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

*Response is in blue & Sentences or paragraphs after revision are in red

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 have made specific changes to the manuscript in response to each of his comments, and believe the manuscript has been 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 previous iteration of the manuscript was heavily focused on the Dolan wildfire burn scar case study, and that the manuscript is better served now because we put a greater focus on our modeling advance and its potential for future applications and advancements according to Dr. Santi's suggestion. In our revision, we re-focused the latter half of our abstract to better emphasize the advance and potential of using WRF-Hydro in debris flow studies and forecasts.

The abstract now reads:

"In steep wildfire-burned terrains, intense rainfall can produce large volumes of runoff that can trigger highly destructive debris flows. However, the ability to accurately characterize and forecast debris-flow susceptibility in burned terrains using physics-based tools remains limited. Here, we augment the Weather Research and Forecasting Hydrological modeling system (WRF-Hydro) to simulate both overland and channelized flows and assess postfire debris flow susceptibility over a regional domain. We perform hindcast simulations using high-resolution weather radar-derived precipitation and reanalysis data to drive non-burned baseline and burn scar sensitivity experiments. Our simulations focus on January 2021 when an atmospheric river triggered numerous debris flows within a wildfire burn scar in Big Sur – one of which destroyed California's famous Highway 1. Compared to the baseline, our burn scar simulation yields dramatic increases in total and peak discharge, and shorter lags between rainfall onset and peak discharge, consistent with streamflow observations at nearby U.S. Geological Survey (USGS) streamflow gage sites. For the 404 catchments located in the simulated burn scar area, median catchment-area normalized discharge volume increases nine-fold compared to the baseline. Catchments with anomalously high catchment-area normalized discharge volumes correspond well with post-event field-based and remotely-sensed debris flow observations. We suggest that our regional post-fire debris flow susceptibility analysis demonstrates WRF-Hydro as a compelling new physics-based tool whose utility could be further extended via coupling to sediment erosion and transport models and/or ensemble-based operational weather forecasts.

Given the high-fidelity performance of our augmented version of WRF-Hydro, as well as its potential usage in probabilistic hazard forecasts, we argue for its continued development and application in post-fire hydrologic and natural hazard assessments."

As requested, we also acknowledged (and cited) 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. Rather than focusing on our case study results at the start of the discussion section, we provided a more general summary emphasizing the advances provided by using a model like WRF-Hydro to investigate debris flows.

The first paragraph of our revised discussion section in lines 639 – 656 now reads:

"Given the historic and growing frequency of wildfires in the western U.S. (Williams et al., 2019; Goss et al., 2020; Swain 2021) and globally (Flannigan et al., 2013; Jolly et al., 2015), developing tools to investigate, better understand, and potentially predict changes in burn scar hydrology and natural hazards at regional scales is critical. Here, we demonstrate the first use of WRF-Hydro to simulate the susceptibility of a burn scar to postfire debris flows during a landfalling AR. We augmented the default version of WRF-Hydro to output overland flow and to replicate burn scar behavior by adjusting vegetation type and infiltration rate parameters. WRF-Hydro simulations were validated against PSL soil moisture and USGS streamflow observations before we used simulated streamflow and overland flow volumes to characterize debris flow susceptibility. A comparison between baseline and burn scar simulations demonstrated that changes in hydraulic properties of burned areas causes drastic changes in surface flows, including faster discharge response times, and greater peak discharge and total volumes, consistent with findings from previous postfire hydrology studies (Anderson et al., 1976; Scott, 1993; Meixner & Wohlgemuth, 2003; Kean et al., 2011; Kinoshita & Hogue, 2015; Brunkal & Santi, 2016; Williams et al., 2022). At the catchment level, for the 404 catchments located within the Dolan burn scar, median catchment area-normalized volume increases nine-fold relative to the baseline. In addition, Mill Creek, Big Creek, and Nacimiento basins were simulated to have high-to-very high debris flow susceptibility, corresponding well with identified debris flow occurrences."

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 discussed the application of WRF-Hydro in other regions, and provided a roadmap for global community usage.

In lines 758 – 775, we add a paragraph in the discussion section that reads:

"Lastly, the above discussion of potential WRF-Hydro applications and advancements speaks to the adaptability and customization of this open-source numerical model. An additional layer of WRF-Hydro's adaptability concerns its geographic focus. While we calibrate and use the model over a central California domain, the choice of geographic footprint is limited only by the availability of requisite initial and boundary conditions, environmental observations for calibration, and computational resources. For use in non-central California domains, we recommend calibration beginning with the default version of the model. Given the ecological and geological diversity of locations that experience wildfires and debris flows, it is likely that calibrations distinct from those reported here will be needed in different regions. For example, soil sealing effects, infiltration, and runoff in wetter and more vegetated locations, such as Oregon, USA, behave differently than those in central California (Palmer, 2022). As such, calibration of relevant model parameters (e.g., saturated hydraulic conductivities) should be based on a physics-informed approach that accounts for local environmental conditions and hydrologic behaviors. Indeed, given the ability to simulate large heterogeneous geographic domains, it is likely that different regions within a given domain may require different calibration schemes. As WRF-Hydro is fully distributed, spatially heterogenous calibrations are non-problematic. This spatial adaptability may prove particularly helpful in post-wildfire debris flow hazard assessments when considering multiple generations of wildfires and variable degrees of burn scar severity and recovery."

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 explained this terminology and altered its usage throughout the manuscript and in the title.

In lines 81 – 90, we explain the difference between "susceptibility" and "hazard": "However, due to long-standing terminology ambiguity in the natural hazard community (Reichenbach et al. 2018), we first begin with a definition of terms. In this study we demonstrate the use of a new physics-based tool to map postfire debris flow susceptibility at regional scales. We follow the guidance of [Reichenbach et al. (2018) & references therein] and define susceptibility as the likelihood of debris flow occurrence in an area, and hazard as the probability of debris flow occurrence of a given magnitude within a specified area and period of time. In other words, debris flow susceptibility does not estimate debris flow size or consider the timing or frequency of the debris flow occurrence. Rather, it focuses on locating areas prone to debris flows considering local environmental factors (Brabb 1985; Guzzetti et al., 2005)."

We also discussed methods that could be employed to create hazard assessments using WRF-Hydro, which would facilitate an apples-to-apples comparison with the USGS product.

