**Review response to 11-Nov-2025 (RC1)**

**Reviewer Comments**

Black – reviewer; Blue – response

Thank you for this opportunity to review this manuscript. This study details a remote sensing and field-based approach to understanding how post-wildfire debris flow susceptibility changes through time and is subject to seasonal variations following fire in five headwater basins burned in the Lake Fire (2020) located in the Transverse Ranges of southern California, USA. The authors combine information from field visits four years following fire where the authors were able obtain only outlet grain size measurements immediately following the fire (Year 0) and then also obtain grain size and soil hydraulic properties in two visits four years following the fire. In intervening time periods between Year 0 and Year 4, the authors were able to estimate dNDVI using higher-resolution spatiotemporal data from PlanetScope data to assist with estimating vegetation recovery at seasonal timescales and possible surface property changes related to sediment deposits. These data and some assumptions based on similar previous datasets collected in the region (particularly for Year 0 where field data was limited to only grain size) were then used to calibrate a process-based debris-flow likelihood model. From the authors' modeling approaches their main claims are that grain size and seasonally-variable recovery of hydraulic conductivity were greater controls on debris flow susceptibility than vegetation recovery.

I have many concerns with this study stemming from issues related to basic information on the data presented, likely flawed assumptions, and whether model results (which run contrary to established conceptual understanding of post-fire hazards) can be trusted given a lack of information on presence or absence of debris flows to validate model outputs given the overall low number of ground-based observations (only three observation campaigns over a four-year period). Therefore, I do not believe that this manuscript is ready for publication and opt for a recommendation to reject it in its present form. Below I provide my reasoning based on the information presented in the manuscript and in the event the authors revise for a future resubmission, I also provide additional detailed line comments.

Thank you for the review. We address the reviewer's comments below in blue.

Specific comments:

Potential flaw #1: Unclear if data presented are reliable and possible flawed assumptions

The authors were able to visit the field site a total of three times, in Year 0 immediately following fire and Year 4 in the wet and dry seasons. For each visit, no information is provided about how and what number of sediment samples were taken (i.e. multiple samples, a composite of outlet sediments, landform position?). Also, because soil infiltration properties are extremely variable, it is also important to describe the study design of mini-disk measurements – what is the sample size and spatial coverage of measurements (authors seem to indicate it may be higher up in the watershed given accessibility concerns in Year 0?). For hydraulic conductivity (I'll herein refer to as Ks – as the authors should also in line with previous literature they cite) – what value is eventually used as an input to the model (e.g., median, mean) and how variable are these values? A mountain of evidence provides information on just how variable this can be (see McGuire et al., 2018 and Ebel, 2022 for some examples). From the standpoint of Ks this becomes important given the finding of its importance in model outputs for these basins from sensitivity testing. Scientific reporting should always include this set of basic information (e.g. summary statistics).

Thank you for the comments. For the Ks measurement, we will clarify the procedure of obtaining our Ks in the main text. We only report the initial Ks measurement from locations at the outlet of the basin. We did perform a 2nd Ks measurement (up to several meters from the first measurement) but because these measurements were located further away from the basin outlet, we only used them for comparison. We understand that most studies perform multiple measurements within the basin and average those measurements. But we believe that the best way without disturbing the topsoil or flattening the topography is to collect measurements and sample for grain size analysis at the outlet of the basin where basin inputs are integrated. We will also add more description of the potential limitations of this method in the discussion.

Additionally, there are some assumptions that are potentially flawed: Let's begin with Year 0, where grain size values were measured at the outlet of the watershed under a shaky assumption that grain size measurements taken at an outlet should be representative of the full distribution of grain sizes from colluvial sources located upstream and reference a study (Santi et al., 2008) that I do not believe makes this claim. Firstly, there is a lot of dry ravel as indicated in Google Earth historical imagery of the site and Figure 2 – Panel A that are common to the Transverse Ranges (see Lamb et al., 2011). Given that field sampling occurred following very limited rainfall and associated runoff (Figure S2) it is very unlikely that a grab sample at the outlet is representative. Runoff events such as a debris flow could act to mix these sources of sediment sources but this did not appear to be the case here (and a judgment of the representativeness of such sampling is also hampered by a lack of information on sample collections indicated above). There is less concern for this particular issue in Year 4 samples following runoff, however, even in the event of a debris flow, there can be large variations in grain size that can occur depending on where sampling takes place – such as very coarse clast-supported deposits such as boulder trains or levees versus more fine-grained poorly-sorted matrix-supported portions of a debris deposit.

Thank you for the detailed comments. We responded below to individual points from this comment. We acknowledge these issues and agree that the lack of debris flows may result in less mixing, especially at the basin outlets and that collecting upslope measurements would be the best representation of the basin and debris flow source. However, given the difficulty accessing upslope, we believe this method could provide a better representative of the basins. We will include this in the discussion.

