

From: em.fws.0.5f524a.5fd4dbf4@editorialmanager.com on behalf of [Freshwater Science](#)
To: [Marcus Beck](#)
Subject: Freshwater Science Submission MS# 2018118
Date: Saturday, November 17, 2018 10:56:32 AM
Attachments: [Beck et al 2018118 Freshwater Science Reviewer1.pdf](#)

CC: fws.editor@freshwater-science.org, dcarlisle@usgs.gov

Ref.: Ms. No. 2018118

Prioritizing management goals for stream biological integrity within the developed landscape context
Freshwater Science

Dear Marcus Beck,

We have recently received the reviews of your article Prioritizing management goals for stream biological integrity within the developed landscape context. These reviews, attached below, indicate that your paper has the potential to be of interest to our readers, but substantial work remains to be done before we can consider publishing the piece.

Should you further develop and clarify the ideas presented in the article, we would be happy to take another look at it. We would treat this revision as a new submission, subject to further review.

To submit a revised version of your article, please go to <https://fws.editorialmanager.com/> and log in as an Author. You will need to submit your revision as a new manuscript, and it will be assigned a new tracking number. It will be assigned to the Associate Editor who handled the original submission.

With best wishes,

Daren Carlisle
Associate Editor

Freshwater Science

ASSOCIATE EDITOR COMMENTS

I believe the manuscript is of potential interest to many Freshwater Science readers. I like the fact that the paper is more than just another model that predicts biological outcomes using GIS data. If that were the case, I could not recommend the paper for publication. Reviewer 1 thought the manuscript suffers from trying to do too much, and recommended refocusing the paper to either the web tool or the model. Although I agree with Reviewer 1 that both sub-topics need more explanation, having both a model and its application in a single package is what, in my opinion, makes this paper unique.

These strengths notwithstanding, the manuscript suffers from substantial inadequacies. Both reviewers (as did I) had difficulty grasping your objectives and following the methods. Both had difficulty with your interpretations and characterization of how the model (and associated tool) were used in the case study. I describe several general concerns below, followed by detailed comments. Both reviewers also provide many comments that should improve the manuscript.

My decision is to invite the authors to resubmit the manuscript after making the revisions suggested in this review. Additional peer review is likely necessary to ensure the manuscript has been improved, particularly in clarity.

ASSOCIATE EDITOR GENERAL COMMENTS

The Introduction is too lengthy and most of the paragraphs need clarification. Reviewer 2 had the same concern. The first paragraph is about the need for context in interpreting biological assessment data and is rather succinct.

The second paragraph is less clear, but I think you're laying the foundation for the need for managers to prioritize restoration. But the paragraph doesn't strongly convey that need. Paragraph three is also a bit blurry. It starts with a focus on multiple stressors but appears instead to be making a case for using land use (isn't it really just land cover?) as predictors of biological condition. Again, the paragraph needs to be brought into sharper focus, although I'm not sure you need to justify using geospatial data to predict biological condition. So perhaps this paragraph can be eliminated outright. Most of the subsequent paragraphs also need more focus.

The Methods section needs substantial work. I don't think you need two long paragraphs about California's environment past and present. The reader immediately gets bogged down in extraneous details. More detail is needed about the design of the bioassessment data collection. If the bioassessment data was not collected from a random sample of the NHD population, then you need to discuss the implications of this on your predictions. The justification for your selection of GIS predictor variables is insufficient, as both reviewers also pointed out. Given the extent and severity of hydrological modification in the state, it seems strange that you would exclude this type of alteration from your model.

After reading through the paper carefully, it is still unclear to me how the model and associated web tool are supposed to be applied. A diagram may help here. Is the model supposed to provide context for actual CSCI scores from a sample? Or is it meant to provide a prediction of what's possible biologically at unsampled sites? Figure 6 shows a nice example of how the model puts actual sampled sites into "context," but it would be more helpful if it also illustrated what inferences were possible as a result of this added context. In your text about the case study, sometimes you discuss all the segments, other times you discuss CSCI scores—many of which were used in model development. Line 439 states that the model was used to help interpret CSCI scores for decision making. But throughout the Methods and Results sections you talk about segments, and draw attention to patterns in predictions of segments. Then in line 489 you say the stakeholder group focused on the entire landscape rather than individual sites. So I get confused about what the emphasis is of the model. I also agree with Reviewer 1 who has some pointed questions on how you are characterizing the SCAPE tool.

ASSOCIATE EDITOR SPECIFIC COMMENTS

Line 63: This sentence is unclear to me. How do site-specific conditions place limits on scales?

