Responses to comments from reviewer

This manuscript presents a rich dataset supported by substantial and careful work. The resulting anthropogenic phosphorus input dataset will help advance our understanding of the global phosphorus cycle. However, I still have several concerns regarding the data-processing methods and some important details of the dataset development. I therefore recommend major revision and hope that the authors can address these concerns in a revised manuscript.

Response: We thank the reviewer for the positive assessment of our work and for recognizing the value of the HaPi. We also appreciate the constructive suggestions regarding data validation, uncertainties, and methodology.

General comments

Comment 1: Validation is still weak (Lines 443-456). The comparison mainly focuses on global totals, and most reference datasets use the same FAOSTAT sources. This does not validate the 5-arcmin spatial patterns. I would like to see comparisons with independent national or subnational statistics for several representative regions.

Response: We thank the reviewer for this important suggestion and we agree that evaluation against independent national or subnational statistics is desirable for assessing spatial reliability beyond global aggregates. Currently, independent P fertilizer datasets at national and subnational scales are still very limited, and FAOSTAT remains the most widely used and consistently available global source. Nevertheless, to strengthen the validation, we have further searched available datasets and identified two independent national-scale data sources: (1) the National Bureau of Statistics of China, and (2) CMH-P, a cropland P fertilizer application dataset for the United States derived from USDA-based statistical records (Cao et al., 2024).

We compared aggregated HaPi estimates with these two independent datasets. The results show that HaPi estimates are generally slightly lower than those two independent datasets; however, the temporal trajectories are highly consistent, and the magnitudes are comparable across the overlapping periods. Spatially, the P fertilizer input patterns in the HaPi dataset closely agree with those in the CMH-P dataset across the United States. These results provide additional support for the reliability of the HaPi dataset at national scales, particularly in capturing interannual variability and long-term trends. We have added the following content in the discussion section 4.1 (L463-467):

Additional validation against independent national statistics from China and the U.S. shows that HaPi captures comparable magnitudes and similar spatial-temporal patterns of P fertilizer use

(Figure S5), although with slightly lower estimates relative to both datasets. These results provide further support for the robustness of national-scale temporal dynamics in HaPi.

