Interactive comment on "Continental drivers of ammonium and nitrate in Australian soil under different land uses" by Juhwan Lee et al.

Anonymous Referee #1

Received and published: 14 May 2018

Overview

In this study the authors use a cubist model to decipher the soil parameters influential to mineral nitrogen (NH_4^+, NO_3^-) in Australian soils under different land-uses. The scientific question is valid and the data approach taken is state-of-the-art, so that the manuscript certainly falls within the scope of SOIL. However, the presentation and interpretation of the data lacks depth and specificity and needs significant revision before it can be accepted for publication.

We have addressed all the comments carefully to improve the overall manuscript. Please refer to our responses below.

Comments

Abstract

L5-6: It is unclear what the agricultural region and non-agricultural regions are. This is a large problem for the paper, because samples taken from within the agricultural region may actually be from natural environments, or vice-verse.

Thank you for pointing this out. Now the "ecological" region has been renamed as the "non-agricultural" region throughout the manuscript. Please note that the sample locations were classified by the broad and detailed land uses from the Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES) land use map of 2016, as shown in Figure 1 and Table S1, which were grouped into two regions. However, you are correct that some of the samples within the agricultural production areas are from a natural environment, and vice versa. Nevertheless, given the large amount of samples in this data set, we are confident that the vast majority of samples from either region are accurately described by the given grouping (Table S1). Furthermore, despite these potential issues, which we discuss in the discussion section (P7 L18-19), we believe these general classifications are able to adequately show differences between the two broad land-use type across the continent.

L9: Normally NH_4^+ introduces H^+ into the soil and therefore affects pH, not the other way around.

NH_4^+ concentrations are also affected by pH, particularly during microbial nitrification (e.g., Cuhel et al. Appl. Environ. Microbiol. 2010 vol. 76 no. 6 1870-1878). In this study, we focused on the role of pH as a potential soil control.

L11: what is the 'other' ecological region?

Please see previous response -- we changed this region to the non-agricultural region to avoid confusion.

Introduction

P2, L1: Definition of mineral N?

The definition of mineral N has been added. The sentence now reads: "In addition, different mineral forms of N **(e.g., NH_4^+ and NO_3^-)** often …"

P2, L6: I think it is the mineral N application, not continuous cropping per se that leads to increased N emissions.

We agree that a level of fertiliser N application is closely related to N gaseous emissions. In addition, diversifying the sources of N fertilization, such as N fixed by legumes, may also result in some emissions from cropping systems. Nonetheless, we have modified the sentence and added that emissions are "**mainly through N fertilizer application**".

P2, L13: Mention some numbers/a range of expected mineral N values in soil.

We have checked the references. Viscarra Rossel and Bouma (Agric. Sys. 2016 vol. 148 71–74) estimated NO_3^- contents in the range of 0–200 mg N kg^{-1} for cropping systems, based on the data from www.bfdc.com.au. Unfortunately, we were not able to find a good reference on the regional/continental scale that is compatible with our study. To point out this lack of information, we have added a sentence: "**However, such large-scale soil information is generally not available.**"

P2L19-24: there are also studies showing that mineral N fertilizer can promote turnover of (and deplete) SOC (e.g. Shahbaz et al, LDD, 2017 or Neff et al., Nature, 2002).

Thank you for the references. We have added a sentence to address a decrease in SOC by added N "**Conversely, there are studies showing that SOC can be depleted by N fertilizer application (Neff et al., 2002; Shahbaz et al., 2017)**".

P2, L34: Is biodiversity an ecosystem function? I would think it is an ecological indicator or even property, not a function.

Thank you for picking this up. We agree that biodiversity is not an ecosystem function. We simply intended to infer the status of biodiversity is related to ecosystem functioning. Therefore, we changed "ecosystem functions" to "**ecological properties**" in the sentence.

Materials and methods

P3, L15-16: What is a unique environmental condition?

Please refer to our response below.

P3, L16-17: How was sampling performed? The text about no further samples being resourced is confusing.

The BASE project performed the soil sampling, and one of their criteria was to choose sampling location across a wide range of environmental conditions and unique combinations of soil, climate and management. To make it clear, we have modified the sentence and added the following short description of soil sampling: **"Each sample was collected from a site that represented a unique combination of soil, climate and management. Specifically, between 9 and 30 soil cores were sampled in a 25 m x 25 m quadrat and split into two different depths, 0–0.1 m and 0.2–0.3 m, respectively, and then combined into one composite sample (approximately 1 kg of soil) for each depth."**

Please note that we have deleted "because no further samples were resourced". This simply means that the BASE project had only sampled once and the project ended.

