REVIEWER COMMENTS AND EDITORIAL DECISIONS FOR POUSSE-BELTRAN ET AL

ROUND 1 REVIEWER COMMENTS

Reviewer #1:

I appreciate the effort made with this manuscript to automate part of the process that is scarp landform analysis. I am not familiar or qualified with the computational methods in use here and so will make no technical comment on those aspects of the work. It seems that not much specific detail is provided regarding the codes structure within the document, that may be customary within the field and may be appropriate given that the codes will be published online. I would be clear, that I hope the codes that will be published include: the convolutional neural network method codes, SimScarp, ScarpLearn, and ScarpLearn_1F. With the manuscript I could access, the specific codes do not appear to have been available to me and the data availability section does not specifically indicate which codes will be contributed. I think the analytical pieces where pertaining to scarps were generally in-line with the current state of knowledge for the basic tectonically-generated escarpment case.

The manuscript does a good job of acknowledging similar publications on scarp analysis and codes that have been developed with similar goals. The quantitative comparisons made in the manuscript are generally sufficient and clear. I hope the few points that confused me on first read were clear in my comments, if I understood what was being indicated, I think some clarification in the text will go a long way.

A final point on clarity, the code does primarily a vertical separation or scarp height analysis for the user, whereas the title employs the broad term 'characterization'. It does not inform lateral variation or necessarily constrain structural dip, etc. It can capture a large number of measurements along and across a complex fault zone with multiple strands, but does not inherently measure other parts of scarps. The title may benefit from being more specific in that.

The synthetic modeling piece of this for the training element is very nice and the filtering element for treating short-wavelength perturbations (i.e. trees) is apparently advantageous. I would be interested to see it tested on a wide variety of perturbations to see what it may or may not be sensitive to across many natural settings (as one example, you note flat-bottom riverbeds near a scarp are a complication).

Additionaly, I think the second code conditioned here, ScarpLearn_1F should be presented as a more integral part of the paper. I think that can be a simple addition, it should be mentioned in abstract (1 fault vs general and multiple scarp case) and with simple additional statements within the introduction and methods sections, it should likely have its own section for 1-fault analysis in the results section, and not as an afterthought in the comparison section. Given the similarity in methods and specific use case of that code, it may not be important to describe in as much detail, but I think it should not be a surprise, niche use case that is not mentioned except for a few statements near the end. It is an impactful result that the initial model does not perform as well on single-scarp (1-fault) profiles as a model that is trained exclusively on them.

This work is encouraging in the quantitative advantage it provides, it is an incremental step toward dealing with certain types of noise in profiles with a novel machine learning method. Further, the computational advantage is a step toward leveraging large data methods in the tectonic geomorphic community, which has taken more time to develop for scarp analysis than for some other applications (channel profile analysis) because of the manual or semi-manual nature of the existing methods. I can imagine similar advantages providing utility in producing scarp datasets for past ruptures that could find applications in probabilistic fault displacement hazard analyses, fault activity rate characterizations, and other, more indirect applications.

However, this manuscript suffers from a significant shortcoming; the english in this piece is often rough, difficult to parse/understand, and results in some ideas and concepts (even simple ones) being presented in ways that are confusing or hard to understand. I began the language early in the paper, but realized it would take a large effort to do that all the way through. I think this paper's technical foundation and the work generated upon that is worthy of publication, but the language hinders its communication to a degree that I think requires additional work. I think the organization of the paper overall is appropriate and intuitive, the presentation of ideas and concepts builds in a way that is easy to follow, but the phraseology and prose are where I find issues. I think the authors would benefit from working with an editor closely to clarify the language. I think doing so will result in a strong publication, ultimately.

In summary, the technical geomorphic background, strategy, and analyses here are appropriate and sufficient. The tools presented here present utility in this type of analysis and the authors have done a good job of comparing them to existing tools and validating them in real use cases. I think the second tool generated (ScarpLearn_1F) should be presented earlier in the manuscript, and addressed more directly throughout. It is a specific, but common use case and if the tool is to be adopted, this will be important for users to know and distinguish up front. The language in the manuscript is a signficiant barrier to the communication of the authors work and concepts in the paper. For that reason, I recommend this paper be resubmitted after significant revision.

Reviewer #2: Ramon Arrowsmith

Review of

Automatic characterization of normal fault scarps using convolutional neural networks

by

Pousse-Beltran, et al.

reviewer

Ramon Arrowsmith

Summary statement

This is an interesting paper to bring the tools of machine learning to bear on the problem of the measurement of fault scarp height. It has a logical sequence developing and testing the tools first on synthetic then on observed fault scarps. I think it should be published after some revisions. I have made many edits and comments compiled below.

