Reviewer 1

Thank you very much for the reviewer's detailed review and recognition of our work. Your suggestions have greatly improved the quality of this manuscript. We have carefully considered all of your comments and have revised the manuscript based on each individual comment. We have done our best to enhance the manuscript and hope that the revised version will meet your expectations. A point-by-point response to your comments is listed below.

Reviewer comment:

This manuscript describes a methodology that created a CONUS scale dataset related to agricultural lime application (mass, area, and rate) from 1930 to 1987 based on state and county-level census of agriculture datasets. Further, the authors analyze some of the environmental predictors of lime application using statistical methods. Overall, this study fills an important gap in better understanding the spatial-temporal variation in agricultural lime application and the authors write in a clear and organized manner. Therefore, I recommend acceptance with minor revisions.

My main concern is related to the scope of ESSD. This is my first review for this journal so I am not sure how the scope is enforced but the statistical analysis aimed at determining the main drivers/predictors of lime application seemed like it could be interpreted as outside the scope. In particular, the journal website states: "Any interpretation of data is outside the scope of regular articles". It appears that many currently published articles could reasonably violate a strict interpretation of this scope statement. But I raise this issue because the statistical analysis presented is entirely separate from the development of the dataset. It is interesting and the methods are sound, but it doesn't directly contribute to the production of the dataset. Therefore, I would appreciate more guidance from the editors to determine whether that part of the manuscript is out of scope or not.

Another broader point that I suggest the authors make is related to the data availability at the county scale. As the authors note, the county-scale data is no

longer available beyond the year 1987. In the interest of better scientific understanding and more open public datasets, I would recommend that the authors highlight the need to restart the collection of this specific county-scale dataset by the USDA. Otherwise, researchers are severely limited as we move forward into the future, especially as liming is potentially seen as a carbon sequestration practice. A good citation to include here would be Rissing et al. (2023) (https://doi.org/10.1038/s43016-023-00711-2) who discuss the importance of USDA data collection policies and the need to include datasets that better help us manage our agricultural landscape in a more sustainable manner.

Response: Thanks for the reviewer's careful reading and consideration of our work. We were also debating whether the statistical analysis should be removed. Thus, since the reviewer raised the concern, we have checked with the editor, and the editor responded that analysis that may help the interpretation of the data is acceptable. Thus, we kept our RF analysis, but we reframed our wordings in the manuscript, which we aim to use RF machine learning model mainly as a sanity check/confirmation if our liming maps are explainable with other relevant predictors. Here our focus is changed to see if known agronomic relationships (eg. Such as the strong link between pH, crop type, and lime use) can potentially help us interpret the pattern, rather than quantitatively assess "what are the causal drivers."

We also thank the reviewer's suggestion. We added this point with the recommended citation in our conclusion section

"We would like to highlight that the liming data being not available in the county level after 1987 greatly limit the progress of the capacity of research. A restart of the collection effort of lime application variables by the USDA would greatly enhance the community's assessment of the climate mitigation of lime addition (Rissing et al., 2023). Future research should build on these spatially explicit data of liming considering the variations in soil types, climate conditions, and other environmental factors to comprehensively assess the viability of liming as a nature-based climate solution."

Specific comments:

Reviewer comment:

Line 53: What time period is associated with this range? Please specify.

Response: Thanks for the reviewers comment. We specified the time range, changing the phrase to "…with a range of 15-25 million tons (Mt) per year between 1954–1987."

Reviewer comment:

Lines 104-106: Need a citation to support this sentence.

Response: We removed this sentence as it potentially leads to confusion.

Reviewer comment:

Lines 106-108: This is confusing. You just said that the historical consensus has been that carbonate amendments are net CO2 sources but now you're citing sources that argue that carbonate amendments "contribute substantially to carbon…sequestration"?

Response: Thanks for the reviewer pointing this out. Our original intention was to mention that historically people perceive liming as a CO2 source, but recent studies have reconsidered it as a CO2 sink/sequestration strategy. We added a new citation to support this point.

Raymond, P., Planavsky, N., & Reinhard, C. T. (2025). Using carbonates for carbon removal. *Nature Water*, *3*(8), 844-847.

Since we have already addressed the change of viewpoint of liming being a

source of CO2 (entirely weathered by strong acids) to being a sink of CO2 (mostly weathered by carbonic acids), we removed the part that mentioned liming originally being considered as a CO2 source. We further condensed this section as below to prevent confusion.