P3 L17-22: 'Spatial distribution. . . at different times of year'. This section isn't methodology, it appears a combination of introductory remarks and discussion. Delete or move to the appropriate sections.

This was one of the important assumptions for our modelling approach so we would rather consider this as a method. One of the common known issues to study large-scale soil mineral N dynamics is a difficulty in representing the study over time because of its temporally variable dynamics. To address this comment, we have deleted the sentence (previously on L22), "However, this source of uncertainty …"

P4 L8-10: It is unclear whether samples from the 'agricultural region' really came from agriculturally used sites or were just located in what the authors appear to have defined to be a region dominated by agricultural use. Looking at the map, there are vast tracts of land that have been defined as the agricultural production region which I would have thought are natural (e.g. nearly all the national parks in the great dividing range). How did you define these 'agricultural' and 'non-agricultural' regions? I am uneasy about such a broad sweeping definition being used to cover

a continent (albeit a 'small' continent). If the samples are only from the agricultural region, does this mean that they were definitely taken from an agricultural land-use? This is a critical point concerning all the results.

We apologise for the confusion. Now the two regions are re-named as "agricultural" and "non-agricultural" to exactly match broad land use labels from the ABARES land use map, published in 2016 (see Supplementary Table 1). Samples from the production region solely came from sites used for any agricultural production activity – i.e., dryland and irrigated cropping and improved and native pastures used for animal grazing – see Table S1. The other samples were collected from sites under conservation and in natural environment, such as national park and residual native cover.

In our revision, we re-defined these regions as: "**(2) the samples from the sites that originate from dryland and irrigated cropping, and from improved and native pastures used for animal grazing (hereafter referred to as the "agricultural" region), and (3) the samples from the sites that were conserved and in natural environments outside of the agricultural production zones (referred as the "non-agricultural" region)**". We believe that these terms are clearer.

The caption of Figure 1 has been changed to: "**Location of 469 sampling sites across Australia, indicated with open circles for soils used for dryland and irrigated cropping, and from improved and native pastures used for animal grazing (160 sites) and closed circles for soils from areas that are conserved and in natural environments outside of the agricultural production zones (309 sites). The dark grey area represents intensive agricultural and plantation production. The light grey area represents agricultural production from relatively natural environments. The white area indicates the non-agricultural region.**"

P4, L13: briefly mention the depths again. Or state 'two sampling depths' or similar.

Thank you. We have stated **'two sampling depths'** as suggested.

P4 L14, L20: I would have thought that a tree-based model such as cubist does not require log-transformation or the assumption of normality.

This is indeed correct. However, a tree-based model consists of linear regression models, which would be more robust when the values are log-transformed.

P4 L20, L23: Which 'selected' soil properties? A table with all the predictors used in the models would be nice.

Thanks for the comment. All the soil properties used in the models are listed in the section 2.1. Nevertheless, it may be a good idea to repeat the list again. We have revised the sentence to, "**The concentrations of soil NH_4^+, NO_3^-, TOC, TN, TP, CEC, and the fraction of sand, silt, and clay, BD, and AWC, except for pH, were …**"

P4 L27: A brief explanation of number of committees and nearest neighbours parameters would be helpful.

We fixed the number of committees to one, although we have tested and reported the importance of soil controls in more complex models in supplementary figures (with 5, 10 and 20 committees). We have added "**with 3-9 neighbors**" to the sentence, "The optimized models ...".

P4 L29: I dislike this use of the term relative importance. Your results show that several variable have a relative importance of 100 % in the same model. This implies that total importance is > 100 %. Relative importance in other models is frequently based upon influence of the predictor on model accuracy/goodness-of-fit, not solely on its inclusion in the model. Perhaps just use 'importance'?

We fully agree and thus have corrected the term as suggested throughout the manuscript. Perhaps, we may have to re-define the term importance of variables from Cubist modeling. Please note that this is how 'relative importance' is currently defined without a full consideration of model coefficients.

Results

P5 L13: 'regional patterns' I do not think this is an appropriate use of 'regional'. Normally, a region is an area which is spatially defined because it is smaller and belongs together (e.g. Gippsland, the Hunter Valley, The West Australian Wheat belt). Looking at Figure 1, both the ecological and the agricultural production regions cover disjoint areas which have completely different climates, geology, vegetation among other things. I do not feel comfortable

clumping these 'regions' together unless the samples in the regions are really defined by land-use, in which case you should refer to land-use categories, not regions.