In lines 714 – 737: "In addition to investigating the operationalization of WRF-Hydro's natural hazard prediction capabilities, we note that our susceptibility-focused methodology could be advanced to hazard assessment, in line with current USGS products. The USGS Emergency Assessment of Postfire Debris-flow Hazard predicts debris flow volume and likelihood. To advance from susceptibility to hazard assessment, our methodology would need to incorporate both debris flow volume estimates and occurrence likelihoods. In the following, we highlight research directions that could help advance our susceptibility-focused methodological framework. WRF-Hydro is a water-only model. While water-only models have been widely used to investigate and better understand debris flow dynamics (Arattano & Savage, 1994; Tognacca et al., 2000; Arattano & Franzi, 2010; Rengers et al., 2016; McGuire & Youberg, 2020; Di Cristo et al., 2021), sediment supply, soil erodibility, and other sedimentological factors play important roles in determining the potential for and severity of mass failure events (McGuire et al., 2017). Developing a runoff-generated debris flow model that couples hydrologic and sediment erosion and transport processes could help to characterize postfire debris flow volumes. Indeed, previous efforts have demonstrated the capacity to couple WRF-Hydro with sediment flux models (Yin et al., 2020; Shen et al., 2021). In addition to sediments, burn scar ash can comprise a substantial fraction of the total debris flow volume (e.g., Reneau et al., 2007). As such, efforts to constrain ash availability and entrainment in hydrologic flows could prove fortuitous in hazard assessment and prediction efforts. If WRF-Hydro is not coupled with sediment models, a domain-specific rainfall ID threshold trained with historic landslide inventory and triggering rainfall events (Tognacca et al., 2000; Gregoretti & Dalla Fontana, 2007, 2008) or a newly developed dimensionless discharge and Shields stress threshold (Tang et al., 2019; McGuire & Youberg, 2020) could provide guidance to help identify debris flow triggering time and location, which in turn may improve WRF-Hydro's debris flow initiation identification."

In addition, we discussed methods for probabilistic advancement including systemic investigation of parameter uncertainties and use of ensemble-based precipitation data.

In lines 739 – 755, now it reads: "In addition to constraining potential postfire debris flow volumes, WRF-Hydro's application in debris flow studies could be advanced via concerted engagement with uncertainties that are both external (meteorological forcing data) and internal (physical parameters) to the model. Previous studies have demonstrated that precipitation is often the largest source of uncertainty in hydrologic predictive models (Hapuarachchi et al., 2011; Alfieri et al., 2012). Engagement with precipitation forcing uncertainties in past, near-term, and future contexts could provide probabilistic nuance to natural hazard investigations. For example, (a) debris flow hindcast studies could use a diversity of precipitation datasets to isolate precipitation-derived debris flow uncertainties in historic events, (b) operational forecast efforts could utilize ensemble-based weather forecast data to inform likelihood statements in debris flow hazard assessments, and (c) probabilistic projections of debris flow likelihood in future climates could assess and partition uncertainties derived from emission pathway, model structure, or internal variability effects on meteorological forcings (Nikolopoulos et al., 2019; Hawkins & Sutton, 2009; Deser et al., 2020). Uncertainties internal to WRF-Hydro are also ripe for investigation. Probabilistic predictions crafted from an ensemble of perturbed model physics simulations have been used to predict rainfall-triggered shallow landslides (Raia et al., 2014; Canli et al., 2018; Zhang et al., 2018). Similar efforts using WRF-Hydro could target post-wildfire debris flows."

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 removed 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 added four zoomed-in maps of stream-level and catchment-level debris flow susceptibility over two other 2020 wildfire burn scars to display more details. we also moved the modified Figure 11 to the discussion section and discussed why regional applications, particularly in an operational setting, provide value.

Revised Figure 11:

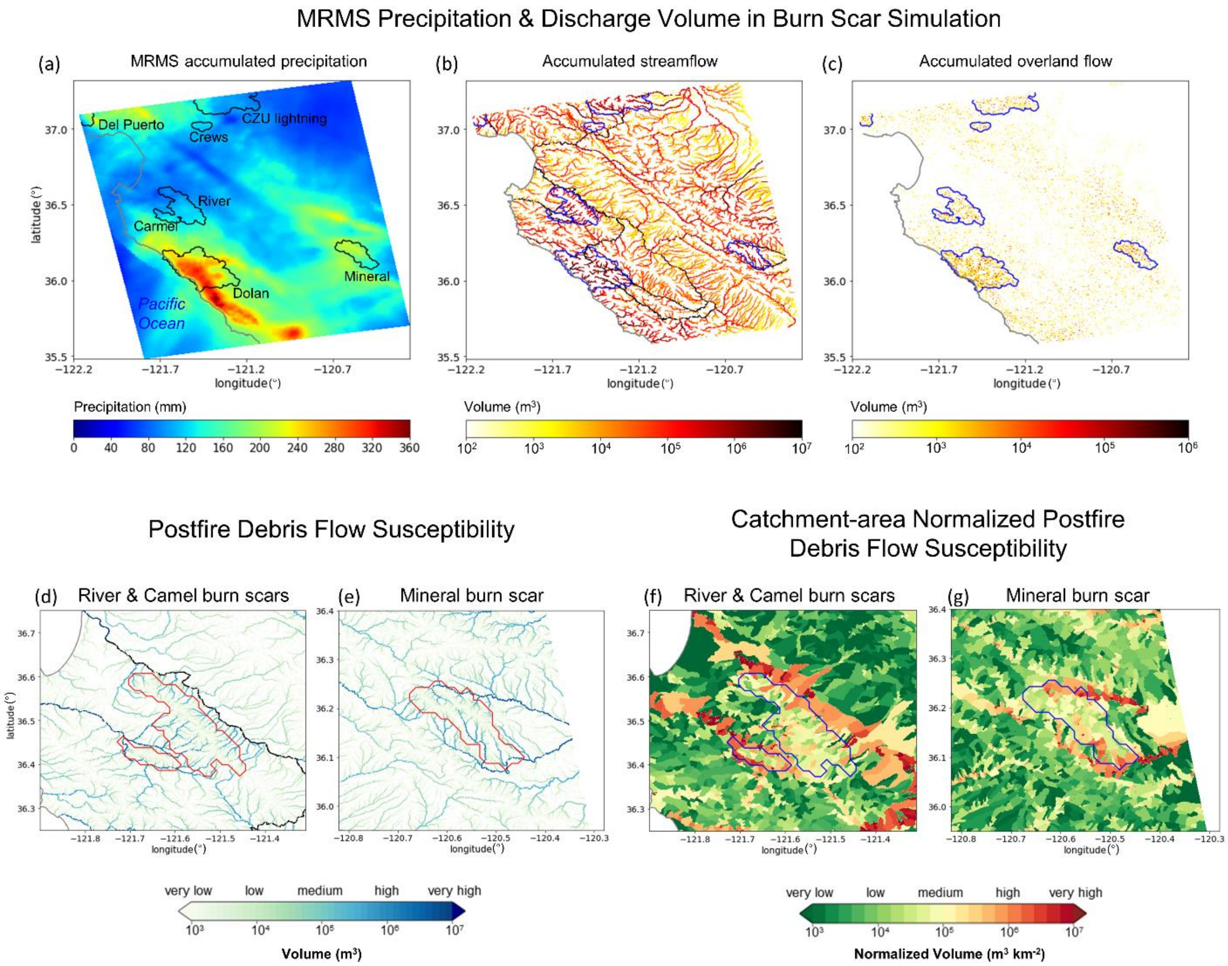

Fig. 11| MRMS accumulated precipitation and discharge volume informed regional debris flow susceptibility. (a) MRMS accumulated precipitation during January 27th 00:00 to 29th 23:00 over the model domain (shading; mm). Names of burn scars are labeled in black. (b) Accumulated streamflow (yellow-to-red shading; m3) and (c) accumulated overland flow from 27th 00:00 to 28th 12:00 over the model domain (yellow-to-red shading; m3). (d)–(e) Stream-level postfire debris flow susceptibility as Fig. 9b but for River and Camel burn scars. (f)–(g) Catchment-area normalized debris flow susceptibility as Fig. 9e but for River and Camel burn scars. Wildfire perimeters of 2020 wildfire season are outlined in black in (a), in blue in (b), (c), (f), and (g), and in red in (d) and (e). The coastline of California is depicted in grey.