“The application of crushed rocks to agriculture fields, referred to as enhanced rock weathering (ERW), has emerged as a promising carbon dioxide removal and climate mitigation strategy (Beerling et al., 2020). Enhanced carbonate weathering, particularly through large-scale agricultural liming, could offer a more immediate approach to atmospheric CO_2 mitigation due to its faster weathering rate (Gaillardet et al., 1999; Liu et al., 2011; Raymond et al., 2025).”

Reviewer comment:

Line 122: "where it can be marketed". I would avoid unnecessarily inserting the idea that this research is essential to develop a carbon market. I recommend just stating something like "it will be important to develop a better understanding of when and where it will likely be a robust and reliable carbon removal practice." or similar

Response: Thanks for the reviewer’s suggestion. We rephrased the sentence below as suggested.

“Since lime can potentially act as both a source and a sink for atmospheric CO_2 depending on the weathering agent, and its carbon footprint can shift over time, often transitioning from a source to a sink depending on environmental and management conditions, it will be important to develop a better understanding of when and where it will likely be a robust and reliable carbon removal practice.”

Reviewer comment:

Lines 135-139: This paragraph is not appropriate at the end of the introduction because it discusses the results. Suggest removing or moving to discussion/conclusion section.

Response: Thanks for the reviewer's suggestion. We removed this paragraph, as most of these points will be discussed later in the discussion section.

Reviewer comment:

Lines 174-177: This description of the pH recommendations is confusing. Where exactly is this information coming from? How exactly were these pH recommendations estimated? Please provide some context on how those recommendations are determined by agronomists.

Response: Thanks for the reviewer's comment. For the pH recommendations data here, we compiled soil pH recommendations for the most commonly grown crop in every county across the conterminous United States. The most abundant crop in each county was determined using the 2022 CroplandCROS layer. In each state, soil pH recommendations are provided by land-grant university agronomic extension agencies. The recommended pH for the most commonly grown crop in each county was then recorded.

We added this information to the methodologies section as below,

"County-level soil pH recommendations were derived from state-level agronomic extension guidelines by assigning each county the recommended pH based on its most commonly grown crop. The dominant crop in each county was identified using the 2022 CroplandCROS dataset (https://croplandcros.scinet.usda.gov), and crop-specific pH recommendations were obtained from land-grant university extension publications for each state."

Reviewer comment:

Line 259: Please add that this was done using Python.

Response: Thanks for the reviewer's reminder. We changed the phrase to "We implemented the model in Python using the RandomForestRegressor class

derived from bootstrapped datasets (Pedregosa, 2011)."

Reviewer comment:

Line 463: Don't you have this total ag land area data compiled and available to better determine this?

Response: Thanks for the reviewer's suggestion. We conducted further analysis on this part, examining the total agricultural area, percentage of cropland that is harvested, and the percentage that is not harvested nor pastured, showing that it is likely that cropland abandonment has taken place, and the region started to transition to other land use.

We added Fig. S16, S17 and rephrased the paragraph as below.

"However, lime application and limed area along the Southeastern coast both shows a decline in the 1980s, which coincides with a decrease in crop production, particularly corn and soybean (Fig. S9-S15). The relative ratio of total cropland in this region did not show a clear decline (Fig. S16). However, the percentage of cropland harvested (Fig. S17) shows a strong decrease in 1987, while the percentage of cropland that is neither harvested nor used as pasture (Fig. S18) shows a clear increase, potentially indicating gradual abandonment of cropland for other uses. The decrease in crop production in this region could be an onset of urbanization and conservation programs (Napton et al., 2010). Additionally, the growth of the regional lumber industry during this period may have further slightly reduced agriculture, contributing to the observed decline in lime application (Prestemon & Abt, 2002)."

Reviewer comment:

Line 524: suggest replacing "cause" with "driver"

Response: Thanks for the reviewer's suggestion, we changed from "cause" to "can explain". As our statistical analysis contains multiple collinear factors and it

is hard to isolate any one cause or driver. We rephrased as “The pH buffering capacity of soils varies systematically between the Midwest and the Southeast, which can also explain the spatial variations in lime application rate and treated fraction.”

Technical comments:

Reviewer comment:

Lines 85-87: Suggest rewording because "depends on the pH of the soil" and "Depending on soil pH" sounds repetitive. Also, "through equilibration of the carbonic acid system" tripped me up while reading and I'm not sure that it's necessary here.

Response: We removed the repetitive depend/depending. We condensed the phrase to “The relative contribution of strong acid derived acidity and carbonic acid depends on the pH of the soil (Plummer & Wigley, 1976). Historically, lime application has been assumed to be a net source of CO_2 to the atmosphere, based on the premise that the majority of lime is weathered by strong acids (Robertson et al., 2000; De Klein et al., 2006).”