We hope we have addressed this concern by re-naming the regions according to broad land uses. All the samples in each of the regions are strictly defined by land use.

P5 L15: 'or the median of' - rephrase.

Based on the comment from the other reviewer, we have reported the median values instead of the mean values. Exception was when we reported mean comparisons. So, this has been deleted.

P5 L19: 'significant regional difference' - be specific.

This has been corrected to "**a significantly higher NO_3^- fraction of TN in the agricultural region than in the non-agricultural region.**"

P5 L21-22: Do you mean that NH_4 and NO_3 differed between land-uses?

Yes, that is correct. We clarified the sentence to: "**The contents of NH_4^+ and NO_3^- differed by both broad and detailed land uses (Table 1).**".

P5 L23-24: 'relatively natural environments' - What specific land-uses or environments are you referring to?

We hope that this has been resolved with our new terms with the new definitions.

P5 L24-25: But the table does show e.g. significantly higher NH_4 in 'habitat/species management'. Make sure the text fits the results you present.

For this land use category, the concentration of NH_4^+ appeared to be high among the detailed land uses. However, it was not significantly higher except for cropping, grazing native vegetation, national park, and other conserved areas. We revised and corrected the text from "little or no differences" to "**no apparent differences**".

P5 L30: Which soil properties, what were the effects? Be specific.

These soil properties were listed as "**TOC, TN, TP, the clay fraction, CEC and pH**". We deleted ", except the sand and silt fractions, BD and AWC" to be specific as suggested.

P6 L1: Once again 'soil properties'. This is too vague as to be meaningful. Be specific.

We revised the sentence from "by soil properties, similar to the controls on soil NH_4^+" to "**by similar soil controls as for NH_4^+**".

P6 L28-31: This is discussion, not results.

Thank you. We agree and we have deleted this sentence form the results.

Discussion

P7 L9: 'may suggest' - or just suggests? No need to be so hesitant to make a statement.

Thank you. We were being conservative in the interpretation of our results because of the limitations of the data, but we agree and have now used 'suggests'.

P7 L14-15: Now you are talking about land-uses, not regions. Please be consistent.
What about the effects of different climate and geology. These have a massive influence on soil forming factors, as well as vegetation. I would have thought that these factors could be accounted for in models. Even if they haven't, they should be given some thought in the discussion.

We revised the sentence "under different land-use conditions" to "**in different regions defined by broad land uses**" for consistency.

We agree with the reviewer's general point about different environmental factors, however, as described in section 2.1, the data from the soil maps, which we used, already accounted for climate, mineralogy, etc. because these covariates were used to derive the soil maps.

P7 L16: This seems to be presentation of new results in the discussion section.

These are not new results. To make this clear, we have revised the sentence to "**We found complex, but consistent regional patterns of soil NH_4^+ and NO_3^- by broad land uses.**"

P7 L18-19: This is the crux of the issue with this paper - it is unclear how you defined agricultural vs. ecological regions, so you are not really comparing land-use effects. In fact, it remains unclear to me what you a comparing, given the very large areas covered by the 'agricultural' and 'ecological' regions, which cover vastly differing climates and site environments.

We hope that our revision of the definition of these regions has been resolved this misunderstanding.

P7 L26: What is a soil disturbance level?

We have revised "at various levels" to "**and various tillage intensities**".

P8 L18: I find this explanation for the lack of climate consideration inadequate. You can easily download at least broad climate data from the BOM and could have considered this in the models. You have clumped sites from Tasmania together with sites from far-north QLD, which have vastly differing climates. Even something as broad as a Köppen climate classification may have been considered in your models, if you do not have access to something more specific.

We disagree with this comment. We know that the continental climate data are available, and we have them and used them to derive the soil maps in a previous work, Viscarra Rossel et al (2015). Here, we are referring to the current BASE data. Please also note that we mentioned that "most of the sites were located in arid or temperate ecological zones". We agree that effects of climate are potentially influential for soil mineral N distribution but probably as distal controls (not as proximal controls). If we had access to soil-water or soil temperature data, this would be definitely be something to consider in our modelling – but we do not.

P8 L30: Sequentially?

We have deleted it.

P8 L31: Do not give examples. You should discuss the actual results.

We have changed "For example," to "**It is well known that**".

