Point-to-point actions in response to Anonymous Referee #1

We sincerely thank the Referee #1 for their thoughtful and constructive comments. We have revised the manuscript accordingly. The detailed point-by-point responses are provided below (blue), and the corresponding revisions in the manuscript are marked in red.

Response to the Reviewer 1:

The manuscript I reviewed is already a revised version. I like it very much. Not only is the research original, but the data products are unique and significant for various stakeholders. To my knowledge no one has published similar datasets in the literature. It's a pleasant read and I also learned from this reading. I have only a few editorial comments, which can be made during galley proof corrections.

Reply: We are very grateful that you took the time to review our manuscript, and sincerely thank your interest in our articles! Your valuable comments provided great assistance for the improvement of our articles. We addressed each comment as follows.

1) Fig. 1. Need to explain (a) in the caption. Also explain the meaning of "frequency of surface water occurrence" and the cross symbols in each panel.

Reply: Thank you for your professional suggestions. Firstly, we have added an explanation for Figure 1(a) in the figure caption. Secondly, we have explained that the frequency of surface water occurrence refers to the normalized count of pixels that were identified as water in Landsat imagery from 1984 to 2021. Finally, we have described the cross symbols in each panel, specifying that they represent field salinity sampling points.

(Page 5, lines 93-97) "Figure 1: (a) Spatial distribution of sampled lakes covered by the dataset, inlet rivers, and sub-basins. (b)-(l) Eleven study lakes in the IMXL from east to west and its surface water occurrence frequency (Pekel et al., 2016), including Lake Hulun, Lake Dalinor, Lake Chagannaoer, Lake Daihai, Lake Nanhaizi, Lake Hongjiannao, Lake Ulansuhai, Lake Juyan, Lake Ulungur, Lake Bosten, and Lake

Sayram. Cross symbols in each lake denote the field salinity sampling points. The frequency of surface water occurrence refers to the normalized count of pixels that were identified as water in Landsat imagery from 1984 to 2021."

2) Fig. 2 caption: Please add one sentence similar to this: "The details of individual steps and the meaning of the symbols are described below."

Reply: Thank you for your pointing this out. We have added it.

(Page 9, lines 181-183) "Figure 2: A brand-new lake water salinity estimation framework by a stacking salinity model, consisting of three steps: step one, data processing and feature construction; step two, model construction and ensemble; and step three, salinity estimation. Details of individual steps and the meaning of the symbols are described below."

3) Figs. 3 and 13. Change "Measure Salinity (ppt)" to "Measured Salinity (ppt)" in all x-axis labels.