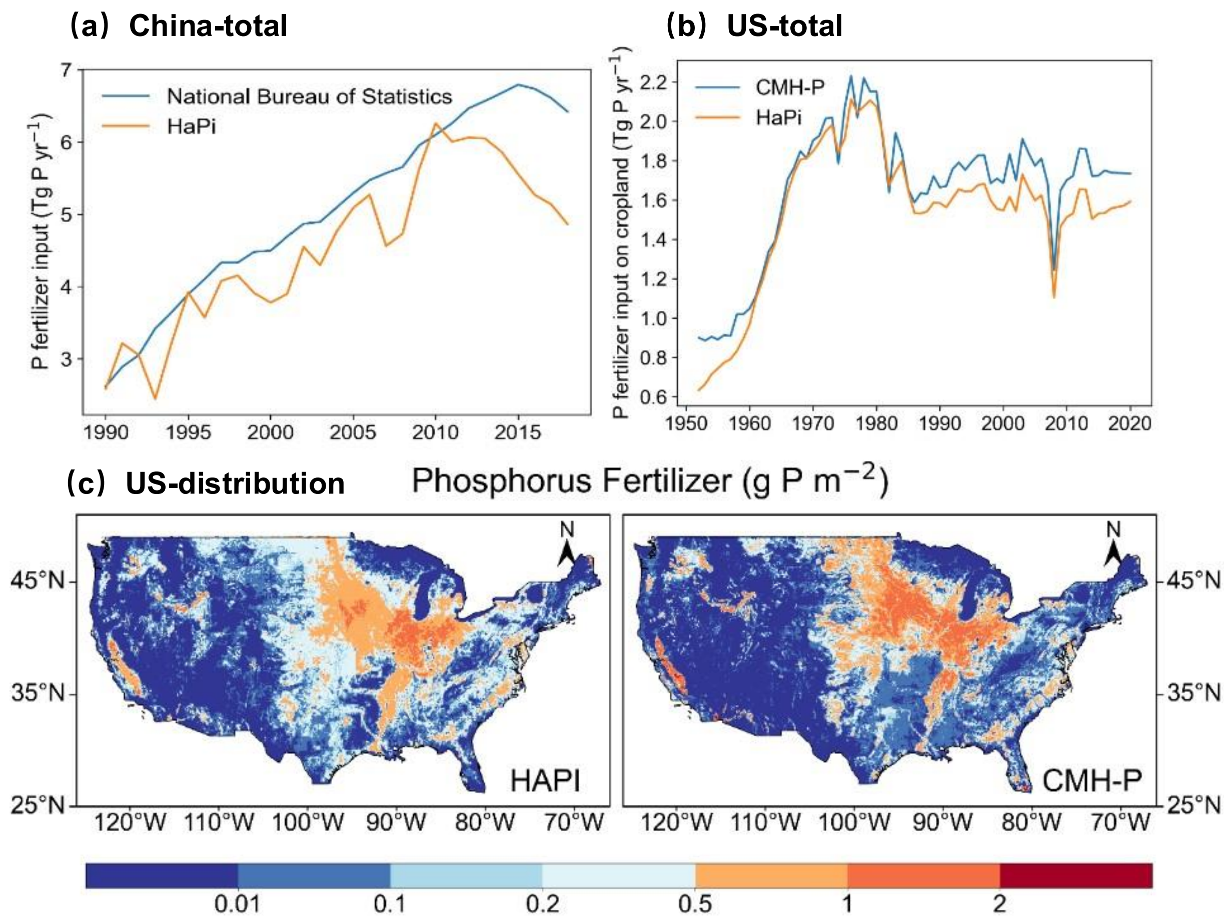

Figure S5. Comparison of P fertilizer inputs from the HaPi dataset with independent datasets: (a) national statistics for China; (b) national statistics for the United States; and (c) spatial comparison of P fertilizer inputs in the United States in 2020.

Comment 2: The estimation of manure-derived P may involve considerable uncertainty. Lines 210-230: Could the authors clarify whether the manure P estimates are sensitive to the fixed animal-specific P:N ratios and the use of GLW3 livestock spatial patterns from a single reference year? Although annual FAO livestock statistics were used to scale national livestock numbers, the within-country spatial distribution appears to remain static, which may introduce uncertainty in regions where pig and poultry production has shifted substantially over time.

Response: Thanks for this insightful comment. The other Reviewer has also raised this issue and we adopted the annual global gridded livestock of the world (AGLW) dataset (Du et al., 2025) to substitute the original static spatial pattern within each country. The AGLW data employs machine learning method to disaggregate FAOSTAT statistics to 5-km gridcells. However, it lacks available livestock density, in particular in the former Soviet Union region. Thus, the original N:P ratio maps were retained for this region. The final product shows high spatial and temporal consistency (Figure S1). All data and related figures have been updated. Comparison of the results before and after updating the manure N:P ratio dataset showed no significant changes in total manure P inputs across regions or over time, suggesting that manure P estimates are relatively insensitive to variations in livestock spatial patterns. This change has been stated in the Methods section (L 235-243).

Livestock distribution data $Anum_{y,g,t}$ were derived by fusing the annual gridded livestock of the world (AGLW) dataset (Du et al., 2025) and a newly downscaled product. The AGLW dataset covers 1961-2021 at a 5 km resolution and was developed based on annual FAOSTAT census information using a machine learning method. But, in some areas (particularly the former Soviet Union region), livestock information is missing. To fill the gaps, we used the national FAO livestock statistics and the Global Livestock of the World 3 (GLW3) database (Gilbert et al., 2018) to generate a long time-series gridded livestock distribution data.

