

Review of manuscript of manuscript NHESS-2013-482:

“How useful and reliable are disaster databases in the context of climate and global change? A comparative case study analysis in Peru”

by C. Huggel et al.

1) General comments

Dear Editor, dear Authors,

This is a well-structured and fairly well written MS on a comparative analysis of three different disaster databases in Peru. The authors detect considerable differences in the contents of the databases namely in the number of reported events, and the “rate” of affected population. The study suggests that a database has to be carefully evaluated before it is applied in a study and that the selection of a database strongly depends on the temporal and spatial scales of interest. The conclusion that reporting criteria in for any database should be clearly defined and the documentation strategy improved is important but also something that has been stated many times in previous studies.

Furthermore, study results show limited changes in event occurrence in two regions of Peru. However, I am not sure if solid statements can actually be made regarding damage occurrence and damage parameters, given the considerable inaccuracy of the applied databases.

All in all this is an interesting descriptive study that will be of interest to the natural hazards community (and should be published) mainly in the fact that damage databases will only improve – especially in developing countries – if they are thoroughly investigated and their data applied.

The MS is adequate in length although rather on the long side for a case study. Especially the discussion section is in some places repetitive and a bit too long for the limited number of statements made.

I suggest the authors thoroughly address all the specific comments and technical details listed below. Given the large number of minor comments, and a few comments with a more “major” character, I suggest that the paper be accepted pending moderate revisions.

2) Specific comments and technical corrections

→ Abstract

- P4332-L1: The term "loss" is used throughout the text in several different ways. What is meant here exactly? Fatalities (loss of life)? Or does "loss" also include financial losses. If it does, what is the difference between financial losses and (financial) damage? Please try to be as precise as possible and consider changing to "Fatalities and damage caused by..." here.
- P4332-L7: Consider changing to "...of different databases...".
- P4332-L13: Is "regional" the adequate term here? There is potential for confusion; maybe you could use "multi-national", "continental" or "regional-continental", or something similar.
The problem arises again in section 3. There you define: [local level < national level < regional level] at line 3 and again line 19 of page 4337. But throughout the MS, the term region (or "Regiones") is also used to describe administrative units within Peru (at sub-national scale).
- P4332-L15: Something seems wrong in this sentence and I would recommend another word choice; what is "the number of disasters occurrence"? Do you mean the number of (single) events/disasters?
- P4332-L20: Consider changing to "...but strong positive trends in people affected..."
- P4332-L21: What is meant by "disaster parameters"? Above at line 16 the authors use "disaster categories and metrics"; This is a bit confusing.

→ Introduction

The Introduction is well organized and the aims of the study are defined.

- P4333-L8: Consider using "supported" instead of "pushed"
- P4333-L8/9: Provide date of "the last international climate negotiations", if possible
- P4333-L12: Consider changing to "Databases are a primary source and tool..."
(Avoid beginning and ending a sentence with the same word)
- P4333-L18/19: Consider providing website "(<http://www.munichre.com/en/reinsurance/business/non-life/natcatservice/index.html>)"
- P4333-L19: Change to "...is the largest global database..."; and delete "global" at the next line
- P4333-L22: Change to "...researchers have increasingly started to..."
- P4333-L28: What do "losses" represent in this statement? Fatalities, financial losses, or both? C.f. comment above

- P4334-L3-5: Delete paragraph mark at line 5, a three-line paragraph doesn't make much sense
- P4334-L8: Change to "Studies on changes of disaster events and losses are much more rare in developing countries than in developed countries..."
- P4334-L15/16: Change to "...insufficient research comparing different ..."
- P4334-L16/17: "...and implications of the respective analysis...": The authors should consider rephrasing this part of the sentence; it is not clear to me; in any case, change "analysis" to "analyses"
- P4334-L25: Consider dividing into two sentences: "This is corroborated by a recent comparative review of country-level and regional disaster databases by the United Nations Development Programme (UNDP). This study found, for instance, that more than 50% of the databases analyzed ...". Try to make short, concise sentences
- P4334-L28/29: This sentence seems rather trivial to me and I do not think you need it here; consider deleting it; accordingly at line 29 change from "This shortcoming is..." to "These shortcomings are..."
- P4335-L5: What is the difference between an "extreme event" and a "disaster" (or "disaster event", as used e.g. at P4333-L11)?
In the Abstract the authors mainly use the term "disaster"; for me a "disaster" is an exceptionally bad (e.g. natural hazards) event. And in my opinion many events such as floods and landslides can create considerable damage costs without really being disastrous. Is there a threshold that has to be exceeded for an event to be a "disaster"? Or can some sort of definition be applied? Does a "disaster" imply at least a fatality? Please clarify.
- P4335-L11/12: Same comment on the definition of "disaster"
- P4335-L11/12: Consider changing to "Finally we analyze the disaster database entries and climatic conditions..."

