March 29, 2020

Dear Editors,

We thank the reviewers and editors for their critical evaluation of the manuscript. Your suggestions proved essential in improving the quality analyses and prose throughout the revision.

In summary, we updated the analytical methods which entailed substantial rewrites of the methods, results, and discussion sections. The revised methods and results provide a clearer picture of the patterns observed along the environmental gradient. In rewriting the discussion, we focused on community-level trends and avoided overreaching claims of causality. The updated R code and supplemental tables were updated and can be found in the appropriate repositories. Due to the extent of revision, we did not track changes in Microsoft Word from the original submission.

In the following, we have compiled the line edits and general comments of the editor and three reviewers (Italicized and colored). We then responded directly to each (normalized font and no background color). We believe the revised manuscript is suitable for publication in PeerJ.

Sean Kinard

PhD Candidate

(804) 684-7309

Virginia Institute of Marine Science

1375 Greate Rd, Gloucester Point, VA 23062

Andrews Hall 314

[skkinard@vims.edu](mailto:skkinard@vims.edu)

*The editor and several reviewers requested major revisions to the methods:*

*Methods are inadequate: Linear regression is misused and NMDS should be replaced by canonical ordination*

* 1. *NMDS insufficient due to instability and mean-variance weaknesses (Legendre 2012, Warton et al 2012).*
  2. *The dissimilary coefficient was not reported for NMDS.*
  3. *Consider the use of data transformation on compositional data such as Hellinger distance recommended Legendre & Gallagher 2001.*
  4. *Consider Model-based ordinations in gllvm and HMSC R packages.*
  5. *Consider visualization via model-based ordination: ecoCopula*

The analysis has been substantially revised to address concerns voiced by the editor and several reviewers. PCA and ordination figures now include proportion of variance explained as well as associated tables with relevant outputs. Regression tables are reported within the text and multiple regressions predicting diversity have associated tables and clearer textual reporting based on AICc and delta values (Burnham and Anderson 2002). All multivariate methods were updated to use statistical terminology before citing the relevant R-package and function names.

First, we put more thought into variable-selection and reduced the number of predictors from 26 to 7 so that there are less predictors than sample sites (10). We prioritized retaining at least 1 variable within the following categories: climate, water quality, flow, geomorphology, and riparian. We also included justifications for the inclusion of each variable in the “Environmental Data” portion of the Methods section.

Second, we included multivariate PCA of environmental predictors in addition to linear regressions to display patterns between regional, local, and climate predictors in this previously un-surveyed study region.

Third, we removed the exploratory NMDS plots in favor of Hellinger-transformed Redundancy Analysis (Legendre et al 2018) of compositional patterns constrained to our environmental gradient of interest (precipitation). This enhanced the precision of the ordination plots (which are no longer rank-based) and better displayed biological patterns in relation to the precipitation gradient.

*Editor: There is no clear integration of biodiversity metrics in the results and discussion*

Following the rework of analyses and figures, the results were rewritten. Fewer environmental predictors and visual improvements in figures present the relevant patterns more clearly. The original results presented disjointed analyses as a textual list. The new results section presents a leaner summary of biota patterns (diversity and compositional), how they relate to the climate gradient, and any other significant relationships with environmental predictors. Lastly, we created and cited tables to be placed within the text with relevant results instead of listing coefficients and test statistics within the main textual body.

**Reviewer 1:**

*24. why "also"?*

Removed

*43. "changes in water temperature "*

Added “changes in …”

*93. "for" evaluating.*

Added

*95. replace "to climate" by "the climate"*

Replaced

*125. "was"*

*Changed ‘is’ to ‘was’*

138. "were"

Changed ‘are’ to ‘were’

*150. "field rinsed"???*

Removed ‘field rinsed’ and included a more detailed description of the invertebrate sampling methods

*195. "we need to know if this was "all" environmental variables or "each" environmental variables*

This section was rewritten and distinguishes which variables are used in the various analyses.

*219. "positively related with precipitation"*

Since the analyses were revised, this section was rewritten with more neutral language regarding the relationships between environmental predictors and patterns in biota.