Also, debris flow deposits can also be reworked by subsequent runoff so this evidence is perishable. Therefore, measurements should account for this depositional variability which is presently not clear in the methods. Other studies have gone beyond this assumption to ensure that d50 measurements are representative through intensive sampling in the watershed (see Tang et al., 2019) but this does not appear to the be case in the present manuscript.

Thank you for the comment. We understand that debris flow deposits can be reworked. However, from our field sampling we did not identify any indication of reworked debris flow deposits at the basin outlet.

Lastly, the vegetation cover approach also remains concerning to me. Were ground measurements using approaches such as point-intercept surveys ever used to validate satellite-based dNDVI approaches? Given that the original parametrization of the dimensionless-discharge model used ground cover data taken in the field, it might be important to be sure that satellite data can adequately represent this parameter through ground-truthing.

We performed a validation using the 20 meter point-intercept survey at along the riverbanks that is accessible outside the burn scar and found an ~81% vegetation cover. The mean NDVI along the same transect on the same day is 0.5191 (max NDVI within image is 0.74). The dNDVI from between the field measurement and 7 months before (Sept 2023) has a mean of 0.034 (indicating not many changes).

We note that the point-intercept transect method is not necessarily a direct comparison to the pixels of satellite imagery. Thus, we used the percentage of vegetation and non-vegetation for comparison to derive vegetation cover of an area from each method. We will address this point in additions to the main text and supplement.

To assess the reliability of the satellite-derived measurements, we also utilized ground surveys within a 3 × 3 m field plot for validation. However, we agree that a more detailed approach is needed in the future to better characterize the box plot method for vegetation cover as satellite imagery captures the canopy cover better than the understory vegetation.

Potential flaw #2: Lack of true validation of debris flow susceptibility modeling

This major comment stems from the fact that the authors seemed to be unable to make unambiguous measurements of runoff response in their basins. Therefore, even with relatively robust field measurements (which is already called into question above), it is difficult to know if the model is adequately representing basin-scale runoff response without independent data on presence or absence of debris flows (or some similar proxy beyond what seems to be ambiguous satellite indices and sparsely-timed field observations). Part of the issue may arise from the fact that rainfall intensity-duration thresholds were never met at all during the study period so it may be that this monitoring experiment was ill-suited to begin with (this is a challenging aspect of post-fire monitoring work where we have little control over what happens with post-fire weather!). The authors do highlight this possibility (of overall low peak rainfall I15) at least, but it does not help the case here.

Thank you for the comment and pointing out the difficulty of post-fire weather. We agree that not having validation of debris flows makes it difficult to validate model results. However, the focus of this study was to understand how different model parameters could influence the PFDF likelihood specifically for Lake Fire. We used the samples and observations from this fire to inform and provide us with reference values for the different parameters. We will include this description in the next iteration of the paper.

Additionally, at face value, the authors' work seems to show debris flow risk sticking around for up to four years, in fact even increasing due to possible seasonal-scale variability in Ks (where drier soils generally, except for Basin A, resulted in large increases in Ks). These findings seem to be contrary to large databases of post-fire debris that find a precipitous drop in runoff-generated debris flow occurrence 3+ years out from wildfire (see Graber et al., 2023), particularly as vegetation recovery takes hold. Therefore, a suggestion of possible increases in runoff-generated debris-flow risk with increasing time since fire (as this study seems to indicate, e.g. Figure 5) is contrary to many studies and a general conceptual understanding of time evolution of runoff-related post-fire hazards (for example guidance on the USGS M1 model used for emergency assessments does not advise using rainfall ID thresholds beyond Year 2 following fire). Less ambiguous and more well-supported evidence to validate such modeling approaches would be needed here to make such claims, which is especially salient in post-fire hazard science where this work is used to inform decision-making related to public health and safety concerns.

Thank you for the comment. We agree with Graber et al., 2023 that post-fire debris flow occurrence will drop years after the fire due to factors such as vegetation, grain size changes, and hydraulic conductivity. We also want to point out that there was significant debris flow in Oak Glen 5 years after the El Dorado Fire in 2020. This event highlights that even though likelihood decreases, the potential for debris flow still exists.

We also understand the confusion from Figure 5 (showing the apparent increase in likelihood). We will add a more detailed explanation ßinto the discussion (see below), along with a note that more studies are needed to focus on seasonal or short-term changes.

We propose that the increase of debris flow likelihood in the wet season in the 4$^{th}$ year could reflect short-term variability such as intense rainfall precipitation immediately before our field collection. Given the changes between the wet and dry seasons in the 4$^{th}$ year, we suggest that this increase might only be temporary or shorter-term rather than a long-term yearly increase of debris flow likelihood. This short-term change in likelihood is also observed in Martinez & McGuire (2025) Figure 5, where they show hydraulic conductivity is reduced in a wetter season.