Reviewer comment:

Line 114: remove the "a" after "high"

Response: We revised and removed the “a” accordingly.

Reviewer comment:

Line 149: change "years" to "year"

Response: We revised accordingly.

Reviewer comment:

Line 338: replace "which" with "with"

Response: We revised the grammar mistakes. We rephrased as “The temporal-spatial pattern of county-level limed area (Fig. 5) approximately resembles that of limed applied mass, with counties that exhibit higher limed areas also concentrated in the Midwest regions.”

Reviewer comment:

Line 400: I would prefer the font size to be uniform across this figure.

Response: We fixed the three panels of Figure 9 to be the same size.

Reviewer comment:

Line 462: missing end parentheses

Response: We fixed this mistake.

Reviewer comment:

Line 535: Please be consistent with capitalizing regions like "Southeastern"

Response: We fixed all the lowercase “south”, “southeastern” into “South”, “Southeastern”. We also fixed “east”, “eastern” to “East”, “Eastern”

Reviewer 2

Thank you very much for the reviewer's detailed review of our work. Your suggestions and concerns have greatly improved the quality of this manuscript. We have carefully considered all your comments and have revised the manuscript based on each individual comment. We have done our best to improve the manuscript and hope that the revised version will meet your expectations. A point-by-point response to your comments is listed below.

Reviewer comment:

This manuscript compiles state-level (1930–1950) and county-level (1954–1987) liming statistics for the United States, linearly interpolates missing census years, and uses Random Forests to identify correlates of lime mass, "rate," and "frequency." While a spatially explicit collation of historical liming could, in principle, be useful, the paper in its current form is methodologically shallow, limited to spatiotemporal gap-filling, and weak in terms of application potential.

There are several fundamental limitations: (1) a serious temporal mismatch between historical lime application data (1978 to 1987) and modern environmental covariates (e.g., SSURGO soils), which undermines the validity of inferred relationships between lime use and soil or climate drivers; (2) The county-level resolution and the historical time span from the 1930s to the 1990s substantially limit the general applicability of the dataset, particularly given that lime is not routinely applied like nutrients (e.g., N and P) and is inherently difficult to predict. As a result, these historical records have limited relevance for understanding or informing current agricultural production.; and (3) an ambitious framing around climate mitigation that is not matched by quantitative assessment, particularly in light of the steep decline in liming since the 1970s. In addition, the manuscript provides limited insight beyond basic interpolation and visualization of known spatial gradients, and the proposed drivers (e.g., corn yield) are not well justified, particularly given the lack of consideration of land-use history. Taken together, the analysis is predominantly descriptive, lacks robustness, and does not yet meet the standards for a data-rich, interpretable, and conceptually well-founded contribution.

Response: Thanks for the reviewer's careful reading of work. We have carefully revised and improved manuscript which we believe is to our best ability to meet and respond to the reviewer's comments. We added an additional figure (Fig. 8) and a new pixel-level data product, reframed the manuscript regarding climate mitigation and RF analysis, and conducted a few analyses according to the reviewer's suggestion. Our response is listed as below.

Major Comments

Reviewer comment:

The central contribution is presented as "the first spatially explicit dataset" of U.S. liming. However, the core analysis consists of (i) linear interpolation and extrapolation of census statistics between reporting years, (ii) county-level summarization of largely static environmental rasters, and (iii) a Random Forest variable-importance analysis. Overall, the conceptual and methodological novelty is modest, and the manuscript currently reads more as a spatial gap-filling exercise than as a substantively new data product or process study.

Response: To address the reviewer's concern of lacking a new data product, we assumed that all the lime is all applied on cropland, and utilizing a temporal varying cropland area dataset, we conducted a transformation of county level data of lime application into pixel-level lime application rate. We added a new section 2.4 describing our work in the methodology, with an additional figure 8 and additional pixel level data product available. This allows potential users of our data to access the dataset much easier and conduct spatial work more directly without further transformation

Reviewer comment:

Temporal scope, interpolation, and data pairing: The final dataset comprises state-level records for 1930–1950 and county-level records for 1954–1987. Missing county values between irregular census years are filled by linear interpolation or extrapolation.

There is a substantial temporal mismatch between lime application data (1950s–1980s) and the environmental covariates used to interpret spatial patterns. Soil properties derived from SSURGO represent relatively recent conditions and are essentially static "one-time" variables, whereas crop yields and climatic variables are treated as time-series. The manuscript does not clearly explain how these differing temporal resolutions are reconciled when identifying "key drivers" of lime application. Using present-day soil data to explain liming patterns several decades earlier is problematic and undermines the validity of inferred relationships between lime use and soil factors. Given these limitations, the resulting dataset primarily reflects historical conditions of lime use. While this has historical value, its relevance for contemporary management and climate mitigation applications is limited and should be framed accordingly.

