Point by point response to Referee#2 comments

Dear Referee#2,

We would like to thank you for the thoughtful comments on our manuscript "**The effect of natural infrastructure on water erosion mitigation in the Andes » (**soil-2021-76). We are pleased to read that you appreciated the systematic review, and the fact that we included supplements that will summarize the underlying datasets.

We take your point that it is necessary to clarify the intervention and control types that we analyzed in this paper, and to introduce them adequately in the framework of Nature-based Solutions (NbS). We will further elaborate the discussion on the NbS, and rework the conclusion so that it is in line with the research questions posed in the introduction.

Please find below a point-by-point response to all the concerns raised and how we will address them. Your original comments are copied below, and we have highlighted in blue color how we will revise the manuscript accordingly. We hope you find our response and changes to the manuscript satisfying and we are looking forward to hearing from you.

Major comments

In general, the manuscript is very well written and is well-structured. The objective and research questions are well-defined, but could be better integrated into the Conclusions (see below). The methodology is clearly described and I'm happy that the authors included several supplements, which gives confidence in the robustness of the study. The main critique I have is the way the different interventions are presented in the manuscript, which is often not very consistent. In the abstract (lines 20-21) and the Introduction (lines 76-81), the authors refer to three types of natural infrastructure, i.e. protective vegetation, soil and water conservation measures and adaptation measures that regulate flow and transport of water. The latter category seems to be related to reservoirs and other hydraulic structures, which is often referred to as gray infrastructure. It would be odd to include such interventions in this study given the focus on natural infrastructure. However, in the Material and Methods, a different classification is given (lines 154-156), which includes protection of natural vegetation, forestation and soil and water conservation measures. These are subsequently used throughout the rest of the manuscript. So at the end, no gray infrastructure is included in the analysis. Please revise the abstract and Introduction so that the same categories are used throughout the entire manuscript. Also please revise the terminology used for the different interventions, because there are some inconsistencies throughout the manuscript. And keep a clear difference between the control and intervention categories, sometimes all six categories are discussed as if they were they were all natural infrastructure or interventions.

We understand that the categorization of the interventions might have been confusing, and we will revise the text. In the original concept of the study, hydraulic infrastructure that regulates flow and transport of water was included because they are de facto considered like "natural infrastructure" in some Andean countries such as Peru. This concerns small pounds (e.g. "qochas") or irrigation systems (e.g. "amunas") often inherited from traditional knowledge. We understand your concern of having them categorized as "natural" rather than "gray" infrastructure.

As only very few of these studies on hydraulic infrastructure included quantitative measures on soil erosion/sediment yield, it was not relevant to keep this category in the quantitative analyses. Therefore, we focused the analyses on the categories for which we had sufficient datasets, i.e. protective vegetation (PRO and FOR) and soil and water conservation measures (SWC).

We will revise the terminology accordingly, and mention only those control and intervention categories for which we had sufficient quantitative data for statistical analyses.

The objective of the study is to quantify the effectiveness of natural infrastructure and in several instances a reference is made to the control-intervention design of the majority of the studies included in the systematic review. With this in mind, I do not understand why the authors present the results of the differences between two control types in Figure 5 (i.e. cropland vs. rangeland, cropland vs. bare soil and rangeland vs. bare soil). These results are not in line with the objective. For clarity and consistency, I suggest to remove those results or show them in a supplement, and focus the results on the difference between conventional agricultural practices and natural infrastructure.

For the analyses of the full-text papers, we used a control-intervention design as is commonly done in environmental assessments. In the Andean region, interventions are often realised on degraded agricultural land. This land can be used for crops or pastures; left as bare fallow or abandoned (Zimmerer, 1993; Henry et al., 2013). We hypothesized that the effectiveness of the intervention will depend on the initial state of the land. That is the reason why we kept traditional agriculture (cropland and rangeland) and bare land (bare) as separate categories of "control" in our analyses.