→ Study region

- P4336-L4: Change to "...from elevations of less than 300 m a.s.l. ..."
- P4336-L17/18: Consider rephrasing this sentence, e.g. "...with poor infrastructure, educational, social and health services, and subsistence farming dominates."

→ Data and methods

- P4337-L3: See comment made above regarding the use of the term "regional" (P4332-L13)
- P4337-L9/11: Please use consistent spelling for "subgroups / sub-groups"

- P4337-L17/19: See comment made above regarding the use of the term "regional" (P4332-L13); you should somehow indicate that "regional" means "multi-national" here.
- P4337-L23: Consider changing to "...and developed the concepts and methods..."
- P4337-L24: Purely stylistic, consider "... relying on existing newspaper..."
- P4337-L26/27: Delete paragraph mark at line 26, dividing information on DesInventar into two paragraphs doesn't make much sense
- P4338-L6: Rephrase please, e.g.: "SINPAD consists of inventories of events since 2001..."
- P4338-L6: Specify "events" here. What is inventoried in SINPAD? Weather and climate related disasters, or is the database broader?
- P4338-L8/9: Consider changing to "categories such as number of fatalities and people affected, infrastructure damage, total surface area affected, etc. (INDECI, 2013)."
- P4338-L17: Please change to "From all three databases..."
- P4338-L20-24: I am skeptical about this categorization. I understand that hail and snowfall can cause direct financial damage, but what kind of damage does heavy precipitation cause which isn't an inundation (due to a flood in a water course or due to overland flow) or a rain-triggered mass movement? Is it really reasonable to have a category defined as "precipitation events"?
Table 3 reveals that e.g. during the year 1973 approx. 20 Rainfall events occurred. I wonder: what happens during a "climatic disaster, of the type Rainfalls"? I would reconsider this definition of categories.
- P4339-L4/5: This sentence ("A geographically referenced...") is somehow repetitive, consider deleting it.
- P4339-L10-15: This sentence is too long. Consider starting a new sentence after "... in statistical terms or not."
- P4339-L18/19: Consider changing to
"For the analysis of damage metrics (people killed and affected) over the past four decades, we used DesInventar and EM-DAT at a national scale for Peru."
- P4339-L24/25: Consider changing to
"...than the multi-decadal analyses described above."
- P4339-L29 to
P4340-L4: Consider changing to "Climatic data for the "Regiones" of Cusco and Apurimac was derived from a portal... ..of the Peruvian Meteorological and Hydrological Service (SENAMHI) (Schwarb et al., 2011).

→ Results

The Results section is well organized and of adequate length although in some parts a bit repetitive (see e.g. comment for P4341-L21-23).

Section 4.1 is based on the number of events (or disasters) occurred in the regions of Cusco and Apurímac. Changes in event occurrence are investigated. I think this is quite problematic, unless every event caused approximately the same amount of damage (which is probably not the case). I am not sure if this comparison or analysis is meaningful at all since e.g. one very large event can easily have more impact than say 10 or even more small events. I understand that the applied damage metrics are not sufficiently well documented to allow for a quantification at the scale of regions or provinces and that data on financial damage is scarce/incomplete. But I think nonetheless that this problem should be briefly addressed here or in the Discussion section.

- P4340-L14: Change to "For the "Region" of Cusco..."
- P4340-L14: Consider deleting "on DesInventar"; it is repetitive (P4338-L25/26)
- P4340-L15: You absolutely have to note somewhere that the last "decade" of your analysis does not include 10 years but only 9 (2000-2009)
- P4340-L18: Consider changing to "No significant increase in disaster events per decade can be noted..."
- P4340-L21: Rather than "analysis" (a term often used in this MS) I would speak of "representation" or "diagram" here.
- P4341-L3: Change to "For the "Region" of Apurímac ..."
- P4341-L8-10: Change to
"Most affected "Distritos" include Abancay (capital of the "Region" of Apurímac), Andahuaylas and Carhuasi which have the highest population density in the "Region" (Fig. 1)."
- P4341-L11/12: This is not completely correct: the first year with noticeably more events is 2000 (Fig. 5) and this year falls into the 1990's and is not part of the 2000's (as defined in Figs. 2 and 4: 1990's = 1991-2000). Please be very careful with the description of your results.
- P4341-L17: Is it really possible that cold spells did not occur or did not cause any losses before 2002? This cluster between 2002 and 2004 is a bit suspect, maybe cold spells were not correctly declared in the first 30 years of the study period.
- P4341-L20: Change to "national-scale"
- P4341-L21-23: This first sentence of this sub-section describes your approach and thus is rather repetitive. Consider shortening it, e.g. "The national-scale comparative analysis for Peru reveals that the number of events reported in DesInventar is... etc. etc."

