3 in ESSD-2016-35?

Thanks for your comment. This lines are illustrated as a relative units, which calculated by the scaled values in each national N fertilizer consumption and national crop area. So, the crossing the lines in this figure does not mean much in this context.

I added the explanation in the caption as follow;

> For both national crop area and N fertilizer consumption, the values are divided by the maximum for this period. So, the unit of fertilizer rates is non-dimension in this figure.

This figure is a just example of procedure and removed all of units for the variables here. For the global average data in ESSD-2016-35, there seem to be no meaningful comparison as far as we thought.

> P18 Figure 3 - Very interesting plot. I believe it conveys a sense of the importance of the covariants because in several cases shown the imputation values clearly do NOT fit the local country time series. The 95% confidence intervals seem quite large in all cases even though of course, by design, they can not exceed values of 1.0; in most cases those 95% CI cover essentially the full range of relative N consumption. Instead of, or in addition to, these specific country examples, could the authors provide a summary of the average error for the 16% of imputation-filled data? This information should help inform the larger uncertainty discussion suggested above?

I'm sorry. We could not understand "summary of the average error for the 16% of imputation-filled data?". Could you clarify this?

> P19 Figure 4 - Not very impressive as log-log plots go. The correlation coefficient and RMSE numbers look good and provide sufficient information. Do we need this plot if we have those numbers in the text or in a small table?

We agreed on your comment. We removed this plot in the revised manuscript.

> P20 Figure 5 - All these numbers presented as absolute, with no uncertainties in either these data or those of Potter 2010?

We clarified the final values for fertilizer map to be obtained the average of the ensemble imputed data in the manuscript. Also, we added the more explanations (how to get, etc) in the caption of this figure.

> P21 Figure 6 - I mentioned already the apparent discontinuity between this data and those presented in the text and in Figure 10. The figure includes some smoothing for each regional data time series, but not explained? Why do so many open circles occur, and why so far above the average, for the North American data which one supposes has reasonably accurate reporting? Explain the boxes: mean plus SE or SD plus max min or quartiles or ...?

Thanks, we revised the discontinuity as explained above.
We added the full explanation for the boxplots in detail.

> P22 Figure 7 - The authors could add a global average number to each panel that should correlate with global data in Figure 6? This figure suggests the small average changes of Figure 6, not the dramatic changes of Figure 10? Pakistan, designated as a hot spot of N use in ESSD-2016-35, does not show up here as particularly important in terms of fraction of $NH_4$ use?

Thanks. We added the global averages in the revised figure.

> P23 Figure 8 - Would we expect to see an offset but repeat pulse in double cropped areas (e.g. of North America or Eurasia) or are the data too smoothed or the application dates too varied? The caption should read "Values represent average $NH_4$ N applied over all crops across each grid cell"?
>
> P24 Figure 9 - Same question as above about repeat pulses of N inputs observed in areas of double crops, perhaps evident here in April / May (first pulse) and August / September (second pulse) for areas of North America and Eurasia? The caption should read "Values represent average $NO_3$ N applied over all crops across each grid cell"?

Thank you for your comments. As refereed in above response, we clarified values are average for the final product in the material & methods section of the revised manuscript.

Regarding this issue, We revised double cropping region according to Dr. Lu's comment. In addition, according to Reviewer 1 comment, we add new figure (Fig. 9 in the revised ma manuscript) to see seasonal variation of N fertilizer application. Then, we added the discussion for seasonal variation as follow;

> Because of the differences in the $NH_4^+/NO_3^-$ ratio among the countries, there are spatio-seasonal differences in the $NH_4^+$ and $NO_3^-$ inputs throughout the course of a year (Figs. 7–9). During February–March, both inputs exhibit a peak in the northern-hemisphere, especially between 30°N and 60°N (Fig. 9). For the tropics (from the equator to 30°N), both inputs are observed throughout the year. In contrast, in the southern hemisphere, $NH_4^+$ inputs are dominantly observed during September and October in Fig. 9. This is because $NO_3^-$ inputs in the southern-hemisphere countries are a small fraction of the total N fertilizer inputs (Fig. 8). For the double cropping regions (Fig. 1), there are second peaks in both the $NH_4^+$ and $NO_3^-$ inputs around 30°N (Fig. 9), particularly in south China and India (Figs. 7 and 8).

P25 Figure 10 - Make clear that this represents cumulative (NO3 plus NH4) total N with fractions of NO3 and NH4 shown by colours. Explain the differences, if any, between data shown here and data shown in Figure 6.

We fixed it in the revised manuscript.

P26 Figure 11 - Make clear that this represents cumulative (NO3 plus NH4) total N with fractions of NO3 and NH4 shown by colours. Authors could add a cumulative number to each panel to give readers a sense of the integrated totals?

We added global total $NH_4^+$ and $NO_3^-$ inputs in each year in the revised figure.

## Reference

Leip, A. (2010). Quantitative quality assessment of the greenhouse gas inventory for agriculture in Europe. In Greenhouse Gas Inventories (pp. 245-261). Springer Netherlands.