Line 67: This sentence is complicated and unclear. After reading subsequent sentences, it's still not clear how data accessibility is relevant to the topic of this paragraph, which appears to be the need for context.

Line 73: "...in need of some level of management." seems like you are advocating here, or at least assuming that government agencies have a recognized duty to actively manage waterways that are in biologically poor condition is a

Line 81: It's not clear why these modifications are "confounding" factors. The subsequent sentences suggest these types of modifications are often the basis for simply redefining management goals. So, perhaps all you need is to improve the connection between these ideas.

Line 87: This sentence is an abrupt transition from the preceding stream of thoughts.

Line 104: Will you be using information about historic landscape changes in your models? If not, I'm not sure why you'd use this statement as a topic sentence.

Line 106: Perhaps the sentence would be more clear if phrased as "...where landscape alteration may seriously limit opportunities to restore biological integrity."

Line 107: Your definition of "constrained" streams seems to have confused Reviewer 2. The other reviewer and I didn't have the same reaction, but consider possible alternatives that could be more clear (e.g., "restoration constrained" or "restoration limited"). Perhaps the reviewer's confusion was due to their familiarity with other definitions of "constrained" in the context of geomorphology.

Line 107: The definition is also a bit vague. How does “large scale” differ from “landscape alteration?” I know what you’re trying to say, I’m just suggesting you refine the wording.

Line 110: Can we assume that “constrain biology” has the same meaning as “constrain the stream?” This entire sentence is a bit vague. How does a stressor originate from scales? Perhaps an example would help?

Line 115: From this sentence to the end of the paragraph is in need of clarification.

Line 117: What is “...variation in bioassessment data?” and “...limits of bioassessment tools?”

Line 120: What is “bioassessment and management potential?”

Line 131: Here, I think you’re attempting to differentiate your effort from previous studies. This point is critically important given the number of previously published modeling papers, but is not given the emphasis it deserves.

Line 136: The sentence starts out great but then whimpers out in the end. May I suggest something like “...to classify and prioritize ecologically impaired stream segments based on the likelihood that restoration efforts would be successful given the degree of landscape alteration...” or something like that?

Lines 155-179: Condense this material to a single, shorter paragraph limited to information about California that is critical to the reader’s understanding of your methods.

Line 190: Is this the only information the reader gets about your GIS predictor variables? Shouldn’t you refer to Table 1 here?

Line 209: It is critical that you describe the design of the bioassessment data collection. Were sites selected to be representative of the entire stream network? Or were sites targeted?

Line 212: Didn’t you already say this?

Line 217: If I follow your argument here, you’re saying that you only selected predictor variables that represent human impacts that can’t be mitigated? I don’t believe this argument. If the argument is true, then why not include NPDES dischargers or major dams / reservoir storage, both of which are widely available GIS datasets and represent human impacts that don’t seem to go away? (with the exception of an occasional small dam)

Line 221: If you’re not going to include NPDES discharges and dams, then address the reasons why here.

Line 221-223: It’s not clear what this sentence means in the context of statistical modeling.

Line 223-226: Rephrase sentence to: “...human activity not related to the predictor variables used in the model...”

Line 239: Unclear what these percentiles refer to. It’s relatively easy to envision quantile regression with a single predictor variable, but most readers (myself included) will need more explanation of how the procedure works in the case of several predictors. Is there a separate RF model developed for each 5% increment of CSCI scores? If so, does each model use the full range of all predictor variables?

Line 244-247: The topic sentence makes no sense (typo?). It suggests that the full CSCI dataset was designed so as to be representative of the state’s regions, but this wasn’t explained earlier. But my reading of your methods here is that you randomly selected calibration and validation subsets from the full dataset—which doesn’t by itself guarantee representation of landscape gradients.

Line 247: What is meant by “this stratification?”

Line 262-276: This is a slightly better definition of “constrained” than was given in the Introduction, but still needs polish.

Figure 2: Needs clarification. The “Segment type” y-axis label is confusing. Aren’t these just examples of individual segments for which a prediction was made? Make it obvious that the symbols on the right-side panel are actual biological samples. Why use the term “Relative...?”

Line 274: Again, not clear what these percentiles refer to.

Line 281: “or certainty?” Not clear how you are defining this term and how it was quantified

Line 279-289: I understand, in general why you would want to do this. But why use CSCI thresholds that are essentially meaningless? Isn’t it enough to say that the 0.79 threshold has a clear statistical and management footing and just stick with that? Maybe I’m missing something here.

Line 291: Why do you need a “relative site score?” Can’t you just say you compared observed bioassessment data to the range of expected scores?

