Response to anonymous reviewer / RC1

We sincerely thank the anonymous reviewer for thoughtful feedback on this manuscript, including both broad topical suggestions and grammatical or word choice suggestions. We especially appreciate these reviews considering this paper presents the first iteration of a new research tool.

A recurring misconception with the review is that HazMapper is a semi-automated method and that we did not provide quantifiable data needed to assess the rigor of a semi-automated approach in identifying natural hazard features. This initial release of the HazMapper application does not semi-automate the identification of natural hazard features. Identification of natural hazard features and assessment of the signal-to-noise ratio in the rdNDVI output is incumbent upon the user. At present, we believe many in the natural hazards research, prevention, and outreach communities will find HazMapper to be a useful utility that will assist in their exploration and characterization of many different types of natural hazard features. Our hope is that future iterations of the tool will focus on individual hazard types (e.g., mass wasting; wildfires; tornadoes) where further analysis and semi-automation will be employed.

Color legend
RC1 - black
Author response - blue

Topical comments

Scheip and Wegmann present an online-GIS to display and analyse perturbations of the Earth's surface. This toolbox allows the user to select among three different satellite missions, to choose a period of interest, and calculate landscape changes using a vegetation index. The authors show five case studies to visualize the detectable impacts from volcanic, coseismic, and rainfall-related mass movements, and burnt areas from wildfires. The authors argue that HazMapper is designed for people with little prior knowledge in a GIS environment, or who could have limited access to powerful computing facilities. This goal seems to be fulfilled given that the interface (Figure 1) is designed in clear and visually appealing fashion. The downside of the presented web GIS is that the possibilities for the time being remain very limited beyond calculating a vegetation index. Clearly, practitioners may benefit from the resulting maps of vegetation change. Yet from a scientific perspective these maps need at least a minimum amount of quality check to judge how useful this maps are. Yet, unfortunately, any measure of accuracy or uncertainty (and discussion thereof) remains elusive in the current manuscript. Some questions (without logical sorting or relevance) that could be answered in more detail are:

Response:
Thank you for the compliments on the web interface. For this initial release of HazMapper, a multi-spectral imagery index is calculated and no semi-automation is employed. That is, this iteration of HazMapper is not a semi-automated tool for identifying individual landslides or burn extents, for example. Instead, a vegetation index is displayed which the user can interpret as to the nature of the landscape change.

In its current state, we believe this tool is useful for a wide variety of researchers without semi-automation. As such, uncertainty is not quantified. Qualitatively, published burn extents and lava inundation extents from the Chimney Top 2 fire and Kilauea example are compared. We agree that uncertainty assessments should follow semi-automation where predictions of disaster occurrence are made but posit that this is premature for the current iteration of the platform.

What can we do in regions with frequent cloud cover?

Response:
From the reviewed manuscript, lines 64-65:
To circumvent potentially opaque atmospheric conditions, HazMapper capitalizes on a technique within Google Earth Engine to generate and perform calculations on a greenest-pixel composite (Figure 2).

and lines 66-67:
...records the pixel with the highest normalized difference vegetation index (NDVI) result, or the "greenest" pixel (Eq. 1).

We will expand on this text to make this more clear. We composite many images together and retain only the pixels with the highest NDVI value from the entire stack. This composite indicates the peak phenological cycle of pre- and post-event conditions. The rdNDVI metric is computed from this peak phenological cycle calculated from the user input duration (time - months) for the pre-and-post event windows.

"How can we detect mass movements that do not cause disruption of the vegetation cover, such as slowly moving landslides or mass movements in arid or un-vegetated (high mountain) regions?"

Response:
We will add text to the manuscript to reinforce that this is a vegetation-based metric and will not be useful in completely un-vegetated regions (e.g. desert, polar). Arid or humid, the key is the presence or absence of vegetation. For example, this summer, we have had success

tracking wildfires across the arid western United States, for example, the Bush Fire north of Phoenix, AZ [https://twitter.com/HazMapper/status/1277988549741207552; https://en.wikipedia.org/wiki/Bush_Fire_(Arizona)]. While this area is quite arid, vegetation is still burning, and, thus, we can register that loss with HazMapper. If a disaster does not impact vegetation, in most cases, HazMapper will not be a suitable tool for detecting it.

"How does HazMapper perform in the era before 2012, when only patchy Landsat 7 images are available?"

