Associate Editor Synthesis of Review Comments on the Revised Proposal Titled "Riparian Characteristics and Shade Response Experimental Research Study -- Draft Study Design"

December 20, 2021

The reviewers were asked to assess whether the concerns they identified on the original version of the proposal were addressed in the revised proposal. Reviewers 1 and 2 (R1 and R2) provided overall responses to this question with a focus on higher level concerns. Reviewer 3 (R3) provided responses to each specific comment in the spreadsheet containing the review matrix.

All reviewers found the revised proposal to be clearer and easier to follow. However, the reviewers identified issues that should be addressed, as summarized in part A, below. My recommendations follow in part B.

A. Overview of reviewers' comments

The most substantive of the reviewers' comments generally fall into three groups: (1) definition and calculation of shade (R1, R2, R3), (2) acquisition of the hemispherical photographs (R1, R2), and (3) study design and analysis (R3).

1. Definition and calculation of shade

(a) R1 and R2 identified a lack of clarity around the definition of the term "shade." As R2 pointed out, this lack of clarity is partially due to an apparent error in the definition of Global Site Factor (GSF) as provided by Roon et al. (2021). From my reading of the Delta-T Hemiview manual, it appears that the software first computes time series of clear-sky global solar radiation ("global" means direct plus diffuse) reaching the ground under the canopy for a specified period, then integrates the time series. I imagine that the computations follow an approach similar to that applied by Moore et al. (2005, Hydrol. Proc., Eq. A1). This quantity is then divided by the time-integrated global solar radiation at an open site to generate an index that ranges from 0 (complete shading) to 1 (no shading).

(b) R3 identified several occasions for which there is a lack of clarity about whether shade or the change in shade was the response variable in the predictive models.

2. Acquisition of the hemispherical photographs

This is the most problematic aspect of the proposed study. As pointed out by R1 and R2, below, the authors have not addressed the comments related to acquisition of the hemispherical photographs in a fully satisfactory manner.

R1 comment. Because hemispheric photography in stream channels is difficult and time-consuming, it is important to conduct some prototype studies and refine the methods from the experience. The authors still don't have a plan for prototyping and testing their photography plan. I infer that none of them have actually taken hemispheric photography photos in a stream under a forest, and they don't appreciate all the problems that go with it. You might have a plan to take photos at set distances, but when you get there, there is a wood jam, or a deep pool, or low woody vegetation, or something else, and you have to set the camera somewhere else. How do you randomize that choice and avoid implicit bias in location? It takes a while to get the camera set up at each spot and pack it up safely for moving to the next spot. The whole process is slow. And then there is weather. If you have a partly cloudy sky, or too much cloud cover, or rain, your photos aren't going to work for estimating canopy cover. It rains or is very cloudy much of the time in western Washington. Photos are better in flat light of morning or evening. Consistently good photography days will be limited to the period from July 5th to mid-September. They investigators really need to do some trial runs and revise the plan based on what they find out. They also need to evaluate the 7.5 foot spacing relative to canopy gap widths and gap frequencies in channels. 7.5 feet is very close together for hemispheric photography, and I'm not sure that a small number of photographs taken at the density will provide a good sample of reach-scale canopy cover. I would also suggest that they scale the distances to the channel width, e.g. take a photo every channel width. This will scale the spacing to the size of the channel. Furthermore, five photographs is not a lot to capture the variability of canopy cover on a forested stream.

R2 comment. I am still concerned that the limited number of hemiphotos used in the analysis will be adequate to capture the variations in stream shade. In the comment matrix it was noted that a statistician with expertise in these types of studies was consulted which is somewhat reassuring. It would however be much more convincing if the specific quantitative details of the guidance were provided in the document (possibly as an appendix) or if this aspect of the study design were supported with specific data either from an existing or pilot study. It is therefore not possible to critically assess whether this aspect of the study design is adequate to address the study objectives.

3. Study design and analysis

Most of the comments on these aspects were provided by R3. That reviewer is mainly satisfied with the authors' responses. However, R3 made some comments that the authors need to consider as the study proceeds and the data are analyzed; see highlighted rows in the spreadsheet that contains the responses.

In addition to the responses by R3, I have a comment related to the normality of the shade values. In response to the reviewer's comment on this point, the authors stated that:

> The text states, "Shade values will <u>not</u> be normally distributed; however, the differences in shade values will be approximately normally distributed." The text has been changed

such that a mean, not median, of the five shade measurements will be taken for each plot. Means are normally distributed and differences in means are normally distributed.

