Review responses and revisions from submission to Ecological Applications

Manuscript title: Local and watershed controls on large wood storage in a mountainous stream network

Authors: Matthew C. Vaughan, Gregory B. Pasternack, Anne E. Senter, and Helen E. Dahlke

Reviewer #1 (Other Suggested Journals):

Geomorphology, ESPL, River Research and Applications

**We have accepted the reviewer’s suggestion to submit a revised manuscript to Geomorphology.**

Reviewer #1 (Comments to Author):

The manuscript entitled, "Local and watershed controls on large wood storage in a mountainous stream network" presents a significant dataset and clear meta-analysis of large wood deposited within a stream reaches of a 6th order watershed. The manuscript aims to add overbank wood into our conceptual model of LW processes in river networks. In addition, the analyses focus on correlating landscape and local variables with LW deposition patterns. The manuscript (data, analysis and interpretation) is a novel contribution to the literature and is well written. Analyses are sound and discussion does not go beyond the results.

**We appreciate the positive feedback.**
I have highlighted a number of issues in the manuscript below that I think will improve the manuscript. My main points of concern are: (1) the lack of any ecological justification and interpretation of the results (important for an ecological journal), (2) the statement that our conceptual model lacks recognition of overbank deposition is an overstated, (3) the results and discussion are concise but potentially detailed beyond necessary for making clear conclusions, and (4) some skepticism on the multivariate results is necessary; that is, a five variable model with an R2 = 0.31 is quite complex and direct driving variables are not always process-based so discussion should include some mention of potential limitations of the study design.
My final point is about the appropriateness of the manuscript in its current form for Ecological Applications. LW indeed has ecological implications of which none seem to be discussed in the manuscript. The way this manuscript is written does not appear to be for an ecological nor applied audience. Though LW studies suffer from being not purely ecological nor purely physical in nature. The manuscript could add these sections to improve its presentation for the EcoApps readership.

1. **This is a fair point. For this reason, we think that the manuscript will be better suited for *Geomorphology*.**
2. **We disagree that this is an overstatement. A thorough literature review was done to consider this question specifically, and we found that overbank LW was largely left out of the conceptual model of watershed scale LW storage. Evidence of this is shown explicitly in Table 1, where the literature is summarized.**
3. **This is a subjective statement, since some readers will undoubtedly appreciate the level of detail the manuscript reports. We are inclined to maintain the level of detail currently included in the manuscript. We are open to paring the results down if this is requested from *Geomorphology* reviewers andeditors**
4. **We agree that a discussion of study limitations is helpful. We have added sentences in the methods, the first paragraph of the discussion, and also in the conclusion in order to be more explicit about the limitations of our approach. On the other hand, we feel that our approach is a significant advancement upon previous studies that only looked longitudinally instead of on a watershed-scale basis. We think our methods and results will be used by future researchers.**

Misc. Strengths:
• The quantitative comparison of these study results to other datasets is helpful.
• LW studies do need to clarify the used of terms and meaning. The introduction helps to define some of the terms.

**We appreciate the positive feedback.**

General Comments

• The ecological motivation for studying LW is largely absent in the introduction and conclusion. Although LW clearly has ecological and geomorphic implications (and therefore indirectly ecological) this manuscript does not discuss the ecological justification for studying LW nor the ecological meaning the results. That is, since more wood is stored in the floodplain than in the active channel, what does that mean for our understanding of aquatic and riparian systems? This could be tied in with the manuscript's claim that our conceptual model of the longitudinal patterns of LW is potentially limited.

**We thought that it was self-evident that a study about the fate of once-living organisms would be relevant to ecology and of interest to ecologists. Ecology definitely includes the study of dead organisms and how they are processed in the environment. Thus, LW is not just the realm of physical scientists. We could easily follow the template of other LW studies and bloat the introduction with ecological justification to negate this criticism, but since the article was rejected by Ecological Applications largely on the basis of this one comment from reviewers and since LW scientists appear to be dominated by physical geographers, we thought we would move forward accepting the perspective that we should submit to a physically-oriented journal. *Geomorphology* is a well established journal in which other studies of LW are published.**

• The lack of significant differences in the volume of wood by coarse variables is interesting but much of the results suffer from the lack of significant patterns. The results are concise but are quite exhaustive. The results and discussion could be focused on ecologically significant emergent patterns rather than a full description of the non-significant parts of the dataset. Same goes for the lack of significant patterns in in-stream LW. However, some discussion on the meaning of the results, significant or not, would improve interpretation of the results. For example, there was an exception of 3rd order stream reaches compared to 1, 4 6 when comparing LW volume. This suggests that something else controls LW capture on floodplains that does not uniformly in the downstream direction. It appears that local-scale variable might be better at predicting LW volume such as capturing efficiency or production (as noted in the results) than coarse driver variables like discharge,
slope or elevation. Repetition of results in the discussion could be trimmed and focused on important findings.