The biggest issues that I identify are associated with error and uncertainty. I think that a more coherent thread on these topics pulled through the document will strengthen it. The general problem with this kind of automated effort is that the geomorphologist (human) would likely discard many of these problematic profiles as not useful. So, there is a tradeoff between larger N (number of scarp height measurements) and noisier or just in appropriate measurements. This is worthy of discussion.

And, ScarpLearn is faster so this can be a powerful reconnaissance tool.

General comments

1) In general SimScarp (and the general conceptual model) seems to capture most of the phenomena of interest. But, I had a few concerns:

I disagree that the fault has to be the same dip as the slope of the scarp. For most normal faults it would be steeper.

Step 4--I don't see that the secondary fault is diffusing?

Step 5--are these perturbations also subjected to diffusion and are they persistent?

I am concerned that the 5m step size does not resolve most of the details of the scarp form, especially for small perturbations and early times and high slip rates

2) Need to cite github or other repository for the codes here and in the text.

3) Scarp height as the main metric seems ok. However, It is important to be clear if that includes the sum across all scarps in the models, and also implied in the scarp height estimate is a fit of the profiles themselves and there is associated uncertainty embedded. Some of this is covered in Table 3 but it is not so evident in the main text.

4) It is very helpful to have the error assessment between ScarpLearn and SPARTA/MCSST. Ameca fault errors (difference) is nearly half the scarp height--does this imply that the tool is not as useful for smaller scarps? But, how useful is this if the errors are substantial relative to the signal? I think one part of the answer is that many scarp height estimates have under accounted error. This is discussed in section 4.1.4 and I think should be a bit more prominent throughout the manuscript.

5) It would be helpful to have a discussion of operator error and where the error sources in MCSST and the training data will come from. I think in general a discussion of the inventory of uncertainty would be helpful. For example, figure 18 and 20 especially panel C is quite interesting. What contributes to those error bars in both axes?

6) For both the Ameca and Bilila-Mtakataka Fault examples (and probably the synthetic cases), it would be helpful in either the main text or the supplement to show several direct comparisons of the profiles and the estimated scarp heights. This should be done for both cases where the errors are small between the

methods, and for where they are large. (e.g., more like figure 11). Figure 11 also highlights a general problem with this kind of automated effort: the geomorphologist (human) would discard many of these problematic profiles as not useful. So, there is a tradeoff between larger N (number of scarp height measurements) and noisier or just in appropriate measurements. This is worthy of discussion. What is also missing is that association between the profile and a detailed surficial geologic mapl

Specific comments tied to the manuscript by element

line 11 Fault marker is unusual usage to me. I prefer geomorphic marker that indicates fault related deformation or offset. Probably not important.

line 12 I disagree that the fault dip is a geomorphological characterization. Usually we assume it or have some external geological constraint

line 19 should be active normal faults not the opposite

line 21 I don't think future tense is correct here. I would just use present tense, "we explore"

line 23 check spelling bias

line 25 check spelling deep-learning

line 35 thanks for the citation; can add "and slip rates" not always easy to get the number events from a composite scarp

line 40 fault plane

line 49 and below: important to take a moment here to remind the reader about the differences between scarp height, vertical offset, and throw, etc.

line 54 complexity also comes from identification and reconstruction of the markers that are offset, that is their projection to the fault to estimate the throw (ultimately)

line 57 these are not "horizontal" lines

line 60 project

line 70 and below: this (and table 1) is a useful summary of the methods

line 93 the uncertainty in height is not directly estimated, but it is indicated by the overall quality of the fit of the template.

lines 95-97: this needs a bit more explanation for those not familiar; it is not clear why this is different than the current study

lines 101-103 this can be stated more subtly. Manual and semi automatic methods are prone to operator error but still have the human in the loop to help with aspects of the estimation. The fully automatic ones may save time but they too have their artifacts which require some awareness to appreciate

line 108 Matteo did not do open fractures and not sure needs to say onshore. THey were mostly opening mode fractures but were also mostly full of vein or crushed material.

line 117 I realize almost all fault scarps have been eroded at least some but why this caveat? eroded or not, right? could this work on the moon or the day after an earthquake on the earth, etc.

line 121 data are plural

line 123 characteristics are plural

line 126 secondary

lines 141-143: this sentence should be rewritten, broken into two or three sentences that will have one thought each

line 144 synthetic

line 147 seems like figure 13 should be figure 3?

line 155-156 Smith and Bretherton ok, but I was looking for Carson and Kirkby plus Culling

line 159 units here ok not correct in table 2

line 160 Looking at Hanks, 2000 for example

Hanks, T. C. (2000). The Age of Scarplike Landforms From Diffusion-Equation Analysis. https://doi.org/10.1029/rf004p0313 in Quaternary Geochronology: Methods and Applications. In AGU Reference Shelf 4 (Vol. 4).

