Reviewer # 1

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.

Response: Thanks for your suggestions, which help improve the quality of our manuscript. We have revised the manuscript according to the comments. The detailed changes are as the following responses.

Specific points to raise :

1. **Depth as a predictor:**

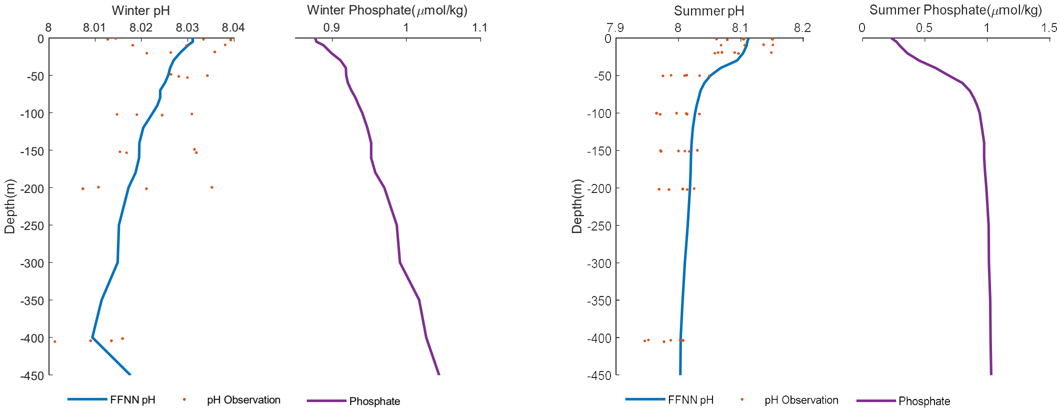
   I apologize for the confusion caused by my earlier comments where I used "pressure" instead of "depth.". I fully understand and agree with the authors' rationale for excluding pressure due to its high correlation with depth. However, my concern pertains to the use of depth as an input predictor, which 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, N_{mon}, 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.

   Reponse: In bioregions depth was not used as a predictor directly, pH at different depths was reconstructed based on input 3D field products of environmental variables used as predictors (e.g., nutrients, DO, TA, and DIC), which contain different values at the 41 defined depth levels. The vertical variability of these variables provided important information on how physical or biological conditions varied with depth. On this basis, the FFNN model learned how pH varied with depth from the vertical pattern of environmental variables used as predictors, which is notably similar to vertical pH pattern in bioregions depth was not used. The gridded seawater pH value at different depths was then accurately retrieved from the FFNN model and 3D field values of environmental variables used as predictors.

   For example, in the Irminger Sea station in the Subpolar North Atlantic bioregion, the phosphate concentration has a notably similar vertical pattern with seawater pH during different seasons (as shown in the following figure). Using phosphate and other 3D field variables (DO, DIC, and Sal) as predictors provided sufficient vertical distribution information in different seasons to reconstruct the vertical profile of seawater pH. Similarly, in the Equatorial Atlantic and Subtropical South Atlantic in the mixed layer, the FFNN model can also learn how pH varied with depths from the 3D fields of nitrate, silicate, and other environmental variables. However, the consistent vertical pattern only exists in specific regions. In the intermediate layers across wild depth ranges, more 3D fields of environmental variables and using depth directly were necessary to provide sufficient vertical distribution information of seawater pH. Therefore, only in certain bioregions in the mixed layers was not used as a predictor, and depth was consistently included in the intermediate layers.



For better clarity, the description about how pH can be accurately retrieved for different depths has been added in the section as the following:

"In addition, depth is important in reconstructing the vertical pH distribution. However, it was not used as a predictor in certain regions of the mixed layer due to the notable similarity between the vertical pattern of pH and particular environmental variables used as predictors, such as phosphate, nitrate, and silicate. In this case, the FFNN model learned how pH varied with depth based on the similarity of vertical pattern between seawater pH and specific physical or biological conditions indicated by input environmental variables, and subsequently reconstructed seawater pH values at different depths using 3D fields of these environmental variables."

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.
Reponse: Thanks for the suggestion. We have added the number of samples used to compute the biases in Table 5 as the following:

Table 5. pH bias by area and depth computed with BGC-Argo and GLODAP dataset.

	Area		0-50 m	50-200 m	200-500 m	500-1000 m	1000-1500 m	1500-2000 m
Pacific	BGC-Argo	bias	0.028	0.016	-0.003	-0.013	0.027	-0.004
		N*	16433	34708	36431	19840	8772	3565
	GLODAP	bias	-0.001	-0.001	0.000	0.000	0.000	-0.001
		N	18687	26629	22746	24843	12613	13817
Atlantic	BGC-Argo	bias	0.018	0.019	0.013	-0.021	0.031	0.068
		N	3285	6832	7152	3565	1622	1288
	GLODAP	bias	0.000	0.000	-0.001	-0.001	0.000	0.000
		N	11808	15894	14330	18056	10686	11780
Indian	BGC-Argo	bias	0.023	0.034	0.025	-0.022	0.000	0.036
		N	407	916	920	491	241	57
	GLODAP	bias	-0.006	-0.001	-0.003	-0.004	-0.004	-0.001
		N	3145	5397	5124	5276	3457	3421
Southern	BGC-Argo	bias	0.008	0.000	0.001	0.015	0.040	0.068

		N	66436	130563	135817	72564	27579	18692
	GLODAP	bias	0.004	0.001	0.001	0.000	0.000	0.000
		N	7983	12268	10457	10341	6169	5800
	BGC-Argo	bias	0.012	0.004	0.001	0.008	0.036	0.057
Global		N	86561	173019	180320	96460	38214	23602
	GLODAP	bias	-0.001	0.000	0.000	-0.001	-0.001	0.000
		N	46415	66635	57491	62447	34994	37008

(*: N is the number of BGC-Argo or GLODAP samples used to compute the biases.)