**Results have been pared down to the most important for us to communicate the science. This includes significant and insignificant statistical results. To say that the results “suffer from the lack of significant patterns” implies a bias on the part of the reviewer that patterns should be present. Zero is a significant number to report in science. Unfortunately, without specific reference to where this reviewer thinks the results are unnecessarily repeated in the discussion, we’re unable to know what they are suggesting. We have read through the discussion again and feel that only important results are discussed and compared to similar results of earlier works.**

• The statistics and explanation of the regression model are sound and well described. As with most LW studies, the results are not altogether clear. Five variables and an R=.31 suggest that LW volume is a complex process to predict. The manuscript notes that more process-oriented variables were selected during the cross correlation process. However, it is unclear how urban area and percent igneous relate to LW production, transport or deposition.

**We appreciate the positive feedback. We agree that it is a complex process. Our discussion offers our best effort to explain the influence of significant variables. We have added language to try and clarify this further for future readers, and added the suggestion that “percentage of contributing stream cells that were over intrusive igneous rock was a significant variable in the model since it is correlated with separate process-based variables that were not considered in the study.” In other words, acknowledging that statistically significant correlation does not prove causation. It does offer a guide for future scientists to look into this more and in other locales, which is meaningful to communicate to the community.**

• The downstream connectivity of the Yuba River seems compromised by multiple dams. A brief justification for why the Yuba River system, despite its historical impact, would help the reader understand if the dams have an impact on the observed patterns.

**Dams are ubiquitous in mountain river systems around the world. To choose a system without dams would limit the study’s comparability to other more typical areas. Meanwhile, the entire North Yuba above New Bullards Bar reservoir is undammed (we have now added a sentence to this effect in the study area section), so in fact our study did investigate and compare LW conditions in both undammed and dammed areas of comparable size, etc.**

• Regarding the argument that our conceptual model has "understated or overlooked" floodplain deposition of overbank wood seems like a straw man argument in this paper. Others, including this reviewer, have stated this previously. Interestingly, the manuscript itself notes at the end of the paper that others have noted this but not quantified it (Line 496). The manuscript would be improved by tempering this statement. The manuscript is correct however, that quantification of in-stream versus overbank wood in the longitudinal direction is a significant contribution of this manuscript.

**It seems from the statement about “our conceptual model” that the reviewer is likely an author of one of the articles we have identified as missing or inadequately addressing this key point we are raising here, so it is understandable but alarming that the reviewer is concerned for other motives that are unfair to us. We should be free to publish our fresh perspective and objective scientific analysis without being blocked by a partisan viewpoint. Let the scientific community weigh the merits of our data and reasoning to decide for themselves. In terms of following the scientific method and applying technical soundness, we made every effort to be transparent and communicate which previous studies included overbank wood, and which did not in Table 1.** **All but two of the 22 studies we were able to find disregarded floodplains in their surveys. That is quantitative, technically sound, and scientifically significant. Unfortunately the reviewer did not specify which specific points in the manuscript created a “straw man argument,” so we cannot further improve a particular sentence the reviewer might be taking issue with, which we would be happy to do. We feel that we are simply stating the results of our literature review, which is that 20 out of 22 studies we found only looked for LW in the stream and not overbank. This creates a unique opportunity for our study to add to the understanding of LW as it was observed in the entire fluvial corridor. Science changes through time and it is not only acceptable, but essential for new manuscripts to review and evaluate past work. We have endeavored to not judge individuals, but keep the focus on the ideas and the science. We ask reviewers to do the same.**

• Multiple statistical test were performed in the analysis. Considering this, the manuscript would be improved by application of a Bonferroni correction to the alpha value. The results appear robust, however, a conservative approach would improve the manuscript.

**We have thoroughly researched the issue raised here. The Bonferroni adjustment is a commonly used method for adjusting the significance levels of individual tests when multiple tests are performed on the same data. It divides the nominal significance level, α, by the number of tests being performed simultaneously to prevent the overall level of significance from exceeding the nominal level, α. In our analysis, however, a stepwise (i.e. iterative) AIC-based model selection was conducted based on the exact AIC. Note, the AIC (Akaike Information Criterion (Akaike 1973, 1974) is calculated as followed:**

|  |  |
| --- | --- |
|  | (1) |

**where *Y* is sample size and is *RSS* / *Y*. *RSS* is the residual sum of squares:**

|  |  |
| --- | --- |
|  | (2) |

**where xy and are the values of the dependent and independent variable for a given data pair, y. *λ* is the number of parameters explicitly appearing in the model plus one, as in this context the fit metric (RSS) is considered an additional estimated parameter. The AIC is a model selection tool. From *M* models, *gm*=*1,M*, each having different parameter values or model structures, the one having the lowest AIC is considered best. “Best” is defined as model *m* gives the optimal combination of performance and parsimony, as AIC increases with both model error (*RSS*) and model complexity (*λ*). That is, AIC-based model selection amounts to a rigorous formalization of Occam’s razor, and by construction avoids model over-fitting. If we are mistaken on this, we are open to further addressing this to the manuscript with more specific guidance, but it seems to us that this comment is technically unsound.**

