Reviewer comment:  
  
  
EDITOR'S COMMENTS TO AUTHOR  
  
Editor: Chapman, Daniel  
Comments to the Author:  
Thank you for submitting to JBI. Two expert reviewers have assessed the manuscript. Both found it a potentially valuable contribution. However, they also raised some questions about the statistical analysis, which will require consideration and potentially some reanalysis. I also had some questions about the analysis, as follows.

I would like clarification of whether transformations to the response variables in your mixed models was applied (e.g. log transform to prevent residual heteroscedasticity and prevent predicted values going below 0) and whether a generalised model (GLMM) might have been more appropriate (e.g. Poisson or negative binomial for richness).

Did you test for interaction terms between survey type and the human development index, to test for variation in response between the habitats?

I also wonder whether the U-shaped relationship between human niche specialisation and human development might in some way be an artefact of the specialisation index being calculated in part from the development index, and it being capped at 0 and 100% (e.g. because species found near 0 or 100% can’t exceed those values and so appear more specialised than species found more often at 50%). Is there some way to test whether your procedure would result in this kind of relationship as an artefact (e.g. using simulations of a null model)?

As there are only 2 figures in the main text, you might consider presenting NMDS plots, which illustrate the clear difference between development levels (notwithstanding the reviewer’s suggestions to improve the analysis).  
  
Dr. Daniel Chapman  
  
--------------------------------------------  
REVIEWER COMMENTS TO AUTHOR  
  
Referee: 1  
  
Comments to the Author  
See attached.

This paper summarises an interesting dataset from the ABMI, investigating patterns pertaining to

human development, species richness, and community structure across wetland community types. I

think the approaches taken are generally plausible, and the results are not too surprising. However,

the Discussion section seems to be missing a part where the authors critically reflect on whether the

combination of their dataset with their chosen modelling strategy has any weaknesses with respect

to the ecological inferences drawn. At the moment the Discussion just takes the results and assumes

that they are a completely correct representation of reality, with no uncertainty attached to them.

This is a missed opportunity to reflect on the modelling approach taken relative to the data, and to

therefore make the paper more interesting and useful to other ecologists. Some points that might

be considered in such a section of the Discussion are given below in the more detailed comments

against manuscript line numbers.

## Specific comments

L94 Confusing, as introducing NNS can also increase local diversity, as you imply in the first

paragraph of the introduction. It’s obviously debatable that NNS always reduce local diversity in

terms of always being the proximal cause (as opposed to passengers of change responding to other

factors).

L123 This sentence should end with “in our study area” to avoid over-reaching.

L154 I think referring to unique visits in time and space as “sites”, which normally implies a location

in space only, is the opposite of clear. Why not refer to them as “samples”?

L165 What’s the lower bound on the area sample for the wetland plots?

L216 The intermediate level here seems rather arbitrary. Looking at Fig 1 there seems to be a good

spread of samples between 25 and 75%, and one can well imagine that at the scale of the remote

sensing data, a site with 25% human-influenced land-cover could be quite similar to one with 75% at

a landscape scale – e.g. perhaps a wetland site is in a similar location to another with respect to

human development, but it just happens that the location of your SPOT6 polygon includes a corner

of some development in one instance, and most of it in another. At the very least it seems to me

that this choice of 45-55% could do with a sensivity analysis, particularly as your choice results in a

small intermediate sample, and PERMANOVA is known to be liable to confounding changes in

variances with changes in locations (i.e. means). (See David Warton’s papers on mvabund as a more

robust alternative to PERMANOVA in some situations).

L228 Including sampling methodology as a fixed effect: does this really deal adequately with the

different totals and types of areas sampled between the two methods? Terrestrial is 1 ha, wetland is

from some unknown lower bound up to 14 plots x 20m2 = 280 m2. Normally the comparison of

differently sized areas in terms of richness requires species-area curve adjustments (e.g. see Palmer

et al. 2008); now, I can see that a fixed effect allows for some constant difference in richness

between method types alongside adjusting for other variables in the model, and maybe this is

enough, but on the other hand your wetland sites apparently have transects that vary in length, so is

this fixed effect approach sufficient? Something for the Discussion.

L229 Leading on from the comment against L154, it’s worth noting here that you do not even follow

your own convention of calling samples “sites”.

L230-L235 You seem to approach the question of non-linearity from the point of view of an a priori

decision to compare a model with a polynomial to one without, but nowhere do you explicitly say

that this was a priori based on some literature evidence or hypothesis. This always makes me

suspicious that the decision was actually based on a lot of data dredging and model

experimentation, and subsequent lying by omission by not stating that this was actually what

happened. If it this modelling strategy was actually decided upon before looking at the data, please

state this clearly, otherwise be honest and state that it was a result of experimenting with the data,

and discuss what implications this has for your inferences (e.g. invalidity of claimed P-values).

L232 Leading on from this, if the approach wasn’t a priori, then why not also look at whether a cubic

term improves fit in terms of AIC. Or use AIC to choose the number of knots in a spline? Even if the

quadratic approach was determined a priori, looking at more flexible forms might reveal other

patterns worthy of discussion in an exploratory mode. I think the AIC-based approach to spline knots

is available in Frank Harrell’s “Hmisc” R package (which accompanies his Regression Modelling

Strategies book). Also, is AIC appropriate, or should AICc have been used?