Line 305-309: Are these details needed? Did the stakeholders limit their deliberation to segments that had actual bioassessment samples? If not, this information on samples is extraneous.

Lines 409-415: It’s not clear to me how these results are relevant to the Case Study.

Line 437: Second half of this sentence is unclear. Also, even if there were few degraded streams in California, wouldn’t these two requirements still be true?

Line 454: What do you mean by “evaluate?”

Line 458-464: This example REALLY helps, but it needs more detailed explanation. If biological samples from a site revealed that the site met biological objectives, it isn’t intuitively obvious that it could be classified as “constrained.” Include in your example an ecologically plausible explanation of why this could be the case.

Line 481: Very awkward sentence.

Line 483: Seems like you’ve made this statement a couple times already.

Line 501: This sentence doesn’t quite make sense to me.

Reviewers' comments:

Reviewer #1: Review of MS2015088R1

Prioritizing management goals for stream biological integrity within the developed landscape context

Overall Comments for Authors:

This paper describes the application of a statistical learning method to model O/E index scores using a small set of landscape predictor variables to identify which samples may have predicted distributions below reference-based thresholds. The application of landscape predictors to bioassessment scores is not novel and can, indeed, trace its history back to the earliest applications of GIS (E.g., Richard et al. 1997 and Johnson et al. 1997 initial work on biota, water chemistry and habitat and landscape predictors from 20 years ago and all the many studies that

followed) and little of this original literature is acknowledged although Allan et al. 1997 (part of that same group that Richards, Johnson, and Host were in), was cited. I think the authors could do more to place their work in the science of landscape modeling of stream characteristics and that would set up the novel application of this quantile regression forests to this end. Right now the paper sits between one that is very management oriented and one that is technique development oriented. As such, I think it under-develops each and it might benefit from deciding to be one or the other. For example, I think the SGR and SCAPE tool are interesting, but they are underdeveloped and more of a distraction from the technical tool description, in my opinion. The details of the SGR exercise are not described in much detail. The SCAPE tool development, testing, and application is also under-described, in my opinion. Likewise, the quantile regression tree modeling is under-described. It is my limited understanding that these methods could choose a wide range of predictors that might be ranked in importance, etc. Why did the authors choose the predictors they did? Which were most influential? Were other combinations considered? Why or why not? Did predictors change by quantile? I think more of the technical details of that model selection, application and testing would be interesting to the readers of this journal - who tend towards the technical rather than management side - and many of whom may not have encountered this method before. I think you could remove the whole application story to the SGR and the SCAPE tool and save that for a second article for a management journal (that could get into more detail on the management implications of your predictors) and instead focus on the technical development of the constraint modeling tool - especially what other approaches might have been tried and rejected. I think the readers of this journal would want to know why you might have rejected a lot of other methods so they could learn from that journey rather than just be told this was the one model and these the only predictors used.

In my opinion, as it stands, this is a very fine and interesting piece of work and I think the writing was generally good (with a few missing words and some sections that seem to vacillate and could benefit from some clearer messaging, see below specific comments). However, I think it is more suited to a management oriented water resources journal at this time because of the reasons stated above. It reads a bit like a technical report that was developed into an article rather than an original scientific contribution to the field of stream ecology developed for that purpose alone. Again, that is not to say this is not publication worthy. I think it definitely needs to be read by the resource management community. I am just not sure FS is the appropriate venue right now (nor one that has that target audience) and I hope that is taken in the constructive light in which it is intended.

Specific comments follow by line number:

Line number Comment

28 Factors constrain in-stream conditions.

37 ...achieve goals for biological integrity...

Intro first paragraph I think the first paragraph has to have the argument laid out a bit more clearly.

63 You mention site specific characteristics, but you are using landscape predictors.

79-80 May want to reference the national rivers and streams restoration synthesis papers of Bernhardt, Palmer and others on how poorly we understand restoration success

83/84 This is only CA centric - and this sentence is very vague.

85/86 This is allowed by the CWA regulation - so cite to the regulation.

87-89 Why is this a priority? This sentence comes at the end of potential arguments for it - so put it up front.

94 Have not has

95-98 I would have expected to see the King and Baker and Cuffney and Qian back and forth arguments on urban thresholds for biointegrity cited here.

102 There are a lot more papers on landscape predictors being used to predict stream condition, dating back 20 years. The question is, what can we do about landscape predictors?? We really can't manage many of them...easily. This needs to be woven in to the introduction a little - that there are limits to what landscape level predictors can help us with at the scale we really can manage streams - reach and segment.