• Background to the sites is quite long and tell of the history of the Yuba River. This is important information but the historical review needs to be tied better to use impacts on LW. How has the use history impact within channel and floodplain LW?

**Because we included many variables related to natural land attribute and anthropogenic impacts in our study, adding a new domain to the LW literature along the way, we find it necessary to provide a relatively thorough description of the Study Area. We report in the results of the manuscript that our testing found some anthropogenic variables such as wildfire to be insignificant, while others such as urbanization appear to have an effect. Thus, we have presented the LW findings that relate to the Study Area text we have. Rather than speculating about causality, we aimed to report the facts and information available, and only made inferences where sensible. If reviewers would like us to shorten the text in the Study Area and Methods sections, we can gladly shift text identified by reviews specifically to a supplemental information section. That is an emerging practice now and we would be willing to do that upon suggestion from reviewers.**

• Wood volume was calculated differently for solitary and log jams. If there is a longitudinal pattern in solitary and log jams, this estimation could impact the results. Was there a downstream pattern in LW solitary pieces and log jams? If so, justify why this did or did not impact the results.

**It is true that volume was calculated differently for these, as reported in the manuscript. This may have impacted the calculation of LW volume at each site, but it is unlikely to impact the downstream trends, since the estimation was consistent at all sites. In other words, if it impacted any, it impacted them all in the same way. The manuscript already has enough objectives and tests as well as length, so we view it as out of the scope of this manuscript to report downstream trends in solitary pieces vs. LW jams. Our experience is that trying to answer all questions in a single monograph-style article is not what journals, reviewers, or readers want in the 21st century. That is something that we envision addressing in future manuscripts along with results for other ways of looking at this large dataset.**

Specific Comments:
• In the results section there is no mention of spatial autocorrelation. The bin-ing of the data might avoid this issue but it might be worth mentioning. It is assumed that the analysis did not take into account that adjacent stream sections might not be statistically independent. Justify this assumption.

**Addressed in paragraph four of section 3.4 paragraph 4.**

• Final paragraph in the study site section does not orient the reader to the dam locations. Reference the study site figure.

**Good suggestion. A reference has been added.**

• Line 251 - Unclear how the impact of the dam was resolved. Specifically, Line 254 states that only ten stream site were affected by the dam. The following text is unclear.

**We changed the language to clarify this statement. It now reads: “In order to reflect the fact that no LW in New Bullards Bar Reservoir was able to be fluvially transported to downstream sites, the amount of contributing drainage area upstream of this reservoir was subtracted from stream sites downstream of the reservoir.”**

• Line 259 - Presumably this is channel or down valley slope. Adjust to be precise. Text also mentions side slope. Unclear why this was measured and how it is to be used in the analysis.

**Modified to be more precise.**

• Line 484 - Use of the term 'Authors' here is unclear.

**Changed to “previous researchers”**

• The information in Figure 2 could be move to text to decrease the number of figure. Although there are not a lot of figures or tables in this manuscript.

**The plot in Figure 2 shows much more information than could possibly be described in the text. As the reviewer points out, the number of figures in the manuscript is very reasonable. Our preference is to illustrate this with a figure.**

------------------------------------------------------------------------

Reviewer #2 (Other Suggested Journal):

Geomorphology

**We agree that this manuscript is well suited to *Geomorphology.***

Reviewer #2 (Comments to Author):

General Comments
1. Importance and interest to this journal's readers
a. While goals are clearly stated, a significant research problem has not been articulated, and the findings do not seem to have specific relevance to environmental management and policy.

**We are confident that readers of *Geomorphology* will be interested in our novel data set, experimental approach, and results. Preliminary results were presented at the Fall 2013 AGU conference, and were very well received.**

2. Scientific soundness
a. I think this part of the paper could be stronger if it was anchored in a conceptual model or tied to a priori hypotheses.

**We have done our best to explain the existing conceptual model in the introduction, point out its flaws, add to it using new information, and discuss it at the end of the manuscript. This is done using a suite of clearly defined hypotheses. We have tweaked the language in the introduction and discussion to try and make these points more clear for future readers.**

3. Originality
a. The study design was one of the more robust that I have seen used in studies of wood abundance and distribution, in terms of site selection and sample size.