*240. "included"*

This section was removed and an appendix was referenced

*286. "Pisidium" in italics*

This section was removed and an appendix was referenced

*290. Oksanen (2013) is not in the references. I guess you mean Oksanen et al. (2019)*

Thanks. All VEGAN citations were updated to Oksanen et al. (2019)

*223-230. Long description*

Agreed, we condensed the variable descriptions and reorganized for improved readability

*239-250. Simplify*

We removed the lists and replaced this section with a summary of the ordination and mention several rarer taxa found in the wetter regions.

*251-260. Integrate with 239-250*

See above comment

*378-384. Long sentence 381-385*

While I prefer brevity, including species characters provides useful context for this species’ ruggedness. Most of the length is parenthetical so a cursory reader should not be bogged down.

*404-406. Necessary?*

This ending was weak. So we have rewritten the discussion with a greater emphasis on broader community patterns. We included some of the sampling shortcomings of this exploratory survey and how our future direction with this research. Lastly, we make some broad inferences using a space for time substitution in the final paragraph to provide the closing narrative of the manuscript.

*Why no trait analysis?*

We have incorporated some sampling context throughout the discussion. Namely, kick-net sampling in slow sandy streams misses a significant proportion of available habitat. The outlier at Tranquitas and the small sample size subverted our confidence in running and interpreting a functional trait analysis.

*409. Acknowledgements look strange*

We added the reasons for which each individual was acknowledged.

*Figure 1 has Figure 2 caption and vice versa*

Thanks. We made new figures and made sure to label them correctly.

*442. Are references complete?*

Thank you, in our revision process, the citation editor dropped some citations. We have double checked all citations in the final draft.

*Consider reference: DOI: 10.1111/j.1365-2486.2007.01375.x*

Thanks for the suggestion.

*Consider reference: DOI: 10.1007/s10750-012-1244-4*

Thanks for the suggestion.

*47-49: Detecting causal relationships with correlations is too ambitious*

Our original introduction was too ambitious for an exploratory study of this scale. We avoid deducing causation in the final revision.

*314-317: correlation is not causality. No evidence for precipitation having direct relationships with conductivity and nutrients. Following sentence indicates 1 biased site*

This section has been rewritten with emphasis on relationships between variables and their implications for biota.

*324-327. Not convincing. Can similar compositional shifts occur in other environmental situations?*

We retracted our overzealous claims and replaced them with a plausible explanation and a call for continual surveys to evaluated seasonal variation.

*258. Why are there the same numbers for different taxa at different sites in the appendix?*

We attribute identical numbers resulting from the following process: Invertebrate samples were picked to a count of at least 300, In cases without 300, the sample was picked to completion. The sum of individuals in each taxon were multiplied by the fraction of unpicked sample and reported as abundance of individuals per square meter. We are confident that the relative abundances of various taxa at each location are conserved. But we refrained from utilizing abundance analyses due to this sampling artifact.

*The paper would benefit from a model of cascading effects*

Agreed, we need more weight of sampling power in the form of replicates and events.

*Which community processes are addressed (dispersion? Reproduction?)*

The revised discussion addresses this question. To summarize, the limited sample size limited our investigation to identifying broad environmental filters and potential biotic interactions along the gradient. Fish communities were likely restricted in arid sites by water quality (and potentially sedimentation) based on the present study. We posited that biotic interactions played a larger role in humid sites, based on community compositional patterns of fish and invertebrates. Disperal and reproduction will be investigated once we obtain a robust sample size through continued sampling efforts to conduct functional trait analysis.

*Main concern: few sites 10 cannot capture all the natural variability of aquatic biota even if the other factors are controlled*

In Frank Harrell’s Regression Modeling Strategies (2015), detecting relationships using regression require a minimum of 10-20 observations. Our sample size meets the minimum requirement, albeit barely. In our revision, we reduced the number of variables used in regression analyses with biodiversity.

*Authors should use corrected p-values to account for false discovery rate*

We substantially overhauled our regression analysis from including over 20 variables to 7. We have also updated our reported tables for multivariate methods.

*Recommend multivariate regression with no more than 3 uncorrelated environmental drivers*

By reducing to seven predictors, we didn’t run into issues with overfitting. Our maximum VIF scores for fish and invertebrates were 14 and 7 respectively.

*Consider table of relationships between environmental variables to substantiate mechanistic explanations for changes in biological patterns*

Great suggestion, we included a table with gradient features and a table containing the seven environmental predictors used to evaluate flow, water quality, and local habitat.