- P4341-L23: You write here "...looked at changes in disaster losses over the past four decades (1970–2010, c.f. Table 1)."
Why do the analyses in sections 4.1 ("1971-2009" P4340-L19) and 4.2 not relate to the same time period?
→And as a matter of fact, the year 2010 is not represented in Fig. 6.
→Is the 2010 data really included in the values given in Table 1?
- P4341-L24: One order of magnitude? Rather two. Please be very careful with the description of your results.
- P4342-L3/4: Consider changing to "Also, EM-DAT is not feasible for sub-national scale analysis due to limited number of events reported (see highlighted Cusco and Apurímac in Fig. 6)."
- P4342-L5/6: This statement applies particularly to the number of people affected; I am not sure if the data for people killed in the two databases are significantly different.
- P4342-L6: Consider replacing "enormous" with "very large"
- P4342-L7/8: Delete paragraph mark at line 7, a three-line paragraph doesn't make much sense
- P4342-L9: Consider changing to "there is a relatively good correspondence between the two databases for the number of people killed and affected (Table 1)."
- P4342-L11: Consider changing to "...while DesInventar shows a reduction of 35% from the 1970's to the most recent decade."
- P4342-L15/16: One order of magnitude only roughly applies for the 1990's and 2000's. In the earlier decades the difference is a factor of 3 to 3.5, considerably less than an order of magnitude.
- P4342-L19: Consider changing to "...was hit by intense rainfall between..."
- P4342-L21: The spelling of "Quispicanchi" here and on page 4343 differs from the spelling used in Fig. 1 ("Quisquipanchi")
- P4342-L23: Rephrase please; the rainfall triggers landslides etc., not the duration. "triggered by the intense and long-duration precipitation."
- P4342-L23: Consider changing to "...along the Huatanay river downstream from the city of Cusco, and in the Urubamba valley along the Vilcanota river..."
- P4342-L27: Consider using "floods" instead of "flood impacts"
- P4343-L2: Consider changing to "...the INDECI disaster database SINPAD indicates..."
- P4343-L7: Consider changing to "...in the first case (SINPAD)..."
- P4343-L9: Consider changing to "...revealed by a comparison of..."

- P4343-L12: Delete "with indication of the number of people affected" because it is repetitive
- P4343-L17/18: Change to "...in the "Provincias" of Anta and Quispicanchi, 34% and 31 % of all residential houses were destroyed, respectively."
- P4343-L25: Do you mean year-to-year variability of meteorological variables in general or of precipitation specifically? Consider changing to "...year-to-year precipitation variability..."
- P4344-L1/2: Delete "and thus represents a new record of January precipitation sum." because it is obvious. Alternatively, you could write "January 2010 precipitation amounted to 269 mm and represents a new long-term monthly maximum".
- P4344-L6: I would delete "events" here; an event represents an entire precipitation episode from its start to its end. Rather use x-day precipitation sum or x-day precipitation value.
- P4344-L9: Change to "...such a 5-day value is..."
- P4344-L11: Consider changing to "...data indicate minimum changes in 1-day and 2-day maximum precipitation [...] on a higher level since then (Fig. 10)."
- P4344-L15-19: This sentence is a bit long, rephrase or consider changing to "Furthermore, hourly data of Cusco airport indicate that during the above mentioned 5-day stretch precipitation was characterized by quite intensive short showers. However, daily precipitation sums were not particularly high, which is confirmed by the 1-day and 2-day data of Granja Kcayra station with approximate return periods of 5 and 20 years, respectively (Fig. 9)."
- P4344-L20: Replace "flood events" with "flood generation"
- P4344-L23/24: Consider changing to "...that the days with most people affected (in SINPAD) coincide with the days..."

→ Discussion and conclusions

The section Discussion is not as well structured as the other parts of the MS. It is partly repetitive and should be shortened.