**We appreciate the positive feedback.**

4. Degree to which the conclusions are supported
a. I found the investigation of variables that may control LW storage to be very exploratory and found the results too ambiguous to be very useful. The tests were not based on explicitly stated, a priori hypotheses about which factors could be most important (e.g., in the context of a conceptual model).

**We agree that they are exploratory, which was by design. We are not sure what this reviewer means by the results being ambiguous. In science, finding a lack of statistically significant relations is equally as important as finding very strong one. Reviewer #1 thought that results were too detailed and specific. Multiple hypotheses were tested using robust statistical tests, and a new conceptual model was communicated in the discussion. We have added language to the introduction and discussion to make our intensions as clear as possible. In applying the scientific method, the manuscript is very straightforward and adds the novel approach of sampling throughout a watershed. Now that we are providing an exploration of numerous variables it points the way for future scientists to either compare against watersheds in different climatic and geologic contexts or focus on specific processes that are most promising for the same climatic and geologic conditions as we studied.**

5. Organization and clarity
a. The paper was well-organized and, for the most part, clearly-written.

**We appreciate the positive feedback.**

6. Cohesiveness of argument
a. See other comments about the need for concept model and articulation of the research problem.

**More language has been added to the introduction make this clearer for future readers.**

7. Length relative to information content
a. Suggest deleting exploratory section relating LW storage to landscape variables.

**At AGU in 2013 we received feedback to do this kind of landscape testing prior to submitting this manuscript, especially given that the watershed has complex geology, land use, land cover, wildfire history, and flow regulation. What if those variables confounded the relations with simple topographic variables more commonly tested in LW studies? We should not hide from such natural complexity, but address it head on. We believe that this is one of the strengths of the paper. We were able to investigate the relationship between GIS based variables on multiple spatial scales, and identify differences between sub-basins of the same system. This type of analysis has not been done before.**

8. Whether material should be moved to the digital appendices
a. I would consider moving the Table 1 summarizing the prior literature to the digital appendix or representing it with a figure (e.g., bar graph) instead.

**This information is at the core of the argument we are making concerning the language and conclusions of prior LW studies. We fear that leaving it out of the main manuscript would welcome criticism from future reviewers. Scientists should not fear pointing out problems with past theories, but instead address the issues head on while also showing due respect for the challenge of science over the ages. We have tried our best to report the truth, while also not making criticism of individuals. Potential scientific disagreements or findings of past faults should not be barred from the literature, especially in such a young field as LW studies.**

9. Conciseness and writing style
a. The paper is well-written and sufficiently concise.

**We appreciate the positive feedback.**

10. Appropriateness for Ecological Applications
a. This paper effectively demonstrates some of the flaws in the widely-embraced assumption that wood storage generally decreases from upstream to downstream in a large watershed. That is a valuable contribution to the literature and to our understanding of large wood dynamics in streams, in general.

**We appreciate the positive feedback.**

b. However, I don't think the findings, as presented, have specific relevance to environmental management and policy. These linkages would need to be drawn out more explicitly, which would require the research problem to be defined more clearly.

**The journal Ecological Applications broadly publishes work on the science of LW among other topics, not just management concepts, so we feel that we chose wisely. Given the comments of these reviewers, if given the opportunity to resubmit there we could easily tie this to both ecological theory and practical application- this independent study was undertaken as a result of divergent findings and allegations in an adversarial environmental management situation, so it would be easy to add text to raise such issues- we just chose to focus on the science. Nevertheless, the reviewers clearly sent the message that this study about the basic science of LW at the watershed scale is highly suited for *Geomorphology*, so we are submitting it to Geomorphology**

Specific Comments (supporting general comments with specific evidence; with line references)
1. Presentation
a. Line 44: Introduce a conceptual model in this section to explain the factors that are known to influence LW storage.

**Added to the introduction**

b. Line 62: LW storage volumes and longitudinal trends are interim research outputs. Need to define the real-world problem and application here, or link this discussion to existing knowledge gaps. Explain the consequences of problems in units and terms for river ecology and environmental management.

**Addressed in introduction and in section 1.3.**

c. Line 114; 615: Goals are stated clearly but it's not clear why it is important to accomplish these goals. For example, what was the problem that needed to be solved? A knowledge gap alone (Line 112) or something else? For example, why is important to know that LW per CW does not show a decreasing trend in the downstream direction? On line 632, you hint that little is known about how disturbances interact to impact the LW budget in the Yuba. Was this the driving question behind the work?