104-107 I agree with you, but you have not spelled out why for the reader. Why can these help prioritize actions? For me, it is because local govts have limited budgets and have to decide where they think they can affect the greatest improvements in condition OR protect the most quality waters from degradation. And see comment above on whether "landscape alteration" is a reality. I don't think it is.

107-109 Not sure even with abundant resources - one may need limitless resources to change "urban land use"... this, again, gets to the issue of how manageable your endpoints are. They are informative, but I would like to have seen an attempt to see if local, reach scale predictors - any - actually mattered. Moreover, we don't know anything about the ranking of these predictors you've used - which are most important? This matters, if riparian zone forest

or road crossings are MOST influential, then maybe these are manageable - at least more than something as nebulous as "Urban land use". This is part and parcel of where I think this paper can flex more technical muscle - this is the type of investigative detail I think the FS audience would expect and benefit from.

109-111 I think this can be placed elsewhere in this paragraph

120-121 "management potential"? You model really focuses on bioassessment score potential not management potential. You have not measured or included management practices as predictors - so management potential only in a speculative sense.

122-123 Here is where you need to cite all those papers who have shown that modeling tools for understanding biocondition along landscape gradients HAVE provided support for management. We have 20+ years of this. Hynes paradigmatic paper "The stream and its valley" - prophetic for most stream ecologists - was about this very theme. So, this is not new.

139 What different priorities?

144 I think 4 is too much for this paper. I would go back and spend more time on development of 1, research and compare with other approaches, and also detail more of the effects of 3. I also think there is not much or anything in this paper on comparing management decisions or options.

152 Overall I think the introduction is a little lengthy and need to focus on the key arguments and specific narrative, which I think can be tailored to a more technical and less management focus if the desire is to publish in this journal.

168-169 Awkward

170-171 Remove showing and remove have

173 Decide on hydrologic or hydrological; insert pre-development after match

176-177 Cite the source of these regionalizations. I think it is PSA?

185-186 StreamCat makes these estimates, not you. So, I think rewrite that StreamCat provides estimates of and use at the....

189 Insert NHD+ between each and segment

193 Insert least disturbed between under and reference and cite to Stoddard et al. 2006

201 I would rewrite as "...and was used herein as a desired target condition". You're not really managing....

206-208 Did sample date not figure into selection? Wouldn't you want samples from a segment closest to StreamCat data?

212 Add "...using StreamCat predictors/" to end of first sentence.

212-214 This sentence is redundant with previous text

216-219 So, how do you figure these are beyond the scope of management and why talk about the value of this model for management if the predictors are unmanageable? Wouldn't we be more interested in constraint that is actually manageable? And aren't landscape interventions like LID/GI/riparian management designed to manage these factors? I think you need to explain this more if you believe it is true. Also, what was the reasoning or testing of alternative predictors to come to this list? Was there any analysis to identify how many of these were necessary to the model performance? Model selection needs a bit more development, imo.

219-221 Awkward construct. IS it biological condition that can vary through time or potential effects?

244 Missing word between gradients and major

264-267 Again, I agree with you but you need to explain more of what you mean here. Give some examples or provide citations to support it.

271 For "most" analyses? Which ones did not use 0.79?

278 Remove parentheses

286-289 I am not clear how this is sensitivity analysis. There is no "truth" (ie typical sensitivity analysis is how a change in a predictor changes calibration success or respond precision;) and it is not clear why the percent in each class is the right response for this.

297 Not sure a management case study is the right topic for FS.

357-367 I admit I am surprised the model performs this poorly given your predictors; I would have expected far better prediction than 50% r-squares. That is a lot of O/E score not being resolved to advocate for this being an accurate prediction of constraint. I think more discussion of the implications of this error need to be made. There must be error around the percentiles then as well - so how off are the ranges?!?

Also, from my calculations, 0.85 O/E is your inflection where the predicted to observed agree. Below 0.85, the model predicts a lower value than observed - or more constraint than actually exists. If I have this correct, this needs to be explained to readers since it is actually happening where the threshold is.

372-373 Do you have large river bioassessment models? I thought they were for wadeables?

374-375 Isn't this self revelatory - I mean you are only using urban and agricultural predictors, which you admit. SO, this should not be a surprise, right?

382-385 And this is consistent with the model bias (slope and intercept), correct?

387-400 This is where I think you can develop a lot more on model selection and variable selection. How parsimonious was your model? Could you get comparable results with just urban land cover? OR urban and agricultural land cover? OR just road crossings....

401 Again, I think this case study is underdeveloped and not really appropriate for the FS audience.

443-447 You seem to be vacillating. I'd stick to just a support tool...

