Author's Response

Below are the author's response to comments from two reviewers and one community reviewer.

Response to Reviewer #1, January 18th, 2024

We would like to thank this anonymous reviewer for their feedback, the comments have led to a greatly improved manuscript. Below are the reviewer's comments in regular type and our responses in bold type.

General Comments

This is a manuscript that describes several modeling approaches to estimate postfire sediment yields. They compare these yields with anecdotal data (photographs) and truck records that are subsequently used to estimate a sedimentation rate. The figures are generally clear, although somewhat repetitive, and thus there is an opportunity to reduce the current figures. The writing is generally clear, but some methodological details are missing, and some information is not well introduced in the paper (e.g., mention of the Pipeline fire) prior to the discussion.

I have several major suggestions to help improve the structure and clarity of this manuscript. The first suggestion is to better describe the type of model that you are using. Model is a vague term and can mean anything from a process-based numerical model to an empirical equation. Since many of these are previously published you don't need to get into all of the details, but I think you should describe the model type, the inputs, the parameters that you adjust, and the functional relationships. As far as functional relationships, that would mean that if, for example, you model deposition, then you say the key equation/method in which deposition is controlled.

The model types are explained in more detail in this revision. However, the reviewer does correctly point out that there is abundant existing literature on the individual models. The purpose of this manuscript is not to introduce new modeling techniques but to compare existing models between themselves and empirical results. Adding too much detail about each modeling technique would detract from the purpose of this manuscript.

Second, I think that the comparison between model output and observations could be clarified. Try adding a figure that shows a timeline of events that occurred. And then refer to the timeline when you say what you are modeling and how you are comparing models to observations. Right now this is confusing. In section 4.2.1 you are modeling four flood events, but were there more storms than that? Were those just the four largest storms? In addition to the timeline, I think you need to have some sort of figure that shows rain events as a function of time in the study. It is not always clear how what subset of postfire time you are comparing to the model.

More description of the flood events (there were four) is provided in this revision. We could provide a timeline but that would be redundant with the explanation of events and the modeling domains. A new subsection, 2.1, was added to provide more clarity.

Finally, organizationally, there are a few items that don't make sense. You start talking about the Pipeline fire in the discussion, but it wasn't mentioned prior to that in the study site section. Why do you model 4 events in 4.2.1 and three events in 4.2.2. How do the subwatersheds highlighted in

figure 5 relate to “work areas” in Table 3? Should work areas be a single point rather than a watershed? And if not, then I think you need to state that a work area is a full subwatershed. And there could be an overall improvement of descriptions. For example, in the caption for Figure 10 it says “note the spread of flow and subsequent drop in water velocity”, however you are not actually showing anything that indicates velocity. Other examples can be seen in the specific line comments below.

These are good comments and are addressed in the line comments (which are the same as the paragraph above). Figure 10 has been updated to include a velocity map.

Line Comments

1. Here you use “post-wildfire” but on line 50 you write “post-fire”. Try to use a consistent term throughout the manuscript.

We have replaced “post-fire” with “post-wildfire” throughout the manuscript, thank you for catching that discrepancy.

1. Paradise/sunnyside is not shown on the map in figure 1. Only Paradise is shown. Can you fix this?

This is fixed in this iteration of the manuscript, thank you for catching that omission.

1. Add a reference at the end of this sentence to support the statement

Unsure which sentence, missing a reference to the line number in this comment form.

1. after the word “debris” can you reference a news article to support this?

A link to the NOAA webpage for the flood events (https://www.weather.gov/fgz/FlagstaffJuly2021) is now included.

1. when you say “increase in runoff” what is the baseline for the increase?

“Surface water” has been added for clarity. The basis for the modeling and empirical observations of flooding are provided in the reference in the sentence (Schenk et al. 2023).

113-114. When you say that the watershed was divided into sub-watersheds, say the criteria you used for the watershed delineation.

“Based on USGS National Hydrography Dataset” has been added. With a citation of: USGS, (U.S. Geological Survey), 2019, USGS 3D Elevation Program Digital Elevation Model, accessed October 31st, 2021

1. Sentence ending on 127, support claim with a reference.

More description of the highly heterogeneous rainfall distribution in the American Southwest was added to this sentence.