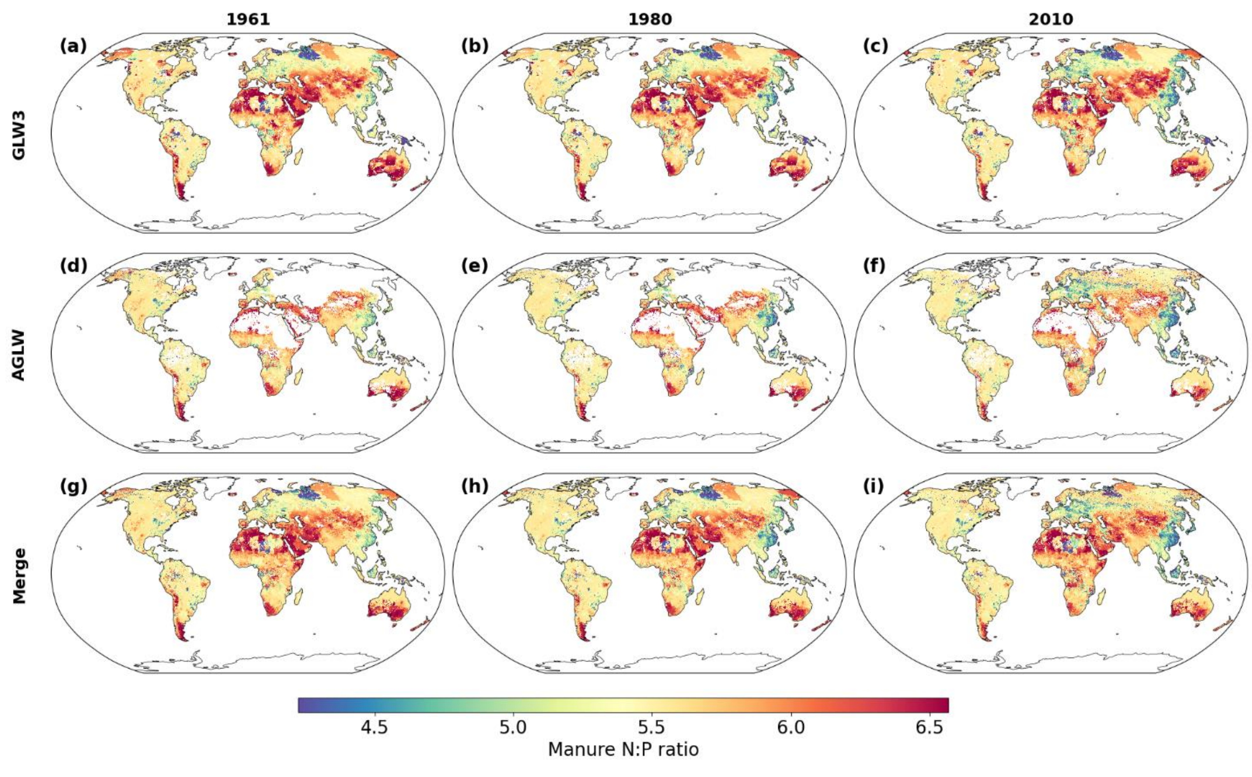

Figure S1. Comparison of manure N:P ratio maps derived from three datasets in 1961 (a, d, g), 1980 (b, e, h), and 2010 (c, f, i). The first, second, and third rows show the GLW3, AGLW, and merged datasets, respectively. Columns correspond to 1961, 1980, and 2010.

Comment 3: Lines 188-199: Uncertainty in the pre-1961 reconstruction. As I understand it, the 1961 spatial distribution was used as a reference pattern and scaled backward using the global temporal change rates from Cordell et al. (2009). If so, the pre-1961 maps may mainly reflect changes in global total P fertilizer use, rather than changes in regional or national spatial patterns.

I do not fully understand why this approach is appropriate for extending the dataset back to 1860. In my view, applying the 1961 spatial pattern to nearly 100 years of earlier history may seriously weaken the reliability of the pre-1961 data, because crop composition, cropland distribution, fertilizer accessibility, and agricultural practices likely changed substantially during this period. I strongly suggest that the authors more carefully address the uncertainty introduced by this backward extrapolation and provide clearer information in the dataset, such as reliability classes, or confidence flags for the pre-1961 period.

Response: Thank you for raising this important issue. As correctly noted by the reviewer, the spatial distribution of fertilizer inputs in 1961 was used as the baseline map, and annual fertilizer inputs before 1961 were reconstructed by scaling this baseline using the historical global fertilizer consumption trajectory reported by Cordell et al. (2009). We acknowledge that this approach cannot explicitly capture historical changes in crop composition, fertilizer accessibility, and agricultural management practices prior to 1961. Consequently, uncertainties in the spatial distribution of fertilizer inputs are inevitably higher during this period. Although the pre-1961 estimates are associated with greater uncertainty than those after 1961, they remain valuable for Earth system and land surface modeling studies. Many land surface models require continuous forcing data from pre-industrial conditions to quantify long-term ecosystem responses and legacy effects of anthropogenic nutrient inputs.

It is important to note that globally consistent, spatially explicit fertilizer datasets before 1961 are extremely limited. Under such circumstances, the use of a baseline spatial distribution combined with historical fertilizer consumption trajectories provides a pragmatic first-order reconstruction that enables the extension of the dataset back to the pre-industrial period. Similar reconstruction strategies have been adopted in many global nutrient input datasets when long-term spatial observations are unavailable (Batool et al., 2025; Xu et al., 2019). Critically, the global consumption of inorganic P fertilizer was limited before the 1960s. Prior to the 1960s, P fertilizer use was negligible in most regions, meaning that the spatial pattern observed in 1961 largely reflects the regions that had both the accessibility and the demand for P fertilizers at the time when large-scale fertilizer use began. This historical context supports the assumption that these early-adopting regions are representative of the priority areas for P fertilizer application in preceding decades.