In lines 685 – 702, we discussed why a regional application of WRF-Hydro is important: "In addition to the above results focused primarily on the Dolan burn scar, a key feature of WRF-Hydro is its ability to simulate the land surface hydrology of expansive geographic domains, e.g., NOAA runs the National Water Model over the entire continental U.S. Development of tools capable of regional susceptibility assessments is crucial, particularly in a wildfire-prone region like California, due to the large spatial scale, diverse morphology, and often tight spatial gradients of precipitation events and their interactions with geographically widespread wildfire burn scars. For example, landfalling ARs are often long (1000s of km) filament-like systems with heterogenous intensity gradients along their length. As a demonstration of wide geographic applicability, we assess susceptibility over our full model domain which includes more than 10,000 catchments and a number of 2020 wildfire burn scars in addition to the Dolan burn scar (Fig 11). The domain-wide analysis reveals elevated discharge volume, i.e., elevated susceptibility, in areas of high precipitation and in burned terrains (Figs. 11a–c). We highlight channelized and catchment-area normalized debris flow susceptibility in non-Dolan burn scar sites in Figs. 11d–g. In an operational forecast context, the ability to simulate landslide and debris flow susceptibilities and hazards over numerous catchments at meteorologically appropriate scales represents a step-change in the field. We argue that our demonstration of WRF-Hydro's debris flow susceptibility hindcast capabilities should motivate further exploration and development for potential use in operational hazard forecasting."

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 cited this paper in the discussion section to highlight the similarity.

In lines 647 – 652: "A comparison between baseline and burn scar simulations demonstrated that changes in hydraulic properties of burned areas causes drastic changes in surface flows, including faster discharge response times, and greater peak discharge and total volumes, consistent with findings from previous postfire hydrology studies (Anderson et al., 1976; Scott, 1993; Meixner & Wohlgemuth, 2003; Kean et al., 2011; Kinoshita & Hogue, 2015; **Brunkal & Santi, 2016**; Williams et al., 2022)."

Technical Corrections

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

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

Revised Figure 1:

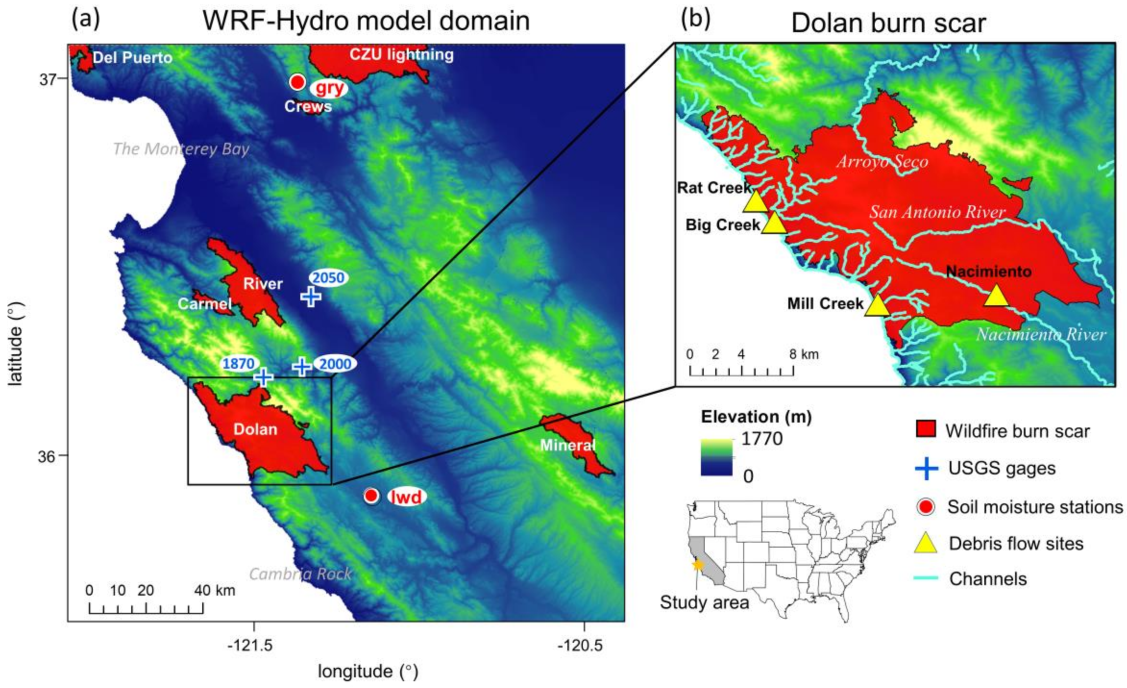

Fig. 1| WRF-Hydro model domain and Dolan burn scar. (a) WRF-Hydro model domain depicting topography, 2020 wildfire season burn scars, and PSL soil moisture and USGS stream gage observing sites. The black rectangle outlines (b) the Dolan burn scar inset, in which debris flow locations and major streams are marked and labeled. The location of the study area is shown in the embedded U.S. map with the state of California shaded in grey.

Revised Figure 7:

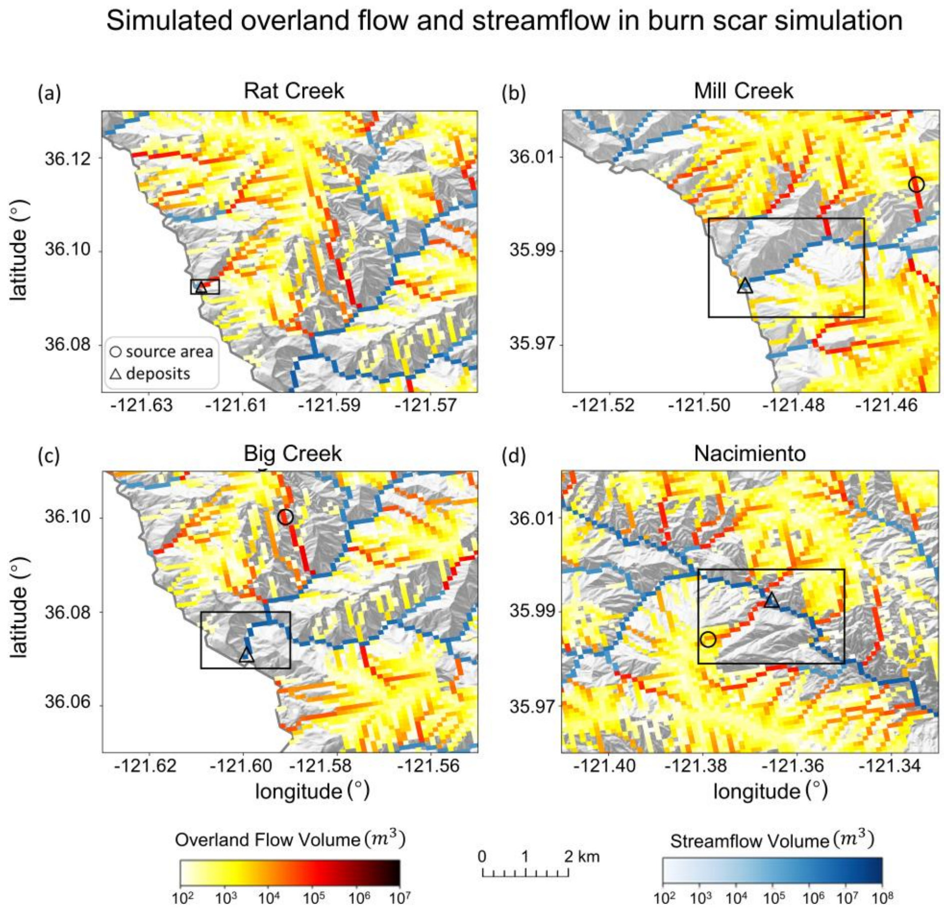

Fig. 7| WRF-Hydro simulated overland flow and streamflow in the burn scar simulation. (a)–(d) Total volume of accumulated overland flow (yellow-red shading) and streamflow (blue shading) between January 27th 00:00 and 28th 12:00 at four debris flow sites draped over a hillshade of topography. Black rectangles correspond to domains in Fig. 3a–d. Black circles and triangles indicate debris flow source areas and deposits, respectively.

Revised Figure 9:

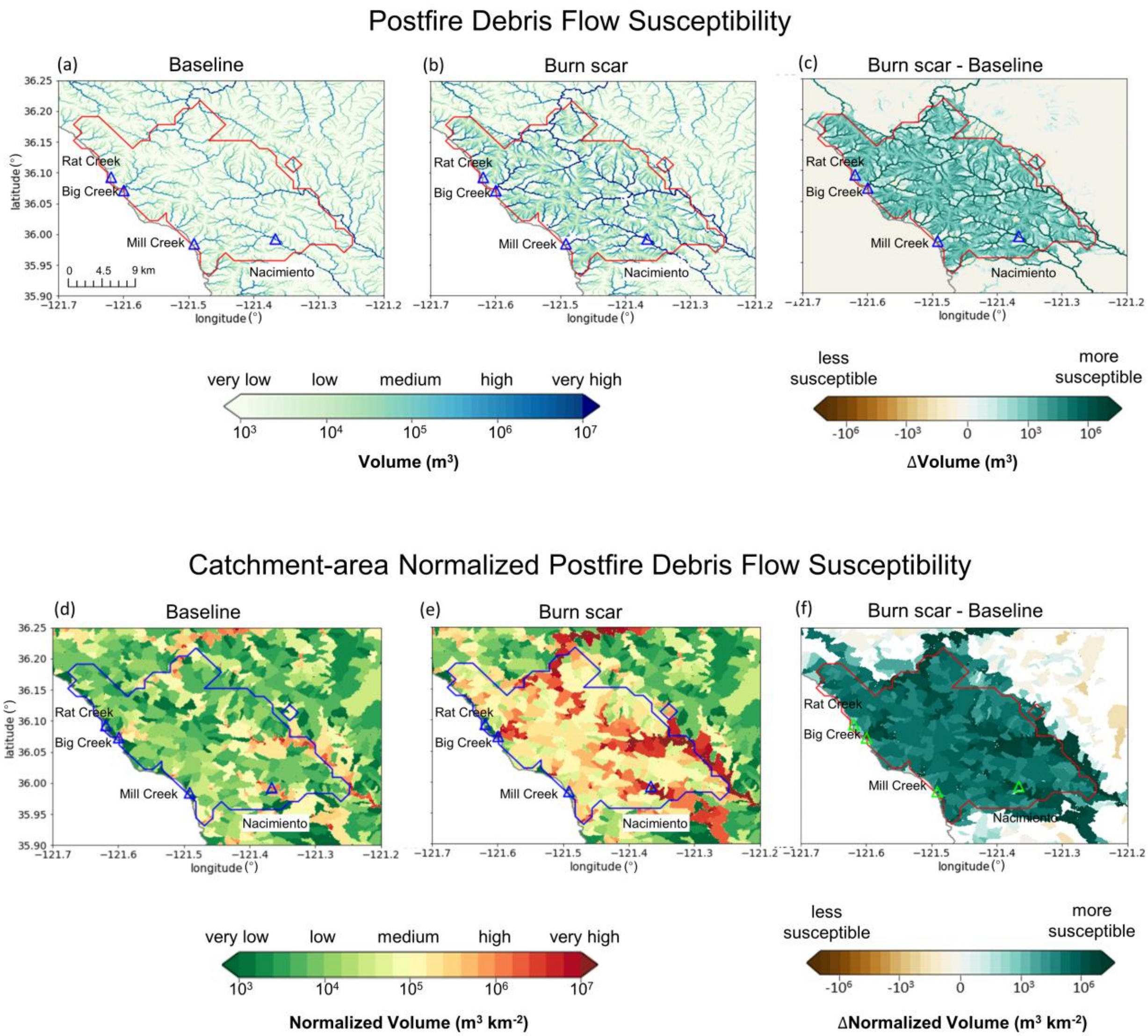

Fig. 9| Discharge volume-based postfire debris flow susceptibility. Debris flow susceptibility at individual stream level for the (a) baseline, (b) burn scar, and (c) difference between burn scar and baseline simulations. Susceptibility is estimated as total discharge volume from January 27th 00:00 to 28th 12:00. (d)–(f) Normalized debris flow susceptibility by catchment area at catchment level. For each catchment, the susceptibility is determined by total discharge volume at the catchment outlet from January 27th 00:00 to 28th 12:00 divided by catchment area.

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. In the revision, we revised the legend to make this clearer. Please see above for the revised Figure 9.

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

References

Alfieri L., Salamon P., Pappenberger F., Wetterhall F., Thielen J. (2012). Operational early warning systems for water-related hazards in Europe Environmental Science & Policy 21:35-49 doi:https://doi.org/10.1016/j.envsci.2012.01.008

Anderson, H. W., Hoover, M. D., & Reinhart, K. G. (1976). Forests and water: effects of forest management on floods, sedimentation, and water supply (Vol. 18): Department of Agriculture, Forest Service, Pacific Southwest Forest and Range Experiment Station, Berkeley, CA.

