

# Authors Response to Second Review

Our responses to the review are written in **blue**.

**Thank you. We appreciate the chance to do a further revision, and hope that these changes are satisfactory.**

## Response to Guest Editors

Address the remaining comments from both reviewers as far as possible. If something cannot be done easily, suggest this for future work.

- Add a color legend to Fig. 5 (and others - as needed).

**The color legend for Figure 5 was included. Also in the schematic diagrams, Figure 7 and Figure 11, we have included one color legend as reference for all the plots at bottom of each diagram. In Section 4 (Interactive Web App) there are some Figures without color legend because are images from the web app available at <https://natydasilva.shinyapps.io/shinyppforest/> and as I mentioned in previous review Class color legends were not included in the interactive web app because the interactivity allows you to get the class name when you mouse over each plot and will be redundant to include them on the shinyapp.**

- Enlarge all figures to full page width, e.g., Fig.12 **Done**

Please make a final review to detect typos, e.g., Line 18 on page 13. It seems to be that "... it ..." should be "... its ...". **Done**

Please review reference style. Carefully review and adjust all your references. There are numerous problems:

**The references were modified as you have suggested**

- Capitalize all journal names **Done**
- Add missing volume, issue, and page numbers **Done**
- Simplify the publisher info for Inselberg to "Springer, New York, NY" **Done**
- Extend "et-al" author lists and list all co-authors **Done**
- Fill in missing information, e.g., for Quach (what is this & where can it be obtained)
- Same for Sievert (2017) **Done**
- Include the location (city, state) for all books & book chapters **Done**
- Remove ISBN numbers and series information (e.g., useR) from books **Done**
- Capitalize R and other proper names **Done**
- Remove months (e.g., in Wickham 2009) **Done**
- Use a simplified URL for Puranen: [http://jse.amstat.org/jse\\_data\\_archive.htm](http://jse.amstat.org/jse_data_archive.htm) **Done**

## Response to Reviewer 2

The updated version of your paper addresses all of the concerns that I brought up in my original review. Both the added functionality of the package (e.g., flexibility in selecting tree nodes for exploration) and the explicit discussion of limitations will hopefully prove useful to the users of your package and readers of Computational Statistics. Moreover, the current version of the paper is cleaner and easier to read thanks to the edits and minor fixes (e.g., typos).

**We appreciate all of your feedback which help to improve our paper.**

The only remaining minor comment that I have is that the side-by-side jittered dotplot in Figure 7 seems to be based on the crabs dataset instead of the fishcatch dataset used in other plots in that figure.

**Yes you are right, we have changed the side-by-side jittered dotplot with fishcatch data and include a general color legend as reference.**

## Response to Reviewer 4

The paper looks better now. Regarding the points I raised in the comments about the dot plots with jittering across the x-axis when the x-axis is categorical, I still think you might want to consider an alternative. I think in many cases, the overplotting of the points makes the plots very cluttered, and in reality you only need to plot the individual points for a few observations. I understand that in this case there is the application of being able to select points and connect between plots. However, I think that this will mostly be of interest for the most extreme points. Here is a suggestion: perhaps in the future the default view (e.g. Fig 3, importance plot in Fig 10), you could use a violin plot with only the most extreme (lowest and/or highest quantiles, with a check to plot only the extreme values that would not coincide with a “fat” point of the violin plot) values plotted (see eg. <https://r-charts.com/distribution/violin-plot-points-ggplot2/>). There could be a button to toggle on/off the option to see all points as a scatterplot rather than violin plot. In the cases of say Figure 3 and 10 where you want to highlight one of the points in red corresponding to a selected tree, you could only plot that point (rather than all) in red (in addition to the outliers) when it is selected, but not all the time. Perhaps you could also do something similar for the side by side plots (e.g. Fig 8) where you have violin plots by color rather than the points (or at least have the option to do so if there are few colors).

**Thank you for your comments, based on your suggestions I have explored two alternatives: the violin plot and also a boxplot without whiskers (just the interquartile range) adding the points using transparency. At least for the case examples we are including in the paper the result was not clearly better For the fishcatch example the importance have an structure (groups of specific values for each variable ) and when we use violin plot that structure get lost. The best solution could be to include a button to toggle on/off the violin option or other, this will introduce an additional level of reactivity in the second tab of the application. The second tab will be more complex because there will be a lot of reactivities to control (error tree in the boxplot, node selection, importance variable values and finally plot options for the importance variable) and it is not clear how to order the reactivities sequence. This is something for sure to explore in the future. I have included a comment about alternative visualizations for importance plot after Figure 3.**