Suggestions for revision or reasons for rejection

(visible to the public if the article is accepted and published)

The revised manuscript is substantially improved compared with the previous version. I appreciate the authors' efforts to clarify the historical data sources, provide a clearer description of the dataset construction, and moderate several of the broader claims. The compilation itself represents a valuable contribution. The organization of state-level agricultural liming data for 1930–1950, county-level records for 1954–1987, and the resulting 1 km gridded product provides an important historical baseline that will be useful for future studies of agricultural management, soil acidification, carbon cycling, and Earth system modeling.

Despite these improvements, several issues should still be addressed before the manuscript is suitable for publication. Most of these concerns relate not to the value of the dataset itself, but to its documentation, reproducibility, and internal consistency. In particular, the manuscript would benefit from more rigorous documentation of missing-data treatment and uncertainty propagation, a clearer description of county-boundary harmonization, a more carefully framed comparison with previous national estimates, a thorough consistency check of the statistical analyses, and correction of the remaining technical and editorial issues.

Thanks for the reviewer's careful reading of our work, which helped us greatly improved the quality and robustness of this manuscript draft.

Major comments

1. Missing-data treatment and uncertainty propagation require stronger documentation

The revised manuscript now explains the interpolation procedure more clearly than before and explicitly notes that agricultural liming variables were collected only in selected census years and that some counties did not report liming

information. These revisions improve the transparency of the workflow. However, the uncertainty introduced by temporal interpolation and extrapolation remains insufficiently characterized. Figure 1 shows that missing observations are spatially structured in several census years, yet the manuscript concludes that these gaps have only a limited influence on national estimates without providing quantitative evidence. Readers therefore cannot evaluate how sensitive the reconstructed dataset is to different assumptions regarding missing counties.

I recommend adding a table summarizing, for each census year, the number of missing counties, the associated missing land area and cropland area, and an estimate of their potential contribution to national lime application. The released county-level dataset should also include flags indicating whether individual values are observed or imputed. In addition, reporting uncertainty intervals for national or regional totals together with at least one sensitivity or bounding analysis based on an alternative interpolation strategy would substantially strengthen the dataset.

Thanks for the reviewer's comment. We added Table S2 which shows a table summarizing, for each census year, the number of missing counties for lime application weight and limed area. The associated missing land area and cropland area range from 0 to ~20%, except for 1974 which didn't report limed area.

However, based on the spatial distribution lime application, shown in Fig. 1, the missing counties are mainly in the regions of little or no lime application, and thus the potential contribution of missing counties to the uncertainty in national lime application should be very low. To demonstrate this, provide a quantitative analysis of this uncertainty in Fig. S1, which shows the uncertainty introduced by the missing counties and interpolation. We show that the difference between 1. national total amount of lime applied accounting post-interpolation values and 2. pre-interpolation with missing values is very small (uncertainty within 1%

in most years).

We revised our text, so it includes all the information that the reviewer requested.

“Table S2 shows the number of independent cities and missing counties that did not report lime application weight and limed area, and the associated missing land area and cropland area. The percentage of associated missing land area and cropland area range from 0 to ~20%, except in 1974, which didn’t report limed area. However, as shown in Fig. 1, most counties missing original lime application data are situated in areas with limited liming activity, particularly in the arid and desert regions of the western United States. We also provide a quantitative analysis of this uncertainty introduced by the missing counties in Fig. S1, demonstrating that the total lime application estimates are only slightly affected by the interpolation and data-filling for the missing counties each year (uncertainty are within 1% in most years between pre- and post-interpolation values; Fig. S1).”

We also revised our released county-level dataset, so it now includes flags indicating whether individual values for lime applied, limed area, and cropland area are observed or imputed. Below is information in the revised shapefile, which we added in the data availability section.

LimeTons: Total lime applied (metric tons)

TonsStatus: Whether LimeTons was observed or imputed

LimeArea: Area limed (hectares)

AreaStatus: Whether LimeArea was observed or imputed

CropArea: Total cropland area (hectares)

CropStatus: Whether CropArea was observed or imputed

2. Historical county-boundary harmonization should be documented more explicitly

The manuscript now explains that county-level statistics were joined to a U.S. county shapefile and briefly discusses the treatment of independent cities that were not reported separately in the agricultural census. This is an important improvement. Nevertheless, the description remains incomplete for a historical spatial dataset spanning more than three decades. The manuscript does not specify which county-boundary vintage was used, whether historical administrative changes were harmonized, whether GEOID/FIPS crosswalks were applied, or how county splits, mergers, identifier changes, and independent cities were handled. Because county boundaries directly affect both the spatial products and the area-normalized variables, these details are essential for reproducibility.

A dedicated subsection describing the spatial harmonization procedure would greatly improve the manuscript. This section should identify the shapefile source and vintage, describe the GEOID/FIPS matching procedure, explain how unmatched administrative units and independent cities were treated, clarify whether county areas correspond to historical or modern boundaries, and document any harmonization decisions. Ideally, the released dataset should also retain both original and harmonized identifiers together with matching and boundary-quality flags.