We acknowledge that this might have caused confusion, and will revise the text accordingly. In addition, we will reorder the categories in the figures, so that we have the different interventions (PRO, FOR, and SWC) grouped in one part of the graph, and the three different control types (RANGE, CROP, BARE) in the other part. We will also give the average value for all interventions, and all control types. In Figure 5, we will better differentiate intervention from control.

The authors mainly focused on the 6 indicators as shown in the figures and tables. Apart from that, the authors recorded some other information regarding climate, soil type, land use, among others. It would be interesting to see how these variables affect the results. This means an additional analysis, but could explain some of the uncertainty. Since there are not many studies to perform such an analysis, the authors could aggregate all categories into two categories, i.e. natural infrastructure and control.

We agree with the referee that it is worthwhile to analyze whether effect sizes (Figure 5) vary with climate, soil type, and land use history. In the original study, this was not done because of the limited number of observations per category that prohibited subdividing the dataset per climate, soil type or land use (history) classes. For some indicators (like Sloss, SOC or BD) where more datasets are available, such an analysis might be possible when we group interventions and control types. We will elaborate this point in the revisions.

Regarding the discussion of the results. The interventions studied in the current study are, of course, not only being studied in the Andes, but also in other geographical areas. I suppose similar studies have been performed in other regions. It would be interesting to discuss how the effectiveness of these interventions in the Andes compare to similar interventions in other geographical regions.

We agree, and will include in our discussion studies from other geographical regions with similar environmental conditions.

Regarding the Conclusions. The authors have defined three research questions in the last paragraph of the Introduction. This is very convenient for the Conclusions, because you can just literally answer these questions here. However, I have the feeling that the Conclusions mainly focus on the second research question. The other two questions are somewhat discussed, but could get some more attention. I suggest to reduce the conclusions regarding the second research question, mainly focusing on the most relevant results, and answer the other research questions, at least more than has been done now, i.e. in lines 390-394 and 412-414.

Thank you for this suggestion. We will rework the conclusion so that the three research questions are adequately addressed in the text.

Specific comments

Line 17: I suggest to first introduce the term "natural infrastructure" before using them in these research questions.

Ok, we will do so.

Line 24: Do these two values (1.3 and 2.8) belong to, respectively, the two categories protective vegetation and soil and water conservation measures? Please clarify in the text.

We will clarify this.

Line 25: Again a range of values is provided, please be more precise about where this range is based on.

We will clarify this.

Line 73: Add "as" between "such" and "peatlands".

Ok, we will do so.

Lines 73-74: According to the website of the IUCN, natural infrastructure or natural water infrastructure is considered to be a Nature-based Solution (NbS). I suggest to introduce NbS here, because it is an emerging topic that most readers are familiar with. Lines 74-76: This is actually how Cohen-Shacham et al. (2016) refers to NbS. See also previous comment.

We agree with these statements, and will refer to NbS following Cohen-Shacham et al. (2016).

Line 110: Please specify which search fields were used, e.g. Article title, Abstract, Keywords, etc.

Ok, we will do so.

Lines 129-132: The second criteria should explicitly suggest that modelling studies are included or excluded, now it is a bit unclear if modelling studies are considered or not.

We will clarify this in the text. Modelling studies were only included in the analyses when the models had been fully calibrated and validated using field experiments from a similar geographic and environmental setting.

Line 137: So from 813 studies the authors went to 190 studies based on the inclusion and exclusion criteria (lines 129-134), but how were the 53 studies excluded. Here it seems that these were excluded based on the same criteria. Please clarify in the text.

From the 190 studies, 53 were excluded as they studied only landslides or landslide erosion, and did not contain measurements of soil erosion by water or soil quality. We will clarify this in the text.

Line 145: Why is altitude not numbered?

We will correct this. We considered it to be part of the coordinates (X,Y,Z) but will make the necessary changes.

Lines 151-152: The latter two categories overlap, i.e. large catchment (> 1000 km2) and landscape scale, which is not defined by a study area size. Please be more specific about the difference between these two classes.

We will clarify this in the text. We made the difference between studies that are organized per plot, per drainage basin (small and large catchment), and studies that contain field measurements that are taken over a larger geographical area. The latter can include measurements from different catchments.