**Language was added to the beginning of this section to explain motivation for these goals.**

d. Line 145: Compare vegetation in upland and floodplains, and explain longitudinal variation in riparian forest structure from low- to high-order streams. This is important for understanding variation in wood supply.

**We have described forest vegetation on lines 145-152. Reviewer #1 asked for fewer details in this section. We are open to adding more details if requested by *Geomorphology* editors and reviewers. Again, we could move some of this Study Area text to supplemental Materials if reviewers provide specific guidance on what they think is non-essential, as we see it all as essential given the tests we ran.**

2. Length
a. Line 153: Shorten this entire section to several concise statements. The main point is that the Yuba River has mostly second-growth forests of a certain age, and the stream channel contains an overabundance of alluvial sediments owing to century-old mining impacts. Add a statement to explain how flows are regulated and the primary ecological impact.

**We fear that shortening this section and leaving information out will likely raise questions from future readers. Again, we could move some of this Study Area text to supplemental Materials if reviewers provide specific guidance on what they think is non-essential, as we see it all as essential given the tests we ran.**

3. Methods
a. Line 182: Justify the use of half log-scale drainage area bins instead of preserving the continuous data you worked so hard to obtain. The reasons become apparent later (Lines 416-420), but the question arises immediately.

**This is explained in the same sentence that introduces the bins in section 3.1. It was part of the stratification sampling scheme to “yield equal effort sampling spanning all scales and some basic characteristics of streams in each bin.” This section introduces the field sampling methods only, and says nothing about whether continuous or discrete data was used. Because the reviewer was unfamiliar with stratified sampling and the need to stratify on the basis of the log of basin area, we have added more text and citations to buttress this key aspect of our experimental design.**

b. Line 192: Not clear how bankfull channel width (BFW) was measured. Explain and distinguish 'active floodplain' from more commonly-used 'active channel' terms.

**This is explained in section 3.1 paragraph 2**

c. Line 194: These study reaches are too short, especially for wider channels (i.e., Bins 4-8). I expect this contributed to the very high levels of observed variability in estimated LW storage and reduced your statistical power substantially. Study reach length should have been scaled to channel width (e.g., 10-20 times the bankfull width) to ensure each sample was more representative of wood storage at the reach scale instead of the scale of only a few channel units. This is a hard problem to overcome.

**More data is always better, and scaling to channel width would have been ideal. Practically speaking, the reviewer’s suggestion would have been impossible with any reasonable amount of time and resources. Since channels were 40 – 60m wide, surveying 10 or 20 times this length would have taken multiple days for each site. Instead of having fewer sites with more surveyed at each site, our research questions focused on the watershed and sub-watershed scales, so we opted for 50-100m surveys, and many observations. Also, considering the total length of streams mapped in each drainage area bin, the amount of mapping was significant and appropriate.**

d. Line 204: Explain how area was delineated (e.g., lateral/longitudinal extent of visual assessment). If you used an established technique for measuring canopy cover (vs. canopy closure), specify which one. This would be hard to repeat, as described.

**Added language to this section.**

e. Line 214: Not sure what 'functionally connected by a significant amount' means. I think you mean that a continuous line of racked small debris was in contact with each piece in the jam.

**Language was tweaked to be more clear.**

f. Line 217: Cite reference for survey methods if possible. For example, LW should equal or exceed 10 cm in diameter - did you measure to tip of logs, regardless of diameter, or stop at the 10 cm mark? Please cite reference for RW method, as well or justify the approach.

**We are not sure why this method is unclear to the reviewer. We already addressed this when we wrote, “….the diameter at each end and rootwad diameter (if present) were measured with large forester calipers.” The reviewer prefers the word “tip” to “end,” but these are synonymous. We are unclear on what this reviewer means by “stop the 10 cm mark.” The rootwad methods are not based on another study since they are usually neglected, but we have clearly outlined how we went about estimating rootwad volume, so they are repeatable.**

g. Line 222: Rootwads often excluded from analysis in prior literature. Need to acknowledge this potential reason for a difference in your estimates and in others.

**Added language to section 3.4 paragraph 4.**

h. Line 234: This is a weak assumption that probably contributes to substantial error and uncertainty in your storage estimates. You need to convince the reader this was legitimate, either with photos or descriptions of a 'typical' jam in the Yuba river (e.g., see typology by Abbe & Montgomery 2006 or similar).

**Since it is an estimate, it will undoubtedly introduce error. However, the estimates are based on the results of a rigorous study that investigated LW jam density specifically already. We are building on that solid foundation relative to the claim of this comment. We feel that photos and qualitative descriptions will not add more defense of our estimates than the referenced manuscript already does.**

i. Line 276: The cell/pixel size was not stated. Add this info, as I presume the stream is being represented as a series of rasters that have been attributed with these various characteristics.