Thanks for the reviewer's comment. Our county boundary was based on the 2012 U.S. county-level shapefile. We added a subsection on how we process data merging, GEOID/FIPS matching procedure, and treatment of unmatched administrative units and independent cities, and the information the reviewer requested.

"For visualization and analysis, the county-level data were joined with the 2012 U.S. Census Bureau TIGER/Line county-level shapefile (https://www.census.gov/cgi-bin/geo/shapefiles/index.php). We opted for a modern version of a county-level shapefile since an earlier version was not available before

the year 2007. The joining was executed by matching the unique 5-digit Geographic Identifier (GEOID) in the county-level shapefile, with the corresponding Federal Information Processing Series (FIPS) state-and-county codes embedded in the county-level census data. Due to missing land area for some counties in the original census data, the county-level land area was recalculated based on the more modern shapefile geometry, providing a more consistent basis for further rasterization of area-normalized values. We excluded all counties outside the contiguous United States (e.g., Alaska, Hawaii). A number of independent cities in the shapefile, including Baltimore, St. Louis, and several independent cities in Virginia, are absent from the census reports. These cities are administratively distinct from counties in the U.S. Census framework and are not included in county-level reporting. Consequently, no lime application data is available for them across any census year. Given that these independent cities are highly urbanized, with little to no agriculture, we reasonably assumed that no agricultural lime was applied in the areas. We treated two counties affected by historical name or administrative boundary changes separately before merging the data. Further details on the changes to county identifiers and the merging process are documented in the Supplementary Information (SI)."

We also revised our revised dataset which included the original FIPS, and the GEOID of the modern county-level shapefile. Below are added information in the revised shapefile, which we also added to the data availability section.

GEOID_shp: County GEOID associated with the modern U.S. county shapefile

FIPS_cens: County FIPS code in the historical census reports (Cities/modern counties that are not included in the historical census reports will show missing values)

3. The comparison with West & McBride requires clearer unit definitions and a more cautious interpretation

The newly added comparison with West & McBride is a useful addition, and the discussion of the uncertainty associated with Mineral Yearbook-derived estimates provides valuable context. However, the comparison currently mixes

different quantities without sufficiently clarifying their units. The manuscript refers to values of approximately 35 Tg C yr^{-1}, whereas Figure 10 and the accompanying discussion appear to compare total lime mass. It is therefore unclear whether the reported values represent carbon mass, applied limestone mass, $CaCO_3$-equivalent mass, or another metric. In addition, agricultural census data describing farm lime application and Mineral Yearbook production or consumption statistics may not represent exactly the same quantity, making direct comparison less straightforward.

The manuscript should clearly define the units used throughout Figure 10 and the corresponding discussion, provide any conversion factors that are applied, and explicitly distinguish among Tg C, Tg of applied lime, $CaCO_3$-equivalent mass, and metric tons of agricultural limestone. Unless a comprehensive uncertainty analysis is provided, I also recommend replacing statements suggesting that the new estimate is "more accurate" with more neutral wording indicating that it represents an alternative estimate derived from agricultural census records.

Thanks for the reviewer's comment. We apologize for the typo in "Tg C/year".

The mixed use of Tg and Mt in weight units all refers to the "total mass of lime applied". Here, in our revised manuscript, we unified all the units into Mt, as it is more consistent with the discussion in our other paragraphs.

We also revised our wording of our discussion into a neutral tone, according to the reviewer's suggestion.

"Therefore, we believe that our compiled dataset potentially offers a more robust representation of the total amount of lime applied in the US. The range of 15-25 Mt per year derived from the agricultural census records presents an alternative estimate of total lime application, compared with the 25-35 Mt per year suggested in West &

McBride (2005).”

4. The statistical analyses require an internal consistency check

The revised manuscript provides a clearer description of the Random Forest analysis than the previous version. However, several inconsistencies remain between the Methods, Results, and Discussion. For example, the Methods indicate that Total Cropland was removed after correlation and variance inflation factor filtering, whereas the Results and Discussion continue to describe Total Cropland as one of the most important predictors. Similarly, the Results first state that cropland harvested percentage is negatively correlated with lime-treated fraction but then interpret higher harvested percentages as corresponding to higher lime-treated fractions. These inconsistencies make it difficult to interpret Figure 9 and reproduce the analysis.

I recommend including a final table listing every predictor considered in the Random Forest model, together with its data source, temporal coverage, units, any transformations applied, VIF status, and final inclusion or exclusion status. This table should correspond exactly to the variables presented in Figure 9. The reported Pearson correlation signs and the interpretation of the partial dependence plots should also be carefully verified, particularly for cropland harvested percentage.