Response: Thanks for the reviewer's comment concern. The SSURGO soil property data is a composite of data from the past century, averaging everything in the same location, while for the crop and climate data are temporal varying. However, we would like to point out that in our machine learning model, we also averaged the crop and climate data as well as the liming variables across 1977–1987, not the entire dataset dating back to the 1950s. Thus, there is no "time-series" component in the model, only averaged soil, climate, and crop properties, as we are only aiming to understand the spatial variation of lime application. Although the time of the soil properties is different from the average of crop and climate (1978-1987), there should be some overlap of time between the two, thus not an entire temporal mismatch.

Our original sentence "The agricultural data corresponds to the specific years of the lime application records, while the climate, soil, lithology, and pH recommendations lack temporal resolution and are treated as static across all years." may have caused this confusion".

We replaced this with

"We further average the agriculture (crop, fertilizer) and climate data across 1978-1987, corresponding to the last three census years (1977, 1982, and 1987) of lime application data. Yearly varying soil, lithology, and pH recommendations are not available. Therefore, we assume the long-term average is similar with this temporal period, which we acknowledge the slight temporal mismatch is a limitation of this

analysis."

We acknowledge the reviewer's point that the temporal mismatch in the predictors, as well as the fact that liming is a human behavior (not entirely dependent on natural variables) prevents us from exactly drawing strong causal "driver". We reframed our wordings in the manuscript, which we aim to use RF machine learning model mainly as a sanity check/confirmation whether our liming maps are explainable with other relevant predictors. Here our focus is changed to see if known agronomic relationships (eg. Such as the strong link between pH, crop type, and lime use) can potentially help us interpret the pattern, rather than quantitatively assess "what are the causal drivers."

We changed some of the wording in the abstract, changing the framing of the predictors as explanatory rather than causing, or driving the pattern.

"…We used a machine learning framework to assess if the spatial variation of lime application from our dataset can be explained by agricultural and environmental factors with known agronomic relationships. Our results show that the total amount of lime applied in each county, as well as the application rate and treated fraction, can be explained with regional climatic conditions, soil properties, and agricultural management practices (such as crop production, fertilizer use, and soil pH recommendations). "

We changed also some of the wording in the introduction:

"Use machine learning approaches to verify if agricultural and environmental factors can explain the spatial pattern for lime applications. "

We added some further clarification in the Methodologies:

"… Since liming is ultimately a human decision made by farmers, we can only determine the biophysical factors that might lead to higher probability of liming practice, while the actual occurrence of lime application still depends on individual

management strategies. Thus, in the context of our study, the RF model is not intended to extrapolate temporal trends beyond the observed data range or to quantify specific causal drivers. Instead, it is used to assess whether the observed spatial patterns of lime application can be explained by environmental and agricultural factors with well-established agronomic relationships. …"

We also removed words in the manuscript such as "causal driver", "predictors", into more explanatory terms, "relevant factors", "may explain", as our main goal here is not to compare say, if crop yield or soil pH is a more dominant predictor, but if we can explain our liming pattern with these factors. We added and changed the wording of the results as below.

"… Overall, we confirmed that our estimates of intensively limed areas correspond with areas of lower pH, productive cropland with high fertilizer application etc. The average total lime applied in counties can largely be explained by agricultural management practices, including crop production, fertilizer application, and soil pH recommendations, as well as climate. The average application rate and treated fraction can also be explained by regional climatic conditions and soil properties. …."

Reviewer comment:

Climate-mitigation framing and lack of quantitative support: Lines 53–55 state that the study aims to "assess the potential of agricultural lime application as a climate mitigation strategy." In its current form, the manuscript does not provide sufficient evidence or quantitative analysis to support this claim. The Introduction briefly discusses enhanced carbonate weathering and potential CO_2 removal, but there is no explicit evaluation of carbon budgets, the fraction of lime weathered by strong versus carbonic acids, or the net CO_2 impact of liming across regions. Furthermore, national lime application has already declined sharply since its peak in the 1970s, as documented both in this dataset and by West & McBride (2005). Given that lime is not routinely required once soils reach a suitable pH for crop growth, it is not clear that historical liming patterns, especially where current use is low, represent a high-impact pathway for future climate mitigation. The climate-mitigation framing is therefore overly ambitious and should either be substantially toned down or supported by a dedicated geochemical and carbon-accounting analysis.