Lines 154-157: These are the three natural infrastructure categories as defined earlier (lines 94-96)? It seems that the previous three categories were defined differently, i.e. from the previous categories PRO and FOR would be included in category 1, SWC in category 2 and category 3 is not included here. Please clarify in the text.

We agree that this was confusing, and will rework the text accordingly. See our reply to the main comment above.

Lines 182-183: Did the authors use specific software to extract the data from figures? I suppose that tables were also used.

To extract data from figures, we used the software "PlotDigitizer". Information from in-text tables and supplementary material was copied and tabulated in spreadsheets.

Line 217: Please replace "both" with "all three of them", or similar.

Ok, we will do so.

Lines 217-220: The most southeast located study area doesn't seem to be located in the Andes. Please clarify why this study is nevertheless included. Or show in the map what area is considered to be the Andes.

A similar comment was made by Referee#1. We agree that this study comes from an area that is far from the other study sites (and might represent another environmental context). We have removed this site from the systematic analysis.

Lines 243-244: Here the conventional/traditional agricultural practices are also included under natural infrastructure. To prevent confusion, please rewrite this sentence.

Correct, we will rephrase

Lines 243-246: This sentence needs to be revised. I suppose the authors are referring to the results of the Kruskal-Wallis test, as shown in Figure 4, not in Table 1. Please be more specific about which results is discussed and especially refer to the statistical test.

We're referring to the results of the Kruskal-Wallis test that are shown in Figure 4. In the text, we will better differentiate the results of the Kruskal-Wallis test, and the posthoc comparison test (shown in Table 1).

Lines 261-262: Where is this results shown? Pastures and native grasslands are included in any of the categories? Please clarify in the text.

These results are shown in Figure 5 where we report the effect size for restoration and protection of native vegetation (PRO) compare to rangelands/pastures. The category "PRO" includes native grasslands. We acknowledge that this might have been confusing, and will rephrase the sentence accordingly, and refer to Figure 5.

Lines 277: With natural vegetation the authors mean the protected areas (i.e. PRO)? Please be very consistent with using the names of the categories. To prevent confusion, I suggest to always use the same name and/or use the abbreviation.

Correct, with areas covered by natural vegetation, we mean the protected areas (PRO). We will revise the text, and systematically use the same name/abbreviation for the categories.

Lines 279-280: Is there a separate statistical test performed on the differences between the intervention and control categories? I could not find this in the manuscript. Please clarify in the text.

This information was not included in the original manuscript, as we focused on the comparison between the six categories. We will add the results of the comparison [all intervention/ all control] in the text.

Lines 304-305: But Figure 5 includes comparisons of 1 and 2 independent case studies. I suggest to indicate in Figure 5 which comparisons are considered and which not. For instance, the comparisons that are considered could be indicated with a darker color and the ones that are neglected with a lighter color (or with some level of transparency).

Thank you for this suggestion. We will adapt the figure, and show the results that are based on < 3 case studies in lighter color.

Line 326: Why are two values shown here, to what are these two values referring to. Please clarify in the text.

These values are the effect sizes of SWC compared to croplands and rangelands under traditional agricultural management. We will clarify this in the text.

Line 329: Where are SWC compared to grasslands? In Figure 5 the categories are compared with either cropland or rangeland, not with grassland. Please clarify in the text.

Many thanks for noting this error, as it should read as "rangelands" and not as "grasslands". We will make the necessary correction in this text.

Lines 300-343: In Figure 5 the comparison is made between the categories and cropland (upper) and rangeland (lower). This subsection is mainly focused on the differences with cropland and rangeland is only mentioned in one paragraph (lines 307-312). I was under the impression that the authors would

compare the intervention categories (PRO, FOR, SWC) with the control categories (CROP, RANGE, BARE). Which is also much better in line with the objective of the study. So why did the authors compare, for instance, rangeland and bare soil with cropland? And why are the intervention categories not compared with bare soil, in a similar way as has been done with cropland and rangeland?