I would argue that k=10 m^2/kyr is more semi arid to humid temperate (California and Michigan), not tropical

line 169 I am concerned that the 5m step size does not resolve most of the details of the scarp form, especially for small perturbations and early times and high slip rates

Section 3.2 this is fairly well explained thank you

lines 191-192 Not clear what is this error. Is it the fit to the profiles (RMS) or is it relative to the input throw or?

line 197 remind the reader how exactly the process works to interpret the profiles in the other tools SPARTA and MCCST and what is compared? Is it only Scarp Height?

line 206 You could cite recent papers talking about repeatability of offset measurements, e.g.,

Salisbury, J. B., Haddad, D. E., Rockwell, T., Arrowsmith, J R., Madugo, C., Zielke, O., Scharer, K., Validation of meter-scale surface faulting offset measurements from high-resolution topographic data, Geosphere, v. 11, no. 6, p. 1–18, doi:10.1130/GES01197.1, 2015.

That is for strike-slip I know but there might be some similar ones and similar concepts for vertical offsets

lines 216-219 Thank you for clarifying what is the simple and complex model

line 220 I think more often it will sound better to say Synthetic Cases (singular). The individual realizations are the synthetics (plural).

line 223 What is not clear is the relative error. 3.9 m relative to a 20-30 m high scarp?

line 231 And that sounds like a somewhat expected intuitive outcome

lines 234-239 Would be interesting to have a sense of the internal operator error of the MCSST method

lines 263-268 This is very good summary

line 272 I am not sure that bias is the right word here.

line 290 More information about the DEM is needed. What is citation for those data? Are they lidar or photogrammetry-derived?

line 293 Are these single pixel derived profiles or is their any cross swath averaging?

lines 293-304 This is a nice discussion.

line 307 Refer to figure 9 here. More information about the DEM is needed. What is citation for those data? Are they lidar or photogrammetry-derived? Are these single pixel derived profiles or is their any cross swath averaging?

line 312 obtains

lines 312-320 This is a nice discussion.

line 342 I don't know what climatic changes are over time as they refer to the change in form. Is it a change in surface process or rate?

line 343 I disagree that non constant diffusion will be a problem for the scarp height. It will certainly affect an age estimate, but that is not being done here. Recall that the morphologic age is k*t.

I was also under the impression that some of the Bilila-Mtakataka scarps are bedrock which won't diffuse in the way that simply idealized here.

lines 363-370 I agree with this paragraph! I think you could point to this as both a challenge and opportunity from the introduction.

line 376 similar results as

line 378 less than

line 378 elaborate on t-test. there are many scarps "in tension"

line 381 field not flied?

lines 384-386 and so it can be useful to quickly scan an area and identify targets for more detailed work. I think that could be stated more directly

line 401 secondary

lines 402-403 need some more details on the "diffusive erosion". I presume it is a finite difference solution? Need to indicate method, stability constraint (relative to dx and dt, etc.)

lines 430-435 I think some of the detailed information about the topographic data should be in the text as well.

Table 1 ok

Table 2

I don't know what this means: "Diffusion Uniform 0.1 m 10 m" if it is diffusivity it should have units of [L/T^2]?

Not clear from table along what the parabolas are. Caption is maybe too short; can at least say to see text for explanation

Table 3 is helpful

Table 4 Ameca fault errors (difference) is nearly half the scarp height--does this imply that the tool is not as useful for smaller scarps?

Table 5 is helpful, but what is the difference from Table 2? Caption is too short.

I don't know what this means: "Diffusion Uniform 0.1 m 10 m" if it is diffusivity it should have units of [L/T^2]?

Table 6 I don't know what this means: "Diffusion Uniform 0.1 m 10 m" if it is diffusivity it should have units of [L/T^2]?

Table 7 is good

Figure 1 a and b ok

Figure 1 c: I disagree that the fault has to be the same dip as the slope of the scarp. For most normal faults it would be steeper.

Also important to take a moment here to remind the reader about the differences between scarp height, vertical offset, and throw, etc.

Figure 2: I disagree that the fault has to be the same dip as the slope of the scarp. For most normal faults it would be steeper.

