Below are our responses to the reviewer’s comments, with their initial comments in black, our responses in red and quotes from the manuscript indented.

The manuscript is indeed very interesting and presents to my knowledge an innovative approach to GNSS-R. The authors use several low-cost antennas together in kind of an array. They use the well established inverse-modelling technique for their GNSS-R analyses. It appears that using several co-located antennas mounted vertically above each others significantly reduced the SNR noise and thus produces more precise water level results. I think the manuscript is already in a good shape and can be accepted after a minor revision.

We thank the reviewer for their positive feedback and thoughtful suggestions that have helped to improve the article.

I would like the authors to address a few issues in their minor revision:

1. The used term "array" implies a common analysis of the data received with the individual GNSS antennas. However, it seems that the four GNSS-R sensors are analyzed completely independently, i.e. not as an array, and then the B-spline coefficients are simply averaged. Please try additionally to analyse the four sensors in one combined inversion directly. Does this improve the performance even more.

The reviewer's understanding is correct that the data from each antenna is initially analysed independently and then the b-spline coefficients are averaged to produce the final time series. We have also tried, as the reviewer suggested, to do a combined inversion using more than one antenna. The results are similar when doing the inversion using more than one antenna compared to doing the inversion with each antenna separately and then averaging the b-spline coefficients. We have added the following text to the supplementary material section S1 where we discuss this combined inversion approach:

"Instead of using the methodology described in Section 2 in the main text for combining data from co-located antennas by averaging b-spline scaling factors, we also tried performing inverse analysis using data from all co-located antennas simultaneously. This combined inversion is performed by adapting equation (2) in the main text, such that $h$ is written in terms of the reflector height for a chosen reference antenna and the predetermined vertical separation between each antenna and the reference antenna. The RMSE increases from 1.02 to 1.24 cm at Trois Rivières when using the combined inversion approach with all four antennas as opposed to averaging b-spline scaling factors from different antennas. This result suggests that there is no advantage in performing a combined inversion using data from multiple co-located antennas."

And we reference this additional section in the main text in Section 2 of the main text:

"It is also possible to use SNR data from all four antennas simultaneously as part of the inverse modelling to retrieve a single set of b-spline scaling factors. However, as discussed in the supplement section S1, we found this approach to be less effective."
2. Concerning the temporal resolution, one question is whether the sampling needs to be synchronized, and/or whether there could be advantageous by purposely sampling at different epochs, in particular when doing a real combined analysis (see question above).

As explained in Section 6.5, we found a limit of 15 seconds for the temporal resolution, passed which (i.e., for greater resolution) results did not significantly improve. We have also found that there is no significant difference between taking a single SNR value every 15 seconds or taking an average of 15 values (one for every second) to produce a single value every 15 seconds. These results suggest that sampling SNR data at different times for different antennas is not likely to improve the results, although it may make the inverse analysis more efficient. The efficiency of our technique is not the primary motivation for this study but should be studied more in the future.

3. Is there any benefit in additionally installing a horizontal array, so kind of a cross installation with both vertical and horizontal extension?

We did not explicitly test a horizontal array, however as explained in Section 5.1, we analysed data from a four-day period where there were two vertical antenna arrays installed several meters apart at Trois-Rivières. Our results suggest that there is no clear advantage in using sets of antennas from both vertical arrays at the same time (i.e., spaced horizontally and vertically).

4. There is at least one typo that I found, line 167 "geodet-standard", but there might by further typos that I missed. So please check the manuscript carefully with a spell checker.

We thank the reviewer for pointing out the typo and have corrected it. The second reviewer noticed several small typos that we have also corrected.