Meanwhile, according to your suggestion, we have revised the manuscript to clarify the reconstruction procedure, expanded the discussion of uncertainties associated with the pre-1961 period, and added temporal confidence classes in the dataset documentation to distinguish between the lower-confidence reconstruction period (1860–1960) and the period constrained by annual national fertilizer statistics (1961–2020).

Although this approach provides a temporally continuous long-term dataset, it cannot explicitly represent historical changes in crop composition, fertilizer accessibility, and agricultural management practices, and it ignores regional variations in the changes of P fertilizer usage across different countries before 1961. Consequently, uncertainties are substantially higher in the pre-1961 period, and the resulting maps should be interpreted primarily as broad-scale reconstructions rather than precise representations of historical subnational fertilizer distributions. (L563-570)

Table R1. Confidence flags of HaPi data in different periods

Year	Confidence Flag	Description
1860–1960	1	Lower confidence: reconstructed using historical global fertilizer consumption trajectories and the 1961 baseline spatial pattern
1961–2020	2	Higher confidence: constrained by annual national fertilizer statistics

Minor comments

Comment 4: Lines 34-55 and 233-246: HaPi is described as an "anthropogenic" P dataset, but atmospheric deposition includes dust, sea salt, volcanoes, and other natural sources. The terminology should be reconsidered.

Response: Thank you for pointing out this inconsistency. We use the term anthropogenic P inputs because the dominant P inputs represented in the HaPi dataset are anthropogenic, whereas natural atmospheric deposition contributes only a small fraction of the total P inputs. Moreover, using the term total P inputs would also be inappropriate because another major natural P source, weathering, is not included in the dataset. Therefore, we retain the term anthropogenic P inputs while clarifying in the revised manuscript that it primarily refers to the dominant anthropogenic input pathways, and that the atmospheric deposition component includes both natural and anthropogenic sources.

Although the HaPi dataset primarily focuses on anthropogenic P inputs, it additionally includes both natural and anthropogenic sources. (L606-607)

Comment 5: Lines 145-154: The manuscript states that 18 crops were included, but only 17 are listed. Please check the crop list.

Response: Thanks for your careful check, and 17 is the correct number.

Comment 6: Lines 182-185 and Table 1: The land-use data are reported to end in 2019, while HaPi continues to 2020. How was the 2020 land-use map produced?

Response: We used the HYDE 3.3 and LUH2 land use datasets in our analysis. We apologize for the earlier misstatement in which we incorrectly referred to HYDE 3.3 as HYDE 3.2. We confirm that HYDE 3.3 was the version actually used in our calculations.

For the period prior to 1949, HYDE 3.3 (https://landuse.sites.uu.nl/datasets/) provides land use maps only at decadal intervals (i.e., one map every ten years). To generate continuous annual data, we applied a temporal downscaling approach: the LUHv2 data for each year within a given decade were used to proportionally interpolate the HYDE 3.3 map from the beginning of that decade. This method preserves the decadal baseline provided by HYDE 3.3 while incorporating the interannual variability captured by LUHv2.

We have corrected the dataset version reference in the revised manuscript and clarified this reconciling procedure in the Methods section (L193-201).

Comment 7: Lines 233-263: The atmospheric deposition period is inconsistent: 2001-2019, 2000-2019, and extension back to 1900 are all mentioned. The treatment of 1860-1899 and 2020 should be explained.

Response: In the atmospheric phosphorus deposition dataset provided by Ringeval et al. (2024), global P deposition shows relatively stable values during the early part of the twentieth century (approximately 1900–1960). Based on this pattern, we assumed that atmospheric P deposition remained approximately constant during the earlier period of 1860–1899. Therefore, the deposition values for 1860–1899 were extrapolated using the same deposition levels as those estimated for the early twentieth century. We acknowledge that this assumption introduces some uncertainty because direct estimates of atmospheric P deposition prior to 1900 are limited. To improve transparency, we have the following content in the Method Section (L301-303).

P deposition for the period 1860–1899 was assumed to remain constant and equal to the early twentieth-century levels.

Comment 8: Lines 247-263: Downscaling 2.5° model output to 5 arcmin does not necessarily provide real information at 5-arcmin resolution. The authors should avoid overstating the effective resolution.