Response: Thanks for the reviewer's suggestion. The original intended framing of the climate mitigation is for this study to be a useful observational dataset for future works. The carbon accounting and geochemical analysis is outside our original scope in this study. We changed the framing of the climate mitigation aspect and explicitly mention that these are meant for future work.

We toned down the framing of the climate-mitigation in the introduction, changing the ending of our abstract to "…. This study establishes a critical observational baseline of historical agricultural lime application and highlights the need for further research into its long-term environmental impacts."

and changed the ending of the introduction to "… A spatially-explicit dataset of historical lime applications across regions is critical for future studies to further explore the potential of enhanced lime application as a climate solution….Quantifying the net climate mitigation impact of lime application is beyond the scope of this study and would require additional analysis in future work."

Reviewer comment:

Choice and interpretation of drivers (soil properties, crop yield, climate): the manuscript treats soil properties as static variables while crop yield and climate are time-varying. This mismatch raises questions about how the model disentangles spatial from temporal patterns when identifying key drivers. The role of crop yield, especially corn yield, as a major driver of lime application is not convincingly justified. Lime is primarily applied to correct soil acidity, not primarily to boost the yield of a specific crop, unless clear evidence is provided that different crops require systematically different lime rates or have distinct acidity threshold.

Response: Thanks for the reviewer's comment. The SSURGO soil property data is a composite of data from the past century, averaging everything in the same location, while for the crop and climate data are temporal varying. However, we would like to point out that in our machine learning model, we also averaged the crop and climate data as well as the liming variables across 1977–1987, not the entire dataset dating back to the 1950s. Thus, there is no "time-series" component in the model, only temporally averaged soil, climate, and crop properties, as we are only aiming to understand the spatial variation of lime application. Thus, although the time of the soil properties (past century) is different from the average of crops and climate data (1978-1987), there should be some overlap of time between the two, thus not an entire temporal mismatch. Nevertheless, we agree that due to this uncertainty, and the fact that there is no temporal component in the model, prevent us to discern a causal driver. Therefore, we tone down the RF analysis from "identifying key drivers" as a verification/confirmation of our liming data is explainable in relation with other relevant predictors.

Regarding the fact that specific crops emerge as dominant predictors, we agree with the reviewer that lime is applied mainly to correct soil pH, and even if different crops maybe be able to tolerate slightly different threshold, it is not the major reason that corn and soybeans stand out as strong predictors. The two stand out to be the main predictors because of the greater intensity of their farming practices in the U.S, which

more heavy fertilizer input. We agree with the reviewer that in this case it is hard to identify whether corn/soybean production, or fertilizer, or soil pH is the driver of lime application, thus we rephrased all our wording of “driver”, or “causal” in our analysis and tone it down to more explanatory terms (such as “relevant factors”) given that these are all essentially tied to soil pH and farmers decision.

“Since lime is mainly applied to correct soil pH, and not to boost the yield of specific crop type, the higher importance of corn and soybeans in the model of the amount of lime applied (e.g., hay, wheat, cotton, tobacco) (S11-S15), is mainly related to the intensity of associated farming practices. Corn and soybeans constitute a significantly larger portion of the US’s harvested grains (Zulauf et al., 2023) and are often grown with much higher intensity relative to other crop types (Annan et al., 2024). The intensive agriculture, including frequent fertilizer inputs, acidifies the soil, necessitating lime application.”

We changed in the discussion.

“Taken together, our results suggest that while broad patterns such as the association between lime use and corn-soybean agriculture are evident, the factors that influence lime application could be complex, regionally dynamic, and shaped by interacting environmental, agronomic, and economic factors. The lack of sub-county–level lime use data and the absence of crop-specific lime recommendations are limitations of our study. Thus, results from the machine-learning spatial analysis should be interpreted only as an explanatory perspective on large-scale spatial patterns and relationships of lime application, rather than as field-scale drivers of liming decisions.”

Reviewer comment:

The claim that liming has “historically occurred almost exclusively in the eastern half of the United States, which is characterized by more acidic soils” (Lines 135–137) is an oversimplification. In addition to soil acidity, the historical distribution of cropland is a major determinant: agricultural land has been heavily concentrated in the eastern U.S.,

especially prior to the 1960s. Thus, soil pH, cropland extent, and land-use history should be jointly and quantitatively assessed rather than attributing liming primarily to acidity alone. Similarly, the concluding statement that spatial variation in lime application is shaped by "agricultural practices, climate, and soil properties" (Lines 137–138) is too vague. The conclusion should clearly state the specific key predictors identified by the analysis—for example, precipitation and corn yield (though the latter requires further justification)—rather than restating broad categories.