Thanks for the reviewer pointing out these two errors. In our revised manuscript, Total Cropland and Harvested were not removed. We revised the mistake

Regarding the cropland harvested percentage relation to lime-treated fraction, it is a positive correlation. Both of which are shown in Figure 9. We revised the text as,

“For lime-treated fraction, soil pH shows the highest permutation importance, exhibiting a strong negative correlation. Other important predictors included mean precipitation and sand percentage (positive correlations), and carbonate mineral stocks (negative correlations), cropland harvested percentage (positive correlations)

(Fig. 9c). Partial dependence analysis shows that counties with lower soil pH, greater rainfall, higher sand content, lower carbonate minerals, and higher harvested cropland percentage exhibited a higher lime-treated fraction among croplands (Fig. S9)."

We revised and added information to Table S1, so it includes all the information the reviewer recommended us to provide.

The revised text for this relevant section currently reads as "To mitigate multicollinearity, we also examined pairwise correlations among predictors and removed variables from highly correlated pairs (Pearson's r > 0.90), including *N fertilizer (kg)* vs. *P fertilizer (kg)* (r = 0.91), and *N manure (kg)* vs. *P manure (kg)* (r = 0.98). Based on domain relevance and predictive strength, we retained one variable from each pair and removed the others. We calculated the Variance Inflation Factor (VIF) for all remaining predictors. Variables with high VIF values were selectively removed to reduce multicollinearity and improve model stability. Table S1 shows all the predictors considered in the analysis, together with their data sources, temporal coverage, units, transformations, VIF status, and final inclusion or exclusion status. The final list of features excluded is: *P manure (kg)*, *P fertilizer (kg)*, *Silt (%)*, and *Available Water Storage (cm)*. This process yielded a final set of 32 predictors with acceptable VIF values (range: 1.12 to 7.32), which were used for statistical and machine learning modeling (Table S3). This step ensured a more robust and interpretable analysis of environmental controls. "

Minor comments

The manuscript is much improved in terms of presentation, but several unclear points should be corrected before publication:
Page 2, lines 32–34: Revise "Previous works on agricultural liming has..." to either "Previous work on agricultural liming has..." or "Previous studies on agricultural liming have...".

Thanks for the reviewer's suggestion. We have revised our text and have changed to "Previous studies on agricultural liming have...".

Page 2, lines 46–48: Remove the unnecessary comma before "to assess".

We removed it according to the reviewer's suggestion.

Page 3, lines 71–73: The phrase "where access is available" is repeated and the sentence should be rewritten for clarity.

Thanks for the reviewer pointing this out. We changed to "In response, lime application has become a common strategy to mitigate the issue of soil acidification in croplands where access is available."

Pages 7, 11, and 12: Several references to Figure 1 appear to be incorrect. These likely should refer to Figure 2 or Figure 3, as Figure 1 presents county-level missing-data maps rather than state-level estimates.

Thanks for the reviewer pointing this out. We revised these errors.

Page 16: The statement that lime-treated areas peaked in 1970 is inconsistent with the reported agricultural census years. Please revise this to the appropriate census year or use a broader description such as "the early 1970s" if that better reflects the data.

Thanks to the reviewer pointing out our error. We revised to "in the later 1970s."

Page 19: Correct grammatical errors such as "R^2 values drops" and "For the remaining of the discussion.

Thanks to the reviewer. We revised the two statements to "The performance R^2

values drop from a range of 0.66-0.73 to 0.46-0.66" and "For the remainder of the discussion…"

Page 20: Correct wording such as "thought relative ranking" and "which all shows."

We revised into "though the relative ranking.." and "which all show".

Pages 22–23: Revise the sentence "The required lime application rate and lime-treated fraction is..." to use the correct plural verb ("are"), and remove the repeated "Both processes..." statement.

Thanks to the reviewer pointing this out. We revised according to the reviewer's suggestion.

Page 24: Revise "temperature precipitation" to "temperature and precipitation."

Revised

Page 29: The conclusion should consistently describe the temporal coverage of the dataset. It currently refers to 1910–1987, whereas the manuscript elsewhere describes the principal data products as covering 1930–1987 or 1954–1987. The final sentence of the conclusion would also benefit from revision for improved clarity and readability.

Thanks for the reviewer pointing this out. We changed it to 1930-1987, as it includes both the state level from 1930-1950, and county level from 1954-1987. We also revised the final sentence as

"Future research could potentially build on these spatially explicit data of lime applications by accounting the variations in soil types and climate conditions. Comparing these trends against historical climate-relevant data (eg. soil carbon stock, river alkalinity) during this period will enable a comprehensive assessment of further agricultural or climate-mitigation questions. Future research could potentially build on these spatially explicit data of lime applications by accounting the variations in soil types and climate conditions. Comparing these trends against historical climate-

relevant data (eg. soil carbon stock, river alkalinity) during this period will enable a comprehensive assessment of further agricultural or climate-mitigation questions."