Reply: Thank you for your professional suggestions. We have revised it.
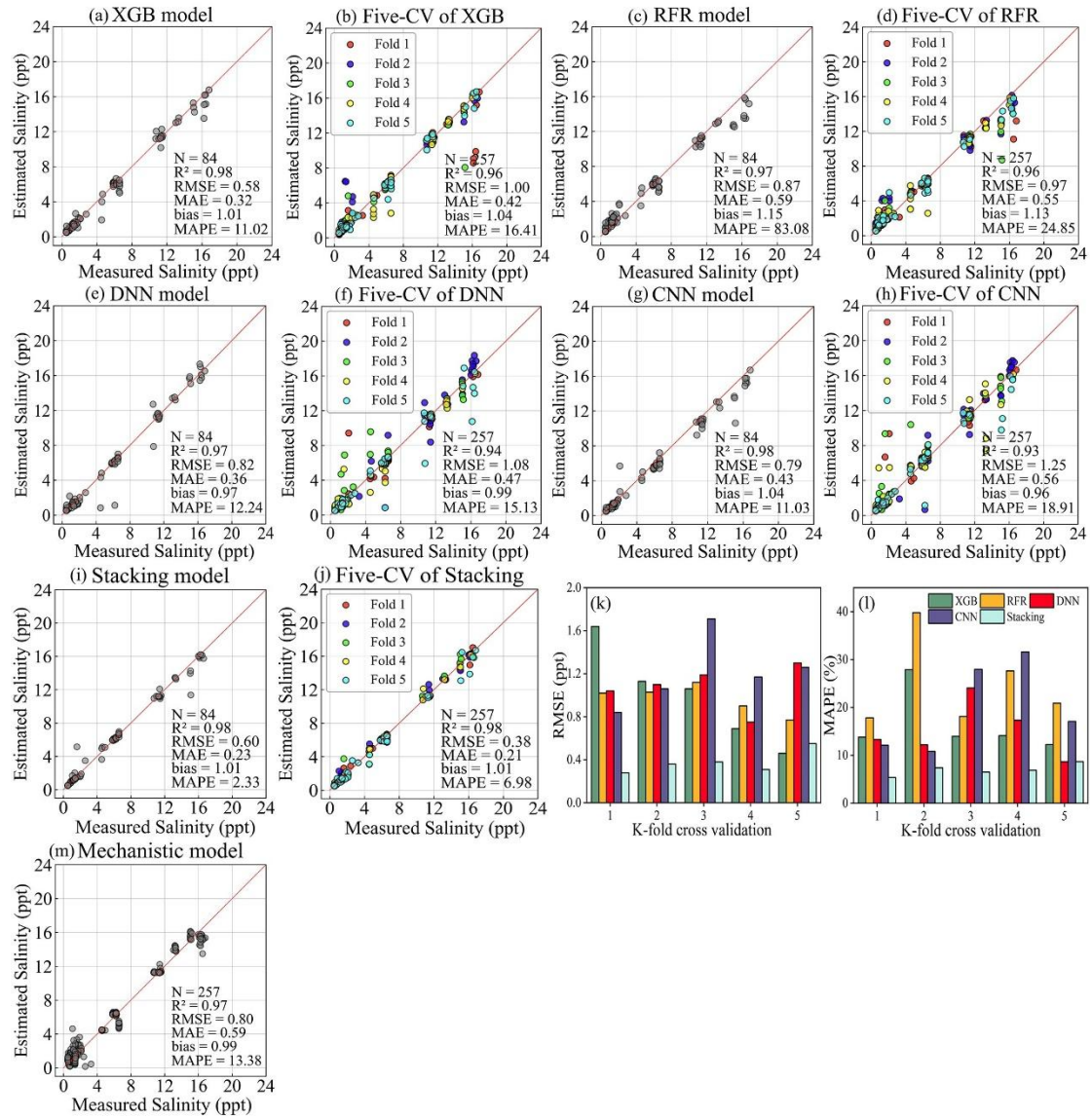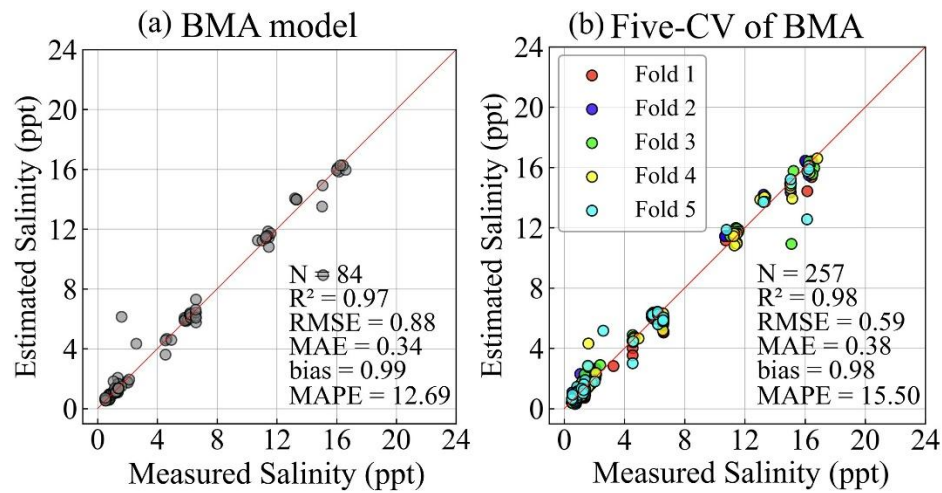**(Page 14, lines 298-301; Page 27, lines 498-499)**

Figure 3

Figure 13

4) Fig. 6. I like this figure as it makes visual inspection straightforward. The lower salinity around the land-lake boundaries may be due to runoff, but what caused the high salinity points (reddish colors) in c7?

