This paper presents Jingwei-Nutrients, a global monthly four-dimensional reconstruction of dissolved nitrate, phosphate, and silicate spanning 1965–2023 at 1°×1° horizontal resolution and from 0 to 2000 m depth.

The core methodological contribution is the application of a Transformer-based Multi-Task Learning (MTL) framework that simultaneously reconstructs all three nutrient fields, exploiting their biogeochemical stoichiometric coupling rather than treating them as independent targets. The authors demonstrate that the MTL approach consistently outperforms Single-Task Learning (STL) baselines in a rigorous time-based K-fold cross-validation, yielding roughly 24 % RMSE reductions for nitrate and phosphate and ~9 % for silicate. The validation framework is comprehensive, including comparison with WOA23 climatology, time-series station data like HOT and KERFIX, and specific GO-SHIP cruise sections.

The study addresses a genuine gap: whereas WOA products aggregate decades of observations into static climatology, this data product provides a continuous monthly time series enabling study of multi-decadal biogeochemical variability. The MTL concept is well-motivated. However, several methodological concerns must be addressed before publication.

Major Comments:

- Table 1 shows that CCHDO data contribute 69,355 nitrate, 98,652 phosphate, and 73,165 silicate profiles to the training dataset. I think that the GO-SHIP cruises used for independent spatial validation in Section 3.4 (P16N 2015, P16S 2014, P06E 2017) are disseminated through CCHDO. So, it looks like these cruise data were included in the training set. If it is the case, the spatial validation in Figs. 13, 14 and 15 is not independent. The authors must clarify whether these specific cruises were withheld from training, and if not, please provide alternative independent spatial validation (e.g., GEOTRACES sections or GO-SHIP lines not present in CCHDO).
- Both HOT and KERFIX are long-running time-series stations whose data are archived in WOD, which contributes the majority of training observations. The authors do not state whether profiles from these two stations were explicitly withheld from training. If HOT data (running from the late 1980s to the present) were included in WOD training profiles, the apparent high R^2 values at HOT (≥ 0.994 for all three nutrients) may partly reflect data memorisation rather than generalisation. Please clarify whether HOT and KERFIX were excluded from training and, if not, recompute station-level metrics using withheld data only.
- The primary motivation for the MTL approach is that jointly learning the three nutrients encourages physically consistent elemental ratios. However, the paper contains no direct evaluation of whether the reconstructed N:P or N:Si ratios are consistent, and whether they are more consistent in the MTL product than in the STL baseline. Given that this is the central scientific claim, the authors should compute and map Redfield ratios for example N:P across depth levels and regions and compare them with observations.
- BGC-Argo contributes 1,367,531 nitrate measurements. BGC-Argo nitrate is measured by ISUS/SUNA sensors that require careful calibration and are known to drift, particularly in deep-water deployments. The QC section (2.4) applies generic boundary and statistical checks but does not mention float-specific calibration procedures (e.g., cross-referencing with GLODAP climatological values at depth, as recommended by the BGC-Argo QC

manual). How was BGC-Argo nitrate specifically validated, and were adjusted (DOXY-corrected) profiles used?

- I can't find this reference: Broullón, D., Pérez, F. F., Velo, A., Hoppema, M., Olafsson, J., and Ríos, A. F.: A global monthly climatology of macronutrients in the surface ocean, Earth Syst. Sci. Data, 13, 2593–2612, https://doi.org/10.5194/essd-13-2593-2021, 2021. A comparison with relevant products is expected. At minimum, provide surface R^2, RMSE, and bias,and discuss where Jingwei-Nutrients improves and where it does not.

Minor Comments

- Please make the model code (training scripts, architecture, inference pipeline) publicly available (e.g., GitHub with a Zenodo DOI) and cite the repository.
- Section 5 (Data availability) is brief. Can you adapt it to include technical description about the data product: variable names, units, fill values, coordinate, depth..., and the structure of the ensemble uncertainty files? Please add a Data Format subsection or expand Section 5.
- The WOD23 dataset is cited as Garcia et al. (2026). I can't find it in the list of references, I recommend a full revisions of reference list.
- The R^2 for surface phosphate at KERFIX is 0.758, attributed to iron-driven N:P decoupling. Given that the Southern Ocean is both undersampled and biogeochemically important, users would benefit from a regional uncertainty map or bias analysis for the Southern Ocean surface layer, beyond the depth-averaged uncertainty shown in Fig. 16.

Minor language and typographical issues

The following minor issues should be corrected:

- Line 75: "they lost temporal variance" should read "they lose temporal variance".
- Line 105: "four-dimensional" is hyphenated here but not at line 19; please be consistent, and change to data product not a dataset. Data products are generally derived from datasets;
- Line 383: "anomalies are generally low" consider "anomalies are generally small" (concentrations, not anomalies, can be low).
- Figure captions for Figs. 13–15 do not align with the order of discussion in the text (Fig. 13 caption references P16N, but P16S is discussed first).