128-139: Why is this paragraph not in section 3.4?

This can be explained either in the Sediment Modeling section or the Sediment Yield section, providing the method twice is redundant.

1. I think you need to add some details to explain how "an analysis of sediment transport across a conceptualized design channel" was done. This is really vague.

We took cross sections in their condition at our proposed fan work areas and modeled them as 2% consistent grade as a proposed condition. We compared the modeled cross section with the current 2021 condition.

1. Say how these data are used to estimate bank erosion. What is the method?

Bank erosion was quantified in the field in 2021 and estimated in out-years using channel slope, expected annual flow rates (from FLO-2D modeling completed outside of this study; Schenk et al. 2023), soil condition, and channel cross-sectional area.

185: I think you mean model instead of modeled

Thank you, corrected.

189: State explicitly why you considered 2021 as the second year post fire.

The fire began during the monsoon season of 2019 which would make 2020 the first full year post-fire.

Say why you chose one inch, two inch, and three inch precipitation events. Also say the duration? Is this over an hour, a day, etc?

"in one hour" was added to provide the appropriate detail.

1. Say why you are using those K values. Also, did these vary by year?

We chose those K values based on field observations. The MUSLE modeled watersheds were larger than the ERMiT watersheds (because we were looking at different metrics and different timescale…annual average sediment yield vs event-based sediment yield). The condition of the soil pre fire was sandy loam with a modeled K value of approximately 0.29. Post fire, we anticipated that the K value was ~0.80. Averaging these two balanced the initial condition vs the post-fire condition.

1. Did your c value vary by year?

No, it did not change by year because a) it is a dimensionless number used to denote a landscape scale practice (in this case it was forest area, doesn't change per MUSLE year to year)) and b) because we only completed MUSLE for current conditions as an event-based modeling method (not annual for multiple years.

1. State error associated with photo measurements.

The manuscript explains that these were qualitative assessments, there is no known error to the qualification nor are there repeatable confidence limits for this method. The landfill scale measurements were over several weeks and likewise do not have confidence intervals or a standard for error measurement.

1. I think you actually need to describe the model details here. Is this just an equation, a numerical simulation, a spreadsheet with a GUI wrapper, something else? Show what calculations are being performed.

This section has been expanded to describe the modeling technique and inputs.

The following was added to the manuscript:

"FLOWSED models the total annual sediment yield, both suspended and bedload, using flow-duration curves and their corresponding sediment yields. The dimensionless flow-duration curve is developed from representative watersheds in the region using USGS stream gage data. The POWERSED model compares sediment transport in various configurations of channel geometry."

1. Say how FLOWSED/POWERSED estimates rebuilding of alluvial fans. What equations/methods are used to estimate deposition.

FLOWSED/POWERSED are sediment supply and sediment transport models. They are used to determine the amount of sediment moved in a year by the channel. In this case a narrow, gully channel with no floodplain (existing condition) is compared to a valley wide channel (proposed alluvial fan). The difference in sediment transport is compared across these cross sections to show how the sediment transport competency changes with channel geometry.

1. Where channel types updated? If they were not, why not?

The channel types presented in Section 4.1 were assessed after the 2021 flood season and already account for changes in post-wildfire geomorphology.

1. Sqy what you used to estimate the transport capacity and then how that was taken into account.

This is addressed in previous comments on the FLOWSED/POWERSED modeling.

1. You looked at 2021 but used 2022 modeled data. Explain why here.

We took the current condition during the fieldwork period (Fall 2021) and used this period of time to model for 2022. The model uses the stream banks and hillslope in the current condition, considers average annual precipitation for the calendar year, and then models sediment output for a year. Since we completed the analysis and modeling in late fall 2021 (after the 2021 monsoon season), this analysis and modeled results are predictions for sedimentation for the entire 2022 winter runoff season and 2022 monsoon season. We didn't use other years (2019 or 2020) because we didn't have relevant field or geospatial data for these time frames.

1. Change "calculates" to "estimates"

Done

1. Cite “a table or figure after “erosion rates.”

Figure 7 and Figure 8 are provided, and cited, for displaying the data from this section.

1. Say the reasoning behind the different precipitation events you modeled in the methods. Also, state why you didn’t just use the observed precipitation.

