Referee comment on "Rainbow color map distorts and misleads research in hydrology – guidance for better visualizations and science communication" by Michael Stoelzle and Lina Stein, Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-118-RC2, 2021

General comments

This paper tackles the question of color palettes, which is important for decodability of figures by individuals with color vision deficiency and more generally, for scientific interpretability.

The authors find that "36% of publications in HESS in 2020 had visualisations that were not scientifically correct, not perceptually uniform, and difficult to access for around 4% of the readership due to color vision deficiency". I particularly like the way the authors evaluate reviewer comments, and highlight the lack of awareness among both reviewers and editors.

Overall, I fully support this paper. It raises an important issue; the analyses are novel and technically sound; the paper is well written, and it provides helpful solutions. Figure 7 is particularly useful as it highlights examples of "poor" and "good" color palettes clearly.

Specific comments

- The paper focusses mainly on color vision deficiency (CVD) and people with low/reduced vision, however one might also argue that good visualisation and labelling is equally important for people with other cognitive differences, such as (I'm guessing) dyslexia. Has there been research on this? If so, this aspect might be worth including in your literature review.
- The discussion of color palette type (negative-to-positive, strictly negative, or strictly positive) comes a little late in the manuscript (Figure 7c-d). It might be worth describing the type of color gradient that is most suited for negative-to-zero, negative-to-positive, and zero-to-positive scales sooner; e.g. a red-white-blue palette, which is currently missing from Figure 1.
- It might be helpful to provide the readers with a "checklist" of items to verify when creating a readable scientific figure (e.g. “the data-ink ratio”; “a white mid-point at zero for negative-to-positive palettes”).
- Some repetition could be avoided, e.g. section 3.4 also contains some repetitions about
CVD etc; perhaps it could be condensed a little.

- It was useful to read about the colorblind options in R packages. Are there similar options for Python users?
- I wondered if the paragraph about preprints (l.89-95) was really useful. It seemed to me this was a small sample compared with the analyses in subsequent paragraphs, so the utility wasn’t entirely clear.
- Lines 212-218 and elsewhere mention various types of visualisation (e.g. heatmaps at l.252), but it might be helpful to see examples (especially examples of good hydroclimatological visualisations).
- Finally, the title focuses on the hydrologic community but there were large parts of the text that were not specifically hydrological. Perhaps this could be strengthened a little. For example, Figure 1 could provide examples of hydroclimatic variables (i.e. highlighting which types of palettes are particularly suitable for specific variables).

**Technical corrections**

1.14 raise awareness how -> raise awareness of how

1.14 the rainbow color maps still is -> the rainbow color map still is

1.23 we sketch a way to improve the communication of rainbow flaws -> we outline an approach to ?? (unclear what is meant by ‘improve the communication of rainbow flaws’)

1.31 10 millions of unique -> 10 million unique

1.35 In terms of correct encoding, (comma needed for meaning)

1.36 “we are stronger in encoding..”: meaning could be clarified

1.42 “uses” -> “used”

1.47 the word “shares” is used instead of “percentages” (here and elsewhere); perhaps consider replacing for clarity

1.61 the term “perceptual uniform” needs to be explained. I would recommend replacing “perceptual uniform” with “perceptually uniform” throughout the paper. It is explained at line 309, but this comes too late.

1.87 notable -> notably

1.114 “a graph with two lines encoding continuous variables over time without any annotations... is classified as rainbow-related”: worth providing examples alongside A-D for clarity?

1.127 “a vision deficiency scale” – terminology could be clearer.

1.139 “two cross checks... led (not lead) to minor deviations”: if this information is included, then it might be worth specifying what “minor deviations” means and how many people are in the cross-checking and original reviewer teams.

1.149 It might be worth justifying the choice of journals – why were Sci Rep and NComms selected?

1.168 “a current redistribution of disappearing black and white papers into papers with and without color issues”. I think this means something like ‘coincidence between the decline
of black and white papers and the emergence of papers with color issues’

l.186 less -> fewer

l.190 73-92% of how many? A little unclear why two numbers here.

l.193 four “suggestions” perhaps

l.233 Scientists

l.237 “a pointedly use of” is unclear

l.246 “luminance” is unclear (also used elsewhere in manuscript). Does it mean transparency? Shading?

l.84 rise awareness -> raise

l.289 parts of science -> areas of science

Figures

Figure 1. “The same delta changes in values”: this could be rephrased for clarity; it is not entirely clear what the “+1” on the figure or in the caption refer to. Also, is “perceptual uniform” “perceptually uniform”?

By this point in the manuscript (line 70), I think if would be helpful to distinguish the colors used for scales that range from negative to positive (e.g. “red white blue”) and those that are “strictly positive” or “strictly negative”.

Figure 2. I wonder if it might make more sense to show the % of red-green or rainbow color maps as a fraction of the total number of papers (instead of just the papers with color issues).

Figure 4 seems clear to me.

Figure 5 is a little unclear. I wonder if examples (of alternatives, properties, tools etc.) could be provided for clarity?

Figure 6.

Dark background – is this supposed to be easier or harder to read?

Panel c. is it brightness or shades?

Panel i. beyond the graph title, good labelling can also be helpful. Historically, many journals have discouraged the use of labels on figures; but for some people, clear panel/facet labelling can help greatly. Perhaps this is worth a mention.

Also worth making sure that all panels are referred to in the main text.

Figure 7.

OrangeRed and Batlow are almost too small to read; would recommend deletion.

Do you mean “white strokes decrease the data-ink-ratio” (rather than increase)?
P.S. Is the correct technical term “color map”, “color scale”, “color palette”, or “color gradient”?