*Why isn’t disturbance mentioned in the introduction*

Good point, we incorporate 2 variables to evaluate variation in flow to approximate flood and drought disturbance. We revised our second paragraph to more thoroughly orient the reader to the various mechanisms in which precipitation interacts with stream community processes.

**Reviewer 2:**

*The analyses cannot fully show what the authors claim during discussion.*

We removed declarations of causality and focused instead on how the observed patterns have implications for bottom-up and top-down controls along the environmental gradient.

*The discussion itself is too focused on subjective links between species and habitat characteristics.*

In our rewrite, we discuss broad community characteristics with more emphasis on placement within current aquatic community literature. For example, we discuss the effects of ambient salinity and NH4+ on fish communities within existing literature rather than apply deductive lines of reasoning on based on several species and their traits.

*Recommend talking about communities in general terms and making inferences about future impacts of climate change*

Thanks, see the above comment.

*Recommend: intermittent stream literature: T. Datry, N.Bonada*

Thanks, I came across several of their papers in the literature review.

*1. North-American?*

This was removed. The description of the study region within the abstract is now limited to “the Texas Coastal Prairie”.

*5. This may make sense for North Americans, but not everyone*

We removed this from the abstract.

*6. Replace measured with gathered?*

Thanks for the suggestion, we prefer to use of ‘measured’ to describe the various collections performed.

*10. Low flow pulse percent is undefined*

We included a brief description in parentheses for LFPP in the abstract to aid cursory comprehension.

*11. It isn’t clear how bottom-up and top-down controls are at work here*

We immediately follow this statement with observed patterns that support either hypothesis across the gradient.

*43. Replace ‘changes’ with ‘changes in’*

Done.

*59. Define hierarchical?*

While ‘hierarchical’ appears like jargon, it refers to the organization of successive filters elaborated in the citation and throughout the ensuing paragraph.

*68-71. This last sentence is disjointed from the paragraph. You are talking about abiotic and biotic drivers and clse with spatial scales and biogeography*

Thanks, we removed it.

*80. Not only biogeographical studies. The whole metacomuunity and priority effect studies also consider these processes*

True, but we are trying to constrain the flow of ideas towards observational studies at this point in the introduction.

*88-89. Nice*

*94: When using ecotone, include phytophysiognomies. Recommend explaining the area in terms of geology, geomorphology, and general geographic terms.*

Great suggestion, I have included a more thorough setting description to help the reader familiarize themselves with the study region.

*108. Please define USGS*

Thanks, suggestions like this help the paper reach a broader audience. I included “(U.S. Geological Survey)”

*119: The last 2 paragraphs are mixed with methods. Recommend focusing on the idea and general hypothesis without focusing too much in the region. You have an outdoor natural experiment with a good range of precipitation with few confounding effects over a small spatial extent. Otherwise, I have the impression that you are only giving part of the information*

We reorganized the last 2 paragraphs to improve the separation of these two topics: Efficacy of the study region and our hypotheses.

*136-137. So, every stream were sampled according to the stream width, or do you consider the value of 4.1 x 25 for all streams?*

The previous language was too casual. We clarified this with “Each reach length was 25 times the average stream width”

*138. Explain low variation. Avoid subjective adjectives*

We removed this statement which read more like a vague results statement than methods, which were taken from U.S. EPA recommended protocols.

*139. Improve clarity changing subjective adjective ‘high success’*

See the comment above.

*194-195. Linear regression is generally used among variables with causal links. Shouldn't you use a simple correlation matrix? Or a PCA?*

We revised our analyses in several ways; we reduced the number of environmental predictors, performed PCA, included a multiple regression, and changed to canonical ordination with one axis fitted to precipitation.

*196-197. Several linear regressions or a multiple regression?*

We ran linear regressions to test our hypotheses and ran multiple regression to examine gradient patterns more properly.

*196-197. Have you tested for correlations among the variables?*

Yes, we included tests for autocorrelation among environmental predictors in Table 3.

*201. Indicate the dissimilarity metric and if data were transformed*

Thanks, we revised ordination methods to “Redundancy Analysis (RDA), constrained by precipitation, on Hellinger-transformed community data for each group”.

*204-206. Is clustering related to NMDS? Explain the aim of each analysis*

We removed cluster analysis, colored sites by precipitation and examined species and predictors using vector analysis instead.