The reasoning has been explained in the Methods for each of the models as well as using the four observed flood events.

330-331. Is this the full inventory of storms? I think you need to actually show the storms that occurred, and so I’d suggest creating a new figure showing that.

A sentence was added to clarify that these storms were the only ones to generate flood runoff in 2021 within the study area.

1. This sentence doesn’t make sense here. Topic sentence of paragraph is about precipitation, but here you switch to a totally different topic of sediment removal.

This sentence was moved to the next section on sediment transport and retention.

1. here you are talking about a result, but you don’t mention the mechanism that triggers more sediment retention?

More description is added to this section.

356-357. Here you talk about 2022 flow events. This makes me think that it might be helpful to have a timeline so we can better keep track of events that you are investigating.

The sentence explains that the 2022 flow events were minor with no flooding. More context is added to this section. During the study there were only four large flood events, which are described in the manuscript.

1. Sentence that starts with: “Observations on …” doesn’t seem relevant. What about this statement makes it go with the rest of the paragraph? It seems like a random statement about a different fire.

Added context, the Pipeline Fire watersheds had similar sediment retention structure designs (“alluvial fan restorations” or “work areas”).

1. Here you start talking about the Pipeline fire again. But this is not mentioned in the study Site section or the methods. If you want to talk about it in the discussion, it needs to be introduced in those sections beforehand.

The Pipeline Fire is not part of this study, discussion of a nearby fire, outside of the study, is appropriate in the Discussion section. Adding external discussion pieces to the Study Site Section or Methods would not make sense. A description of the Pipeline Fire has been added to the Introduction.

1. Did you see 1-2 orders of magnitude change in runoff? Say how your observations were similar or different than this prediction.

Added a reference to hydrologic observations from the Museum Fire watershed.

1. Reference a table or figure after 2021 on this line.

The flooding events were described previously in the manuscript.

1. Define “complacency”

Complacency has been indirectly defined already in the Discussion, more information is available in the references provided (Stempniewicz 2014; Fulé et al. 2023).

Be more specific about what empirical estimates and “other factors” you are talking about here.

Unsure what specificity is requested. The sentence already provides the assumptions and unknowns that might lead to error: “Other factors likely include uncertainty in the empirical estimates (both over-estimating due to water volume in the sediment/debris loads as well as under-estimation due to floodplain areas not addressed by flood cleanup efforts), as well as WARSSS and WEPP model limitations for rill and gully erosion processes (hillslope incision).”

420-421. You don’t talk about different channel designs in your methods so it is out of context to bring this in now.

The Discussion section is the perfect location to provide a discussion of the results, techniques, and methods. The “Rosgen” style of channel design, also known as Natural Channel Design, is commonly used in the USA but is not without its controversy and detractors. Lines 420 through 429 provide discussion on the use of this channel classification and design method in this study and how it compares favorably with a more generic WEPP model as well as empirical results.

1. What evidence suggests that the channels are evolving to a stable form?

The sentence in line 435 is prefaced with the channel type and based on “natural channel design” process interpretation. A reference has been added to the appropriate paper explaining this process (Rosgen 2009).

1. Here you suggest eliminating the gully, but one popular definition of a gully is that it is not a feature that can be easily reworked by machinery. Once established it is likely to come back.

A sentence has been added explaining the use of rock sills as grade control to maintain the alluvial fan plan view (lines 457 to 459).

1. Say the recurrence interval.

Line 454 now explains that this a 4% annual exceedance probability

Figures and Tables:

Figure 1. Make the tick labels bigger. If you are running out of room on the latitude tick labels, you can rotate them. Label Spruce Wash. Use a (b) to label and refer to the inset. And in the inset say what the shapes represent, I assume they are county boundaries? Add a polygon showing the fire

perimeter. Also the line type for the forest boundary and Mt. Elden Lookout road are difficult to distinguish.

Figure 1 has been updated.

Figure 3. I think figure 3 can be combined with figure 1 because most of the information is the same. Also, I see some cross sections near the label for "Lower North Trib" that aren't crossing the stream lines. Why is that?