Response:
This is an excellent question! We will add text and a figure (draft figure included below) to the manuscript to discuss this. When using Landsat 7 data, short pre- and post-event windows will tend to return outputs with stripes. However, by increasing the window lengths, these artifacts are reduced via our greenest pixel compositing methods described in the manuscript.

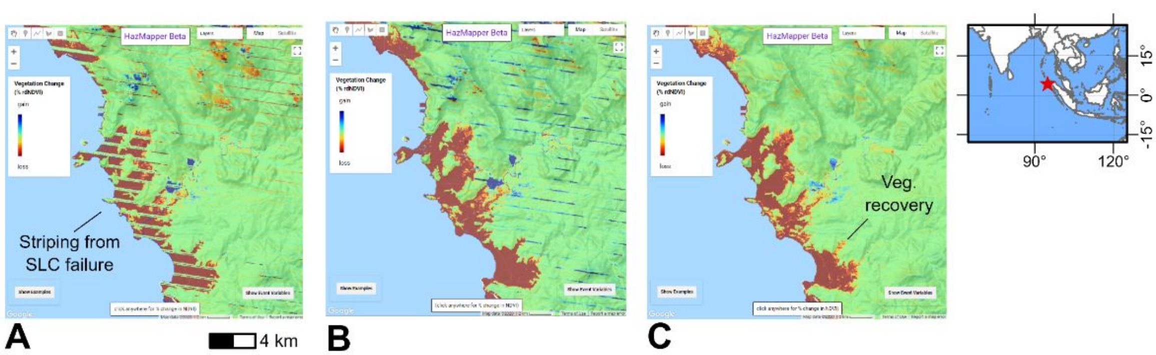

Figure 8. $rdNDVI$ change detection images and greenest pixel composites following 26 December 2004 tsunami in Indonesia resulting from the Sumatra–Andaman earthquake. Parameters - Dataset: Landsat-7, Event Date: 26 December 2004, maximum cloud cover: 100%, and slope threshold: 0.01°. Pre-Windows: 1, 2, and 12 months, Post-Window: 1, 2, and 12 months for panels A, B, and C, respectively. $rdNDVI$ across the event illustrates tsunami inundation zone and resulting loss in vegetation. Striping in the data from the scan-line corrector failure in Landsat-7 is evident in a short look window (e.g. 1 month, A), but these artifacts are reduced by increasing the pre- and post-event windows (B and C). By 12-month pre-post periods (C), the striping is significantly reduced in results, however, vegetative recovery is also present in this longer post-event cycle that captures the first growing season following the tsunami. Screenshots of HazMapper application, example accessible at https://go.ncsu.edu/hazmapper-indonesia-tsunami.

How does the resolution of the sensor affect the minimum size of detected disturbances?

Response:
Sensor resolution is one variable affecting the minimum size of detected disturbances, finer resolution can detect finer disturbances and should be considered in the analysis. For example, in North Carolina, where the average landslide width is ~10m, Sentinel-2 would be more suitable for analysis compared to Landsat.

Are rdNDVI values comparable across the three different sensors? What are the optimal thresholds to set during analysis?

Response:
No, they are not, nor are they expected to be. Differences in pixel size (e.g. 30m vs. 10m) will result in different NDVI values, and therefore, rdNDVI values. Further, no rdNDVI thresholding is available in HazMapper. This is something that the user could choose to do after downloading data for further analysis.

How can we make sure that the automatically detected changes come from the same trigger?

Response:
The pre- and post-windows restrict data used for analysis. Therefore, changes can be confidently constrained to the time frame of the selected pre-post analysis window. This type of analysis provides more confidence in timing compared to traditional field visits that may occur weeks to months after an event and may not be accompanied with strong confidence in pre-event conditions. Further, this provides at least the same level of confidence as traditional single image comparison studies. In those studies, for example, a landslide present in the post-event image but not the pre-event image is typically assumed to have occurred during the event under consideration.

I highly appreciate the goal of the authors to help non-experts in doing rapid post-event analysis, but I found few information that guides these non-experts through their analysis. Limited knowledge about the regrowth rates, for example, could lead to large misestimates of detected changes, if the window is not set accordingly during the analysis. It was therefore surprising to see that the current manuscript offers no discussion section where such issues are considered in detail.