*204-206 Indicate how you standardized data to make fair comparisons among variables*

We revised ordination methods to RDA, constrained by precipitation, on Hellinger-transformed community data.

*207. Refer to the statistical analysis before followed by the R function that performed it.*

We reorganized this method description with the R function listed last.

*Explain the decision for alpha = 0.10*

In our revised methods, we reverted to the conventional alpha = 0.05.

*212: Why not using direct examination of the relationship among multiple variables and a multivariate response matircs, like RDA, and db-RDA?*

Great suggestion. We revised to canonical ordination as suggested.

*222: Define acronyms in each section*

Done.

*230. This info is hard to follow and summarize. What is the point of the regressions? A PCA would potentially summarize it better*

This section was rewritten with emphasis on organization by topic and then analytical method.

*232-238. Linearly and perfectly? You can show it in scatterplots*

Linear relationships are more clearly written with references to scatterplots with regression lines in Fig. 3 and Fig. 4.

*232-238. Why some relationships are demonstrated with linear models, while others only the ranges are shown?*

These were removed and a supplemental table is referenced.

*239. This info is hard to follow. Perhaps a general rank-abundance using all data would be nice to see abundance pattern at the total spatial extent*

By revising methods to RDA with vectors, descriptions of the analyses naturally described the visible patterns more clearly.

*262-270. This is confusing. You are mixing general biotic results with regressions and then get back again to general biotic results. I would describe the general patterns of richenss, diversity, and taxonomic identity first. Then, I would describe correlations with environmental variables.*

We reorganized this section to describe sampling scope, diversity regressions, and compositional patterns.

*271-272. A rank abundance plot would do the job. You can have several colors of bars for each taxonomic group. Also, you can have one rank-abundance for each of the three climate clusters. Similarly, an old but gold Simper analysis would be informative in terms of assemblage changes with climatic changes.*

Thanks for the suggestion, while this would be a great tool in data exploration, our goal was to use regression analysis in conjunction with the natural precipitation gradient which is statistically more robust and more broadly applicable than comparisons among the rank abundance curves for several clustered groups.

*302. Remove ‘extreme’?*

We considered other terms here like “exemplary” or “extraordinary” but felt that it resulted in the discussion beginning with a conceited tone. The use of “extreme” could be misconstrued as sensational, but our intention was to include an enthusiastic reference to the ‘exceptionally great degree’ of the gradient.

*311-317. Nice*

*327-328. Can you explain the mechanisms to which precipitation drives increased diversity? Does it increase habitat stability (perennial streams)? Or is it related to stream heterogeneity (mor allochthonous material and nutrients)?*

The sampling sites across the region are perennial. The revised discussion contends that aridity imposes abiotic filters via seasonal drought effects and that humidity increases habitat stability and connectivity for in-stream biota.

*331-345. This paragraph is not well connected to the central topic of the study*

This paragraph is off-topic, so we removed it in the rewrite.

*350. High-flow?*

We didn’t explicitly include a high-flow hydrological measure in the revised analysis. We did however include a measure of flashiness which has similar implications for habitat stability.

*354-355. A lot of discussion around taxonomic composition, but we cannot see these results in the NMDS. Perhaps a simpler Simper or IndVal would help substantiate this claim*

We revised the discussion to constrain our topics obvious patterns in the RDA.

*366-368. Perhaps too speculative*

This was removed in the rewrite.

*373-375. I would fit this discussion considering that you found strong compositional changes, so you have losses and gains in invertebrates, but mainly losses in fishes. To properly tackle this point, I recommend you discriminate compositional changes in terms of turnover and nestedness components.*

Although we didn’t use quantitative methods to ascertain turnover and nestedness with our limited dataset, we discussed the patterns involving gains and losses of species observed in the ordinations.

*377. Write our Relative Humidity to improve readability*

This variable was removed in the revision since it is derived from annual precipitation.

*395-406: Too speculative with weak empirical evidence. I would not conduct the discussion in this direction. You could focus on the differences among biological groups. Also, I would like to see a caveat paragraph. You have an interesting natural experiment, but we would like to know what the challenges were when analyzing your data*

In our revision, we reorganized these topics with a greater emphasis on discussing them as possibilities indicated by the current body of literature rather than trends that our data empirically suggests. Most difficulties arose from the paucity of sample events in this data set which is mentioned towards the end of the revised discussion.

