

Response to Referee #1

Please see our responses to each referee comment (in blue):

Anonymous referee #1: This paper investigates post-wildfire erosion using multitemporal lidar over time. I think some of the methods used by the authors are quite unique as compared to similar post-wildfire lidar studies (e.g. Pelletier and Orem, 2014 and Orem and Pelletier, 2015). In particular, I like their approach for removing DEM pixels that were determined to be disturbed by the canopy. That technique is novel, and I think it will provide a robust new method that will be embraced by the community. I also applaud the authors for their use of radar estimated rainfall and the approach used to correct it based on local rain gauge data. That allowed the authors to analyze spatially continuous rainfall data, which was useful for their overall analysis. I think that the general thrust of the paper is unique and I think that this manuscript is close to being ready for publication. However, there are a few general suggestions that I will make, in addition to many specific suggestions below.

We thank the referee for reviewing and provided general and specific suggestions. We have responded to each of your suggestions below. In most cases we agree to the suggested changes, and either have made or are making the changes in the manuscript.

Consider re-writing section 4.2. Section 4.2 is a long chronological narrative, I understand the temptation to write it this way, because that is the way it unfolded in time. But it is really boring to read and fails to convey the salient points well. Consider organizing it in terms of drivers and response. This will help to generalize the paper beyond a case study.

We appreciate this comment, and we (the three authors) had discussed how to best organize the paper. We understand the potential value of organizing the paper by processes (“drivers and response”), but this proved unwieldy and even more difficult to follow given the diverse responses over time, space, and between watersheds, and the different drivers in terms of convective storms, snowmelt, the large mesoscale flood, and the reduction in post-fire effects over time. We are revising Section 4.2 to eliminate some of the specific details and make it more succinct. This will lead to a greater focus on the key points and processes as suggested by the reviewer.

Section 5.2 also needs some attention. The way that you break up the paragraphs is a little strange. I suggest abandoning the enumeration that you use “First, second, third.” For example, on P17 line 19, why does that paragraph start with “third”, and contain the “forth” point, but the next paragraph doesn’t start with “Fifth”? I think the same points can be conveyed without this type of enumeration, and then paragraphs can be grouped by similar ideas.

We will revise this section to either make the enumeration more consistent, or eliminate the enumeration altogether. We do feel that by listing the points it is easier for the reader to keep track of the larger organization and context with respect to the series of points that we are making.

I was also confused by your use of the term “sediment availability” in the discussion. At present, I don’t see how your data speak to the sediment availability at all, and yet it is invoked as an explanation. I would consider either adding in data that relates to sediment availability, or rephrasing the sentences in which you point to sediment availability.

Our interpretation of “sediment availability” is based on: a) visual observation of sediment deposition in the valley bottoms and channels; and b) the estimates of deposition based on the lidar differencing. The underlying concept is that the post-fire sediment is much more readily erodible than the older sediment that is protected by vegetation. We will revise the manuscript to more clearly define what we mean by “sediment availability” and explain why the post-fire sediment was more accessible to erosion, especially during the mesoscale flood.

Lastly, I think it would be really helpful to synthesize these really unique results that moves beyond the case study. This could just be a paragraph in the discussion, but consider helping readers to see how the erosion/deposition sequence could be converted into something that might lead to more insight at different sites in the future.

Thanks for the suggestion. We will consider including text to better synthesize the results and use this to help expand the discussion.

This paper is really interesting, and despite my detailed comments, I enjoyed the approach, and I think that this manuscript will be a nice addition to the literature.

Thank you for this very positive overall summary of your evaluation of our manuscript. We have recently presented this story at several conferences, and have received a very positive response from the audience as this article does break some new ground in terms of comparing fires and floods, and also focusing on larger scale effects.

Specific Comments:

P2. L14: Rengers et al. modeled basin scale post-wildfire runoff. doi: 10.1002/2015WR018176

Thanks, we will add this citation.

P3 L7-8: remove “very”

We have deleted “very” as suggested by the reviewer.

P3 L21: replace “small- to moderate-sized watersheds” to “stream channels” because it seems like all of your analysis is in the channels, not in the larger watershed.

We agree, and have revised the manuscript to read “in the valley bottoms of small- to moderate-sized watersheds”.