Response:
Thank you for this excellent perspective. Our initial thought was the event by event discussion captured much of what a traditional Discussion section would, however, the suggestion to add a Discussion section is well received and the revised manuscript will include a Discussion section.

No training materials have been developed yet, however, we have begun work on https://hazmapper.org/learn where we anticipate providing YouTube videos, cookbook examples, and other training materials for exactly this purpose. We do not feel this type of media content is suitable to submit for publication as a research article.

Grammatical and technical suggestions such as word changes or further explanations required

"L14: Is it important to distinguish between 'developed' and 'undeveloped' countries here?"

Response:

We will remove differentiation between developed and undeveloped countries.

L17: What is 'significant' here?

Response:

We will remove word 'significant'

L18: 'characterization': more specific?

Response:

Characterization of natural hazards such as the number or spatial distribution of landslides, progression, and final burn extent of a wildfire, for two examples. We will add clarifying text to the manuscript.

L20: 'typically persist in vegetated landscapes': Could the authors briefly explain why that is the case?

Response:

We will add "due to the constant regrowth cycle of vegetation in temperate environments". For example, a landslide scar in the humid highlands of Papua New Guinea will become covered with vegetation in the subsequent growing seasons (see Figure 4D).

L23-24: 'provides a single time-stamp of ground conditions': People familiar with dating would argue that one can read out way more from a landscape than a single time-stamp from one field visit.

Response:

We will remove the sentence.

L24: Why should people with no 'interest' perform field work?

Response:

Our intent was to suggest limited agency budgets often force project prioritization. With competing interests vying for agency funds, only those of high enough priority, or interest, can be investigated with resource-and-time-intensive methods like fieldwork. Additionally, in the time of a global pandemic, many universities and public science and land management agencies are minimizing, discouraging, or not allowing travel by their employees to minimize exposure risk to the COVID-19 virus. We will remove the word "interest".

L26-27: ‘observe, monitor, and track‘: suggest using only one of these terms
Response:
We posit these have different meanings in the remote sensing literature and would like to leave as is. For example, we may observe vegetation change resulting from a wildfire long after the fire burns, but during the burn, we may be monitoring the fire for expansion, or we may be tracking the progression of the fire once it begins to expand.

L29: How do the authors define ‘increasingly complex’?
Response:
We will replace “complex” with “advanced” to indicate that new satellites and payloads are more advanced than older platforms, for example, increased resolution, precision, and capability.

L35: ‘obvious advantages‘: such that? Could there also be some ‘not so obvious’ disadvantages, for example in the field of data protection regulations?"
Response:
We will remove the word “obvious”. We will need further clarification on what data protection regulations are the concern of RC1. HazMapper is a view-only platform at present, in that users view data but are not currently loading any data into Google Earth Engine.

L55: Use the more familiar ‘GeoTIFF’ instead ‘geoTIF’?
Response:
We will replace geoTIF with GeoTIFF in manuscript.

L58: ‘Can be’ instead of ‘is’?
Response:
We will replace “is” with “can be” in manuscript.

L59: ‘other opaque atmospheric components’: such as? Maybe haze and dust?
Response:
We will include “atmospheric aerosols” as this term is more inclusive than haze or dust, which refers to dry atmospheric particles. Aerosols also include liquid droplets like water vapor.

L58-63: Not sure whether this motivation is useful at this stage, because all these arguments call for using non-optical data sets such as radar, given that they suffer less from atmospheric disturbances"
Response:
The intent of this comment is to set up the next paragraph where we explain how the compositing method partially overcomes the optical data limitations.

L64-69: Are the satellite images radiometrically corrected?
Response:
Datasets currently used in HazMapper are corrected to Top of Atmosphere (TOA).

L98: Would be good if the curated examples are publicly accessible without having a Google account
Response:
We agree fully with this comment. The entire HazMapper platform, including using the web interface, the curated examples, and the source code, will be released with the initial publication to users with or without a Google account.

L112: What if internet access is limited or unavailable in 'regions with less adequate resources'? Would this rule out the use of the HazMapper?
Response:
Yes, this is a web-based tool and cannot be used without the internet.

L112: Please avoid subjective terms such as 'incredibly'.
Response:
We will remove subjective terms such as "incredibly." Thank you for the suggestion.

L113-115: These two sentences are slight repetitions from previous arguments, for example in L21-24. Consider shortening (or deleting)."
Response:
We will delete this repetition.

