Review of: "A global monthly 3D-field of seawater pH over 3 decades: a machine learning approach" by G. Zhong et al.; submitted to Earth System Science Data

First of all, I would like to thank the authors for the careful and thoughtful responses to my comments and suggestions. I believe their revisions have significantly improved the manuscript, and the new details make it clearer to understand.

There are two minor comments that I would like the authors to address before publication.

Specific points to raise :

1. **Depth as a predictor:**

   I apologize for the confusion caused by my earlier comments where I used "pressure" instead of "depth.". I fully understand and agree with the authors' rationale for excluding pressure due to its high correlation with depth. However, my concern pertains to the use of depth as an input predictor, which is not applied consistently across all bioregions. From the authors' first response, I understand that depth is used as input predictor to estimate pH at specific levels (e.g., one of the 41 defined depth levels). However, I remain unclear how pH at different depths is estimated in certain bioregions where depth is not included as an input (particularly in the mixed layer).

   For example, in the Subpolar North Atlantic bioregion, pH in the mixed layer is estimated using predictors such as Phosphate, DO, Nmon, DIC, Sal, and Bathy, but depth is not explicitly included. None of these environmental predictors can fully substitute for depth. This issue also applies to the Equatorial Atlantic and Subtropical South Atlantic in the mixed layer. By contrast, for intermediate layers, this concern does not arise as depth is consistently included.

   The paragraph the authors added regarding longitude, latitude, and time being replaceable by other environmental variables is very useful and improves clarity in the text. However, this point does not address the specific issue of how pH can be accurately retrieved for different depths when depth is not used as an input predictor.

2. **Validation using BGC-Argo data:**
   Considering the spatial distribution of BGC-Argo data, which is concentrated mainly in the Southern Ocean, I think it would be valuable to include the number of points (or profiles) used to compute the biases presented in Table 5. This information would help clarify the representativeness of the validation results.