I see scarp height and vertical offset. There are conditions when they are not the same. Heave and width are similar but not always the same

Figure 3: This is helpful illustration.

I disagree that the fault has to be the same dip as the slope of the scarp. For most normal faults it would be steeper.

Step 4--I don't see that the secondary fault is diffusing?

Step 5--are these perturbations also subjected to diffusion and are they persistent?

Figure 4: Looks good. Do we need the scarp learn icon there too?

Need to cite github or other repository for the codes here and in the text.

Figure 5: do Loss and accuracy have units?

Figure 6: height is mispelled

I think x axis label should be "Synthetic ..." not plural

Is SH scarp height? inconsistent axis labels

No SPARTA uncertainties?

In general not large differences between mean and median relative to 1 sigma uncertainties.

Explain Simple and Complex here or at least refer to text for more information.

Figure 7: Caption needs to indicate what is topographic map data source.

Are these single pixel derived profiles or is their any cross swath averaging?

It would be helpful to label the fault and some of the elements of the scarps in B (while somehow also not obscuring the data).

Figure 8: This is a helpful and interesting figure. Refer also to figures 17 and 18. and in the caption is misspelled. Provide citation for DEM.

Figure 9: Caption needs to indicate what is topographic map data source.

Are these single pixel derived profiles or is their any cross swath averaging?

Figure 10: This is a helpful and interesting figure. Refer also to figures 19 and 20.

Figure 11: I was looking for this kind of figure. It might be nice to also include one where the estimates were quite similar. The figure does highlight a general problem with this kind of automated effort: the geomorphologist (human) would discard many of these problematic profiles as not useful. Explain that the lines are showing the marker ties.

Figure 12: I disagree that the fault has to be the same dip as the slope of the scarp. For most normal faults it would be steeper.

I see scarp height and vertical offset. There are conditions when they are not the same. Heave and width are similar but not always the same.

Figure 13 this is ok; seems a bit arbitrary but I suppose is the consequence of the parameter selections as defined in Table 2. Caption should refer to Table 2.

Figure 14: Useful comparison. I think x axis label should be "Synthetic ..." not plural. Height is misspelled.

Figure 15: This needs a more complete caption to let the reader know what is different about the estimates shown here versus elsewhere. I think x axis label should be "Synthetic ..." not plural. Height is misspelled.

Figure 16: This needs a more complete caption to let the reader know what is different about the estimates shown here versus elsewhere. I think x axis label should be "Synthetic ..." not plural. Height is misspelled.

Figure 17: This is a helpful set of comparisons of uncertainty, etc. and in the caption is misspelled. I like that the parameters versus scarp height are plotted.

Figure 18: This is interesting. What is the beige/yellow banding on the plots. If intentional needs to be explained in the caption.

Figure 19: This is a helpful set of comparisons of uncertainty, etc. I like that the parameters versus scarp height are plotted.

Figure 20: This is interesting. What is the beige/yellow banding on the plots. If intentional needs to be explained in the caption.

EDITOR DECISION FOLLOWING FIRST ROUND OF REVIEWS

Dear Léa Pousse-Beltran, Theo Lallemand, Laurence Audin, Pierre Lacan, Andres David Nunez-Meneses, Sophie Giffard-Roisin:

We have reached a decision regarding your submission to Seismica, "Automatic characterization of normal fault scarps using convolutional neural networks: ScarpLearn". We feel that this manuscript may be suitable for publication in Seismica following revision.

I have now received two reviews for your manuscript. As you will see, both reviewers are generally positive regarding the potential of your work to make a good contribution to Seismica. Common themes of criticism between both reviewers reflect the need for a better description and discussion of uncertainties associated with an automated method such as that you employ, but also some work that needs to be done to improve language clarity and polish. I concur with these issues as raised by the reviewers, and I think consideration of their detailed comments below in a revised version will make for a substantially more impactful contribution.

When you are ready to resubmit the revised version of your manuscript, please upload:

A 'cleaned' version of the revised manuscript, without any markup/changes highlighted.

A pdf version of the revised manuscript clearly highlighting changes/markup/edits.

A 'response-to-reviewers' letter that shows your response to each of the reviewers' points, together with a summary of the resulting changes made to the manuscript.

If you deem it appropriate, please check that the revised version of your manuscript recognises the work of the reviewers in the Acknowledgements section.

Please note that Seismica does not have any strict deadlines for submitting revisions, but naturally, it is likely to be in your best interest to submit these fairly promptly, and please let me know of any expected delays.

I wish you the best with working on the revisions. Please don't hesitate to contact me with any questions or comments about your submission, or if you have any feedback about your experience with Seismica.

