The authors did a great job illustrating the importance of ground cover over infiltration capacity in triggering PFDFs. I believe the study approach they applied will provide a template and foundation for other scientists in the field to emulate and build from, respectively. I appreciate the concise, easily-digestible writing style they applied. In addition, I thought the figures were well thought out and do a great job presenting the data. I provide minimal comments in the attached pdf for the authors to consider.

Line 426: I recognize the USGS suggests using P=0.5 to assess PFDF hazards under emergency conditions; however, because you are ultimately comparing the M1 model results to actual debris flow results, I believe it would be good to also show measured triggering I15 against the M1 model results with P=1.

R: Thank you for this suggestion. As the reviewer notes, it is common to set p=0.5 when using a logistic regression model as a classifier (i.e. to determine whether or not a debris flow will initiate). Staley et al. (2017) set p=0.5 when using the M1 model to determine a rainfall ID threshold in their study. Setting p=1 would result in an undefined rainfall ID threshold, but we can compare the observed debris flow triggering intensities with the rainfall intensity required to achieve p=0.9. Defining an ID threshold based on when the M1 model returns a likelihood of 0.9 would be considered very conservative. Below, we summarize a few results focusing on rainfall thresholds computed for a 15-minute duration using p=0.5 and =0.9.

Defining the threshold based on when p=0.5 and p=0.9 results in mean M1 modeled thresholds of 17 mm/h and 28 mm/h, respectively. On average, the M1 modeled threshold defined by p=0.5 is 36 mm/h less than the observed triggering intensity for the 9 debris flows at our site where we can constrain the debris flow timing within storms. Similarly, the M1 modeled threshold defined by p=0.9 is 24 mm/h less, on average, than the observed triggering intensity for the 9 debris flows at our site where we can constrain the debris flow timing within storms. This supports our conclusion that the M1 model underestimates debris flow triggering intensities at our site. Since determining debris flow initiation thresholds using p=0.5 and p=0.9 does not change the relative susceptibility ranking of watersheds (i.e. which watersheds have lower or high M1 thresholds remains unchanged regardless of the choice of p), these calculations also do not change our conclusion that the M1 model does well at assessing the relative susceptibility of different watersheds to debris flow responses. Given the consistency of results regardless of the choice of p, we have decided not to make a change to the manuscript and only presents results using the more standard value of p=0.5.

Line 498: It would be good to identify the range of grain sizes in the pebble count and the sieve analyses. At one point I believe the point count was performed on grain sizes >2mm. However, it appears here that the point count was done on grain sizes >20mm. Some clarification would be good.

R: The D50 referenced in line 498 was determined from sieve analyses of debris flow sediment. We used sieve sizes of 32 mm, 16 mm, 8 mm, 4 mm, and 2 mm. The D50 ranged from < 2 mm to roughly 20 mm. We have added information about the sieve sizes to the methods section: "These samples were air dried and sieved, using sieve sizes of 32 mm, 16 mm, 8 mm, 4 mm, and 2 mm…"

Line 570: "The temporal distribution of rainfall within rainstorms that produced debris flows (red lines) are similar, with the majority of rainfall occurring during the second quarter of the storm duration (0.25 ≤ normalized time ≤ 0.5)." This is interesting. This implies that there may be a linkage between initial abstractions before runoff occurs that then trigger PFDFs. This is partially suggested by the equally steep curves left of the red curves that have similar peak depths. The only apparent difference between the two sets of curves is the presence of low rainfall in Q1 for the red curves to infill voids prior to initiating overland flow. Moody, 2012, covers this a little by assuming some initial rainfall goes towards saturating surface depressions and near surface voids before runoff develops. This may be a compelling observation that may warrant additional consideration for future studies dedicated to triggering rainfall events.

R: Thank you for this comment and the opportunity to clarify some aspects of the SRP results. When interpreting results of the SRP analyses, it is important to keep in mind that only the normalized rainfall depth is used. We have added a clarifying statement to the caption of Figure 9: "Note that the standardized rainfall profiles are plotted using normalized rainfall depth so the curves do not provide information of the absolute value of rainfall depth during different portions of the rainstorm."

Since the SRP analyses do not depend on the absolute value of rainfall, it is challenging to use them to assess the role of initial rainfall abstractions in the debris flow generation process. There are grey curves to the left of the red lines in Figure 9 that are equally steep, but this only provides information about the temporal distribution of rain within the storm and is not indicative of the overall depth of rainfall or the depth of rainfall that occurred early in the storm. The SRPs are helpful for a quick visual assessment of the type of rainfall event (e.g. convective vs frontal system) that did or did not trigger a debris flow. This technique was recently applied by Esposito et al. (2023) in their study of PFDFs in Italy. We think that analysis of SRPs is one method that may help assess similarities and differences among the types of rainfall events that trigger PFDFs in different regions. We added the following to the discussion section:
"Standardized rainfall profiles of debris-flow producing storms generally plotted above the 1-1 line (black line in Figure 9b), which is a characteristic associated with convective rainstorms (Esposito et al., 2023). This finding is consistent with the timing of debris flows during the summer months shortly following the fire when convective rainstorms associated with the North American monsoon are common in the region. Esposito et al. (2023) similarly

found that storms that produced PFDFs in Italy had SRPs consistent with convective rainstorms rather than frontal systems."