Arattano, M., & Franzi, L. (2010). On the application of kinematic models to simulate the diffusive processes of debris flows. Nat. Hazards Earth Syst. Sci., 10(8), 1689-1695. doi:10.5194/nhess-10-1689-2010

Arattano, M., & Savage, W.Z. (1994). Modelling debris flows as kinematic waves. Bulletin of the International Association of Engineering Geology 49, 3–13. https://doi.org/10.1007/BF02594995

Brabb E.E. (1985) Innovative approaches to landslide hazard and risk mapping. In: International Landslide Symposium Proceedings, Toronto, Canada. pp 17-22

Brunkal, H., & Santi, P. M. (2016). Exploration of design parameters for a dewatering structure for debris flow mitigation. 208, 81-92.

Canli E., Mergili M., Thiebes B., & Glade T. (2018) Probabilistic landslide ensemble prediction systems: lessons to be learned from hydrology Nat Hazards Earth Syst Sci 18:2183-2202 doi:10.5194/nhess-18-2183-2018

Deser, C. et al. (2020) Insights from Earth system model initial-condition large ensembles and future prospects Nature Climate Change 10:277-286 doi:10.1038/s41558-020-0731-2

Di Cristo, C., Iervolino, M., Moramarco, T., & Vacca, A. (2021). Applicability of Diffusive model for mud-flows: An unsteady analysis. Journal of Hydrology, 600, 126512. Doi:https://doi.org/10.1016/j.jhydrol.2021.126512

Flannigan, M., Cantin, A. S., De Groot, W. J., Wotton, M., Newbery, A., & Gowman, L. M. (2013). Global wildland fire season severity in the 21st century. Forest Ecology and Management, 294, 54-61.

Goss, M., Swain, D. L., Abatzoglou, J. T., Sarhadi, A., Kolden, C. A., Williams, A. P., & Diffenbaugh, N. S. (2020). Climate change is increasing the likelihood of extreme autumn wildfire conditions across California. Environmental Research Letters, 15(9), 094016. doi:10.1088/1748-9326/ab83a7

Gregoretti, C., & Dalla Fontana, G. (2007). Rainfall threshold for the initiation of debris flows by channel-bed failure in the Dolomites. Debris-Flow Hazards Mitigation: Mechanics, Prediction, and Assessment, 11-22.

Gregoretti, C., & Dalla Fontana, G. (2008). The triggering of debris flow due to channel-bed failure in some alpine headwater basins of the Dolomites: analyses of critical runoff. Hydrological Processes, 22(13), 2248-2263. doi:https://doi.org/10.1002/hyp.6821

Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., & Ardizzone, F. J. G. (2005) Probabilistic landslide hazard assessment at the basin scale 72:272-299

Hapuarachchi, H.A.P., Wang, Q.J., & Pagano, T.C. (2011) A review of advances in flash flood forecasting. Hydrological processes. 25:2771-2784 doi:https://doi.org/10.1002/hyp.8040

Hawkins, E., Sutton, R. (2009) The Potential to Narrow Uncertainty in Regional Climate Predictions Bull Amer Meteor Soc 90:1095-1108 doi:10.1175/2009BAMS2607.1

Jolly, W. M., Cochrane, M. A., Freeborn, P. H., Holden, Z. A., Brown, T. J., Williamson, G. J., & Bowman, D. M. J. S. (2015). Climate-induced variations in global wildfire danger from 1979 to 2013. Nature Communications, 6(1), 7537. doi:10.1038/ncomms8537

Kean, J. W., Staley, D. M., & Cannon, S. H. (2011). In situ measurements of postfire debris flows in southern California: Comparisons of the timing and magnitude of 24 debris flow events with rainfall and soil moisture conditions. Journal of Geophysical Research F: Earth Surface, 116(4). doi:10.1029/2011JF002005

Kinoshita, A. M., & Hogue, T. S. (2015). Increased dry season water yield in burned watersheds in Southern California. Environmental Research Letters, 10(1), 014003. doi:10.1088/1748-9326/10/1/014003

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.

McGuire, L. A., & Youberg, A. M. (2020). What drives spatial variability in rainfall intensity-duration thresholds for post-wildfire debris flows? Insights from the 2018 Buzzard Fire, NM, USA. Landslides, 17(10), 2385-2399. doi:10.1007/s10346-020-01470-y

Meixner, T., & Wohlgemuth, P. M. (2003). Climate variability, fire, vegetation recovery, and watershed hydrology. In Proceedings of the First Interagency Conference on Research in the Watersheds, Benson, Arizona (pp. 651-656).

Nikolopoulos, E.I., Schwartz, C., Zhang, X, & Anagnostou, E.N. (2019) Rainfall estimation uncertainty and early warning procedures for post-fire debris flows. In: Geophysical Research Abstracts.

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

Raia, S., Alvioli, M., Rossi, M., Baum, R. L., Godt, J.W., Guzzetti, F. (2014) Improving predictive power of physically based rainfall-induced shallow landslide models: a probabilistic approach Geosci Model Dev 7:495-514 doi:10.5194/gmd-7-495-2014

Reichenbach, P., Rossi, M., Malamud, B.D., Mihir, M., Guzzetti, F. (2018) A review of statistically-based landslide susceptibility models Earth-Science Reviews 180:60-91 doi:https://doi.org/10.1016/j.earscirev.2018.03.001

Reneau, S.L., Katzman, D., Kuyumjian, G.A., Lavine, A., & Malmon, D.V. (2007). Sediment delivery after a wildfire. Geology, 35(2), 151-154. doi:10.1130/G23288A.1

Rengers, F.K., McGuire, L.A., Kean, J.W., Staley, D.M., & Hobley, D.E.J. (2016). Model simulations of flood and debris flow timing in steep catchments after wildfire. Water Resources Research, 52(8), 6041-6061. doi:https://doi.org/10.1002/2015WR018176

Scott, D. F. (1993). The hydrological effects of fire in South African mountain catchments. Journal of hydrology, 150(2-4), 409-432.

Shen H., Lynch, B.., Poulsen, C.J., & Yanites, B.J. (2021) A modeling framework (WRF-Landlab) for simulating orogen-scale climate-erosion coupling Computers & Geosciences 146:104625 doi:https://doi.org/10.1016/j.cageo.2020.104625

Swain, D.L. (2021). A Shorter, Sharper Rainy Season Amplifies California Wildfire Risk. Geophysical Research Letters, 48(5), e2021GL092843. doi:https://doi.org/10.1029/2021GL092843

Tang, H., McGuire, L. A., Rengers, F. K., Kean, J. W., Staley, D. M., & Smith, J. B. (2019). Developing and testing physically based triggering thresholds for runoff-generated debris flows. Geophysical Research Letters, 46, 8830– 8839. https://doi.org/10.1029/2019GL083623

Tognacca, C., Bezzola, G.R., & Minor, H.E. (2000). Threshold criterion for debris-flow initiation due to channel-bed failure. Debris-flow hazards mitigation: Mechanics, prediction and assessment (pp. 89-97).

