Reviewer’s comments to „Spatial distribution of oxygen-18 and deuterium in stream waters across the Japanese archipelago“, version 2, by Katsuyama et al.

Dear editor,

Dear Drs. Katsuyama, Yoshioka and Konihira,

with pleasure I reviewed the 2nd version of your abovementioned manuscript. This manuscript includes several substantial improvements compared to the first version, including the revision of a number of tables and figures and the addition of text to better explain the phenomena observed. As such, the manuscript now provides a fairly coherent view on stable water isotopes in Japanese stream waters, particularly the temporal component and relationships to precipitation are now much better linked to additional data. The inclusion of the raw data as annex is not only a very valuable contribution to the scientific community but also serves as a role model for other researchers, and is highly appreciated.

As a general comment, I wish to encourage the authors to screen their text for occurrences of ‘smaller’ in terms of isotopic delta values. I suggest to replace ‘smaller’ by ‘lower’ or ‘depleted’ (and vice versa in the other direction). This better fits the isotope terminology.

I’d like to make some comments on some paragraphs:

- Page 4, L 23 – grammar: change to ‘verified the in-existence of…’ (otherwise sounds like you searched explicitly for anthropogenic influences)
- Page 7, L8-12: I tried to cross-check data for each individual prefecture of these 2 regions (not shown) and it seems that the regression lines vary strongly even within each of these regions. In Chugoku region, e.g. slopes of 5.2, 2.3, and 3.8 (2x) were found. Slopes for some prefectures in Kinki region were as low as 1.8. Clark and Fritz (1997, p.43) show that the hypothetical slope of an evaporation line at 0% humidity was 3.9. I have the impression that in these two cases, the low intercepts are partly due to the small overall range (usually between -8 and -10‰ $\delta^{18}O$) but two effects may have entered as well:
    - Lower slopes are (though I don’t claim profound knowledge on geothermal features) often associated with hydrothermal alteration (slopes as low as 0, with various degrees of mixing with meteoric waters or incomplete exchange for slopes > 0). These waters may have entered into the dataset (based on the methodology which only explicitly excludes anthopogenically altered sites).
    - I have the impression that the forecast regions may be subject to dual-moisture-source problems – if the data is plotted for each prefecture individually, some prefectures seem to have their regression line start from LMWLs (albeit uncharacterized) of different intercept.
    - I suggest breaking up these two regions and screening for these effects, verifying whether my findings may hold true and if so, explaining them in the text. If quantitative data on hot spring occurrence is available on a prefecture level, I suggest to cross-check these and mention these occurrences in the text. To make it clear: A final solution of this question won’t be achieved with the dataset available

and is beyond the scope of the paper. But a more detailed explanation, or at least hints to possible causes would be desirable.

- Page 7, L15-17: This is a bit speculative. There are whole regions in which the LMWL differs substantially from the GMWL (e.g. the Mediterranean) – the LMWLs approaching the GMWL at national level cannot be presumed. Suggest re-phrasing: Your data is different from the GMWL but this is not unexpected because the latter is 'comprehensive' and 'global'.
- Page 10, L3: When describing the spatial distribution is 'robust', please give a statistical indication of this 'robustness' (e.g. a measure of the spatial autocorrelation).
- Page 13, L24-33: The isoscape models cited indeed rely mainly on GNIP data but each of them exploits spatial autocorrelation to make predictions especially for the data-deficient areas. Certainly, any regionally adjusted prediction model (e.g. Liu et al. 2008 for China) will be able to characterize local precipitation better than globally parameterized ones (Terzer et al. 2013); however this does not render globally adjusted models unable to adequately predict (within stated limitations). You may compare your results to the models you cite (though that likely goes beyond the scope of the manuscript); however I wouldn't dispute their validity and usefulness as such but rather underline how your work can fill their spatial gaps.
- Figure 4: I'd recommend to invert the colour scale (depleted – blue to enriched – red) to better reflect colour theory (warm colours – warm regions [and more enriched isotope values] whereas cool colours reflect cool regions and depleted isotope values).

After all, I'd like to thank you for the improvements you have already made to the manuscript and I'm looking forward to see it evolve further. Thanks for the interesting read.

References (only those not cited in the manuscript):

Liu, Z., Tian, L., Chai, X., and Yao, T. D.: A model-based determination of spatial variation of precipitation δ18O over China, Chem. Geol., 249, 203–212, 2008

Terzer, S., Wassenaar, L.I., Araguas-Araguas, L., and Aggarwal, P.K.: Global isoscapes for δ18O and δ2H in precipitation: improved prediction using regionalized climatic regression models, Hydrol. Earth Syst. Sci., 17, 4713–4728, 2013