L115-116 & L118-119: These arguments have also been brought up before. Could it make sense to redistribute the content of Chapter 3 into Chapters 1 and 2? Most of these argument could strengthen the overall motivation of this paper in the introduction. I do not see too much additional value in a chapter on its own that compares different GIS environments.
Response:
This is a great suggestion. We will remove the existing Chapter 3 *Earth Engine vs. Traditional GIS Environments*. That content is better suited for Chapters 1 and 2, as suggested by RC1.

L128-135: These sentence largely contain arguments from previous sections, and could be more useful to expand the line of arguments (or number of references) there.
Response:

We will move this text and use it to expand the number of references, as suggested by RC1.

L134-L137: These two sentences should go into Section 2.
Response:
We will move this text to Section 2, as suggested by RC1.

L138-139: Repetition, consider deleting.
Response:
Repetition will be deleted.

L151-154: Again repetitions from previous sections. By the way, I'm not sure whether downloading 'one to a few pre- and post-event images' demands 'high-powered computers and large digital storage capacity'. The authors may acknowledge that downloading one Landsat scene before and one after a landslide, each ~800 MB large, and loading them into memory can be done with the bulk of post-2010 computers, no?
Response:
Repetition will be deleted. Yes, modern computers can download "one to a few images" but it becomes increasingly difficult to download and process many dozens of images or to expand spatial extent after processing static images. Furthermore, HazMapper works on mobile devices like Chromebooks, tablets, and smartphones, which in most cases do not have the required software nor processing speed to perform a similar analysis. We have added text to this effect in the (new) Discussion section.

L157, L159: What can Huffman's paper from 2014 tell us about a debris flow that happened in November 2019?
Response:
We are happy to modify this database reference and agree it can be confusing. We are following the publisher's' citation recommendation (available at https://gpm.nasa.gov/data/policy), which suggests a 2014 citation year, consistent with the database release, and including a note of the access date, which in our case, reads 2020-01-24.

L165-173: What is the reason for this mini-review on the local geology / geomorphology? It feels like this paragraph dissects a bit the logical flow between the preceding and the following paragraph in this chapter.
Response:
We will modify this text to make this read a bit more logically. Thank you for the suggestion.

L185: 'Fatalities from coseismic mass wasting events can increase significantly': from which baseline do fatalities increase and by which rate? And how do the authors define ‘significantly’ in this regard?

Response:

We removed the word “significantly.”

L197: ‘when expanding the analysis window to the predicted 300 km maximum distance’: What did the authors prevent from not considering the full radius of 300 km from the beginning?

Response:

It can be difficult to view small landslides when viewing 300km of data, so we start the analysis zoomed in, and then expand out to identify total impacts.

L198-199: How did the authors make sure that these landslides were generated from the same earthquake?

Response:

Lines 198-201 of the reviewed manuscript reply to this comment:

Furthermore, we noted possible coseismic slides and flows as far as several hundred km west of the epicenter in the Maoke Mountains of Indonesia. Mass wasting is common in the region and these events could have unique triggers, however, restricting pre- and post-event time windows to as little as 2 months bracketing the Mw 7.5 mainshock demonstrates consistent timing with the 25 February 2018 earthquake.

We mention “possible” coseismic slides and flows, and agree that they “could have unique triggers.” However, we can be confident that they occurred within the period or two months before the mainshock to two months after the mainshock because of restrictions in the pre- and post-event window lengths. Please let us know if additional clarification is required.

L199-201: Structure of the sentence is not fully clear. Please elaborate, possibly splitting this sentence into two.

Response:

We will divide this into two sentences, thank you for the suggestion.

L201: ‘future’ is a bit odd for a tool that uses historic images. . .

Response:

We agree with this comment and will remove“future” and re-word this sentence.

L202: What is the content of this ‘robust spatial and temporal catalog’? Figure 4 shows no more than a rdNDVI map, without explicitly digitizing individual landslides or debris flows, measuring their areas, estimating their volumes, their runout paths, spatial density, potential different time stamps of occurrence, and so on. Also, how do the authors define ‘robust’ here? So far, the authors show no accuracy assessment, in terms of how much of the total

area (as one measure of accuracy) is correctly classified by the rdNDVI. What are the commission / omission errors, compared to manually mapped landslides? It is hard to believe that this approach picks mass wasting events error-free, especially in tropical regions with rapid regrowth rates along river channels.
Response:
We will remove this sentence. HazMapper does not "pick mass wasting events" as it is not currently a semi-automated method. Because it's not picking anything, evaluating commission/omission errors is premature. This type of analysis will be necessary once we implement the semi-automation of the platform, which will be hazard-specific and not a broad platform like this iteration of HazMapper.

