Responses to comments from reviewer

We sincerely thank reviewer for your time and the comments on our work. We have carefully revised the manuscript accordingly to address all the points raised. With your help, the overall quality and readability of this manuscript have now been significantly improved.

Comment 1: Foundational Flaws in Land Use/Land Cover Data

Response: We acknowledge the reviewer's concern regarding the potential biases in historical land-use datasets such as HYDE 3.2 and LUHv2. These uncertainties are well recognized in the literature and are indeed an important source of uncertainty in global-scale reconstructions. However, the primary objective of this study is to develop a long-term, internally consistent phosphorus input dataset (1860–2020). At present, HYDE and LUHv2 remain the only datasets that provide globally consistent, long-term, and harmonized land-use information over this extended historical period. In contrast, remote sensing-based land cover products (e.g., ESA CCI-LC, MODIS) are only available for recent decades and therefore cannot be directly used to construct a temporally continuous dataset spanning more than 150 years. Moreover, large discrepancies exist among the remote sensing products (Gang et al., 2023). Therefore, rather than introducing inconsistencies by merging fundamentally different land-use products (which is obviously out of scope of this study), we chose to rely on a single harmonized dataset to ensure temporal consistency. Nevertheless, we agree that this choice may introduce spatial biases, particularly at regional scales. The above statements have been included in the discussion part to comprehensively clarify uncertainties around our products.

Comment 2: The Unrealistic Temporal Invariance of Fertilizer Application Rates

Response: We would like to clarify that our study does not assume temporally constant fertilizer application rates. Instead, the crop-specific fertilizer application rates from circa 2000 are used only as a baseline spatial pattern, which is then scaled annually using FAOSTAT national fertilizer consumption data. This approach allows fertilizer application rates to vary over time at both national and grid levels, thereby capturing long-term trends in fertilizer use. Due to the lack of global gridded dynamic crop-specific fertilizer use maps, we have to adopt the static relative spatial pattern. Such a trade-off is often seen and accepted in practice and previous studies (Lu and Tian, 2017; Nguyen et al., 2024).

Comment 3: Lack of Spatially-Explicit Validation Against Remote Sensing or Independent Data

Response: While remote sensing proxies (e.g., NDVI, chlorophyll-a, night-time lights) are valuable, their relationship with P inputs is indirect and influenced by multiple confounding factors (e.g., climate variability, management practices, hydrology). Therefore, such comparisons may not provide a direct or unambiguous validation of P input distributions. Instead, we have adopted a commonly used approach in global dataset development by comparing our estimates with independent datasets and national-scale statistics, which ensures consistency in magnitude and temporal trends.

Comment 4: Overly Qualitative Uncertainty Analysis

Response: In this study, uncertainties arise from multiple sources, including land-use data, fertilizer allocation, manure estimates, and atmospheric deposition. Due to the lack of globally consistent uncertainty information for each input dataset, it remains challenging to perform a fully quantitative, spatially explicit uncertainty propagation at the global scale. But we have discussed the potential uncertainty sources, from the perspectives of both input data and the methodology, in section 4.4 Uncertainty.