Reply: We greatly appreciate your recognition, and thank you for your thoughtful comments. Through observation and analysis of MSI natural color composite images and salinity maps derived from various models, we found that the high salinity in the land-lake boundaries of Lake Bosten was caused by DNN and CNN overestimates of salinity influenced by mixed pixels. This high value was not fully corrected by the integrated model. Additionally, we supplemented the explanation for the high salinity pixels appearing in Figure 6(c7) in Section 3.2.
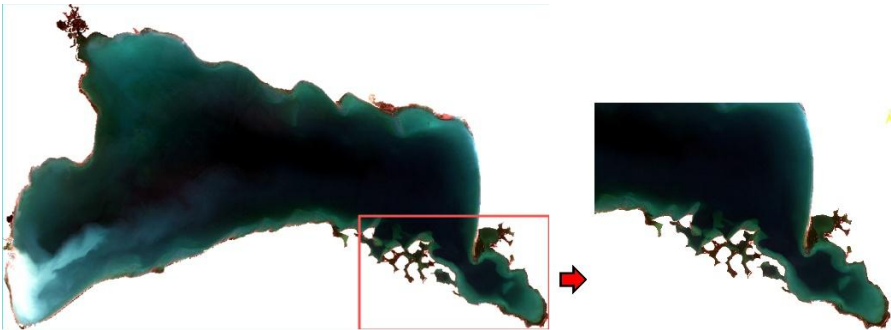
MSI natural color composite images
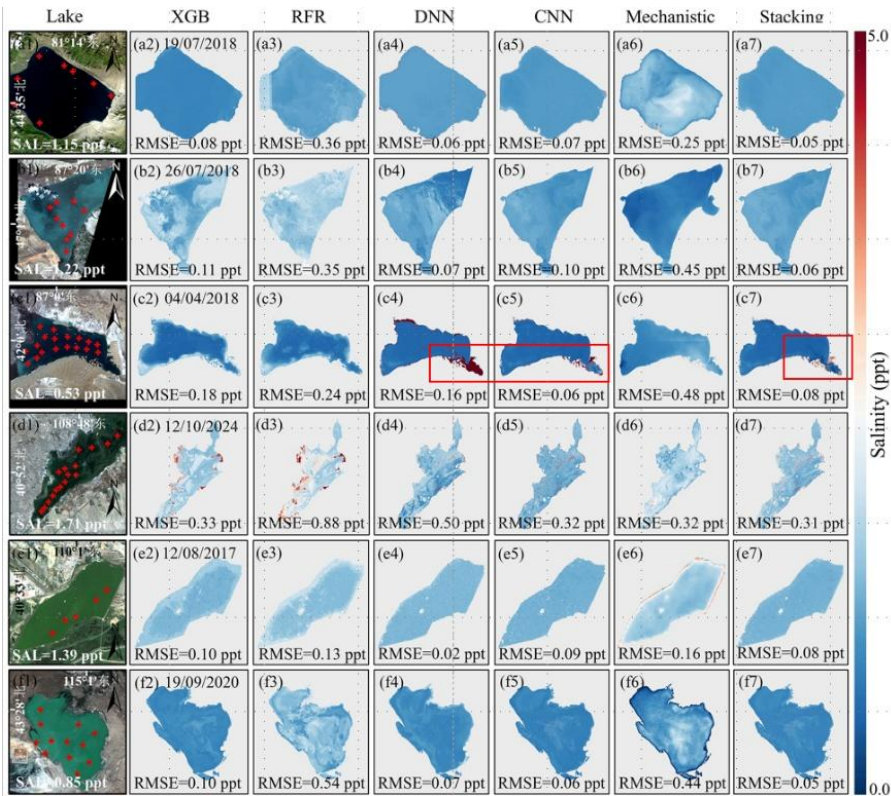
Salinity maps derived from different models

(Page 16, lines 314-316) "However, salinity values in Lake Bosten's nearshore pixels were not fully corrected because the DNN and CNN models overestimate salinity influenced by mixed pixels (Figure 6(c4,c5,c7))."

5) Fig. 11. Put some space between column 3 and column 4, and between column 6 and column 7.

Reply: Thank you for your pointing this out. We have revised it.
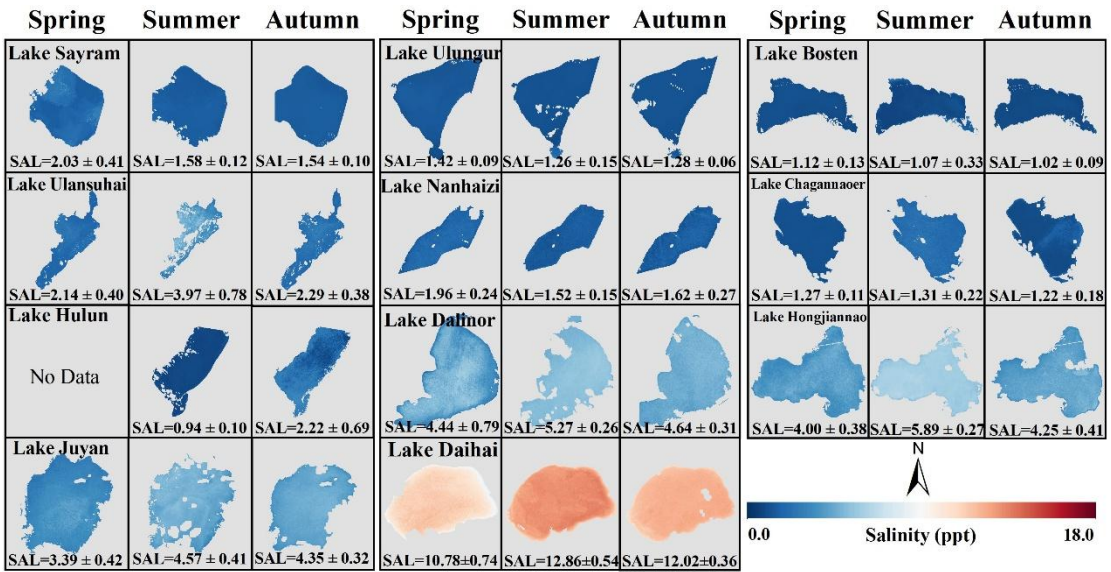
(Page 25, lines 404-406)

Figure 11

6) Eq. (8) – what's the unit of g? Need to list.

Reply: Thank you for your thoughtful comments. We have clearly specified that the units of integrated model estimates are aligned with the input data with units of ppt.

(Page 12, lines 257-258) "where g(E(Y)) represents the predicted salinity with a unit of ppt"