Thus, we provided a series of model parameter tests and found Ks to have a strong influence in our modeled PFDF likelihood.

**Here are some specific line comments:**

Line 50: Tillery and Rengers, 2020 reference – I don't think this study supports the idea of post-fire debris flow risk persisting 10 years or more following fire, this study looked at adjacent burned vs unburned debris flow initiation mechanisms. I would opt for Graber et al., 2023 instead for a rich discussion of this in addition to DeGraff et al., 2015.

Thank you for the suggestion; we will change this accordingly.

Lines 218-219: "Similarly, we assumed no runoff generated debris flow occurred during this period due to the lack of high-intensity rainfall from our simulations" is an unfortunate choice in phrasing that implies you are only relying off simulated rainfall for making this assumption. I would think it would be better to say based on the identified rainfall-ID threshold of __mm/hr, we found that it was unlikely debris flows were initiated because rainfall values never exceed this threshold (or something similar – needs to be supported by observations).

Thank you for the suggestion; we will change this accordingly.

Line 266: Related to the idea that basin topography could be an important control on rainfall ID threshold variability across basins – this could be easily assessed by looking at basic watershed morphology (listing basin mean slopes or channel slope variation for example). In fact, I would report these somewhere as standard practice in many post-fire geomorphic studies.

Thank you for the suggestion. We will add this to the supplement.

Line 277: Related to the comment about opening macropore spaces in drying soils - but what about the literature that supports dry soils promoting enhanced soil water repellency? (see Ebel & Moody, 2013)

From analysis of our field measurements, we did not see dry soil repellency as a major influence potentially due to the lack of soil development and coarser grain size at our site.

Line 318-320: I'm glad the authors acknowledge this limitation here and something to consider more, especially for the Year 0 sampling where there did not seem to be evidence of debris-flow deposits being sampled.

Lines 321-330: The guesses here about observed shifts in grain size are interesting but I think a more careful consideration of exactly where deposits were taken could help with this. For example, did the authors see evidence of stratification of where deposits were taken (e.g., fines sitting over coarser material supporting theory #2) or inferences type of flows depositing the sediments based on sedimentary characteristics (e.g., debris flow deposits are generally more poorly sorted versus alluvium which is more well-sorted and normally-graded). If such information was recorded in the field, it could help with constraining this and may also help to elucidate concerns brought up earlier related to representativeness of grain size estimates for susceptibility modeling. If not, this is a learning opportunity for the authors to more carefully plan sampling designs and collect this information in the future (as they indicate in the future work statement at the close of the paragraph).

Thank you for the comments. We performed a soil core up (~30 cm) and did not observe any clear indication of stratification or debris flow deposits. We will add more information to the supplement.

Line 360-361: Variability in rainfall patterns influencing debris flow triggers: what variability type is being referred to here? Spatial or temporal variability? Not really sure how useful this sentence is.

Thank you for the comment. We meant by temporal variability instead of a gaussian rainfall. We will clarify this in the main text.

Lines 368-371: I agree here, I think the inferences made from satellite data without ground validation are too generous from satellite data alone (as outlined in major comments).

Line 381: Typo to fix: “larger to not be affected” should read “large enough to not be affected.”

Thank you for the suggestion, we will fix it accordingly.

Lines 423-424: Related to the primary succession of grasses: Does your modeling approach of relying on satellite dNDVI account for the growth of these grasses?

Thank you for the clarification. Yes, it does. Our field observation confirmed grasses in large areas adjacent of the river is captured by the dNDVI.

Also for some of the figures: I would recommend changing the scale bar to be more legible (i.e. add a white background) as they can be a bit hard to read.

Thank you for the suggestion, we will fix it accordingly.

**References cited in this review beyond those cited in study:**

Ebel, B. A., & Moody, J. A. (2013). Rethinking infiltration in wildfire-affected soils. Hydrological Processes, 27(10), 1510-1514.

Ebel, B. A. (2022). The statistical power of post-fire soil-hydraulic property studies: Are we collecting sufficient infiltration measurements after wildland fires?. Journal of Hydrology, 612, 128019.

Graber, A. P., Thomas, M. A., & Kean, J. W. (2023). How long do runoff-generated debris-flow hazards persist after wildfire?. Geophysical Research Letters, 50(19), e2023GL105101.

Lamb, M. P., Scheingross, J. S., Amidon, W. H., Swanson, E., & Limaye, A. (2011). A model for fire-induced sediment yield by dry ravel in steep landscapes. Journal of Geophysical Research: Earth Surface, 116(F3).

McGuire, L. A., Rengers, F. K., Kean, J. W., Staley, D. M., & Mirus, B. B. (2018). Incorporating spatially heterogeneous infiltration capacity into hydrologic models with applications for simulating post-wildfire debris flow initiation. Hydrological Processes, 32(9), 1173-1187.