We left Figure 3 in "as-is" since it would make Figure 1 very busy and difficult to read to add the cross-sections on top of the other neighborhood labels and delineations of the sub-watersheds. The cross sections may not line up with the stream lines if the National Hydrography Dataset (NHD) has mapping errors. The NHD is the base layer for the stream lines shapefile.

Figure 4. Can you put the alphabetical labels close to some of the stream segments on the map to better see how the segments are related to the legend?

We have updated the figure.

Figure 5. Say what sets the categories for unit erosion

Unsure what the reviewer is asking for, the units are presented in the figure and the figure legend.

Figure 7. The labels in Figure 7a are too small. In caption, consider changing "major channels" to "tributaries"

This figure has been improved accordingly.

Figure 8. In the caption say why you chose the 100 year forecasted hillslope annual yield

This is a common WEPP model output and is explained in the WEPP model literature cited in this manuscript.

Figure 9. Your arrow is attempting to indicate two things. So I suggest you actually swap the vertical error for a vertical bar. Then have a horizonal arrow going in one direction to indicate 6260 Mg before August 17, and a second horizontal arrow in the opposite direction to indicate 3760 Mg after August 17.

Unsure of what the request is for this figure. The current figure shows the amount of sediment removed from the neighborhoods by day and also delineates the difference between the first three storms and the larger August storm.

Figure 10. make a figure 10 showing the observed precipitation.

The four flood events are described in the manuscript, a more in-depth discussion is provided in the cited hydrology report (Schenk et al. 2023). The study area is in a semi-arid climate with

infrequent rain and flow events. Flow events that did not trigger a flood are extremely minor with little sediment mobilization.

Table1. Add name of model in column labels

The model used for each column is already expressed in the table caption. Adding this detail, and another explanation, in the table would be duplicative and make the table less concise.

Table 2. Put the K value in the column titles.

Added:

K-Values:

Low K: 0.29 (sandy loam)

Medium K: 0.545 (sandy loam + high burn)

High K: 0.80 (High burn severity K Value)

Table 3. Indicate a column title for the last unlabeled column

The header has been fixed

Response to Reviewer #2, August 8th, 2024

We would like to thank this anonymous reviewer for their feedback, the comments have led to a greatly improved manuscript. Below are the reviewer's comments in regular type and our responses in bold type. We expect to update the manuscript to the EGU website after this Response to Reviewer.

Schenk et al investigate sediment source and their transport in a region of Arizona that has been impacted by wildfire. Their work is strongly framed around modelling approaches and comparisons. They also use empirical information to validate their model outputs and discuss mitigation efforts. Considering the context of such processes with respect to the occurrence of landslides processes such as debris flows; the present study is relevant to the audience of NHESS. However, I think that there is room for improvement, notably via an effort at presenting this work for a broader audience. Too often the manuscript reads as being very case-study focussed and the connection with other studies (either modelling or result based) is missed.

Overall I also agree with the comments of the first reviewer who points out to several key aspects. I have additional specific comments to hopefully improve the manuscript.

Thank you, please refer to our responses to Reviewer #1 from January 2024. We improved and clarified the manuscript based on their comments and suggestions.

Abstract:

The abstract sounds very technical, especially for readers that are not directly familiar with the models. Abbreviation are usually to be avoided here. Quantitative values, if mentioned, should ideally be compared/discussed with the broader literature.

We understand the need to avoid abbreviations in the abstract but unfortunately these model name acronyms are analogous to the model name (nobody references these models by their full names and each of these are discussed widely by their acronym as a surrogate for the model name). Attempting to write out the full model names in the abstract would be cumbersome and reduce the readability. The quantitative values provided in the abstract are directly comparable between observed rates of sedimentation and the different models. Providing a comparison with the broader literature, in the abstract, would not be practical since there are multiple variables that go into an event-based sediment laden flash flood event. The purpose of the paper is to discuss the precision and accuracy of different sediment models and not provide a comparison of sediment flux between disparate studies of post-wildfire runoff. The numbers provided in the abstract are directly comparable between different models and the observed sediment deposition.

Introduction:

Overall, this section could be improved thinking about the broader audience. It needs to provide a better justification of the models used. The study area must also be better justified. . See specific comments below:

Lines33-34: the focus seem to only concern the American West. I would welcome an introduction that goes broader. In other words, can a researcher from Spain, Greece or Mexico for example be interested in this research as well?