Williams, A.P., Abatzoglou, J.T., Gershunov, A., Guzman-Morales, J., Bishop, D.A., Balch, J.K., & Lettenmaier, D.P. (2019). Observed Impacts of Anthropogenic Climate Change on Wildfire in California. Earth's Future, 7(8), 892-910. doi:https://doi.org/10.1029/2019EF001210

Williams, A.P. et al. (2022) Growing impact of wildfire on western US water supply. 119:e2114069119 doi:10.1073/pnas.2114069119 %J Proceedings of the National Academy of Sciences

Yin D., Xue Z. G., Gochis D. J., Yu W., Morales M., & Rafieeinasab A. (2020) A Process-Based, Fully Distributed Soil Erosion and Sediment Transport Model for WRF-Hydro 12:1840

Zhang S., Zhao, L., Delgado-Tellez, R., Bao, H. (2018) A physics-based probabilistic forecasting model for rainfall-induced shallow landslides at regional scale Nat Hazards Earth Syst Sci 18:969-982 doi:10.5194/nhess-18-969-2018

Response to referee #2 document for NHESS manuscript 2021-345 by Li et al.

*Response is in blue & Sentences or paragraphs after revision are in red

The manuscript deals with assessment of the debris flow hazard in burned areas through simulations that used high-resolution weather radar-derived precipitation. The manuscript has several interesting points, and overall is well written. It is certainly worth to be considered for publication, but I have a couple of points which need to be clarified.

Response: We thank referee #2 for providing valuable comments for the improvement of the manuscript. We have made specific changes in response to each of these comments, and believe that the manuscript has been significantly improved as a result of these changes.

The first (and main) one regards the terminology used. I am afraid that, throughout the article, the term hazard is not used correctly. In my opinion, Authors are rather talking about susceptibility, and not hazard, the difference being that hazard should depict the probability of occurrence of a certain phenomenon not only spatially but also temporally. This latter issue (time) is not considered in the study. I suggest go back to the original definition by Varnes (1984) and UNESCO, and in later works as well, to clarify the meaning of susceptibility and hazard, and to change accordingly the terms in the manuscript.

Response: We thank the referee for highlighting the issue of our loose use of terminology, and we agree with the referee's assessment that our manuscript and methodology are primarily focused on susceptibility, rather than probabilistic hazard assessment. In our revision, we adhere to the definition of terms summarized in Section 2.1 of Reichenbach et al. (2018), *A review of statistically-based landslide susceptibility models*, Earth-Science Reviews. Accordingly, we changed the word "hazard" to "susceptibility" throughout the manuscript and in the title, and we added a paragraph in the introduction section to define the terms.

In lines 81 – 90, now it reads: "Due to this increasing threat, the development of tools to assess postfire debris flow susceptibility and hazards is critical. However, due to long-standing terminology ambiguity in the natural hazard community (Reichenbach et al. 2018), we first begin with a definition of terms. In this study we demonstrate the use of a new physics-based tool to map postfire debris flow susceptibility at regional scales. We follow the guidance of [Reichenbach et al. (2018) & references therein] and define susceptibility as the likelihood of debris flow occurrence in an area, and hazard as the probability of debris flow occurrence of a given magnitude within a specified area and period of time. In other words, debris flow susceptibility does not estimate debris flow size or consider the timing or frequency of the debris flow occurrence. Rather, it focuses on locating areas prone to debris flows considering local environmental factors (Brabb 1985; Guzzetti et al., 2005)."

We added a paragraph to the discussion section which discusses methods by which WRF-Hydro could move beyond "susceptibility" assessments to probabilistic "hazard" quantification. Firstly, we can estimate debris flow volume by coupling WRF-Hydro with a sediment erosion and transport model, which has been proved successful in previous studies (Yin et al., 2020; Shen et al., 2021). If WRF-Hydro is not coupled with any sediment models, rainfall intensity-duration thresholds (Tognacca et al., 2000; Gregoretti & Dalla Fontana, 2007, 2008) or dimensionless discharge and Shields stress thresholds (Tang et al., 2019; McGuire & Youberg, 2020) can provide information on identifying debris flow triggering time and location.

In lines 714 – 737, now it reads: "In addition to investigating the operationalization of WRF-Hydro's natural hazard prediction capabilities, we note that our susceptibility-focused methodology could be advanced to hazard assessment, in line with current USGS products. The USGS Emergency Assessment of Postfire Debris-flow Hazard predicts debris flow volume and likelihood. To advance from susceptibility to hazard assessment, our methodology would need to incorporate both debris flow volume estimates and occurrence likelihoods. In the following, we highlight research directions that could help advance our susceptibility-focused methodological framework. WRF-Hydro is a water-only model. While water-only models have been widely used to investigate and better understand debris flow dynamics (Arattano & Savage, 1994; Tognacca et al., 2000; Arattano & Franzi, 2010; Rengers et al., 2016; McGuire & Youberg, 2020; Di Cristo et al., 2021), sediment supply, soil erodibility, and other sedimentological factors play important roles in determining the potential for and severity of mass failure events (McGuire et al., 2017). Developing a runoff-generated debris flow model that couples hydrologic and sediment erosion and transport processes could help to characterize postfire debris flow volumes. Indeed, previous efforts have demonstrated the capacity to couple WRF-Hydro with sediment flux models (Yin et al., 2020; Shen et al., 2021). In addition to sediments, burn scar ash can comprise a substantial fraction of the total debris flow volume (e.g., Reneau et al., 2007). As such, efforts to constrain ash availability and entrainment in hydrologic flows could prove fortuitous in hazard assessment and prediction efforts. If WRF-Hydro is not coupled with sediment models, a domain-specific rainfall ID threshold trained with historic landslide inventory and triggering rainfall events (Tognacca et al., 2000; Gregoretti & Dalla Fontana, 2007, 2008) or a newly developed dimensionless discharge and Shields stress threshold (Tang et al., 2019; McGuire & Youberg, 2020) could provide guidance to help identify debris flow triggering time and location, which in turn may improve WRF-Hydro's debris flow initiation identification."

In addition, we also discussed methods for probabilistic advancement including systemic investigation of parameter uncertainties and use of ensemble-based precipitation data.

In lines 739 – 755, now it reads: "In addition to constraining potential postfire debris flow volumes, WRF-Hydro's application in debris flow studies could be advanced via concerted engagement with uncertainties that are both external (meteorological forcing data) and internal (physical parameters) to the model. Previous studies have demonstrated that precipitation is often the largest source of uncertainty in hydrologic predictive models (Hapuarachchi et al., 2011; Alfieri et al., 2012). Engagement with precipitation forcing uncertainties in past, near-term, and future contexts could provide probabilistic nuance to natural hazard investigations. For example, (a) debris flow hindcast studies could use a diversity of precipitation datasets to isolate precipitation-derived debris flow uncertainties in historic events, (b) operational forecast efforts could utilize ensemble-based weather forecast data to inform likelihood statements in debris flow hazard assessments, and (c) probabilistic projections of debris flow likelihood in future climates could assess and partition uncertainties derived from emission pathway, model structure, or internal variability effects on meteorological forcings (Nikolopoulos et al., 2019; Hawkins & Sutton, 2009; Deser et al., 2020). Uncertainties internal to WRF-Hydro are also ripe for investigation. Probabilistic predictions crafted from an ensemble of perturbed model physics simulations have been used to predict rainfall-triggered shallow landslides (Raia et al., 2014; Canli et al., 2018; Zhang et al., 2018). Similar efforts using WRF-Hydro could target post-wildfire debris flows."