Response: Thanks for the reviewer's comment. We removed these few vague sentences, as we first found it is not appropriate to briefly mention some of the points of the results at the end of the introduction. All these points will be further discussed in the Results & Discussion section.

Reviewer comment:

Predictability of lime use and suitability of metrics and predictors: lime is not a plant nutrient like nitrogen or phosphorus and is not applied in a regular, predictable manner. It is typically used episodically when soil acidity becomes sufficiently severe to limit productivity. This makes liming inherently harder to predict than nutrient application. The manuscript defines a "frequency" metric that is effectively the fraction of cropland area treated with lime in a given year, but then interprets this as a temporal rate (e.g., "once every N years"). The metric should be described and used as lime-treated area (or fraction) rather than as a true application frequency per field. Converting treated fraction into an implied interval assumes uniform and repeated application across the entire cropland base, which is unrealistic given the episodic nature of liming.

Response: Thanks for the reviewer's suggestion. We changed all the 'liming frequency' in the text and figure captions into 'lime-treated fraction'.

To address the non-uniform application practice, we rephrased our original paragraph into "If assuming a uniform application across cropland, this corresponds to an average liming interval of once every 20 years in the Midwest (Oh & Raymond, 2006), versus once every 5 years in the Southeast. However, the actual liming frequency is

difficult to estimate, as lime is often used episodically, when soil acidity becomes sufficiently severe which crop growth becomes limited."

Reviewer comment:

In Section 3.2, the emergence of corn yield as a key predictor of lime use is not adequately explained. Given the strong decline of liming after the 1970s (Figure 9) and the central role of soil acidity, the identified "key factors" may not be truly causal or robust predictors, especially in the absence of explicit soil pH reconstructions or crop-specific liming recommendations. The additional restriction to counties with >10% cropland by area also needs stronger justification in interpreting factor importance.

Response: Thanks for the reviewer's comment. There might be a slight misunderstanding that our statistical analysis is done on a time series basis (We changed the wording of a few of these sentences to prevent confusion). We only conducted the analysis on an average spatial average (1978-1987, thus there is no temporal component as temporal variation is already taken as an average. However, we agree with the reviewer that the "key factors" here are not strictly causal drivers. Thus, we change the framing wordings in the manuscript, which we aim to use RF machine learning model as a sanity check/confirmation if our liming maps are explainable with other relevant predictors. The main point is to test if known agronomic relationships (eg. Such as the strong link between pH, crop type, and lime use) can potentially help us interpret the pattern,

As mentioned in an earlier response, we added some further clarification in the Methodologies:

"… Since liming is ultimately a human decision made by farmers, we can only determine the biophysical factors that might lead to higher probability of liming practice, while the actual occurrence of lime application still depends on individual management strategies. Thus, in the context of our study, the RF model is not intended to extrapolate temporal trends beyond the observed data range or to quantify specific

causal drivers. Instead, it is used to assess whether the observed spatial patterns of lime application can be explained by environmental and agricultural factors with well-established agronomic relationships. …"

In terms of the 10% threshold of cropland area in the analysis, we added a more explicit explanation in the methodologies.

"For the analysis of lime application rate and lime-treated fraction, we excluded counties with negligible agricultural activity to avoid undue influence from areas with little or no farming. Specifically, the analysis was restricted to counties in which cropland accounted for at least 10% of total land area. This threshold was chosen to ensure that the results reflect variation in liming practices among counties with meaningful agricultural activity."

Reviewer comment:

Cropland expansion and total cultivated area across the conterminous U.S. (CONUS) reached their maximum in the 1970s, which largely explains the observed peak in lime application during that period. Once cropland expansion stabilized and soils in many regions were sufficiently ameliorated, lime use declined, as reflected in both this dataset and the trends reported by West & McBride (2005).

Response: Thanks for the reviewer's insight. Our original manuscript did not touch on the reason for the potential decline. Indeed the slowing down of cropland expansion in the 1970s coincide with the peak of lime application. We further analyzed the harvested cropland percentage (Fig. S17). We rephrased our paragraph and added this point in the revised manuscript.

"Based on our extended analysis using the methodology of West & McBride (2005), we show that total annual lime application over the past two decades varied between 10–25 Tg yr^{-1} but declined from its peak in the 1970s. A potential explanation for this is that cropland area expansion and harvested cropland percentage reached their maximum in the 1970s (Fig. S16-S17), which largely explains the observed peak in lime application during that period. Once cropland expansion stabilized and soil in

many regions were sufficiently ameliorated, lime use declined."