I would note that, according to the central limit theorem, the sampling distribution of a mean approaches a normal distribution as the sample size increases. If the population being sampled is highly skewed, a sample size of five may be insufficient to reduce the non-normality of either an individual mean value of shade or the difference between two mean values to a level sufficient to meet the assumptions of a statistical technique.

4. Other comments

Reviewer 1 provided the following additional comments.

(1a) The proposal is written as if hemispheric photography is easy and quick. In my previous review, I made the point that hemispheric photography is a real pain in the neck, and most people are going to do something simpler, and that it is important to compare hemispheric photography results with simpler techniques like densiometer measurements. This would be a simple way to make the study more useful, but the authors also ignored this suggestion.

(1b) It's OK that they don't want to run shade models, but other people are going to want to use their data to test shade models, so they should make sure to collect all the data needed for common shade models. It would be a lost opportunity if such data weren't collected.

Reviewer 2 provided the following additional comment.

(2a) L1254: Does this really mean to read ">" in this sentence? It seems that "<" would be more appropriate. Please also provide a justification for why this site constraint is included. In the response matrix it says that sites will not be excluded based on overstorey density thresholds so this addition seems in conflict with the response matrix.

B. AE recommendations

1. Definition and calculation of shade

As an overarching comment, I recommend that the authors use the term "shade" in a generic sense and adopt the use of a more specific variable name to describe the index of shade that is actually used in the analysis. These terms should be clearly defined near the beginning of the proposal and used in a consistent manner throughout.

Specific recommendations following from the comments summarized in part A1, above, are provided below.

(a) I believe that GSF as described in the Hemiview manual is the relevant index to use in the calculation of shade rather than an index representing canopy cover or gap fraction. As

recommended by R2, the authors should review the Hemiview manual to clarify their understanding of how GSF is calculated, and to ensure that the analyses and interpretations are consistent with the physical meaning of GSF.

(b) The authors will need to clarify the definition of the response variable for each analysis – i.e., whether it is the actual value of "shade" or the change in shade associated with the treatment.

2. Acquisition of the hemispherical photographs

I agree with R1 and R2 that five photographs per plot may be insufficient to capture the range of variability along a stream reach. In addition, taking the photographs at a fixed distance from the bank would not capture across-stream variability, and could introduce bias, at least for "wider" streams.

I do not necessarily agree with R1's recommendation to scale the longitudinal sampling distance to the stream width, especially for narrow streams and small sample sizes.

I have the following two recommendations:

> (a) Ideally, a pilot study should be conducted to document the spatial variability of shade both along and across a reach as a basis for designing a sampling scheme to generate unbiased estimates that have sufficient precision for the study purposes.
>
> (b) The study should include sampling at multiple locations across the stream in addition to along the stream, at least for "wider" streams, as indicated by the analysis of results from a pilot study.

3. Study design and analysis

Most of the comments on these aspects were provided by R3. That reviewer is mainly satisfied with the authors' responses. However, R3 made some comments that the authors need to consider as the study is implemented and the data analysed; see highlighted rows in the spreadsheet.

The authors will need to acknowledge that, due to the small sample size, the mean shade values and their differences will very likely not follow a normal distribution. The authors should ensure that their analysis is robust to deviations from normality in the response variable, if normality is indeed an underlying assumption. However, as R3 pointed out, the key assumption underlying linear mixed effects models is that the error terms are normally distributed. This assumption will need to be checked as part of the analysis and appropriate remedial measures implemented if the assumption is clearly violated.

4. Other comments

(1a) I recommend that the authors include measurements using a densiometer to allow an evaluation of the potential utility of a simpler and less costly methodology for shade evaluation in an operational context. These measurements could perhaps be made at a subset of the sample sites.

(1b) This study will involve a major investment of funds and human resources, and it makes sense to make additional measurements that will maximize the utility of the data that are collected. Therefore, while I appreciate that while testing shade models may be outside the scope of the proposed study, I recommend that the field crews collect the data that would allow shade models to be tested in follow-up analyses, as long as those additional measurements do not compromise the collection of the core data.

(2a) I found it difficult to understand the logic underlying this assumption and the threshold chosen for "taking action," which I presume from the following sentence refers to inclusion of an interaction term in the analysis. The authors need to clarify the criteria for site selection to ensure that they are consistent with the intended analysis. It is unclear why this assumption is numbered "1" when it appears to be the only assumption listed.