Another point which needs more details is the description of the debris flows. Authors talk about several debris flows that occurred, and start to cite them in section 2.1. However, a clear description of the events, in terms of geology, morphology, morphometry, volumes is never properly given. This should be done the first time debris flows are mentioned (possibly in section 2.1) to let the reader understand the main characters of the events. For instance, were these debris flows individual phenomena, or did they start from multiple source areas? Further, were they channelized or openslope? More geomorphological info would be useful to understand the conditions under which the debris flows initiated and developed. Only at page 18 some info are provided, but these should appear much before than that, and be well organized, rather than distributed in different parts of the manuscript.

Response: We sympathize with the referee's desire for more data on the debris flows highlighted in our manuscript. However, at present, there have been no systematic studies of these debris flows, so while we are quite confident that debris flows occurred at these locations based on field observations of the deposits and our remote sensing analyses, information about source areas (which are in extremely inaccessible locations) and volumes are not well constrained. We do know that David Cavagnaro et al. (https://agu.confex.com/agu/fm21/meetingapp.cgi/Paper/921613) have undertaken a huge effort to map debris flows in the Dolan Fire burn scar. I suspect this forthcoming work will be able to answer some of the referee's concerns in greater detail.

In our revised manuscript we added a paragraph to section 2.2 to describe the debris flow geologic setting. We moved the descriptions on the debris flows in section 5.2 to section 2.2. Our information on the geology come from the USGS geologic map (https://mrdata.usgs.gov/geology/state/state.php?state=CA). We also calculated mean and maximum slope values for the triggering and deposition sites of the four debris flow events using the slope calculation function in ArcMap and the USGS 30-m DEM.

In lines 235 – 262, it now reads:

"**2.2 Debris flow geologic setting**

According to the USGS National Elevation Dataset 30-m digital elevation model (DEM), the Rat Creek debris flow sits at the base of a 1st order catchment with a drainage area of 2.23 km^2. Mill Creek, Big Creek, and Nacimiento debris flows were initiated within extremely steep, intensely burned, 1^{st} order catchments, but were deposited in 2^{nd}, 3^{rd}, and 3^{rd} Strahler stream order channels, respectively. All four debris flows were channelized. Rat Creek, Mill Creek, and Big Creek debris flow deposition sites have elevations ranging from 20–60 m, while Nacimiento debris flow deposited at an elevation of ~440 m above sea level. We calculate catchment slopes using the DEM and the slope calculation function in ArcMap. The average slope of the catchments containing Rat Creek and Mill Creek debris flow deposition sites is ~25°. The average catchment slope of Big Creek deposition site is ~28° and Nacimiento is ~21°. For debris flow source areas, the average and maximum slopes of Mill Creek are 23° and 39°, 21° and 43° for Big Creek, and 24° and 41° for Nacimiento. According to the Soil Survey Geographic Database and California geologic map data, surface soils at the three coastal debris flow sites (i.e., Rat Creek, Mill Creek, and Big Creek) are texturally classified as loam with underlying Franciscan Complex sedimentary rocks of Jurassic to Cretaceous age. Soil at Nacimiento is classified as sandy loam with underlying Upper Cretaceous and Paleocene marine sedimentary rocks from the Dip Creek Formation, Asuncion Group, Shut-In Formation, Italian Flat Formation, Steve Creek Formation, and El Piojo Formation. Mill Creek, Big Creek, and Nacimiento were relatively large debris flows with runout lengths between ~2–5 km, while Rat Creek occurred in a smaller catchment and had a runout length of ~300 m. The difference in runout length and debris flow size is primarily controlled by upstream catchment size, however for the three coastal debris flow events at Rat Creek, Big Creek, and Mill Creek, also constrained by the downslope ocean. We note that there were likely more debris flows triggered during the AR event. The four debris flow events highlighted here were identified during brief post-event field excursions due to their intersection with major roadways. Given that our primary goal here is to demonstrate the utility of WRF-Hydro – a comprehensive catalogue of debris flows is beyond the scope of this study, although underway by other researchers (Cavagnaro et al., 2021)."

Other issues:

Figure 1 definitely needs a location map, showing where we are in California, and in USA. Authors give for granted that anybody knows the site, but for an international journal a location map is always necessary.

Response: In our revision we added a location map of the USA which depicts the locations of California and the burn scar/debris flow region.

Revised Figure 1:

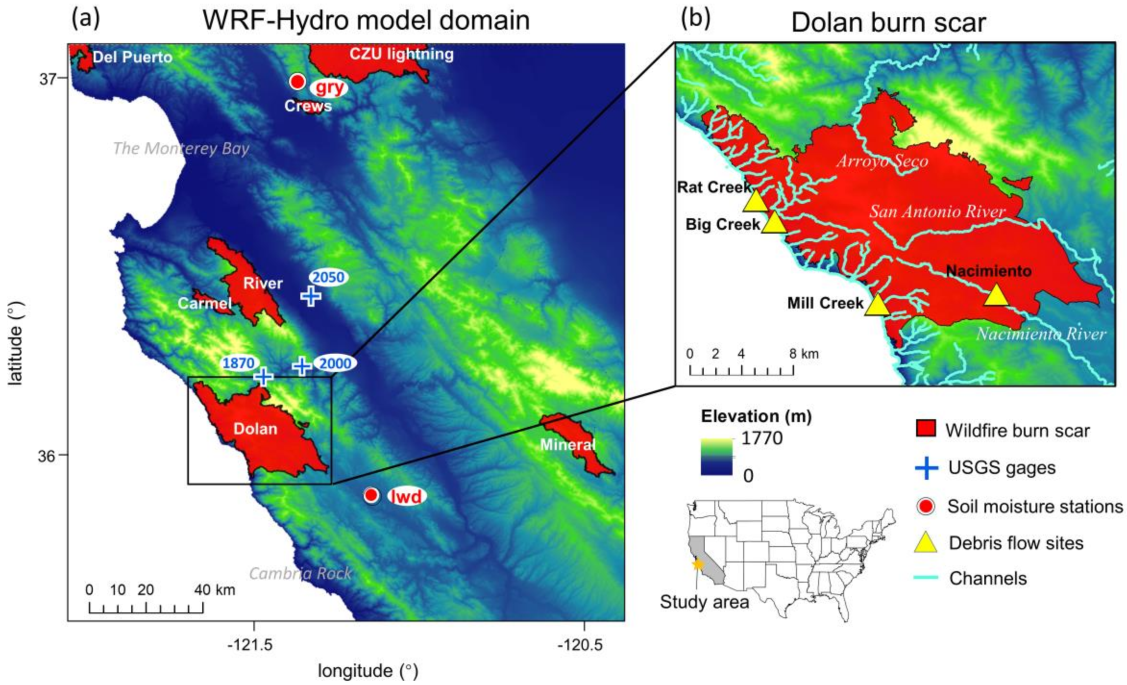

Fig. 1| WRF-Hydro model domain and Dolan burn scar. (a) WRF-Hydro model domain depicting topography, 2020 wildfire season burn scars, and PSL soil moisture and USGS stream gage observing sites. The black rectangle outlines (b) the Dolan burn scar inset, in which debris flow locations and major streams are marked and labeled. The location of the study area is shown in the embedded U.S. map with the state of California shaded in grey.