Reviewer comment:

In the Discussion (e.g., Lines 638–639), the manuscript appears to suggest that lime and fertilizer follow similar trajectories. However, the figures indicate that fertilizer use remains relatively stable while lime application decreases substantially after the 1980s. This inconsistency should be acknowledged and discussed explicitly. Lime use is arguably a more meaningful indicator of long-term soil management and pH adjustment than fertilizer use, and the contrasting dynamics of these two inputs deserve clearer treatment.

Response: Thanks for the reviewer's comment. We agree with the reviewer that lime and fertilizer application does not have to follow the same trajectory, as the correction of soil acidity is also dependent on other soil acidification sources (such as atmospheric acid deposition). We mentioned the trend of fertilizers because West & Mcbride (2005)'s methodology has quite a bit of uncertainty, which we mentioned in the manuscript, thus it is hard to estimate if liming went up or down in the recent decades. The fact that fertilizers also did not increase in the recent decades is evidence that liming likely did not increase during this period.

We added a few sentences to clarify this.

"We acknowledge that due to the uncertainty associated with the methodology of West & McBride (2005), it is not entirely certain whether the total lime application decreased after our dataset ended in 1987. However, the total nutrient inputs via fertilizers have not seen any significant increase in nitrogen and phosphorus inputs since 1980 (Fig. S25). Although liming and fertilizer application do not always follow the same trajectory, a significant portion of lime is used to reduce soil acidity resulting from fertilizer use. Therefore, trends in fertilizer application provide some evidence that agricultural lime application likely did not increase after our dataset ended in 1987. As such, we believe the temporal range of our dataset already captures the period during which lime application reached its full extent in the United States."

Additional Analyses Suggestions:

Reviewer comment:

1. Clarify the “frequency” metric.

Response: Thanks for the reviewer’s suggestion. We changed all the original ‘liming frequency’ in the text and figure captions into ‘lime-treated fraction’. In the revised version of the manuscript, we only mentioned frequency term here, “If assuming an uniform application across cropland, this corresponds to an average liming interval of once every 20 years in the Midwest (Oh & Raymond, 2006), versus once every 5 years in the Southeast. However, the actual liming frequency is difficult to estimate, as lime is often used episodically, when soil acidity becomes sufficiently severe that limits productivity.”

Reviewer comment:

2.Crop-specific and grid-level liming (if possible).

If crop-specific lime application rates and treated areas are available, combining them with pixel-level land-use datasets would enable estimation of grid-level liming rates or amounts. This would add an important spatial dimension and significantly enhance the dataset’s utility.

Response: Thanks for the reviewer’s suggestion. However, beyond the county level scope, the crop specific lime application rates and treated areas are not available from the original source which we collected the data. We assumed an equal application rate over cropland within the same county, and combined our county level data pixel-level cropland datasets and derived a pixel level dataset of lime application. As mentioned earlier in the response, we added a new figure (Fig. 8), a new section in the methodology (2.4 Rasterization to pixel-level dataset), and uploaded this pixel-level data to Zenodo.

Reviewer comment:

3.Link liming trends to land-use change.

In the southeastern U.S., declines in lime application appear to coincide with reductions in annual limed area (Fig. 5) during the 1980s. It would be valuable to explicitly examine whether these patterns reflect land-use transitions (e.g., cropland abandonment or shifts to less acid-sensitive systems) and to discuss this in the context of long-term agricultural change.

Response: Thanks for the reviewer's suggestion. We conducted further analysis on this part, examining the percentage of cropland that is harvested, and the percentage that is not harvested nor pastured, showing that it is indeed likely that cropland abandonment has taken place, and the region started to transition to other land use. We added Fig. S16, S17 which shows this phenomenon. We also added and rephrased the paragraph as below.

"However, lime application and limed area along the Southeastern coast both shows a decline in the 1980s, which coincides with a decrease in crop production, particularly corn and soybean (Fig. S9-S15). The relative ratio of total cropland in this region did not show a clear decline (Fig. S16). However, the percentage of cropland harvested (Fig. S17) shows a strong decrease in 1987, while the percentage of cropland that is neither harvested nor used as pasture (Fig. S18) shows a clear increase, potentially indicating gradual abandonment of cropland for other uses. The decrease in crop production in this region could be an onset of urbanization and conservation programs (Napton et al., 2010). Additionally, the growth of the regional lumber industry during this period may have further slightly reduced agriculture, contributing to the observed decline in lime application (Prestemon & Abt, 2002)."