Kind regards,

Randy Williams

randolph.williams@seismica.org

ROUND TWO REVIEWER COMMENTS

Reviewer #2: Ramon Arrowsmith

Second review of

“An automatic characterization of normal fault scarps using convolutional neural networks: ScarpLearn" by L. Pousse-Beltran, T. Lallemand, L. Audin, P.

Lacan, A. David Nunez-Meneses & S. Giffard-Roisin.

Thank you for the nice revisions.

I have looked at the response to reviewers and the revised manuscript. Thank you for your thorough effort to address my suggestions.

In particular, I think that you have done a good effort to address the issues of error throughout. I think it will read much better and have a greater impact.

I have a few comments below as I reviewed it. I would suggest a final clean up before accepting it. I don't need to look at it again.

I could not access the code on this web site:

https://filesender.renater.fr/?s=download&token=2da4eeda-38e0-4fba-aa0d-b3de2ba919ba

where I received this error:

L'application a rencontré une erreur lors du traitement de votre requête

Dépôt expiré, tous les fichiers ont été supprimés, aucune restauration n'est possible

However, the zenodo link did work and the code was well available.

I note the repetition of this text which reads somewhat awkwardly in successive paragraphs (lines 484 to 493):

"For real cases, operator error

may be mitigated by cross-verifying hillshade DEMs and geomorphological mapping."

and

"it is mitigated by thorough verification of hillshade DEMs and

geomorphological mapping for each profile"

Figure 12 is ok with the caption as is, but I would expect that the fault dip is more like the theoretical expectation of 60 degrees and projects to the mid point of the scarp. Non diffusive processes will quickly adjust the slope to lower more like angle of repose. Same comment holds for figure 1. Thank you for changing it in Figure 2.

Line 455 and later: I don't agree with these statements about the change in slope profiles above and below: It seems to me that changes in glacial to interglacial or in diffusion rates will occur over longer wavelengths than a few tens of meters of a fault scarp. And, in my opinion, diffusion is the process where transport rate is proportional to local slope. The only way that "diffusion changes" is if the slope changes or the diffusivity (rate constant) changes spatially in this case.

Line 515: I agree about the need for careful mapping. But that does mean that the use of ScarpLearn has to be done with caution: operators cannot work blindly but need to understand the context.

Figure 4: "The link to download ScarpLearn is commingsoon." typo

Figure 6: please do not use contractions: "does not"

Figure 10: spell out Table

EDITOR DECISION FOLLOWING SECOND ROUND OF REVIEWS

Subject: [Seismica] Editor Decision

Dear Léa Pousse-Beltran, Theo Lallemand, Laurence Audin, Pierre Lacan, Andres David Nunez-Meneses, Sophie Giffard-Roisin:

I hope this email finds you well. I have reached a decision regarding your submission to Seismica, "Automatic characterization of normal fault scarps using convolutional neural networks: ScarpLearn". Thank you once again for submitting your work to Seismica.

Based on reviews I have received, your manuscript may be suitable for publication after some minor revisions.

I have now received comments on your revised manuscript. As you will see below, the reviewer seems to be broadly satisfied with your responses to their comments and associated revisions to the manuscript. They do offer a few relatively minor comments for consideration and additional cleanup. I am requesting that you implement those comments before we proceed forward with copy editing.

I do not intend to send the manuscript back out for additional review. As such, it will save some time if you are able to also upload separate, full-resolution files for each figure in the manuscript. Please also note the reviewer's comment about the availability of the repository for code, etc. You are not required to host the files for this work multiple places. Zenodo alone would be sufficient, but to the extent you do intend to host the material in multiple locations, we would like to see them all functional before the manuscript is made available to the public.

When you are ready to resubmit the revised version of your manuscript, please upload:

A 'cleaned' version of the revised manuscript, without any markup/changes highlighted.

A pdf version of the revised manuscript clearly highlighting changes/markup/edits.

A 'response-to-reviewers' letter that shows your response to each of the reviewers' points, together with a summary of the resulting changes made to the manuscript.

If you deem it appropriate, please check that the revised version of your manuscript recognises the work of the reviewers in the Acknowledgements section.

Please note that Seismica does not have any strict deadlines for submitting revisions, but naturally, it is likely to be in your best interest to submit these fairly promptly, and please let me know of any expected delays.

I wish you the best with working on the revisions and I look forward to seeing this in print. Please don't hesitate to contact me with any questions or comments about your submission, or if you have any feedback about your experience with Seismica.

Kind regards,

Randy Williams