We take your point, and appreciate this suggestion. In fact, we made the comparison BARE-CROP and BARE-RANGE, as this is often done in soil erosion studies based on e.g. Wishmeier-type experiments (De Noni et al., 2001). The number of case-studies that make this comparison is rather high, which allowed us to do statistics on the effect sizes.

However, we agree that it would be more relevant for this study to analyse the effect size of the three intervention types compared to the control types, even if this is based on a lower number of case-studies. We will re-organise the presentation of the results, and make the comparison between the three intervention categories (PRO, FOR, SWC) and the three control categories (CROP, RANGE, BARE). This will result in a new figure with 9 panels (3 by 3).

Line 369: With "both" the authors mean "simultaneously"?

Correct, we mean studies that have looked at water erosion at the catchment scale, and that at the same time realised measurement of soil erosion at the plot scale in the same catchment. We will adapt the text to clarify this.

Line 383: I think that this should be 43%, instead of 40%.

We wrote "more than 40%", but will replace this in the text by "43%" which is the exact number.

Lines 392-393: These two values (i.e. 85 and 125) are referring to the previous sentence? Please combine these two sentences into one sentence to know where the values are referring to.

We take your point, and will combine the information of the two sentences into one sentence.

Lines 403-407: The authors frequently indicate the effect size results as a range between the minimum and maximum value of the error bars in Figure 5. I think it would increase the readability if the authors would indicate the mean response ratio in the text, rather than the spread of the error. For instance, as has been done in the subsequent two sentences.

The two values that are reported are the mean effect sizes compared to cropland and rangeland under traditional agricultural management. We acknowledge that this might have been confusing, and will make the necessary corrections. In the revised text, we will better separate the effect size compared to cropland, rangeland and bare soil; and discuss them separately as the effect size depends on the initial state of the land.

Figures and Tables

Figures: Please increase the font size of the smallest font.

Ok, we will do so.

Figure 3: I suggest to include the names of the countries in the map, for those readers that are not too familiar with the topography of South America. Please, include a reference to the DEM used as background for the map.

Ok, we will do so.

Table 1: Please replace the last sentence with something like this: “The box-plots followed by a common letter are not significantly different by the Dunn’s posthoc test at the 0.05 level of significance.” See Piepho (2018) for an interesting discussion about the meaning of these letters and for suggestions on how to refer to them in table and figure captions.

Thank you for this suggestion, and for the reference to this paper. We will make the necessary changes to the text.

Supplement

Supplement C: I highly encourage to include the underlying data as a supplement to this study. However, it seems that some information is missing in this spread sheet. It would be useful to include a separate sheet where the different codes (e.g. for the ecosystem services, natural infrastructure and treatment) are explained. Also the actual data is missing, there are no values included on the right side of 5. Indicators. Please explain why these have not been included.

We will provide the supplementary material with the underlying data when the paper is accepted for publication.

References

Piepho, H.-P.: Letters in Mean Comparisons: What They Do and Don’t Mean, Agron. J., 110(2), 431–434, doi:10.2134/agronj2017.10.0580, 2018.

Thanks for this suggestion. We will check the document.

References

Cohen-Shacham, E., Walters, G., Janzen, C., and Maginnis, S.: Nature-based Solutions to address global societal challenges, Gland, Switzerland: IUCN, 97pp., 2016

De Noni, G., Viennot, M., Asseline, J., and Trujillo, G.: Terre d’altitude, terres de risque. La lutte contre l’érosion dans les Andes équatoriennes, IRD éditions, Collection Latitudes 23, Paris, France, 2001.

Henry, A., Mabit, L., Jaramillo, R. E., Cartagena, Y., and Lynch, J. P.: Land Use Effects on Erosion and Carbon Storage of the Río Chimbo Watershed, Ecuador, Plant and Soil., 367(1–2), 477–491, https://doi.org/10.1007/s11104-012-1478-y, 2013.

Zimmerer, K. S.: Soil Erosion and Labor Shortages in the Andes with Special Reference to Bolivia, 1953-1991: Implications for ‘Conservation-with-Development’, World Development., 21, 1659–1675, https://doi.org/10.1016/0305-750X(93)90100-N, 1993.