We changed the sentence to be more inclusive to humans living at the WUI for semi-arid forests anywhere globally. At the time of the initial writing the WEPPcloud model was only available for the USA, it has now been expanded for soil classification databases on multiple continents making it of a more global use.

Line 60: three models acronyms (please explain the acronyms when mentioned the first time) are introduced without any justification. As it is the introduction, these models should be backed-up via the state of the art.

A sentence is added in the Introduction that provides the long form name of each model, the discussion of the modeling techniques is already included in the Methods including the rationale for using each of the techniques. We feel this justification is more applicable to the Methods section than the Introduction.

Line 21: the study area is mentioned without any broader context. In other words why is that study area of interest for the international audience. Why is that an ideal case study? The goal here is not to repeat what is proposed in section 2, but instead make sure that a reader from, for example, China, finds Flagstaff a place of interest.

This comment could be said for any study. We revised Line 21 to further explain the context of the fire area and study (removed the name of the watershed and explained the geologic underpinning of the burn scar).

Study area:

make sure that all the local names are relevant. Overuse of such names are not ideal for the understanding of the research.

We agree, however the use of local names in the Study Area section is intentional, it provides needed context if a researcher was to attempt to replicate this study or confirm the work that we are presenting. If a reader finds the detail distracting they can read ahead to the next section.

Method:

Overall, there is a lack of method justification with respect to the literature. See specific comments below.

Line 104: what does the acronym FLO-2D stand for?

While written like it is an acronym it is the actual name of the model, a search of the internet and a literature search on Google Scholar did not find a different name other than "FLO-2D".

Line 105-110: What Lidar data are used? Provide source, resolution information, etc.

Both lidar datasets (2015 and 2019) are available on the USGS National Map server. This information has been added to the manuscript including a link to the National Map server. The horizontal and vertical resolution vary between both datasets but are sub 10 cm accuracy or better.

Why (based on what physical criteria?) these grid scales in the flood modelling ?

The initial grid size was selected based on emergency conditions during the wildfire (the ability to provide quick results to the emergency management teams). The more refined grid element in 2019 was selected to provide a better resolution of rainfall-runoff. The relative trends for sub-watershed runoff did not change with the change in grid size. The overall watershed is sufficiently large that either grid size provided relatively precise results for runoff and flood events.

Lines 151-156: can you clarify on the sediment transport analyses carried out? Can you the back this up with literature?

This is a similar question to Reviewer #1, we have improved the manuscript after those initial comments to better frame how FLOWSED/POWERSED operates and how it has been used previously in the scientific literature. Thank you for sharing a similar concern as Reviewer #1, we hope the revised text provides needed detail. Please refer to the Response to Reviewer #1 for more information.

Lines 184-186: any reference for these modelling approaches?

References to both modeling approaches are included in the manuscript.

Lines 200-205: any of the values used in the model can be justified from the literature?

The values used for the MUSLE model were predicted based on field conditions, this is now clarified in the manuscript. More information about determining the K value, C value, and P factor are available in the reference provided in the manuscript.

Lines 207. CoF staff?

CoF is defined on Line 82 (City of Flagstaff).

Lines 207-212. Reference(s) to support these methodological choices?

The rationale for the methodology of the empirical measurements is included in Line 212 (observations were noted for Federal and State disaster declaration reimbursements). Due to the emergency nature of the urban flooding there was no opportunity to collect more holistic measurements using precision surveys. Cleanup operations commenced during the falling limb of the flood, the most appropriate way to capture sediment deposition in the urban environment was the Federal and State disaster reimbursement paperwork submitted from the landfill and through flood observer photos.

Results:

General comment: can the erosion/sediment values obtained in this study be compared to other cases? That could help to make the discussion even more interesting/of a broader interest. Providing quantitative values without putting them into perspective is not always relevant.

The purpose of this paper, as stated in the Introduction, is to provide an example of three different sediment modeling techniques and compare the precision and accuracy with empirical results. Comparing gross sediment values to other studies would not be very interesting based on the number of confounding variables that make comparison difficult. This could include the number of flood events, rainfall hyetographs, burn severity comparisons, slope, geologic provenance, watershed recovery/vegetation recovery, and other variables.