Response: Exactly as the reviewer pointed out, downscaling coarser-resolution model outputs to a finer grid does not necessarily generate meaningful information at the finer resolution. In this study, the 5-arcmin product was generated to provide a spatially consistent input dataset compatible with other high-resolution forcing datasets. The downscaling approach and resulting products were rigorously evaluated against site-level observations in You et al. (2026), demonstrating reasonable performance at the annual timescale.

Comment 9: Lines 525-537: Could the authors briefly justify the use of a uniform 20% uncertainty for FAO-based data? Data quality may vary across countries, time periods, and variables, so a short explanation would be helpful.

Response: Thanks for raising this issue. Dr. Francesco N. Tubiello, one of our co-authors and the developer of the FAO survey data, suggests that while it is correct that uncertainty of FAO activity data vary by country or at least region (and theme), FAO does not have a firm hold on these numbers. FAO routinely use a 20% uncertainty across the board as a default assumption, which is recommended by IPCC to use for national greenhouse gas inventories (NGHGIs) using FAO activity data as input (Tubiello et al., 2021). In addition, a more exhaustive assessment of uncertainty in FAO land use data by country and region –resulting in about 25% uncertainty globally, is given in Tubiello et al. (2023). We acknowledge that a uniform 20% uncertainty cannot fully capture the spatial and temporal heterogeneity in data quality. Nevertheless, it provides a pragmatic first-order approximation of the uncertainty associated with the underlying statistical data.

Comment 10: Line 378: "Consumer of manure P" is not appropriate because manure application and deposition are inputs rather than consumption. "Manure P input" would be clearer.

Response: Thanks for your suggestion. We have revised this sentence as "China has surpassed Europe as the region receiving the largest manure P inputs since the 2000s, as manure P inputs in Europe have declined since the 1980s."

References

Batool, M., Sarrazin, F. J., and Kumar, R.: Century-long reconstruction of gridded phosphorus surplus across Europe (1850–2019), Earth System Science Data, 17, 881–916, https://doi.org/10.5194/essd-17-881-2025, 2025.

Cao, P., Yi, B., Bilotto, F., Gonzalez Fischer, C., Herrero, M., and Lu, C.: Crop-specific management history of phosphorus fertilizer input (CMH-P) in the croplands of the United States: reconciliation of top-down and bottom-up data sources, Earth System Science Data, 16, 4557–4572, https://doi.org/10.5194/essd-16-4557-2024, 2024.

Cordell, D., Drangert, J.-O., and White, S.: The story of phosphorus: Global food security and food for thought, Global Environmental Change, 19, 292–305, https://doi.org/10.1016/j.gloenvcha.2008.10.009, 2009.

Du, Z., Yu, L., Zhao, Y., Li, X., Liu, X., Li, X., Hao, P., Chen, Z., Guo, Z., You, L., Ma, X., and Wang, H.: Annual global grided livestock mapping from 1961 to 2021, Earth System Science Data, 17, 5543–5556, https://doi.org/10.5194/essd-17-5543-2025, 2025.

Gilbert, M., Nicolas, G., Cinardi, G., Van Boeckel, T. P., Vanwambeke, S. O., Wint, G. R. W., and Robinson, T. P.: Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010, Sci Data, 5, 180227, https://doi.org/10.1038/sdata.2018.227, 2018.

Tubiello, F. N., Conchedda, G., Wanner, N., Federici, S., Rossi, S., and Grassi, G.: Carbon emissions and removals from forests: new estimates, 1990–2020, Earth System Science Data, 13, 1681–1691, https://doi.org/10.5194/essd-13-1681-2021, 2021.

Tubiello, F. N., Conchedda, G., Casse, L., Hao, P., De Santis, G., and Chen, Z.: A new cropland area database by country circa 2020, Earth System Science Data, 15, 4997–5015, https://doi.org/10.5194/essd-15-4997-2023, 2023.

Xu, R., Tian, H., Pan, S., Dangal, S. R. S., Chen, J., Chang, J., Lu, Y., Skiba, U. M., Tubiello, F. N., and Zhang, B.: Increased nitrogen enrichment and shifted patterns in the world's grassland: 1860–2016, Earth System Science Data, 11, 175–187, https://doi.org/10.5194/essd-11-175-2019, 2019.

You, S., Ma, J., Zheng, H., Liu, X., Tian, H., Wu, Y., Zaehle, S., Wang, Y., Shi, H., Li, R., and Shi, H.: Human Activities Dominate Global Trends in Terrestrial Phosphorus Deposition, Environ. Sci. Technol., 60, 10850–10860, https://doi.org/10.1021/acs.est.5c16208, 2026.