Throughout the manuscript, references should be listed in chronological order when more than two references are cited. Some incomplete or wrong references are present in the list. Please check at this regard the attached file. Eventually, some minor issues are indicated in the accompanying file.

Response: In our revision we reordered the in-text citations chronologically and corrected the references. We thank referee #2 for their very careful examination of references. We have checked the attached file and addressed those minor issues accordingly.

Overall, I evaluate positively the manuscript, which however needs to clarify the points outlined above, and recommend minor revisions.

We thank referee #2 again for their careful review and positive comments.

References:

Alfieri L., Salamon P., Pappenberger F., Wetterhall F., Thielen J. (2012). Operational early warning systems for water-related hazards in Europe Environmental Science & Policy 21:35-49 doi:https://doi.org/10.1016/j.envsci.2012.01.008

Arattano, M., & Franzi, L. (2010). On the application of kinematic models to simulate the diffusive processes of debris flows. Nat. Hazards Earth Syst. Sci., 10(8), 1689-1695. doi:10.5194/nhess-10-1689-2010

Arattano, M., & Savage, W.Z. (1994). Modelling debris flows as kinematic waves. Bulletin of the International Association of Engineering Geology 49, 3–13. https://doi.org/10.1007/BF02594995

Brabb E.E. (1985) Innovative approaches to landslide hazard and risk mapping. In: International Landslide Symposium Proceedings, Toronto, Canada. pp 17-22

Canli E., Mergili M., Thiebes B., & Glade T. (2018) Probabilistic landslide ensemble prediction systems: lessons to be learned from hydrology Nat Hazards Earth Syst Sci 18:2183-2202 doi:10.5194/nhess-18-2183-2018

Cavagnaro D et al. (2021) Variability in hydrologic response to rainfall across a burn scar: observations from the Dolan Fire, California. AGU abstract. https://agu.confex.com/agu/fm21/meetingapp.cgi/Paper/921613

Deser, C. et al. (2020) Insights from Earth system model initial-condition large ensembles and future prospects Nature Climate Change 10:277-286 doi:10.1038/s41558-020-0731-2

Di Cristo, C., Iervolino, M., Moramarco, T., & Vacca, A. (2021). Applicability of Diffusive model for mud-flows: An unsteady analysis. Journal of Hydrology, 600, 126512. Doi:https://doi.org/10.1016/j.jhydrol.2021.126512

Gregoretti, C., & Dalla Fontana, G. (2008). The triggering of debris flow due to channel-bed failure in some alpine headwater basins of the Dolomites: analyses of critical runoff. Hydrological Processes, 22(13), 2248-2263. doi:https://doi.org/10.1002/hyp.6821

Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., & Ardizzone, F. J. G. (2005) Probabilistic landslide hazard assessment at the basin scale 72:272-299

Hapuarachchi, H.A.P., Wang, Q.J., & Pagano, T.C. (2011) A review of advances in flash flood forecasting. Hydrological processes. 25:2771-2784 doi:https://doi.org/10.1002/hyp.8040

Hawkins, E., Sutton, R. (2009) The Potential to Narrow Uncertainty in Regional Climate Predictions Bull Amer Meteor Soc 90:1095-1108 doi:10.1175/2009BAMS2607.1

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.

McGuire, L. A., & Youberg, A. M. (2020). What drives spatial variability in rainfall intensity-duration thresholds for post-wildfire debris flows? Insights from the 2018 Buzzard Fire, NM, USA. Landslides, 17(10), 2385-2399. doi:10.1007/s10346-020-01470-y

Nikolopoulos, E.I., Schwartz, C., Zhang, X, & Anagnostou, E.N. (2019) Rainfall estimation uncertainty and early warning procedures for post-fire debris flows. In: Geophysical Research Abstracts.

Raia, S., Alvioli, M., Rossi, M., Baum, R. L., Godt, J.W., Guzzetti, F. (2014) Improving predictive power of physically based rainfall-induced shallow landslide models: a probabilistic approach Geosci Model Dev 7:495-514 doi:10.5194/gmd-7-495-2014

Reichenbach P, Rossi M, Malamud BD, Mihir M, Guzzetti F (2018) A review of statistically-based landslide susceptibility models. Earth-Science Reviews. 180:60-91 doi:https://doi.org/10.1016/j.earscirev.2018.03.001

Reneau, S.L., Katzman, D., Kuyumjian, G.A., Lavine, A., & Malmon, D.V. (2007). Sediment delivery after a wildfire. Geology, 35(2), 151-154. doi:10.1130/G23288A.1

Rengers, F.K., McGuire, L.A., Kean, J.W., Staley, D.M., & Hobley, D.E.J. (2016). Model simulations of flood and debris flow timing in steep catchments after wildfire. Water Resources Research, 52(8), 6041-6061. doi:https://doi.org/10.1002/2015WR018176

Shen H., Lynch, B.., Poulsen, C.J., & Yanites, B.J. (2021) A modeling framework (WRF-Landlab) for simulating orogen-scale climate-erosion coupling Computers & Geosciences 146:104625 doi:https://doi.org/10.1016/j.cageo.2020.104625

Tang, H., McGuire, L. A., Rengers, F. K., Kean, J. W., Staley, D. M., & Smith, J. B. (2019). Developing and testing physically based triggering thresholds for runoff-generated debris flows. Geophysical Research Letters, 46, 8830– 8839. https://doi.org/10.1029/2019GL083623

Tognacca, C., Bezzola, G.R., & Minor, H.E. (2000). Threshold criterion for debris-flow initiation due to channel-bed failure. Debris-flow hazards mitigation: Mechanics, prediction and assessment (pp. 89-97).

Yin D., Xue Z. G., Gochis D. J., Yu W., Morales M., & Rafieeinasab A. (2020) A Process-Based, Fully Distributed Soil Erosion and Sediment Transport Model for WRF-Hydro 12:1840

Zhang S., Zhao, L., Delgado-Tellez, R., Bao, H. (2018) A physics-based probabilistic forecasting model for rainfall-induced shallow landslides at regional scale Nat Hazards Earth Syst Sci 18:969-982 doi:10.5194/nhess-18-969-2018