Response to Dr. Paul Santi document for NHESS manuscript 2021-345 by Li et al.

General Comments

The paper is very well written and easy to follow, and it is a nice integration of modern modeling techniques and data for use in debris flow analysis.

Response: We thank Dr. Paul Santi for his constructive comments of our manuscript. We plan to make specific changes to the manuscript in response to each of his comments, and believe the manuscript will be significantly improved as a result.

I think a slight change in the declared focus of the paper will better highlight its value. Allow me to explain. At many points in the paper, the authors have gone to a lot of trouble to set up, run, and calibrate models that basically demonstrate the same things that have been said (and quantified) in other papers using much simpler analyses: debris flow volume and discharge increase multifold in burned areas, the hazard is concentrated in stream channels, and there is a lag between rainfall peaks and flow events, for example. However, the authors' analyses provide some information that has not been clearly shown before. Importantly, they are able to create calibrated time graphs of streamflow and discharge. Also, they are able to compare their models with the USGS post-wildfire assessments to show differences (they refer to this in lines 652-656, but don't give details of the analysis). I think the paper would be stronger if they acknowledge early on that other research has demonstrated (and quantified) changes in volume, discharge, and lag. Then they could focus on the advantages offered by a more sophisticated, calibrated model.

Response: We thank Dr. Santi for his suggested re-framing of the manuscript. We agree that the current iteration of the manuscript is heavily focused on the Dolan wildfire burn scar case study, and that the manuscript would be better served if it had a greater focus on our modeling advance and its potential for future applications and advancements. In our revision, we will re-focus the latter half of our abstract to better emphasize the advance and potential of using WRF-Hydro in debris flow studies and forecasts. As requested, we will also acknowledge (and cite) early in the manuscript that others have demonstrated similar debris flow hydrologic behaviors using numerical models in limited domains, while emphasizing the value added by our regionalized and fully-distributed hydrologic simulations. In addition, rather than focusing on our case study results at the start of the Discussion section, we will provide a more general summary emphasizing the advances provided by using a model like WRF-Hydro to investigate debris flows.

I think the discussion should also include a section on applying the model elsewhere. Is it realistic to do this for other sites, or is it too dependent on specific calibration parameters? How could a practitioner do this type of analysis? What does it offer a scientist that they do not already know?

Response: This is another excellent suggestion, and has parallels with a recent news piece found in *Nature* (Palmer, 2022). WRF-Hydro can indeed be applied elsewhere. The model has been used to study a diverse range of hydrological processes in domains of varying size across the world. Use of WRF-Hydro and choice of spatial resolution is dependent on the existence of requisite boundary conditions, forcing files, and observational constraints. For studies focused on burn scar-debris flow dynamics, the model is again readily adaptable. In our approach, we demonstrate that the burn scar characteristics of a land surface can be set in the land surface model (i.e., reduced canopy height, overland roughness, carboxylation rate, and infiltration rate). Since parameter values will vary based on a myriad of factors (i.e., geography, climate, biome, soil properties, etc.), a major advantage of WRF-Hydro is the ability to modify and calibrate the underlying physical parameters as appropriate for each location. In our revised version of the manuscript, we will discuss the application of WRF-Hydro in other regions, and provide a roadmap for global community usage.

The discussion could also compare their model to the USGS model, using a modification to Figure 9, for example, to demonstrate and explain the important differences.

Response: Referee #2 has mentioned a key point that differentiates our susceptibility assessment from the USGS' hazard assessment. That is, the USGS statistical models are able to predict both probability and magnitude of debris flows, which makes them "hazard" assessments, whereas our model is focused on predicting which areas are subject to higher likelihood and should be referred to as "susceptibility" assessment. In the revision, we plan to explain this terminology and alter its usage throughout the manuscript and in the title. We will also discuss methods that could be employed to create probabilistic hazard assessments using WRF-Hydro, which would facilitate an apples-to-apples comparison with the USGS product.

Specific Comments

Section 5.4 - I don't feel that this is a strong section. It concludes that the hazards are greater in the burned area, and mostly in the channels, and that streamflow is elevated downstream in burned areas, which are not unique findings. Likewise, Figure 11 doesn't come across as strongly as previous figures. I suggest dropping this section.

Response: We somewhat agree with this suggestion. In our revision we will remove the regional discussion and Figure 11 from the results portion of the manuscript. However, one of the values added by WRF-Hydro is regional prediction and projection, which differs from more traditional single catchment simulations (e.g., McGuire et al., 2016, 2017). To highlight these capabilities, particularly from a future usage perspective, we will move Figure 11 to the discussion section and discuss why regional applications, particularly in an operational setting, provide value.

line 489 ff - an interesting note, your modeled discharge increases by 3 or 4 fold matches field measured changes published in Brunkal and Santi for large drainage basins (I could not find the area for your drainage basins, since you include normalized values, but I assume they are more than 5 km^2) (Brunkal, H. and Santi, P., 2017, "Consideration of the Validity of Debris-Flow Bulking Factors," Environmental and Engineering Geoscience, DOI: 10.2113/EEG-1774). See Figure 3 of this paper.

Response: In the revision, we will cite this paper in the discussion section to highlight the similarity.

Technical Corrections

Figures 1, 7. and 9 could benefit from a bar scale.

Response: In the revision we will add scale bars to Figures 1, 7, and 9.

Figure 9 - the legend is hard to understand. I assume the first bar is volume and the second is normalized volume?

Response: Yes, the first bar is volume and the second is normalized volume. We will revise the legend to make this clearer.

We thank Dr. Paul Santi again for his careful review and constructive comments.

References

McGuire, L. A., J. W. Kean, D. M. Staley, F. K. Rengers, and T. A. Wasklewicz (2016), Constraining the relative importance of raindrop- and flow-driven sediment transport mechanisms in postwildfire environments and implications for recovery time scales, *J. Geophys. Res. Earth Surf.*, 121, 2211– 2237, doi:10.1002/2016JF003867

McGuire, L. A., Rengers, F. K., Kean, J. W., and Staley, D. M. (2017), Debris flow initiation by runoff in a recently burned basin: Is grain-by-grain sediment bulking or en masse failure to blame?, *Geophys. Res. Lett*., 44, 7310– 7319, doi:10.1002/2017GL074243.

Palmer, J. (2022). The devastating mudslides that follow forest fires. *Nature.* https://www.nature.com/articles/d41586-022-00028-3.