P3 L28: I didn’t see any hypsometric curves

Yes, we did not include the hypsometric curves as we are only trying to indicate that the watersheds are very similar. The hypsometric curves are not critical to interpreting or understanding the results, and so we provided this fact but since we already have 12 figures we did not want to add yet another figure when this is not critical to our study. When describing a study area it is common to provide these kinds of descriptive statements without providing all the underlying data, and we believe that providing a statement about the hypsometric curves without providing the data is consistent with standard practices.

P3 L30: Use a more specific term than “evergreen”

We will add a sentence to indicate that the forest cover was primarily mostly ponderosa pine with some lodgepole pine and douglas fir at higher elevations and north-facing slopes.

P4 L3: Say how straw and wood mulch were applied

We have added “from helicopters” to clarify this. Wood chips were spread manually to a very small area (less than one hectare) that was mostly on a ridgetop and accessible by road.

P4 L6: add “channels” after the word combined

We presume that the reviewer means “confined” rather than “combined”, and we have added the word “channels” to make this more explicit.

P5 L4: Are you going to post the python scripts anywhere?

The scripts referenced here piggyback off the FluvialCorridor ArcGIS add-on, and they were written specifically for our analysis so they would not be immediately useful to others. If a reader is interested they can contact us at the email address listed in the manuscript, and we will be happy to provide them with the scripts and any additional information, but they will have to make some modifications.

P5 L18: Can you explain why you used the 50 m sections? You could have just analyzed the lidar on a pixel-by-pixel basis, so why create short reaches to analyze? This would benefit from some more explanation.

The two main objectives of our paper were to: 1) characterize the spatially-explicit changes in sediment deposition, erosion, and net change over time throughout the channel network; and 2) relate these changes to both the morphometric characteristics and the characteristics of the contributing area (e.g., precipitation and burn severity). We therefore needed to do the analyses on larger segments rather than at the pixel scale as suggested. The segment length of 50 m is somewhat arbitrary, but we chose 50 m because: 1) this is an appropriate length to characterize the local morphometrics (i.e., channel slope and valley width) as well as the rate of change in these morphometrics given our valley bottom widths (i.e., long enough to minimize local noise but short enough to be relatively homogeneous); 2) the 50-m segment length matches up with the 50-m long longitudinal profiles that we were surveying at each cross-section; and 3) a rough rule of thumb is that longitudinal profiles should be about 10 times the channel width, and after the

2013 flood our channels ranged from about 2-10 m wide, so a 50-m long segment is “about right”.

Given this comment and a similar comment from the second reviewer, we will insert one or two sentences in the text to explain and justify why we divided the channel network into 50-m segments.

P5L28-29: area-maximum maximum? Is that just a typo or does the second maximum go with the 30-min. rainfall intensity? Maybe rephrase so easier to understand.

This is not a typo, as we took the maximum value from the maximum values for all of the pixels. We have revised the wording to make this less confusing.

P5 L30: Did you generate the burn severity map, if so mention that, if not say where it came from.

At the end of the sentence we provide the reference for the burn severity map (Stone, 2015).

P6L5 is that 7am on one day and 7am the next day? Maybe make that more clear

On line 2 we state that the radar data were corrected with daily rain gage data, and on line 5 we state that the radar precipitation were summed from 0700 to 0700 to match the daily rain gage data. It also is standard observing practice that daily rainfall is measured from 0700 on one day to 0700 the following day, so we don't think that this needs any additional clarification.

P6L17: for the Moody 2013 ref, you should also ref. Kean et al. 2011 doi:10.1029/2011JF002005 See their figure 8 and reconcile that with your current statement.

It appears that Kean et al.'s (2011) Figure 8 indicates that I15 is most closely in phase with peak stage, but the lag between I30 and peak stage is only a few minutes. We can add a reference to Kean et al.'s paper here.

P6L26: state goodness of fit for correlation

We presume that by “goodness of fit values” the reviewer wants us to provide either the correlation coefficient (r) or the coefficient of determination (r^2). First, this would be a result, not in methods. Second, we cannot provide a goodness of fit values in the text because of the very large number of correlations that we compute for our results. We explicitly state on lines 25-26 that we collected a number of morphometric measurements (valley bottom, channel, contributing area), and on lines 26-27 we state that these data were correlated to the calculated volume changes. As shown in Table A1 and A2 we have 9 different morphometric characteristics and 5 different catchment characteristics, and when these are correlated to the volume changes in each watershed, we end up with 182 different correlations (91 for each watershed). The r values from the analysis are displayed graphically in Figure 10. Hence it is not possible or appropriate to specify the correlation coefficients here.