L283 The first sentence of the Discussion illustrates my point that you are lacking some critical

discussion of your own methods and the potential influences of these on your ecological inferences.

For example, you say here “we found maximum richness in wetlands surrounded by intermediate

extents of human development… etc.”, and then proceed as if this has been proven beyond

reasonable doubt. But even your own fitted model in Fig 1a shows that this is barely true overall, or

for wetland sites in isolation, and this before even considering uncertainty in these estimates, or in

model selection (in terms of whether other plausible models could have been used – note that this is

not just about AIC-based model selection, which is itself dependent on certain assumptions about

the data and of course on the precise dataset in hand). To be honest, I find the whole discussion

rather dull, as it is just trotting out literature-based arguments, without any critical discussion about

what was actually done with this dataset in this paper.

#Section B

Given that there are only two figures in the paper, I don’t really see why so much has to be in

supplementary material. I would, for example, have thought that the NMDS plots were central

enough to actually be in the paper.

#Ref

Palmer, M.W., McGlinn, D.J. and Fridley, J.D., 2008. Artifacts and artifictions in biodiversity

research. *Folia Geobotanica*, *43*(3), pp.245-257.  
  
Referee: 2  
  
Comments to the Author  
This concise and very well-written manuscript describes a study using large and valuable environmental monitoring datasets to understand broad patterns in plant diversity in wetlands. The results are intuitive and interesting. However, bringing together different datasets does bring with it some issues, and I therefore have a few comments and concerns regarding how these datasets were treated in the analysis. I think that considering these issues is important for ensuring the robustness of the results.  
  
- First, there is a very large mismatch between the area covered by the inventories (10 000m2 in the terrestrial and up to 280m2 in the wetland protocols). I would be inclined to use species richness per unit area as the response variables in the models, as well as including protocol as a fixed effect because of other differences in habitat and e.g. the central area vs. transect methodologies.  
- The mismatches in sampling areas probably also make Jaccard distances quite unsuitable for use in the community analyses. Instead, it would be better to use Raup-Crick or Chao distances that try to account for missing species and uneven sampling. These can be implemented in the metaMDS function.  
- Again relating to the central area vs. transect surveys. Coming from a landscape ecology background, I can get annoyed by sometimes lazy referee questions of ‘why choose this size buffer?’, but in this case -and coming from a landscape ecology background- a 250m buffer seems very small to evaluate the effect of human development around a wetland. This is especially true when the terrestrial wetland surveys are already 100m across (if square) – were the 250 metres measured from the centre or the edge of the plot? In the wetland wetlands, the 250m buffer was measured from the edge of the open water zone, while the vegetation transects along the moisture gradient of up to 350m (14 plots \* 25m intervals) or 412 m (including a square plot of 20m2). It therefore seems likely that the transects overlap the buffer in the wetland sites, which would have further implications for including both terrestrial (effect of surrounding human development) and wetland (sometimes including areas of human development) inventories. Is this the case? Either way, I think it should be more clearly described.   
- Including duplicate and triplicate surveys of some of the sites seems to me to be unnecessary. Were the same plots visited again, or were they new plots/transects? Either way, having more than 1500 separate sites, it just seems unnecessary to keep the duplicates and make more complex models, rather than picking e.g. the first or most species rich survey, or the one that falls closest to the date of the explanatory variable.  
- I like the idea of including the effect of non-native species on increasing local richness. I just wonder, because of results later in the paper (Fig 2a) that proportion of non-natives is related to human development, if the correlation between the two might affect their suitability for both being included as-is in the same model.  
-Finally, I wonder about potential spatial autocorrelation. I do not know Alberta, but it seems large enough to have quite large variation in topography, land use, land-use history, soil, bedrock and climate. Therefore I would expect 1585 wetlands across the state to have a moderately large chance of having more similar vegetation to those that are relatively nearby than those that are far away, with differences controlled by more than just very local measures of human development. I strongly recommend that geographic location is controlled for in some way in the modelling.  
  
Smaller comments.  
  
Abstract. Perhaps a short sentence or short parenthesis about what is meant in this study by niche specialisation?  
  
  
Introduction: I like the introduction as it is, but if it were possible to slot in something about nonnative species and niches, it would round off the scope of the study nicely. Also, it would be good here to define fragmentation, as it every-increasingly seems to mean different things to different people (and different things to the same people, over time).  
  
101- myriad ways? I don’t think the ‘of’ is needed.  
  
Methods (or SI)  
It would be nice with a figure showing wetlands with high, intermediate and low human development. I think these things can mean very different things in different regions.  
  
Results.  
For models that don’t include polynomial fits (and because it doesn’t say that variables were scaled), it would be nice to include effect sizes in the results, with plain English explanations, for example ‘for each 10% of human development surrounding the wetland, the proportion of non-native species increased by NN%.  
  
  
SI line 61: It is very difficult to see the overlap between terrestrial and wetland wetlands, because clusters of wetland sites make it impossible to see if there are any terrestrial sites there.