Bold text indicates the original reviewer comment, and plain text indicates the response.

Response to RC1

General comments:

This study presents the WEST volume model, an updated postfire debris-flow volume model that uses an expanded dataset, including events from northern California and Washington. The model also introduces a rainfall ratio variable that normalizes rainfall intensity by the 1-year recurrence interval, improving model performance and transferability across diverse geographic regions. The model evaluation is extensive, assessing performance against the full volume database, regional subsets, prior models, and data-limited regions across the western United States, with performance metrics clearly presented and effectively illustrated. I particularly liked how the authors considered the prior sediment volume models and made a careful effort to build upon the previous work. The improvement in predicting sediment volumes is a significant contribution for postfire hazard assessments. Below I have listed comments that I think can be easily addressed.

Specific comments:

1. **Regarding the final model selection, I noticed that models #1 and #3 in Table S3 include the same predictor variables with identical performance metrics, differing only in the use of the i30 vs. i60 rainfall ratio. Table 2 also indicates that both rainfall ratios have the same correlation coefficient (-0.13). Consider clarifying whether the selection of the i30 model was primarily based upon its sub-hourly duration. Additionally, how should practitioners proceed in situations where i30 data are unavailable, or where i60 may be more appropriate for operational forecasting applications?**

   Thank you for this suggestion. We primarily selected the i30 model based on the criteria outlined in Section 3.2.2. In addition to having the best performance metrics, it used predictor variables that were commonly selected by the 29 final models, all three predictor variables had p-values < 0.1, and all three predictor variables had VIF values < 10. The difference in performance between the i30 model and other similarly performing models, however, was marginal, as noted. As a result, it is possible to use one of the alternative models listed in the supplement instead of the WEST model without sacrificing model performance. In that sense, it is reasonable to select the i60 volume model as an alternative if i30 rainfall data is not available.

   We added the following text to Section 4.1 of the Results where we introduce the WEST model to alert readers to the fact that there are multiple models with similar performance metrics:

   "Note that although we selected the WEST model as the final model for this study using the criteria outlined above, many of the final 29 models offer similar performance metrics (Table S3) and may be viable alternatives to the WEST model in some scenarios."

We also added the following text to Section 5.2 of the Discussion, where we discuss potential use cases for alternative models that use i15 or i60 rainfall ratio, instead of i30:

"Finally, although we selected the WEST model using the criteria outlined in Section 3.2.2, several alternative models that we developed as part of this study offer similar performance metrics (Table S3) and may be preferred to the WEST model in some situations. For instance, we determined that rainfall ratio calculated over a 30-minute duration yielded the best results for our dataset, but there may be scenarios where models that incorporate rainfall ratio calculated over 15 or 60-minute durations are more suitable. It may be more practical, for example, to use a model that includes $i60$ rainfall ratio, instead of $i30$ rainfall ratio, to estimate volumes for a mitigation project if only $i60$ design storms are available. Alternatively, it may be easier to implement a model that uses $i15$ rainfall ratio within a hazard assessment framework that also predicts postfire debris-flow likelihood using rainfall characteristics measured over a 15-minute duration (Landslide Hazards Program, 2018). Given their potential applicability in these scenarios, we present alternative models that use $i15$ and $i60$ rainfall ratio, along with the same terrain and fire variables used by the WEST model (Equation 2), to predict postfire debris-flow volume in Table 9. Note that, although it marginally outperforms the WEST model, the $i15$ model does not pass the Anderson-Darling test (p-value < 0.05), so was not considered for final model selection. The $i60$ model, on the other hand, performs similarly to the WEST model and meets all requirements of multiple linear regression (Table 9)."

Finally, we added the following table to Section 5.2 of the Discussion that shows a comparison between the WEST model and alternative volume models that use i15 and i60 rainfall ratios:

Table 9: Equations for volume models that use rainfall ratios calculated at different durations, including 15 minutes ($i15_{rr}$), 30 minutes ($i30_{rr}$), and 60 minutes ($i60_{rr}$). All models predict debris-flow volume (V) as a function of rainfall ratio, watershed area (a), and watershed area burned at moderate or high severity with slopes ≥ 50% (mh_{50}). Performance metrics include R^2 and root mean square error (RMSE), as well as the p-values for the Anderson-Darling (AD) and Brown-Forsythe (BF) tests.