P9 L4: Do not give one example, discuss your findings.

We have removed "(e.g., TP)" as you suggest, and from our point of view have adequately discussed the results based on the objectives of the paper. Here we highlight different soil controls in different regions or by dominant land uses at the regional scale.

P9 L5: The model explained variance was much lower for NH_4^+ in the agricultural soil, potentially indicating that you have not included all the driving factors in your models.

Having more variables in a model might increases a proportion of explained variation, if these are related to the response. However, the lack of explained variance may also indicate that NH_4^+ in agricultural soil is highly variable and at much shorter distances, based on a sample variogram calculated from the data (results not shown). Unfortunately, we were not able to assess whether the problem was with the spatial variability as we had not enough data to derive models of the spatial variation (i.e., variograms).

P9 L5-10: Which effects - specifically state what you think the relationship is? See comment above on relationship between pH and NH_4^+. Make sure you discuss your results, not just reiterate them.

We have added "**positively**" to show how NH_4^+ was related to TN. We have discussed our results whenever reasonably allowed at the regional scale (see on L8-10).

P9 L10-13: 'Was affected by' - this is too vague. Be specific. What was the relationship? Positive or negative? Are you sure this is causation, or is it merely covariance. . .?

We have added "**positively**". For sure, we 'assume' causation in our modelling approach.

P9 L24-25: What about the error arising from a lack of consideration of other factors driving soil processes (climate, geology, topography. . .)?

The factors mentioned by the reviewer and those not mentioned are all potentially important. Please note that soil maps provide an integrated measure of climate, geology, terrain, etc. Here, we have not directly but indirectly accounted for these factors. Please also note that our main objective was to determine continental/regional soil factors. Thus, this point of discussion is out of our current scope.

P10 L9: Where is land-use intensity presented?

We have deleted "intensity of".

Conclusions

I do not think you actually looked at land use and management in depth (except the results presented in Table 1, which for NH_4 appear insignificant across the broad categories). You defined agricultural and ecological zones, but the way this was done is unclear.

All comparisons are based on broad land uses, not so much on detailed land uses. As shown in Supplementary Table 1, we were not able to address at the level of each of the specific land use types.

P10 L34 (35?): For me, your results do not indicate regionally explicit soil controls, but I find your definition of these regions problematic.

In this study the goal was not to understand differences in different geographic regions, but instead under different land-uses (i.e. agricultural versus non-agricultural) across the continent. Hopefully this is more clear given our new terminology.

P11 L4: 'it was probably due to' What are you referring to with it?

Here, "It" was clarified as: "**the effects of total soil elements**".

P11 L5: which complex biophysical properties?

We have modified "biophysical properties" to "**soil properties**".

Tables and Figures

Figure 1: Throughout the manuscript you contrast the agricultural region with the ecological region, but the map shows three regions 'intensive agricultural and plantation production', 'production from relatively natural environments' 'other ecological region'. It is unclear how you have defined your regions.

The caption now include the definition of the "agricultural" and "non-agricultural" regions. These regions have been re-defined in the section 2.2.

Figure 2: I think a box-whisker plot would be much better here.

We have made a new figure as suggested.

Supplementary Figure 1: You state on P4 L27 that the number of committees was set to one to avoid complex models. Why does the y-axis show up to 20 committees? The caption mentions grey bars, but there do not appear to be any.

For this figure, the caption has been changed as: "**Supplementary Figure 1. Importance of soil properties as the predictors of NH_4^+ contents (mg N kg^{-1}). The importance of the predictors is based on the usage of each variable in the Cubist model (black bars). None of soil properties is used to set the rule conditions.**"

There is no grey bar because none of the soil properties was used to set model rules. We reported the results from model with only one committee. Please note that this is the supplementary information, as described in the section 2.1 on P4 L31.

Supplementary Figure 2: You state on P4 L27 that the number of committees was set to one to avoid complex models. Why does the y-axis show up to 20 committees?

Please refer to our response about the comment on Supplementary Figure 1.

Please note that we have made a new Table 1.

Additional technical corrections:

We have added the units of measured soil properties of the BASE data and values from the soil maps in the section 2.1. NH_4^+ and NO_3^- were reported in mg N/kg. Organic C, total N, total P, and texture were reported in %. Bulk density and exchangeable CEC were reported in g/cm^3 and meq/100 g, respectively. Thus, we have corrected the term "concentration" to "content" for soil mineral N as determined by the mass of sample throughout the manuscript.