L234: '17,000 acres': what would be the area that HazMapper predicts?
Response:
HazMapper does not "predict" a burn area because it is not a semi-automated method. The gray line in Figure 5 is the USGS-published burn extent, which measures 17,140 acres.

L237: If HazMapper only uses the rdNDVI, how can it distinguish between burnt areas and a landslide that happened in that area, or clear cutting?
Response:
We agree that non-unique solutions exist. For example, agricultural artifacts are highlighted on Figure 3D (Kenya debris flow example). As HazMapper is not a semi-automated method, at present the suggested interpretation (e.g. burned areas vs. landslide vs. clear-cutting) would be incumbent on the user.

L238-239: How do the authors measure 'the most severe burn', assuming there are different types of vegetation cover with a study region that might have completely different starting NDVI values?
Response:
From equation 2 in the reviewed manuscript, the starting NDVI value is used in the normalization parameter to account for exactly this phenomenon. The more severe burn is the highest rdNDVI value, consistent with Norman and Christie, 2020 (formerly Ambrose et al., 2019). This will be further clarified in the discussion section.

L249: Is there a reference that shows that the annual number of fatalities from volcanic eruptions has increased from year to year in the past 500 years?
Response:
We will add the Auker et al. (2013) reference to the text. This reference demonstrates what the reviewer requests regarding an increase in volcanic-hazard fatalities in the past 500 years. We specifically refer the reviewer to figure 4 of the Auker et al. (2013) paper, copied below.

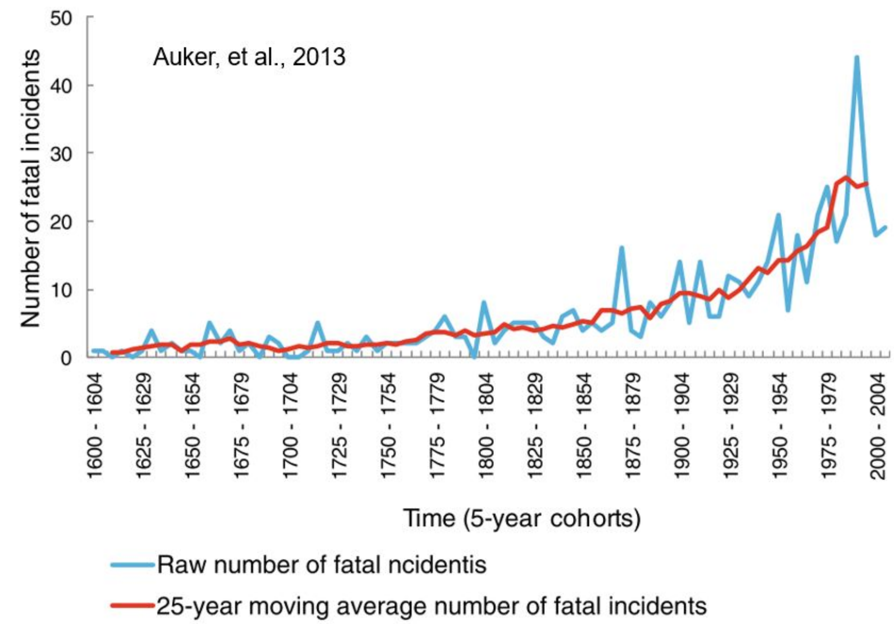

Figure 4 Time series of number of fatal incidents. Blue line shows number of fatal incidents over time; red line shows 25-year moving average of number of fatal incidents over time. Counts are calculated in five-year cohorts.

L254 / L264: ‘downslope hazards‘: more specifically?

Response:

The next sentence in the reviewed manuscript (line 255) defines downslope hazards associated with volcanoes:

Downslope hazards may include lava flows, ballistic projectiles, pyroclastic flows, and lahars (Blong, 1984).

Please let us know if additional clarification is needed.

L265: What do the authors mean be ‘decimated’?

