

RC1

>>> Thank you very much for taking the time to consider our manuscript and for your helpful and constructive feedback. We agree with and appreciate the suggestions offered and believe we have addressed and incorporated each point into the manuscript accordingly. A response to each specific comment is provided below demarcated by “>>>” and bold text, and all changes in the manuscript word document have been marked with tracked changes.

This manuscript presents the Harmonized Land Nutrient Irrigation Dataset (HLNID), providing annual maps of land use–land cover (including crop type), fertilizer and manure nutrients (N and P), and irrigation extent across CONUS for 1938–2020 at flexible spatial resolutions. The study integrates county-level Census of Agriculture (CoA) data with multiple remote sensing and model-derived products using a demand–allocation framework.

The dataset is clearly within the scope of Earth System Science Data. It provides long-term, harmonized, spatially explicit inputs that are highly relevant for ecosystem, hydrologic, and biogeochemical modeling. The workflow is transparent and technically sound.

>>> Thanks so much for the positive feedback.

I recommend minor revision. The requested changes primarily concern clarification of assumptions and improved transparency.

Major Comments

1. Clarification of Dataset Nature

The manuscript appropriately states that HLNID reflects foundational datasets (e.g., CoA and NLCD) rather than providing an independent reconstruction (Section 4.2). However, this distinction should be emphasized earlier in the Methods section. The current framing may lead readers to interpret HLNID as a fully independent historical reconstruction rather than a harmonized synthesis framework.

>>> Thanks for noting this need for clarification. We agree and added additional language at the end of section 1.5 (“Notably, the method presented is not a predictive model resulting in a fully independent historical reconstruction because the county-scale demand is estimated based on available historical datasets; rather, it synthesizes existing datasets to reveal a consistent LULC dataset with high thematic

resolution including major crop types with concurrent estimates of nutrients and irrigation extent.”) and the beginning of section 2.2.1 (“Our method to estimate historical LULC uses a harmonized synthesis framework that leverages existing datasets rather than an independent land change model to reconstruct the past.”)

2. Limited Treatment of Uncertainty

Although data sources and workflow are well documented, structured discussion of uncertainty is limited. Key assumptions that introduce uncertainty include:

- Use of contemporary suitability/frequency maps for pre-satellite periods
- Fixed crop-to-corn fertilizer ratios
- Uniform within-county manure allocation
- Static open-water extent

A full quantitative uncertainty analysis may be beyond scope, but I recommend expanding Section 4.3 to include:

- A concise discussion of primary uncertainty sources
- A conceptual distinction between uncertainty in county-level demand versus spatial allocation

This would improve transparency without requiring substantial additional analysis.

>>> We appreciate this point about improving transparency related to uncertainty in the datasets and the new method presented. We agree and have added more description of the key uncertainties in the first 3 paragraphs of Section 4.3.

3. Pre-1938 Guidance for LSM Applications

Because many land surface models initialize simulations from 1850, the 1938 start year may limit direct compatibility. While the chosen temporal range is justified, it would be helpful to provide guidance for users requiring earlier spin-up.

Without requiring full reconstruction of 1850–1937, the authors could consider:

- Providing an optional coarser-resolution or aggregated pre-1938 extension (e.g., cropland/pasture/forest), clearly documented as lower confidence; or

- Offering explicit guidance on how to backfill or harmonize with existing global historical land-use datasets.

Clarifying this would improve usability for the land modeling community.

>>> We agree that it would be ideal to extend the data as far back as possible to support simulations that start earlier (~1850). We added the following sentence to the second to last paragraph in Section 4.3: “Finally, there is also a clear need to extend these datasets even further back in time to capture important land changes despite the increasing uncertainty associated with datasets like the CoA prior to 1938. However, we recommend using datasets such as the Biophysical Settings in the LANDFIRE dataset (Rollins, 2009) that represents pre-Euro-American settlement vegetation conditions as a starting point to bridge the gap in time.”

RC2

>>> Thank you very much for taking the time to consider our manuscript and for your helpful and constructive feedback. We agree with and appreciate the suggestions offered and believe we have addressed and incorporated each point into the manuscript accordingly. A response to each specific comment is provided below demarcated by “>>>” and bold text, and all changes in the manuscript word document have been marked with tracked changes.