Lines 395. The author refer to gully erosion. This comes as bit as a surprised that this process is not mentioned earlier in the study area section. Something that remains unclear is the origin of these gullies. Where those gullies be already in place before the wildfire? If so, would these gully be associated with earlier wildfires? In some cases, gullies are not to be the consequence of landslide processes. Are their observation of landslides in the regions. Landslides could develop after wildfire

of course, but could also be there as a basic geomorphic agent that bring sediment to the river system. Overall, some clarification (extra relevant information) around these mass movement/erosion processes would be welcome is that helps to better understand the model outcomes.

I believe there may be a definition miscommunication here. The hillslope gully erosion mentioned on this line is for the formation of hillslope gullies and rills through hillslope erosion. We are not talking about mass movement or mass wasting. No mention of landslides or debris flows is provided in the manuscript in terms of modeling or empirical observations. There are plenty of examples of gully erosion definitions in various government reports from agencies on multiple continents, this manuscript subscribes to the common definition of the term and is discussed at length in the Discussions section. Mentioning gully erosion as part of the hillslope erosion modeling is described in the Methods section.

The onset of gully erosion is interesting in this region due to the prevalence of gullying on hillslopes post-wildfire. The process is more fully explored in an earlier paper from the nearby Schultz Fire that is cited already in this manuscript. The reference for that paper is as follows:

Neary, D.G., Koestner, K.A., Youberg, A. and Koestner, P.E.: Post-fire rill and gully formation, Schultz Fire 2010, Arizona, USA. Geoderma, 191, pp.97-104, 2012.

Another reference to the Neary et al. study has been added at this line location.

Note here a reference of gully erosion modelling. Although targeting different scales, that could be useful: Vanmaercke, Matthias, Panos Panagos, Tom Vanwalleghem, Antonio Hayas, Saskia Foerster, Pasquale Borrelli, Mauro Rossi et al. "Measuring, modelling and managing gully erosion at large scales: A state of the art." Earth-Science Reviews 218 (2021): 103637.

Thank you for the reference, the paper is interesting and is now included in the Discussion section under the gully conversation.

Figures:

Figure 1: add elevation quotes.

Unsure of where the reviewer would like elevations called out, the topography is highly heterogenous. Example elevations for Mount Elden, Dry Lake Hills, Mount Elden Estates neighborhood, and Paradise/Sunnyside neighborhoods are now included in the Study Site section. Contour lines, and DEMs, for the area are freely available online at multiple sources (e.g. USGS National Map, Google Earth, City and County GIS portals, etc).

Figure 2. Indicate when the photos was taken. Provide also the geographical coordinates of the photo.

A year and season is now included in the figure caption as is an approximate location that can be compared to the site map.

Figure 10. Indicate when the photos was taken. Provide also the geographical coordinates of the photo.

A month was added to the existing year in the caption, the geographic coordinates for the "Ginger Fan" were added to the caption.

Response to Community Comment #1 (January 2024)

Thank you for a great manuscript. Below are two comments:

1) Lines 359-362 "Observations on the nearby Pipeline Fire...":

What is meant here to have sedimentation in the 70-80% range? Does this mean that the work area captured 70-80% of the sediment mobilized upstream of it? If so, how was this percentage calculated? If pre-and post- storm LiDAR was employed, a figure in the manuscript or supplemental material would be helpful to illustrate what is meant by 70-80% sedimentation.

2) Lines 444-447 "The sediment transport models indicate...":

This statement is not found in Beers et al. (2023). A different source is needed for this statement.

Thank you Rebecca for your thoughtful review and comments.

When we submit our revised manuscript we intend to address both comments. The 70% sediment reduction mentioned in the FLOWSED/POWERSED modeling discussion will be clarified to explain that the modeling results indicate a 70% reduction of sediment based on the difference between the sediment output at the outlet of the "work area" and the sediment input at the inlet of the "work area".

The reference to Beers et al. 2023 has been moved to a new sentence that explains that monitoring in 2023 observed both areas that worked well (per personal observations and comments from the local Flood Control District) as well as steeper slope alluvial fan "work areas" that underperformed due to storms that exceeded the design storm for the project (as monitored and presented in the Beers et al. 2023 reference).