*407. Missing conclusion to summarize results*

The revised discussion includes a summary and the broader implications of the study’s findings.

*409. Include why we are thanking*

We included why they were thanked.

*Figure 1. Wrong caption*

We fixed table and figure references and captions.

*Figure 2. Include p values and R square within plots*

We added R-squared values to revised figures.

*Figure 2. Improve organization of rows and columns*

Table readability was improved by increasing spacing and including row shading.

*Figure 3. Do you have R-square?*

We added R-squared values to revised figures.

*Figure 3. Are the taxa distributed randomly in the figure?*

Revised figures include species vectors and captions that indicate their proximity to their ordinated location.

*Figure 3. Can’t you correlate taxa, sites, and environmental variables using RDA?In this way you would see which sites (type of climate) were correlated to which environmental variables, and which species were associated to which env. Variable*

We used RDA in our revised analysis. Thanks!

**Reviewer 3:**

*Introduction lacks context*

We revised our establish a call to action based on gaps in existing literature regarding climate change and stream community assembly processes which leads to our space for time substitution along the Texas Coastal Prairie to determine precipitation-based influences on stream community assembly.

*Improve hypothesis clarity*

The introduction was reorganized to separate objectives and hypotheses from the study region descriptions.

*Problems with methods*

We revised our analyses, methodological syntax and organization.

*Lack of references*

Thanks for catching this. Zotero citation manager dropped references from the bibliography throughout our editing process. We double checked all of the citations. Now, the revised manuscript includes about 66 cited references in comparison with the original submission’s 35.

*Unsuitable analyses*

We revised our analyses in several ways; we reduced the number of environmental predictors, performed PCA, included a multiple regression, and changed to RDA constrained to precipitation with fitted species and predictor vectors.

*Vague conclusions unconnected to results*

Revised analytical methods revealed clearer patterns and were easier to discuss. We removed overreaching claims of causality, unrelated sections, and focused on relating our findings to our original hypotheses.

*Title. it is in stream? Lake? Ponds? Maybe add something about the climate (arid streams or semiarid streams).*

Thanks, we including streams makes the title more precise which should prove useful for other researcher literature searches. The revised title is, “Effects of a natural precipitation gradient on fish and macroinvertebrate assemblages in coastal streams”.

*18. environmental conditions?*

This section was rewritten with an expedited transition from climate-change processes to stream-related topics.

*25. no numbers in abstract (let it to the results section)*

We removed quantitative results from the abstract as recommended.

*36. maybe a more recent reference is necessary here.*

We retained this citation and added a more contemporary article.

*43-45. A reference is necessary.*

This point was made in the discussion with the included reference. Thanks!

*104 . there is no need to speak of the places’ name only as landscapes they identify. makes the text more scientific.*

We included a more thorough setting description to help the reader familiarize themselves with the ecology of the study region.

*112 a 119. authors who approach work done with fish and invertebrates in the region. they mention only here in the forecasts. to make the reader understand why these predictions, the authors need to deepen the introduction by providing more examples of how these communities reflect changes in precipitation.*

This study conducts the first bioassessments of stream biota in this region. So, we established a framework for discussing abiotic and biotic filtering processes in community assembly, followed up by several expectations based on existing literature and precipitation-driven patterns that can be observed at a glance within the region.

*148. for how long the person who collected did this process?*

We included “(15 s duration)”.

*Line. 153. why 300 individuals? based on any reference?*

We used rarefaction curves on different individual counts. In the revision, we included a reference to this process and the unpublished data from our lab that was used.

*191. Why Shannon? Simpson could be better*

We used Shannon-Weiner index because Simpson diversity performs better with larger samples. Additionally, we wanted to avoid the smaller influence of rare species and large influence of common species on the Simpson index.

*186-188. in that part the letters are in uppercase and in this in lowercase. standardize.*

Thanks, we fixed capitalization for these terms.

*201. NMDS. Did the authors perform any randomization tests to verify whether the distribution is significant?*

We revised our analysis from NMDS to a constrained canonical ordination (RDA) using Hellinger transformed data.

*209. Based on any reference?*

In the revision, we used significant vectors with p < 0.05

*221. write out and in lowercase. Soil*

This variable was removed in the revised analysis.