This study developed a comprehensive land-cover dataset by integrating multiple statistical and gridded data sources. The dataset includes land cover, crop types, agricultural fertilizer application, and irrigation maps. These products provide valuable inputs for ecosystem modeling and facilitate investigations into the environmental impacts of historical land-use change.

>>> Thanks so much for the positive feedback.

I have a major suggestion regarding the data sources used in this study. Since NLCD plays a critical role in generating the final products, it would be worthwhile to consider incorporating the recently released annual NLCD maps (available at <https://www.mrlc.gov/data#>). I suggest consistently using the NLCD dataset for the period 1985–2008 to update and improve the products, which could enhance their temporal consistency and overall accuracy.

>>> We thank the reviewer for suggesting incorporation of the annual NLCD dataset. It is something that we have been contemplating for many months but were waiting for the validation manuscript to be published before proceeding. Since that is now available (Fleckenstein et al. 2026) we have fully incorporated the new annual NLCD dataset into the HLNID development methodology for the years 1985-2020. We agree that this now makes the method less complicated and more accurate using the latest and most consistent LandSat-based land cover estimates over a long period of the record. Note that incorporating the annual NLCD dataset eliminated the need to use LCMAP. Thus, all mention of LCMAP in the manuscript has been removed. The new dataset has been updated at the Dryad repository. All figures have been updated to reflect the change in methodology. Despite the changes, the final datasets and maps have changed very little, which is not surprising because the LCMAP and annual NLCD data products are very similar. However, we still view the decision to change/upgrade the method using the new annual NLCD as an essential way to provide future updates.

Minor comments:

Lines 41: NLCD has updated annual maps from 1985 (<https://www.mrlc.gov/data#>). Please update your maps with the annual NLCD maps.

>>> Thanks. See comment above.

Lines 75: It seems that the citation of Ye et al. (2024) is missing from your reference list.

Also, in the comparison section, could you offer a crop type area comparison between your results and Ye et al. (2024).

>>> Thanks for catching the accidental omission of Ye et al. (2024) from the References section even though we have it cited in the text. We have fixed this. We also added Figure S5 in the Supplement that compares corn and soybean area over the 1938 to 2020 time period between the CropAT_US dataset and the HLNID. They line up very closely, which is not surprising since they both rely on the CoA dataset. We added the following sentence in the Results section: “For specific crop types, the HLNID aligns very closely with the CropAT_US dataset (Ye et al., 2024) as they both rely on the CoA dataset (Figure S5).” (Lines 352-353)

Lines 224-226: Please clarify how discrepancies between the summed crop-type areas and the harvested area are addressed. Specifically, how do you handle cases where the aggregated crop-type area is either greater than or less than the harvested area?

>>> For clarity we added more to that sentence. It now reads “Crop type areas were adjusted so that their sum was equivalent to the total cropland harvested area by multiplying each crop type area by the ratio of the total cropland harvested area to the original sum of all crop type areas.” (Lines 225-227)

Lines 313: Does “290 Mha” refer to 290 million hectares? Please spell out the full unit (e.g., million hectares) at its first occurrence for clarity.

>>> Thank you. This has now been added at the first occurrence. (Lines 315-316)

Lines 423-427: I am interested in how the 1980s crop crises affected land-use change, particularly transitions between cropland and other land-use types. Did this period lead to significant cropland abandonment or conversion to alternative land uses?

>>> The line charts in Figure 4 suggest that cropland declined slightly during the Farm Crisis and some land shifted to grassland and forest. However, due to the limitations of this methodology, it's difficult to determine the extent and location of precise transitions from one land use to another.

Line 445: The statement “it does match well with the CoA and NLCD as designed” would benefit from additional visual support. Could you provide line charts comparing land-use areas from the various data sources? Such figures would help readers clearly visualize the differences and similarities in land-type areas across datasets.

>>> Thanks for pointing this out. We added the reference to Figure 4 at the end of this sentence as we believe that figure clearly shows how well the HLNID matches the CoA and NLCD.