**All rasters were comparable in cell size to the stated 10-m DEM raster. We fear we would bog the reader down with unimportant details if we reported cell size for each of the rasters.**

4. Data Presentation
a. Line 332: LW volumes not normally distributed so median is a better summary stat.

**Mean and median values are both reported, and the full distribution is illustrated in Figure 2.**

b. Line 335: Total volume not meaningful or relevant - delete.

**Done.**

c. Line 338; 394-397; 426-429: Explain how this extrapolation is meaningful to objectives. Qualify estimate by noting that it is highly sensitive to errors in the delineation in the stream network and probably overestimates the actual amount since the mean value was used instead of the median (i.e., because LW volume/channel length not normally distributed at site scale).

**We agree that this result is not critical to the objectives of the study. However, we felt that it was an interesting and potentially useful result, so we have chosen to leave it in.**

d. Line 343: Not clear why bins 1-4 were combined for comparison with bins 5-8; is there a basis for the breakpoint?

**Added explanation in section 3.4 paragraph 2.**

e. Line 359: Show reader values for each sub-basin (and sample sizes).

**Reviewer #1 criticized for results being too detailed and not focused on the main story of our findings. It would not be practical to report this information for all of the many tests that we ran, so we have chosen to include important information for the tests that yielded the most important findings.**

f. Line 453: add ...and regardless of the actual amount of out-of-baseflow channel LW volume per channel length (lines 430-432).

**It seems that reviewer has confused the amount of overbank LW with the percentage of overbank LW at each site. We have strived to make this clear in out methods (3.4 paragraph 5) and results (section 4.4).**

g. Line 526: I think it would be helpful to characterize the places where you observed >90th percentile LW storage and see how these places differ from sites in <10th percentile.

**This would be an interesting comparison, though it would not be a simple one to communicate. We fear it is out of the scope of this paper to go into detailed characterizations of particular sites, and would distract future readers from the main points of the manuscript. We will consider this for future manuscripts that could further mine this large dataset.**

h. Line 540: This explanation should draw on previous studies of wood distribution and trapping sites. It's not complete, as written.

**We are not aware of other studies that investigated the effect of the steepness of side slopes in conjunction with the width of active floodplains.**

i. Line 573-576: Not very informative. Instead, why not compare using your BFW bins for an apples-to-apples comparison.

**While we understand this point, there are two reasons not to make the comparison that is suggested: (1) There were few significant differences in LW volume per channel length between the different bins in our study so making this simple comparison is informative, and (2) comparing it to one of our bins would not be an apples-to-apples comparison either, since sites were broken up differently.**

j. Line 624: What is the claim that you are substantiating with this evidence? Clarify.

**Language changed to clarify.**

k. Line 631: Other kinds of studies are also required.

**Language tweaked to communicate this.**

5. Statistical design and analyses
c. Line 270: This section is very exploratory and unfocused. Need to set the context for these tests with a conceptual model and explicitly stated, a priori hypotheses about which factors could be most important. For example (Lines 384-387), why would you expect the 'upslope percent intrusive igneous rock' to have higher LW volume? What is the mechanism and implication? Same for elevation.

**This portion of the analysis was certainly exploratory, which we feel is a strength rather than a weakness. There is a large gap in the LW literature in this avenue, so naturally there have to be exploratory studies to see what is going on there. Science is supposed to be exploratory, so it is hard to see how this can be bad. Further, the watershed setting in many real watersheds is complex with many overlapping natural and anthropogenic conditions, so they have to be explored. Language was added to section 3.3 in an attempt to explain the reasoning for these variables and make the case for this type of analysis.**

a. Line 252: The reservoir introduces a major complication for interpreting longitudinal patterns in LW storage and I think Bin 8, which presumably contains these 10 largest sites) is not very useful for looking at longitudinal patterns, except to highlight the effects of the dam.

**We recognize this problem and attempt to remove the effect of the dam by subtracting the contributing area in appropriate variables in our continuous data analysis (lines 252-253).**

b. Line 293: "In all cases....robust results"; this sentence needs to be explained as it seems very subjective and not at all clear.

**Sentence has edited to be more objective.**

c. Line 295: Specify how many pairwise comparisons were made. I suggest using a sequential Bonferroni correction to adjust p-values for multiple comparisons and avoid the serious problem of 'false positives' in this part of the analysis.