Response:

We will replace the word “decimated” with “destroyed.”

L273: These ‘analytical false negatives’ urgently demand quantification!

Response:

The reviewed manuscript mentions noise as a concern (e.g. cloud cover, agricultural fields). This is another example of unwanted artifacts or noise. We recognize we should not have used the terms false negative or false positive, which would only apply in a predictive or semi-automated tool. We will revise the text to make this more clear.

L274: How can the authors judge from satellite images that these features are ‘hyper-concentrated flows'? And how they define the 'transition to hyper-concentrated stream flows'?

Response:

The post-event differentiation between hyper-concentrated stream flows and debris flows is difficult to make even in the field, let alone remotely, we recognize. We do not suggest the exact location of this transition because it is unknown, as RC1 indicates. Based on the decreasing stream gradient away from the volcano flanks, at some point, it is a reasonable assumption the transition will occur, and we try to simply state that. Certainly the debris flow conditions do not persist for 60km across gentle slopes near the Pacific Ocean, for example. We will add clarifying text to explain this.

L290: What do the authors recommend for cases where we have persistent cloud cover, possibly over months to years? There are many coastal and mountain regions, and many scientists or practitioners would wish to see a solution to this problem. This calls for a fuller discussion regarding the limitations of HazMapper.

Response:

This problem and the solution is described earlier in the reviewed manuscript:

To circumvent potentially opaque atmospheric conditions, HazMapper capitalizes on a technique within Google Earth Engine to generate and perform calculations on a greenest-pixel composite (Figure 2). The greenest pixel composite is a single composite image generated from all images within the user-defined pre- and post-event window that records the pixel with the highest normalized difference vegetation index (NDVI) result, or the "greenest" pixel (Eq. 1).

We will expand on this in the revised manuscript to add additional clarification and will discuss this issue again in the (new) Discussion section.

As a practical example of this, the reviewer is invited to assess our example from Papua New Guinea (Figure 4), where the problem of persistent cloud cover is a hindrance to remote sensing for natural hazard applications. HazMapper leverages the greenest-pixel composite method discussed in the manuscript to reduce or remove cloud cover from scenes. In some locations, compositing over several months may be required in order to derive a cloud-free greenest pixel composite image for either the pre or post-event window. In the case of the 2018 earthquake in PNG, we composited 12 months of pre-event and 9 months of post-event imagery in order to derive figures 4B, C, and E.

L291: 'Future code modifications': such as? This could be a core problem of HazMapper: How can we map landscape change if there is no vegetation?

Response:

We have initiated collaborative work with others on non-vegetation based metrics, but agree it is premature to include this sentence. We will remove it. We further agree that in lieu of vegetation, a vegetation based metric is not suitable. We make no claim that HazMapper will work in every environment and will ensure the revised manuscript indicates it is a vegetation-based metric that only applies to areas with vegetation.

L297: And what does this ‘comparison’ show? Seismologists tend to use seismic data, and HazMapper uses optical satellite images, both parties probably have a hard time to make their datasets comparable to each other. Where is the overlap?

Response:

The comparison is qualitative and is shown in Figure 7, where the USGS-published lava inundation extents are plotted against the rdNDVI results. Because HazMapper is not a semi-automated routine, no quantitative comparison is appropriate.

L303: No discussion chapter?

Response:

This is an excellent suggestion and the revised manuscript will include a Discussion section.

L317: How ‘good’ is this ‘approximation’ in real numbers?

Response:

We will change the language to represent the qualitative approximation.

Figures 3, 4, 6: What are the slope thresholds good for? Why are they (close to) zero?

Response:

We refer the reviewer to Table 1, which defines the slope threshold and suggests its use.

Slope Threshold	A minimum topographic slope value in degrees, less than which will be omitted from the data visualization. This is helpful to remove water bodies like lakes and adjacent oceans in coastal regions.

In the examples included in the manuscript, a very low slope threshold is useful to omit water bodies (e.g. oceans). Text will be added to the Discussion section to clarify this.

Figure 4: Why do C and D have a different color scale? Could it be that D and E have wrong labels?"

Response:

The labels are correct. C and D are different color scales because C is a figure depicting vegetation loss (negative rdNDVI values) and D is a vegetative gain figure (positive rdNDVI values). Neither D nor E is labeled incorrectly. We will review the caption to ensure it clearly explains the differences.