P6L31 ref a figure after “intervals”

This section provides an explanation of how we calculated the channel slopes and changes in slope (“curvature”). We will improve the wording to better clarify how we calculated the slope and curvature, but a figure would be largely superfluous as the methodology is relatively simple and standard.

P6 L31: I had (have) a really hard time visualizing exactly what you trying to say here in the sentence that begins “Topographic curvature” Can you add a figure that is a schematic of what you are doing? A lot hinges on understanding this process, so I think it will be important that people don’t miss what you are saying here.

As indicated in the previous comment, we are revising the text to make the methodology for calculating curvature more explicit. Curvature did not turn out to be a very important variable, so we would respectfully disagree that “A lot hinges on understanding this process.”

P7L32: Why not just extract a line across the DoD at your X-S location?

That may be another approach to checking the validity of the lidar differencing. However, because the analyses in this paper focus on volume changes in the 50-m segments, we felt it was appropriate to compare the segment volumes to volumes calculated from field data.

P8L11: Cool approach!

Thank you!

P9L23: do you mean “The lowest TOTAL amount...”

We have revised this sentence to clarify that total precipitation was lowest during the T1 period.

P9L29: Does mesoscale refer to the 2013 flood? Make sure that is clear

We have defined the September 2013 flood as the “mesoscale flood” as this is consistent with other published accounts of this flood, and we have not used this term to refer to any other flood event. Inserting “2013” before “mesoscale flood” would imply that there was more than one mesoscale flood. We will closely examine every reference to the 2013 mesoscale flood to make sure that it is consistent, and that we explicitly note that there was only one mesoscale storm and flood.

P9L32: Add this rainfall to table 2

We are adding a short table to summarize the total rainfall and the maximum 30-minute intensities for each watershed and each time period.

P10L4: Do you have a way to estimate the size of the footprint of each laser point on the ground?

We do not have any information on the size of each laser footprint on the ground. Table 2 provides the point densities and the mean absolute error for each ALS dataset in each watershed, and we believe that this is sufficient to document the key point, namely that the data quality generally improved over time.

P10L9: I don't think it is accurate to say that "the ALS data ... generally fall along a 1:1 line". There seems to be a lot of deviation.

We agree that this wording was a bit strong, and we have revised this to note that the data generally plot close to the 1:1 line and then note the exceptions for the first time period in Hill Gulch and one cross section for the second period in Skin Gulch. We also note that one would not expect a perfect match because the cross sectional changes are extrapolated out to 50 m to obtain a volume.

P10L15: reference tables after the word "ratios"

We don't understand this comment, as the tables do not provide any data on the similarity in channel slopes, valley widths, or confinement ratios, and the purpose of this paragraph is to present a summary of these data to: 1) provide a more detailed description of these characteristics in the two watersheds; and 2) show that the two watersheds are relatively comparable with respect to their physiographic characteristics.

P10L18: Did you observe step pools?

Our field observations would suggest that there were a limited number of step-pool channels prior to the 2013 mesoscale flood, but these were generally smoothed out during the 2013 mesoscale flood. Since we have quantitative data on channel slope but only qualitative observations on channel type, we focus on channel slope. We also would argue that channel type is not an important control given the large-magnitude changes induced by the post-fire thunderstorms, snowmelt, and mesoscale flood.

P10L20: add "reaches" after channels

We have added "segments" to address the concern of the reviewer, as this terminology is consistent with our study.

P10L31: add "within our LoD" after "deposition"

By "LoD" we assume the reviewer is referring to our limits of detection. Since this caveat would apply to nearly every result pertaining to our DoD methodology, mentioning it here would imply that we should add it everywhere else. Since we are very explicit in noting that we can only evaluate elevation differences and hence volume changes in the channel and valley bottoms, we think it is best not to mention this caveat here to avoid the potential for confusion when reporting all our other DoD results.

P11L10 :is the net deposition number (19000) from the ALS or your cross-sections? There is so much missing data that it is hard to believe this is a complete number. I am more interested in the longitudinal patterns than the specific volume estimates because of the missing data.

Given the limited number of cross sections and our extended explanation of how we analysed the ALS data, we are confused that the reviewer could think that this volume was somehow derived from our cross section data. We assume by “so much missing data” the reviewer is referring to a net deposition calculation using field cross-sections; however, this number reflects ALS differencing, so we do not believe missing data is an issue here. Altogether our study does include 83% of the total channel length in Skin Gulch and 87% of the total channel length in Hill Gulch as stated in Section 4.1.