**We have thoroughly researched the issue raised here. The Bonferroni adjustment is a commonly used method for adjusting the significance levels of individual tests when multiple tests are performed on the same data. It divides the nominal significance level, α, by the number of tests being performed simultaneously to prevent the overall level of significance from exceeding the nominal level, α. In our analysis, however, a stepwise (i.e. iterative) AIC-based model selection was conducted based on the exact AIC. Note, the AIC (Akaike Information Criterion (Akaike 1973, 1974) is calculated as followed:**

|  |  |
| --- | --- |
|  | (1) |

**where *Y* is sample size and is *RSS* / *Y*. *RSS* is the residual sum of squares:**

|  |  |
| --- | --- |
|  | (2) |

**where xy and are the values of the dependent and independent variable for a given data pair, y. *λ* is the number of parameters explicitly appearing in the model plus one, as in this context the fit metric (RSS) is considered an additional estimated parameter. The AIC is a model selection tool. From *M* models, *gm*=*1,M*, each having different parameter values or model structures, the one having the lowest AIC is considered best. “Best” is defined as model *m* gives the optimal combination of performance and parsimony, as AIC increases with both model error (*RSS*) and model complexity (*λ*). That is, AIC-based model selection amounts to a rigorous formalization of Occam’s razor, and by construction avoids model over-fitting. If we are mistaken on this, we are open to further addressing this to the manuscript with more specific guidance, but it seems to us that this comment is technically unsound.**

d. Line 296: Seems the actual null is that the distributions are equal in both groups.

**This is statement is equivalent to what is in the manuscript.**

e. Line 306: 0.8 is still very high. I suggest using something more stringent like 0.3-0.5 to avoid a real problem. For example, your drainage area (upper end of each bin) and your mean bankfull channel width are related by a power function with an r2 of >0.98.

**In this case, a higher *R* value correlates with a more stringent take on multicollinearity. Reducing the cutoff to 0.3-0.5 would allow more collinear variables into the model, so it does not seem appropriate.**

f. Line 311: I don't think you should say 'predict' but rather, 'explain' or relate; not validating existing MLR on new observations.

**Yes, future predictions are best described with a new dataset, though the term “prediction” is commonly used in statistics to describe the ability of independent variables to explain the variance in a dependent variable. We have found this usage in nearly all the papers we picked at random from our library that use similar regression models (Wohl and Jaeger 2009, Whipple and Tucker 1999, Creed and Beall 2009, Richardson et al 2009, Martin and Benda 2001 to name a few). Given the context of this sentence and the common use of this term, we feel that the meaning is clear.**

g. Line 316: Consider explaining that in MLR, no inferences can be drawn about the relative influence of independent variables in isolation; can't 'unscramble the omelet'. You are pretty careful to walk this line when explaining the results of the test.

**We appreciate the positive feedback on the results section. While this is inherently true for all MLR models, we have added language in the methods so that readers unfamiliar with this type of regression will be aware.**

h. Line 348: This finding is quite possibly an artifact of not correcting for multiple pairwise comparisons.

**This result is the product of one nonparametric test with one statistical hypothesis, where all assumptions were met. Based on our understanding of these statistical tests, we feel that no correction is necessary.**

i. Line 371: R2 value is quite low, meaning the MLR model is not very useful for predictions or explanations -there is a lot of unexplained variability.

**“Not useful” goes too far. Many studies in ecology, health sciences, geomorphology, and beyond have presented similarly low correlations but statistically significant findings. The results are on par with previous LW studies, and we are fair in communicating the limitations of the MLR model. It can explain a fifth of the variance, due to the complex nature of LW storage in a system like this. This is explained in the discussion, not in the results. Again, the reviewers have a biased perspective that a finding of no connection or a weak connection is unworthy of publication, but to the contrary it is essential for scientific studies to be published with results objectively found to have such outcomes. If the field of sediment budgets and watershed scale sediment patterns was subjected to such bias, then most of the literature would have to be discarded.**

j. Line 473: did you try transformations to linearize the predictor variables?

**These analyses were done on transformed data, as explained in lines 301-303.**

k. Line 478: Bin 6 was an outlier. Explain why.

**Reviewer #1 suggested that we go into too much detail in the discussion instead of focusing on important ideas. We do not have a process-based explanation for the outlier, and we feel that explaining this in the discussion would detract from the main story we are trying to tell. It is shown transparently in the results section. If editors and reviewers at *Geomorphology* would also like to see the outlier discussed further, we are open to this suggestion.**

6. Errors in technique, fact, calculation, interpretation, or style.
a. Lines 58-60: Delete or rephrase because this thought exercise is not substantiated and does not consider implications of other changes that happen in river corridor, besides changes in width.