ID	**Model**	**R^2**	**RMSE**	**AD Test p-value**	**BF Test p-value**
$i15$	$\ln V = 7.82 + 0.35 \ln i15_{rr} + 0.76 \ln a + 1.09\sqrt{mh_{50}}$	0.67	1.30	0.03	0.17
WEST	$\ln V = 7.56 + 0.20 i30_{rr} + 0.75 \ln a + 1.11\sqrt{mh_{50}}$	0.66	1.31	0.07	0.17
$i60$	$\ln V = 7.61 + 0.18 i60_{rr} + 0.76 \ln a + 1.10\sqrt{mh_{50}}$	0.66	1.31	0.07	0.15

2. **Consider including Gatwood et al. (2000) in your statement on line 79, "There are several postfire debris-flow volume models that include rainfall variables," especially as Gatwood et al. (2000) uses the maximum 1-hour precipitation.**

   We agree that Gatwood et al. (2000) is a useful reference and have added it where suggested.

3. **I think that the statement surrounding the standard deviation on line 499 is not in agreement with the standard deviation values listed in Table 6. The in-text statement reads "The WEST model slightly overpredicted volumes from data-limited regions but had the lowest standard deviation of the four models (Table 6)." I believe Table 6 lists the lowest standard deviation from the data-limited regions as 1.29 (EAV model) not the WEST model (1.37).**

   Thank you for catching this discrepancy. The data in Table 6 is correct; the residuals of the EAV model had a lower standard deviation than those of the WEST model. We have updated the text so that it now accurately reflects this information. The updated text is:

   "The WEST model slightly overpredicted volumes from data-limited regions but had the second lowest standard deviation of the four models (Table 6). The EAV model had a slightly lower standard deviation than the WEST model but overpredicted volumes from data-limited regions more substantially (Figure 5b). Unlike the WEST and EAV models, the V1 model underpredicted volumes from data-limited regions, on average (Figure 5d)."

4. **"Rainfall ratio" and "rainfall anomaly" are used interchangeably throughout the manuscript and the supplemental. For example, on line 415, "rr" is used in Eq. 2, whereas "ra" is used in Table S3. Although everything is clearly defined, maintaining consistent naming of this variable throughout the manuscript and supplemental may improve clarity for the reader.**

   Thank you for pointing out this inconsistency. We have updated the manuscript and supplement so that we now only use the term "rainfall ratio" in the text and "rr" in equations.

5. **Although Table 1 displays the number of volume measurements geographically, it lacks information surrounding the magnitude (e.g., $10^1 - 10^4$) of the sediment volumes geographically. Consider adding a column with the range of sediment volumes in Table 1.**

   We agree that volume magnitude is an important piece of information that is currently missing from our tables. To rectify this, we have added a column to Table 1 titled "Range of Debris-Flow Volumes" that includes the minimum and maximum debris-flow volumes associated with each site.

6. **While the need for additional volume data from data-limited regions is discussed, it may be helpful to also acknowledge the importance of including larger-magnitude sediment volumes, as the current dataset from these regions contains smaller volumes.**

   This is an important point, so we have added the following text to Section 5.2 of the Discussion:

   "In particular, volume data from larger-magnitude debris flows in data-limited regions are needed to more fully evaluate the performance of the WEST model in these settings, as the median volume of the 19 debris flows from data-limited regions included in this study was 511 m^3 (Table S2), nearly five times lower than the median volume of 2,550 m^3 associated with the entire volume database."

Technical Corrections:

- **Table 1 – consider adding more text for context to the Table 1 caption (e.g., "Fire information associated with measured sediment volumes")**

  We have updated the caption of Table 1 to provide more information about the data contained in the table. The new caption is:

  "Table 1: Fire and associated debris-flow information. Fire information includes the fire name, year of occurrence, and location, including Arizona (AZ), California (CA), Colorado (CO), New Mexico (NM), Utah (UT), and Washington (WA). Debris-flow information includes the number of volume measurements for each fire, the range of associated sediment volumes, and the original sources of the volume measurements."

- **Line 123, add space after "," in "Smith et al.,2021"**

  We have added the space where suggested.

- **Figure 1 – consider simplifying the fire icon to a circle symbol for improved legibility and/or lighten the ecoregion colors to increase contrast between the ecoregion and the volume symbol.**

  We have updated Figure 1 to improve the contrast between the ecoregions and volume symbols. Specifically, we adjusted the color scheme of the volume symbols to better distinguish between the volume categories, as well as between the volume symbols and ecoregions. We also updated the basemap to improve overall readability.