Figures 6 and 7 present the complete data in space for each watershed and each time period, and the reader can use these to draw their own conclusions, and in Section 4.2 we tried to summarize the longitudinal patterns.

P12L26: There is so much missing data that it is hard to feel confident in the total volumes of erosion/deposition values in SG or HG. Consider focusing more on the patterns.

As noted in our previous comment, it is not clear to us what data are “missing”. There are clear limitations on the change that we can detect, but this is true with every paper that attempts to calculate volume change from measured data. Figures 6 and 7 present the longitudinal data, and a close inspection of these figures show that the longitudinal patterns are very complex. A major result of our study is that the correlation analyses show that the volume changes cannot be explained to a high degree of certainty or resolution. We worked hard to try and identify clear and strong patterns, but our efforts had only limited success. Hence we see no way to focus more on the “patterns” as they are not nearly as clear as implied by the reviewer.

P13L5: “this plus other data...” what other data are you referring to?

We agree that this is ambiguous, and the inference was that “other data” was referring back to the list of previously published work on erosion and deposition after the High Park Fire (p. 3, lines 15-21 of our original manuscript). We will revise this to make it more clear that we are referring to other published studies on the High Park Fire.

P13L6: you mention hillslope scale, but I didn't think you had data on the hillslopes

Hillslope-scale erosion data were published previously, and we will revise this to explicitly indicate the source of the hillslope-scale erosion data and that the larger-scale deposition data are coming from the present paper and the recently published paper on our cross-section data (Brogan et al., *Geomorphology*, 2019).

P13L9: not quite a mass balance here, but I understand why

These are all net volume changes, and we will change “net volumes” to “net deposition” to make this more clear (per normal convention, positive values indicate net deposition and negative values indicate net erosion).

P13L16: “highly correlated” with what?

We appreciate the reviewer noting that this statement is ambiguous, and we will add words to note that the independent variables were correlated with volumes of erosion, deposition, and net change.

P13L33: “erosion occurred in the lower gradient” hmm that doesn’t seem intuitive if slope is a major component of the driving shear stress. Can you help to explain why this makes sense somewhere?

Yes, this initially appears counterintuitive. The reason is that there was more post-fire deposition in the lower gradient downstream segments, and because there was much more sediment than there was more erosion. We have added the following wording at the end of this sentence to explain that the greater erosion was associated with greater amounts of post-fire deposition (“...and this is because these reaches generally had the greatest volumes of post-fire deposition and therefore had much more sediment that was readily available for erosion.”)

P14L6: Is BS_m already defined? On page 5 BS is burn severity

We appreciate this comment. BS_h and BS_m refer to the proportion of the contributing area that was burned at high and moderate severity, respectively, and we have now defined these terms in methods.

P14L8: reference a figure after the word “scatterplots”

As noted above, the amount of deposition, erosion, and net volume change was correlated with each of the independent variables for each time period. Hence this result is a more general result, and in the interest of brevity we did not present any of the hundreds of scatterplots in the paper. A table of the overall correlations is included in the supplemental material.

P16L1: What field data shows grain size?

We will include a reference here to Brogan et al. (2019), where field grain size data are presented.

P16L7: Add a ref like Passalacqua 2015 doi:10.1016/j.earscirev.2015.05.012 I also am not sure I agree with the word “recent” We’re going on >20 years of lidar differencing

Thank you for this, and we have revised the text to remove the word “recent” and made the text more explicit. In doing so we have obviated the need for a reference.

P16L11: “the predominant post-fire effect is deposition in the channels and valley bottoms” This is a more general statement than I think you are intending. For example, I don’t think you would argue that this is necessarily true for the Poudre River. That is a channel/valley bottom, but it

sounds like there was not extensive deposition there. So I suggest just refining the language to focus on the spatial scale at which you think it is representative.

We appreciate this comment, but we do not say that deposition is universal, only that it is the predominant post-fire effect. We acknowledge appreciate the importance of spatial scale as suggested by the reviewer, and we therefore inserted the word “downstream” because incision is the predominant post-fire response at the hillslope scale. In confined valleys with large amounts of stream power there may not be widespread post-fire deposition until the channels and valley bottoms widen out, but even in the confined reaches of the Cache la Poudre River large amounts of coarse sediment were occasionally delivered into the river, and these did create relatively persistent alluvial fans. Hence the statement is more generally true, and we provide a series of references to support this statement.