Reviewer comment:

4.Potential extension: estimating current liming needs.

If soil acidity data or crop-specific soil pH thresholds could be compiled, there may be an opportunity to estimate contemporary lime requirements to sustain current crop production. Such an analysis would provide more direct and policy-relevant implications than focusing solely on historical liming patterns.

Response: We agree with the reviewer that an estimate of liming requirements for contemporary studies would be very helpful and provide more direct implications for agriculture. However, we admit that this is beyond our capacity and ability in this study, since we don't have crop-specific pH thresholds to determine this. We conduct a simple comparison between soil pH data and the pH recommendation of each county for the most common crop, and it shows the Southeastern U.S. may have a potential to be further limed. However, due to the time range of the soil pH data (composite over past century), further work is required for a robust assessment of this phenomenon. We briefly touched on some of these points in the discussion section.

"A simple comparison between pH recommendations and averaged soil pH data (Fig. S26) shows that, in many regions of the Southeastern U.S., soil pH is below the pH recommendation for achieve optimal crop yields. As such soil acidity could potentially be further corrected through more lime application. However, the lack of temporally explicit and crop-specific pH thresholds data prevents further quantification of this phenomenon. Further work is required to assess if policy should encourage additional lime applications."

Minor Comments

Reviewer comment:

1. Line 215: The reference to "Fig. 4" appears out of order. Please ensure that figures are cited in the correct numerical sequence throughout the manuscript.

Response: We removed the "Fig. 4" reference as it is not particularly helpful in this context.

Reviewer comment:

2.Units: Line 296: Use a consistent standard format for units, e.g. "t yr^{-1}" rather than "t/yr," and apply this convention uniformly.

Response: Thanks for the reviewer's reminder. We revised and fixed all of this mistake

throughout the manuscript.

Reviewer comment:

Lines 315–316: The text states "exceeding 6 Mt lime year^{-1}," but Figure 3 appears to show a maximum below 6 Mt. This discrepancy should be corrected. Also, choose either "yr^{-1}" or "/yr" and use it consistently to avoid confusion.

Response: Thanks for the reviewer's reminder. We changed "exceeding 6 Mt lime year^{-1}," to "close to 6 Mt lime year^{-1}". We revised the "year^{-1}" to "yr^{-1}"

Reviewer comment:

3.Green Revolution explanation

Lines 310–313: The link between the Green Revolution and increased lime use is not clearly articulated. As written, it does not convincingly explain why Green Revolution technologies would directly drive liming. This section should be clarified or revised.

Response: Thanks for the reviewer's comment. We added more explanation of the indirect like through fertilizer addition and soil acidification.

"A rapid increase in application occurred between 1930 and 1950, likely corresponding to the onset of the Green Revolution in the 1940s, which introduced synthetic fertilizers and modern agricultural technologies (Smil, 2004). The increased use of nitrogen fertilizers likely leads to enhanced soil acidification pressures and the need for pH management, liming is a necessary and more common practice."

Reviewer comment:

4.Assumption of uniform long-term liming

Response: We replaced all the 'liming frequency' that appeared in the text and figures with 'lime-treated fraction', thus the values do not require an assumption of uniform long-term liming. However, we kept a rough calculation of the approximate time scale if lime is applied uniformly, to compare with previous works that mentioned liming interval in the Midwest is 20 years (Oh & Raymond, 2006). We rephrased this section

and explicitly mentioned the difficulty and challenges with this assumption.

“If assuming an uniform application across cropland, this corresponds to an average liming interval of once every 20 years in the Midwest (Oh & Raymond, 2006), versus once every 5 years in the Southeast. However, the actual liming frequency is difficult to estimate, as lime is often used episodically, when soil acidity becomes sufficiently severe, which crop growth becomes limited.”

Reviewer comment:

5.Lines 382–385: The large variation in lime use is likely more strongly related to underlying soil pH patterns and historical land expansion than can be resolved at the county level with the current dataset. This limitation should be acknowledged more clearly.

Response: Thanks for the reviewer’s comments. We changed the phrase and explicitly mentioned these points.

“This contrasting spatial pattern in both Fig. 6 (lime application rate) and Fig. 7 (lime-treated fraction) highlights the spatial differences in lime management, which are largely influenced by underlying soil pH distributions driven by climate and soil properties. We also acknowledge that the large spatial heterogeneity in lime use is influenced by finer-scale factors and historical agricultural land expansion that cannot be fully resolved with the current county-level dataset. As a result, the observed patterns should be interpreted as regional signals rather than precise representations of field-scale liming decisions.”