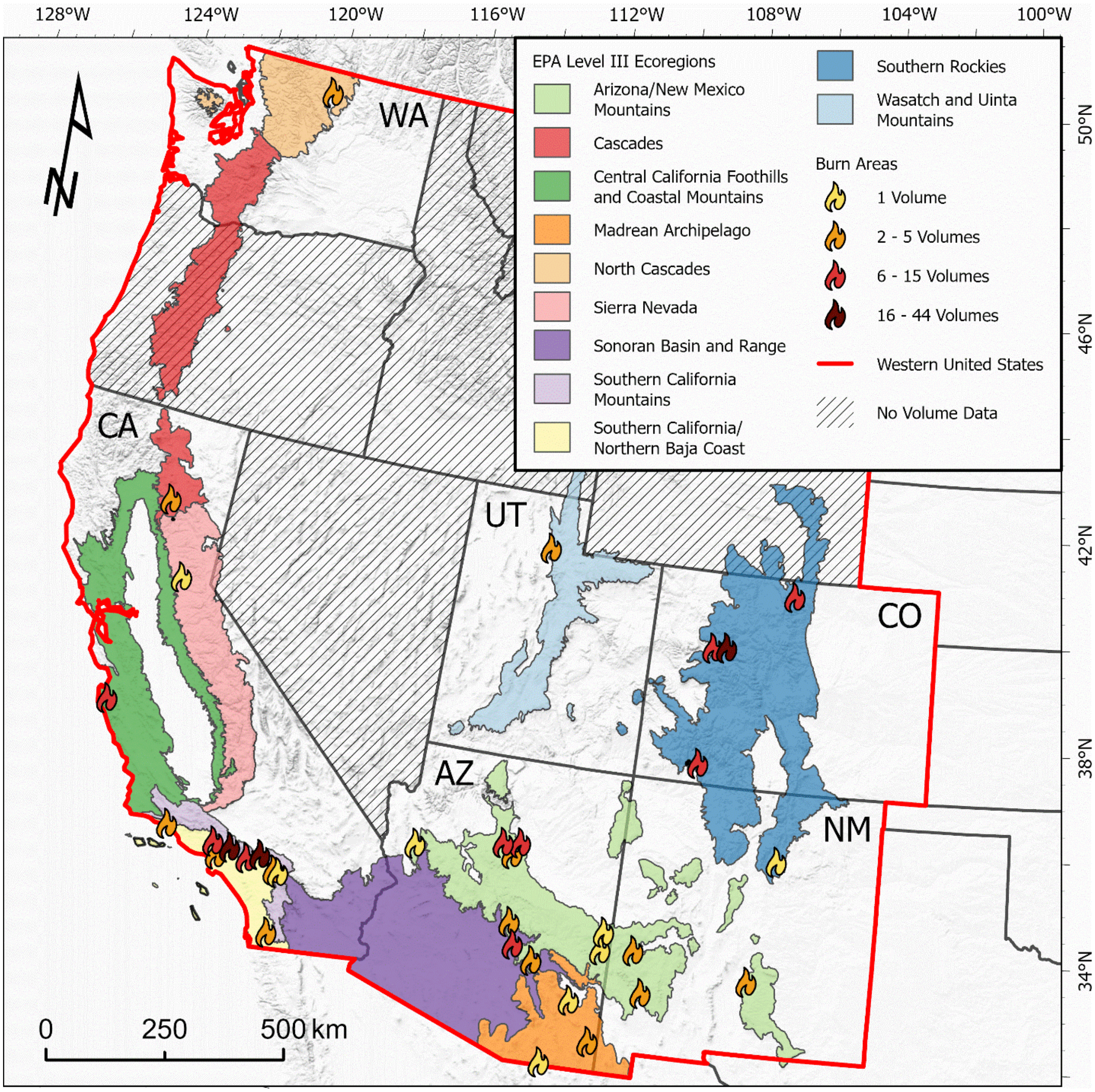

Figure 1: Map of the locations of the 34 burn areas included in this study. The burn areas span six states across the western United States (US), including Arizona (AZ), California (CA), Colorado (CO), New Mexico (NM), Utah (UT), and Washington (WA), and 11 Environmental Protection Agency (EPA) Level III Ecoregions. The names of the ecoregions shown in this figure are derived directly from the EPA (U.S. Environmental Protection Agency, 2013). Basemap credits: United States Geological Survey The National Map: 3D Elevation Program, United States Geological Survey Earth Resources Observation & Science Center: GMTED2010.

- **Line 256 – for increased clarity, consider including the text "for each watershed" in the phrase "We calculated another 17 variables for each watershed related to fire characteristics (Table 4)…"**

We have added the phrase “for each watershed” where suggested.

- **Line 271 – delete “and” directly after “B_k”**

  We have deleted “and” where suggested.

- **Line 415 – consider spelling out the first instance of “VIF” in this section for clarity.**

  We redefined VIF at first mention in the results section, as suggested.

- **Lines 457, 462, and 523 – change “Figure S2” to “Figure S3”**

  We have updated these references from “Figure S2” to “Figure S3.”

- **Line 460 – delete “a” directly after “MAE (Table 7) and”**

  We removed “a” where suggested.

- **Line 489 – change “Figure S5” to “Figure S5a” to maintain consistency with “Figure S5d”**

  We changed this reference to “Figure S5a” as suggested.

- **Line 657 – change “Figure 6” to “Figure 7”**

  We changed this reference from “Figure 6” to “Figure 7.”