- P4344-L27: Consider deleting "furthermore"
- P4345-L8: What do you exactly mean by "this subject"? Differences in the databases? Please clarify
- P4345-L8-16: This paragraph is widely repetitive. I don't think you need to explain the approach of your study again here. Consider deleting it (or shortening it substantially)

- P4345-L20/21: Is it surprising or not? Avoid "maybe" in this context and consider using "rather"
- P4345-L27: Is the reporting process limited or the metrics? Consider rephrasing
- P4345-L2/3: Delete paragraph mark at line 2
- P4345-L3: I would try to keep it simple, consider replacing "disaster loss and damage metrics" by "damage metrics" or something similar
- P4345-L6: Consider using "However, the fact that the number of..."
- P4345-L8: Consider using "...for the decreasing mortality rate are not known in detail for Peru, but we assume..."
- P4345-L12: Again, consider using "damage metric"; in any case be consistent
- P4345-L13: Start new sentence, e.g. with "However, there is little..."; try not to formulate too many long and complicated sentences
- P4345-L17/18: I would make that statement in the paragraph just above (where the mortality rate is discussed)
- P4345-L12-21: I would also state in this paragraph that the rate of affected population is 3 to 7 times larger in EM-DAT compared to DesInventar
- P4345-L23-25: This sentence is a bit confusing, consider rephrasing. Maybe you can replace "loss and damage ratios and rates" by "damage metrics and their rates"?
- P4345-L26: Consider replacing "of people" by "of the society"
- P4346-L5: Change to "...is increasing over time..."
- P4345-L9-10: Consider rephrasing "... at the national scale have significantly changed over the past decades (see also Fig. 10). Hence, vulnerability remains a likely driver of change."
- P4345-L12: Change to "...sufficient to draw definite conclusions and further research is needed."
- P4345-L14: Start new sentence with "But wheter the average..."
- P4345-L25: Delete "in disaster analyses"
- P4248-L3: What is a small-sized disaster? I don't like this word combination because, see also comment P4335-L5. Wouldn't it be better to use something like "...the great majority of small and medium natural hazards events goes unnoticed."?

→ Tables

- Table 1: In the caption, explain accurately what the values for “% No. of events”, “% People killed” and “% People affected” represent in this Table because it is not obvious
- Table 1: Use “ratio” instead of the word “rate” in the caption and throughout the text, because a rate is a measure of change per time and the authors mean a proportion, or ratio, of the population
- Table 1: In the caption, I would use the English translation “Peruvian National Institute of Statistics and Informatics (INEI)” that is used at P4338-L10/11 (or vice versa)
- Table 1: Consider using “Disaster mortality rate” instead of “Mortality disaster rate”; also, give this rate using exponential notation ($1.19 \cdot 10^{-4}$ instead of 0.0000119)

→ Figures

- Figure 2: It might be helpful to indicate the location of the city of Cusco with a white dot, because Figs. 1 and 2 will probably be on different pages
Specify in the caption that the last decade (2000’s) only includes 9 years
- Figure 3: Consider putting “Landslide” to plural like all the other processes; in any case, use the exact same legend in Figs. 3 and 5 (standardize spelling of “Drought/s” too)
- Figure 4: If possible, give the small white numbers in a larger font size, they are barely readable
In the caption, change to “...in the “Region” of Apurímac at the level of “Distritos” over the four...” (analogically to Fig. 2)
- Figure 6: In the caption, consider changing to “Number of climatic disasters in Peru from 1971 to 2009, based on EM-DAT. Events reported for Cusco and Apurímac are highlighted.”
- Figure 7: In the caption, consider changing to
“January 2010 floods with parts of the city of Urubamba (“Provincia” of Urubamba, “Region” of Cusco, c.f. Fig. 1) inundated by the Vilcanota (photo: Municipalidad de Urubamba).
- Figure 7: In the caption, consider changing to
“Comparison between individual disaster event records for the January to March 2010 period as registered by DesInventar and SINPAD for the “Region” of Cusco. The number of people affected (y-axis) only refers to

SINPAD records, DesInventar just indicates the date of occurrence at the top of the graph.”

Figure 8: Caption: red crosses indicate the 5-day precipitation sums and the black dots the 1-day sums!!!

Delete “beginning with the lowest sum and ending with the highest” and change to
“...city of Cusco (Data: SENAMHI), arranged by their respective ranking.”

Figures 9 & 10: Be consistent with the spelling of “1-day / 1 day”, “2-day / 2 days” etc. (legend and caption)