P16L15 what fraction of the channel network does your ALS capture?

The point we were making is that the measured cross sections and longitudinal profiles represent only a small fraction of the channel network, so the DoD of repeated ALS surveys is needed in order to assess erosion and deposition over the entire channel network. This section has been rewritten to better contrast the higher temporal resolution field data with the ALS data, which have lower temporal resolution but can evaluate nearly the entire channel network.

P16L28: Seems like you should mention the coarse substrate and depth to water table before this

We agree that we should have mentioned the channel coarsening when reporting the effects of the mesoscale flood, even though the data on this came from the companion field paper, and we will add this to the results.

We don't have any water table data, but infer this from the coarse substrate and the extensive channel incision that causes the stream water surface to be substantially lower than the elevation of the coarse deposits adjacent to the channel.

P16L28: What exactly do you mean by stripping and coarsening of the channels?

This statement is in reference to the changes induced by the mesoscale flood. As noted in the previous response, we will provide more description of the extensive channel erosion (stripping) and coarsening that occurred as a result of the mesoscale flood.

P17L2: string “large” and add “documented” after “debris flows”

Will do.

P17L15: I don't think you actually mean “allow researchers to be repeated” consider clarifying

Good catch, this will be changed to “allow researchers to collect data at a sufficiently...”

P17L25: you qualitatively describe lidar here, why not just suggest a point density (pt/m²) that you think would be good to shoot for.

Good suggestion; we will change the text to note that the highest mean point density we had from our data was 3.8 pts m⁻². We will therefore recommend a minimum point density of 4 pts m⁻², noting that higher point densities would allow for a more detailed and accurate analysis.

P18L15: What proportion of the area was reduced from this approach?

Vegetation removal reduced the analyzed areas in both valleys by about 2%. We will point this out in the text, as it shows how significant an effect vegetation artifacts in just a small area can have on overall volume differencing calculations.

P18L18: Your 7th point seems pretty obvious, but I guess the people at NEON didn't think about that. I thought it was typically standard practice.

Yes, it should be obvious for volume differencing studies, although researchers with other interests (e.g., vegetation succession) may prefer data collection at different times of year.

P18L24: ref a figure at end of this sentence

We will refer to Fig. 10 here.

P18L25: ref a figure at end of this sentence

We will refer to Fig. 10 here.

P19L2: As far as I can tell, sediment availability is not something you measured (is it measureable?). Your results may allow you to make some inferences about sediment availability, but I don't think that the way things are presented right now allow you to say that the geomorphic changes were largely controlled by sediment availability.

Please see the response to the general comment above related to this topic. Again, we are interpreting sediment availability to observations or measurements of local deposition, which presumably provides sediment available for subsequent erosion and transport.

P19L12-14: Maybe they aren't correlated because you calculated them across 50 m averages.

This may be the case, but we think that 50 m is an appropriate window for computing valley widths, and should be a reasonable approximation of local slope in the absence of a sub-segment-scale knickpoint or other local discontinuity.

P19L16: What data do you have on sediment supply?

This sentence refers to sediment availability – we will change “supply” to “availability.”

P19L21: Am I missing something? How do you know that sediment availability increased? What data are you pointing to for this statement?

This and other comments about sediment availability make it clear that the reviewer had a problem with our description of sediment availability. We will try to make the concept more clear in the revision.

P19L29-30: Spatially explicit models are being used: McGuire et al 2016 doi: 10.1002/2016JF003867; McGuire et al. 2017 doi: 10.1002/2017GL074243

Good point, we will cite these studies here.

P20L27: What makes sediment “available”?

Please see comments on this topic above – we will strive to clarify the concept in revision.

Figure 1: Mark the location of Laramie with a dot. What determines the thickness of the blue lines?

We will change the caption to not mention Laramie, and just say that the diamond represents the radar. The caption already designates the thickness of the blue lines.

Figure 2: How did you calculate the maximum intensity?

This is described in section 3.3 in the methods.

Figure 10: Make sure to say these are “Pearson” correlation coefficients. They are averaged for each time period, right?

Will do. Yes, these correlation coefficients represent relationships derived from ALS datasets collected at either end of the time period.

Figure 11: Consider using equal axes in A and D.

We prefer this figure as presented. Erosion and deposition rates in SG and HG differed enough that showing the plots at the same scale would compress the HG results and make the data more difficult to visualize.