447 I am not sure you can say anything about likelihood of achieving management goals. Your model really doesn't apply to where management is applied or has been applied - it is a largescale landscape model. Can you really say that a GI intervention in a subwatershed that reduces stormflow might not improve biotic condition marginally or reduce stress somewhat? Not sure that is the scale of this work. Moreover, not sure using an "exploratory" tool for management is the right language to use...at least not to managers.

448 This is a definite path to pursue - this tool supports TALU and supporting prioritization among watersheds given current knowledge.

451-453 You should read the Soranno et al. 2010 bioscience paper and Soranno et al. 2008 L&O paper. There may be a lot of ideas in there that would help here - or at least support.

452 I am not sure that selecting appropriate management actions required considering all of these things. Some management actions (e.g., listing, TMDLs) are not dependent on this.

455 - Not sure the CWA allows this sort of flexibility. States can't really choose to do nothing.

461-464 Not considering a site is not an option. What you are doing is providing a tool that helps support TALU structure - which is what should happen for many sites, rather than a life in TMDL purgatory.

472-474 That is TALU - get rid of one size thresholding. This tool will help.

483-485 So now you are vacillating - either advocate for a regulatory support role for this tool or not. In the paragraphs above, you walk right up to saying this tool can support regulatory actions; but here you demur. I say if you're gonna go out on a limb, don't start chopping....

486 UAA's require TALU unless you are saying the ALU should be removed all together, which I do not think you are. Urban streams support ALU - just not a reference one. But if all you have is one ALU, a UAA can only remove it until you establish alternative ALUs, like in TALU.

505 I agree - and this is TALU and regulatory...so it is a regulatory support tool!

523 It'd be nice to back this statement up with data on engineered channels and surrounding land use - I am sure you have the data.

532 What about your predictors individually? Are there sites with high road density that have good biology, for example?

551-553 Did you test that here? If not, why not given it worked for others?

566-568 Again, this suggests more work should have been done or could have been done to address these issues and to explain the predictor selection more.

602-603 Consistently indicated engineered channels were constrained? Where are those data? Did I miss them? Apologies if I did.

605-606 Why the conjecture on habitat limitation and channel modification? Can you cite something or provide supporting data? This has become a paper about modified channels a lot and away from its core focus on constraint in general.

608-609 Doesn't this contradict earlier statements?

628-630 I don't think you can say anything about restoration potential. You can say better biological conditions are unlikely given the current landscape condition. But, you've not really tested whether restoration can help any of these streams or not. How long have we been doing restoration in any of these watersheds, at what scale, and what is the expected scale of recovery? But, you don't really have any predictors on restoration to test this.

Reviewer #2: Review of Freshwater Science "Prioritizing management goals for stream biological integrity within the developed landscape context" by Beck et al. Reviewer 2

Overall, the manuscript addresses an important and practical management question/application and is a strong regional case study with extensive data. However, the Introduction does not clearly present the goals and objectives and needs extensive restructuring and editing (at least it didn't speak to me, perhaps I'm not getting something here). This is key, lines 214-219 give the objectives but it's not clear why these landscape variables were selected and more importantly that they provide a good connection to CSCI, nor why this allows the researchers to make the

connection to constrained vs unconstrained. Because of the problems with the Introduction and lack of clear objectives and appropriate theoretical connections, I had issues with the whole manuscript. Then at the very end of the paper a key statement was found - lines 612-613 "Biologically constrained sites were considered those where present landscapes were likely to limit CSCI scores that describe macroinvertebrate condition." This makes a big difference and helps define what is the modeling goal, that constrained is not necessarily urban or Ag structures but is an indication that IF the land use type, amount and management of the land use doesn't change then the CSCI scores are limited and therefore constrained. I can't emphasize this more, this needs to be better described and the connections made stronger in the Intro. This changes everything.

I've made extensive comments within the PDF document, again, many of these can be addressed with a restructuring and more clear and direct Introduction that clearly defines the outcome of the modeling effort -- what is meant by constrained vs unconstrained and how there is a connection between the landscape GIS variables and the CSCI. Together this will make the paper more direct, clear and stronger. I think the procedures, analyses and overall goals are sound and important.

I hope my comments are not taken as too harsh, but because I was missing these points at the beginning many of the statements did not make sense to me. NOTE: all comments are given with the heading "Reviewer 1" please shift your to Reviewer 2.

There is additional documentation related to this decision letter. To access the file(s), please click the link below. You may also login to the system and click the 'View Attachments' link in the Action column.

<https://fws.editorialmanager.com/l.asp?i=92125&l=NKJCFTDM>

In compliance with data protection regulations, please contact the publication office if you would like to have your personal information removed from the database.