**The statement by the reviewer is technically unsound. There was a typo in this sentence that has been fixed, so maybe that was the problem. The conclusion of the thought experiment follows from simple math – we are assuming that the typo was the point of confusion.**

b. Line 120: replace physics with physical

**Done.**

c. Line 285: predictive capability best tested with cross-validation on new dataset

**Yes, future predictions are best described with a new dataset, though the term “prediction” is commonly used in statistics to describe the ability of independent variables to explain the variance in a dependent variable. It is very common to use all available data to give the most robust predictive equation for future use. We have found this usage in nearly all the papers we picked at random from our library that use similar regression models (Wohl and Jaeger 2009, Whipple and Tucker 1999, Creed and Beall 2009, Richardson et al 2009, Martin and Benda 2001 to name a few). We do not wish to withhold a large amount of our data for validation testing, but to use it all for equation development, as is commonly done in many studies. Future LW studies that follow a similar design can test our predictive equation.**

d. Line 290: 'overbank' is the incorrect term if the boundary was the wetted margin of the channel at baseflow discharge instead of the topographic break at the actual stream bank. This raises a potential problem with your inferences. Need to reframe as comparison between inside-baseflow and outside-baseflow channel area.

**We appreciate this point, and note in the manuscript that language has been a point of difficulty for LW studies. We have decided to change the term in the manuscript to “out-of-channel LW” to more accurately reflect what was measured. We also provide more text explaining the problems with mapping bankfull area at sites spread across the watershed scale, including bedrock, boulder, mixed alluvium, and fully alluvial sites with wide differences in appearance and condition. We accept this as a practical limitation, yet we also feel that our effort to map the entire stream corridor width is very important compared to past studies that did not span whole watersheds or look for wood outside the channel.**

Line 379: Explain discrepancy with Lines 359-361.

**Line 379 refers to the effect of local percent shrub cover in the MLR models, while lines 359-361 refer to categorical differences based on subbasin. There is no discrepancy.**

f. Lines 430-432: Apparent discrepancy with Lines 398-401

**Lines 430-432 compare overbank LW volume per channel length and total LW volume per channel length. Lines 398-401 report results of in-channel LW volume per channel length. There is no discrepancy.**

g. Line 490-491: This is likely true, but this study did not examine floodplains, per se.

**This study included floodplains as part of the whole stream corridor area. The inclusion of floodplains was central to our methods, analysis, and discussion of our results.**

h. Line 495; 619-623: Erroneous conclusions of prior studies may have resulted from a) limiting survey extents to the active channel, particularly in higher-order streams with floodplains, and b) failing to sample larger streams in longitudinal analyses of patterns.

**We agree with point (a) of this comment, and it is reflected in the manuscript (lines 619-621). While point (b) is certainly possible, it is not true for all studies, and is certainly not as important as the simple mathematical mistakes made by many studies (described in lines 53-61). We feel that the most important shortcomings of previous studies are outlined clearly in the manuscript.**
i. Lines 503-506: Seemed that the MLR models were thrown out owing to violations of assumptions, yet they are still used for drawing inferences here.

**Line 501 clearly states that we are referring to the MLR model for total LW volume per channel length, which was valid. The MLR model for in-channel LW volume per channel length was thrown out.**

j. Lines 507-509; 522-526; 537-540: This is a tenuous approach, I believe, because in MLR, you can't determine the absolute contributions of individual predictors - can't 'unscramble the omelet' - because of multicorrelation between predictors. Safer to refer to their combined effect.

**We agree that you cannot “unscramble the omelet” – we also feel strongly that we should discuss what is in the omelet. In each case, we are careful to clarify that each variable has a certain effect when it is combined with the other significant variables. “x was significant when combined with the other four variables” for the first two examples, and “x was a significant contributor to the model” for the third. This is a valid and common way to discuss MLR model results. Thus, our approach is technically sound and we were careful in reporting what we did and found**

k. Line 553: 563-564: Probably the wrong interpretation here. A more likely explanation is that this pattern results from wood removal and channelization.

**We think that it is unfair to call our interpretation “probably wrong” as there is no evidence or oral history of streamwood removal or channelization in the Yuba watershed. Perhaps the reviewer is more familiar with lowland areas prone to clearing for navigation.**

l. Line 603: This may not be representative of areas burned more recently (e.g., 10-15 years BP). See literature on CWD dynamics after forest fires.

**The fires that burned in the Yuba River watershed within the past 10-15 years would not generate variables with enough variance to predict LW volume per channel length. In other words, there were mostly small, scattered fires. We felt that percent contributing stream cells that passed through an area that burned in the past 50 years was a reasonable variable that had the best chance to relate fire history to LW volume per channel length levels in streams that drain the area.**

7. Citations
a. Latterell et al. 2006 was another paper that recently demonstrated floodplains may contain high concentrations of fluvially-derived large wood. Could add this reference to bolster your case if you wish.

**Excellent resource. A reference to this paper has been added.**

8. Overlap
a. Not a problem to my knowledge.